Computational Analysis of Fundus Images
Rule-Based and Scale-Space Models

Ivo Miguel da Fonseca Gravito Soares

Tese para obtenção do Grau de Doutor em
Biomedicina
(3º ciclo de estudos)

Orientador: Prof. Doutor António Manuel Gonçalves Pinheiro
Coorientador: Prof. Doutor Miguel Castelo-Branco Craveiro Sousa

Júri:
Prof. Doutor José Ignácio Verde Lusquinos
Prof. Doutora Ana Maria de Sousa Faria de Mendonça
Prof. Doutor Luís Alberto da Silva Cruz
Prof. Doutor António Manuel Gonçalves Pinheiro
Prof. Doutor Hugo Pedro Martins Carriço
Prof. Doutor João Miguel Raposo Sanches
Prof. Doutora Maria Manuela Areias da Costa Pereira de Sousa
Prof. Doutor Luís Miguel da Luz Caixinha Duarte

Julho de 2022

Acknowledgements

This thesis would not be possible without the help of many people, to whom I am sincerely grateful.

First of all, I gratefully acknowledge the support and guidance from my supervisor, Professor António Manuel Gonçalves Pinheiro. All the discussions we have had throughout these years have contributed to my scientific and personal growth. I learned from him not only the knowledge, but also critical thinking and research skills. His influence, encouragement, and motivation go far beyond the scope of this thesis.

Second, I would like to thank my co-supervisor, Professor Miguel Castelo-Branco Craveiro de Sousa, for the support whenever it was needed.

Third, I express my deepest thanks to my parents and sisters for their support and motivation through the last years.

Least but not last, I would like to thank my wife Patrícia and my son Vasco. Without their support, patience and encouragement throughout these last years, this thesis would not be possible.

Prefácio

”A picture may be worth a thousand words, a formula is worth a thousand pictures.”
Edsger W. Dijkstra

List of Publications

Articles included in this thesis resulting from this doctoral research program

Journal Articles

1. Soares I, Castelo-Branco M, Pinheiro AMG.
   Optic Disc Localization in Retinal Images Based on Cumulative Sum Fields. IEEE J Biomed Health Inform. 2016 Mar;20(2):574-85.
   doi: 10.1109/JBHI.2015.2392712

2. Soares I, Castelo-Branco M, Pinheiro AMG.
   A Multi-scale Based Approach for Microaneurysms Detection in Fundus Images (Submited)

International Conferences

1. Soares I, Castelo-Branco M, Pinheiro AMG.
   A new vessel enhancement transform on retinal blood vessels segmentation. In Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (EG VCBM 18). 2018, 123–127.
   https://doi.org/10.2312/vcbm.20181237

2. Soares I, Castelo-Branco M, Pinheiro AMG.
   Vessel centerline detection in retinal images based on a corner detector and dynamic thresholding, 22nd European Signal Processing Conference (EUSIPCO), 2014, pp. 2020-2024.

3. Soares I, Castelo-Branco M, Pinheiro AMG. Microaneurysms detection using a novel neighborhood analysis. Proceedings of the Ophthalmic Medical Image Analysis First International Workshop, Boston, MA, USA: University of Iowa; 2014, p. 65–72.
   https://doi.org/10.17077/omia.1010.

4. Soares I, Castelo-Branco M, Pinheiro AMG.
   Curvature detection and segmentation of retinal exudates, 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), 2012, pp. 1719-1722.
   doi: 10.1109/ISBI.2012.6235911.

5. Soares I, Castelo-Branco M, Pinheiro AMG.
   Exudates dynamic detection in retinal fundus images based on the noise map distribution, 19th European Signal Processing Conference, 2011, pp. 46-50.

6. Soares I, Castelo-Branco M, Pinheiro AMG.
   Scale-space curvature detection of retinal exudates with a dynamic threshold, 7th International Symposium on Image and Signal Processing and Analysis (ISPA), 2011, pp. 523-528.

Resumo

As imagens do fundo do olho são hoje um dos principais exames imagiológicos da oftalmologia moderna, pela sua simplicidade, baixo custo e acima de tudo pelo seu carácter não-invasivo. A aquisição e armazenamento de imagens do fundo do olho com alta resolução é também relativamente simples e rápida. Desta forma, as imagens do fundo do olho são um exame fundamental na identificação de alterações retinianas, monitorização da saúde ocular, e em programas de rastreio. Considerando o elevado volume e complexidade clínica associada a estas imagens, a análise e interpretação das mesmas por clínicos treinados torna-se uma tarefa morosa e propensa a erros humanos. Assim, há um interesse crescente no desenvolvimento de abordagens automatizadas, acessíveis em custo, e com uma alta sensibilidade e especificidade. Estas devem ser robustas para serem aplicadas à população em geral no diagnóstico e seguimento de doenças retinianas. Para serem eficazes, os sistemas de análise têm que conseguir detetar e distinguir estruturas normais de sinais patológicos.

O objetivo principal da investigação que levou a esta tese de doutoramento é o desenvolvimento de sistemas automáticos capazes de detetar e segmentar as estruturas anatómicas da retina, e os sinais patológicos retinianos associados às doenças retinianas mais comuns. Em particular, estes algoritmos automatizados foram desenvolvidos segundo as premissas de robustez e eficácia para lidar com as dificuldades e complexidades inerentes a estas imagens.

Foram considerados quatro objetivos de análise de imagens do fundo do olho. São estes, a segmentação de exsudados, a localização do disco ótico, a deteção da linha central venosa dos vasos sanguíneos e segmentação da rede vascular, e a deteção de microaneurismas. De acrescentar que usando o método de deteção de microaneurismas, avaliou-se também a capacidade de deteção da retinopatia diabética em imagens do fundo do olho.

Para comparar o desempenho das metodologias desenvolvidas neste trabalho, foi realizado um levantamento do estado da arte, onde foram considerados os métodos mais relevantes descritos na literatura para cada um dos objetivos descritos anteriormente. Para facilitar a comparação entre métodos, o estado da arte foi dividido em metodologias de processamento de imagem e baseadas em aprendizagem máquina.

Optou-se no trabalho de investigação desenvolvido pela utilização de metodologias de análise espacial de imagem em detrimento de metodologias baseadas em aprendizagem máquina. Em particular, as metodologias baseadas no espaço de escalas mostraram ser efetivas na obtenção dos objetivos estabelecidos.

Para a segmentação de exsudados foram usadas duas abordagens distintas. A primeira abordagem baseia-se na curvatura em espaço de escalas em conjunto com a resposta máxima local de um detetor de manchas em espaço de escalas e limiares dinâmicos. A segunda abordagem baseia-se na análise do mapa de distribuição de ruído em conjunto com operadores morfológicos e limiares adaptativos. Ambas as abordagens fazem uma segmentação dos exsudados de elevada precisão, além de lidarem eficazmente com a iluminação não-uniforme e a variação de contraste presente nas imagens do fundo do olho. A localização do disco ótico foi conseguida com uma nova técnica designada por campos de soma acumulativos, combinada com métodos de melhoramento da rede vascular. O algoritmo revela ser fiável e eficiente, particularmente em imagens patológicas. A robustez do método foi verificada pela sua avaliação em oito bases de dados. A deteção da linha central dos vasos sanguíneos foi obtida através de um detetor de cantos modificado em conjunto com filtros binários e limiares dinâmicos. A segmentação da rede vascular foi conseguida com um novo método de melhoramento de vasos sanguíneos em espaço de escalas. Os métodos desenvolvidos

mostraram ser eficazes na deteção da linha central dos vasos sanguíneos e na segmentação da rede vascular. Finalmente, o método para a deteção de microaneurismas assenta num formalismo de espaço de escalas na deteção e na rotulagem dos microaneurismas. Para a rotulagem foi utilizada uma nova abordagem da vizinhança dos candidatos a microaneurismas. A deteção de microaneurismas permitiu avaliar também a deteção da retinopatia diabética. O método para a deteção de microaneurismas mostrou ser competitivo quando comparado com outros métodos, em particular em imagens de alta resolução. A deteção da retinopatia diabética exibiu um desempenho semelhante a outros métodos e a especialistas humanos.

Os trabalhos descritos nesta tese mostram ser possível desenvolver uma abordagem fiável e robusta em espaço de escalas capaz de detetar diferentes estruturas anatómicas e sinais patológicos da retina.

Além disso, os resultados obtidos mostram que apesar de a pesquisa mais recente concentrar-se em metodologias de aprendizagem máquina, as metodologias de análise espacial apresentam resultados muito competitivos e tipicamente independentes do equipamento de aquisição das imagens. As metodologias desenvolvidas nesta tese podem ser importantes na definição de novos descritores e características, que podem melhorar significativamente o resultado de métodos automatizados.

Palavras-chave

Retina; Imagem do fundo do olho; Análise de imagens retinianas; Espaço de escalas; Curvatura; Segmentação; Deteção de características; Doenças retinianas; Retinopatia diabética; Exsudados; Disco ótico; Rede vascular; Linha central vascular; Microaneurismas;

Resumo alargado

Introdução

Esta tese visa o desenvolvimento de algoritmos computacionais que consigam executar a deteção de estruturas retinianas e de sinais patológicos retinianos. Duas estruturas retinianas diferentes vão ser estudadas, nomeadamente a deteção da rede vascular e do disco ótico. Os sinais patológicos estudados são a deteção de exsudados e de microaneurismas.

Motivação e descrição breve do estado da arte

O desenvolvimento de novas tecnologias em processamento e aquisição de imagens, levou a um aumento das suas potenciais aplicações. Sem surpresa, a especialidade médica de oftalmologia tem recebido uma atenção alargada e intensa pela comunidade científica para desenvolver abordagens computacionais automatizadas para o processamento de imagens do fundo do olho. Com a capacidade de analisar um número elevado de imagens num curto período de tempo, estes sistemas automatizados poderão ser particularmente úteis na deteção e diagnóstico de doenças retinianas, e em atividades de rastreio [1].

Existe uma série de doenças oculares e sistémicas que podem ter manifestações patológicas clínicas na retina. Devido à sua prevalência e incidência, as doenças mais estudadas são a retinopatia diabética, glaucoma, retinopatia hipertensiva e a degenerescência macular da idade. Estas doenças têm vários estágios de desenvolvimento caracterizados por alterações retinianas específicas [2].

Existe uma série de dificuldades com que os sistemas automatizados têm que lidar antes que possam ser usados na prática clínica. Em imagens retinianas normais, a análise automatizada e deteção das estruturas retinianas requer algoritmos robustos. Em imagens retinianas patológicas, as estruturas e os sinais patológicos da retina têm uma grande heterogeneidade em tamanho, cor, extensão e forma [3]. Além disso, as imagens da retina também apresentam uma grande intra e inter-variabilidade. A intra-variabilidade das imagens deve-se sobretudo à presença de opacidades, diferentes níveis de refletividade e espessura retiniana variável. A inter-variabilidade imagem deve-se a diferenças entre os mecanismos de aquisição de imagem, iluminação, protocolos de aquisição, e pigmentação da retina [4]. Estas variabilidades de imagem inerentes a todas as imagens retinianas, em conjunto com a heterogeneidade dos sinais patológicos, tornam o desenvolvimento de algoritmos computacionais automatizados uma tarefa complexa e desafiante.

As descrições feitas anteriormente implicam que a deteção e segmentação dos sinais patológicos por sistemas automatizados, devem ser robustas, eficientes e precisas. Desta forma consegue-se determinar um diagnóstico ou seguimento correto assim como permitir uma deteção precoce e terapia eficaz. Nos últimos anos foi publicado um número considerável de trabalhos sobre a deteção automática de estruturas e de sinais patológicos da retina.

No capítulo 3 é descrita uma revisão detalhada das contribuições mais relevantes sobre os tópicos de análise abordados nesta tese. Alguns destes métodos abordam também a deteção automatizada de doenças da retina como retinopatia diabética [5, 6, 7], a degenerescência macular da idade [8, 9], o glaucoma [10, 11] e retinopatia hipertensiva [12]. De maneira geral os métodos abordados no capítulo 3 estão em divididos metodologias de processamento de imagem e aprendizagem máquina, onde se incluem as recentes propostas de uso de metodologias de aprendizagem profunda.

Os primeiros métodos de deteção de estruturas retinianas e sinais patológicos eram baseados em

metodologias de processamento de imagem. Esta categoria de métodos não requerem nenhum mecanismo de aprendizagem, embora seja necessária a sua calibração através de exemplos. Apesar de estas metodologias apresentarem um bom desempenho, podem apresentar dificuldades em superar as complexidades e heterogeneidades inerentes às imagens do fundo do olho normal e patológico [13]. No entanto, muitas características e descritores importantes baseiam-se em metodologias de processamento de imagem.

Para superar as dificuldades dos métodos baseados em processamento de imagem, foram propostas as abordagens baseadas em aprendizagem máquina. Esta categoria de metodologias requer um conjunto de características ou descritores medidos diretamente a partir das imagens do fundo do olho. Em seguida, é realizada uma etapa de aprendizagem que utiliza estas características ou descritores para treinar um determinado classificador. Após a etapa de aprendizagem, o classificador aprende a reconhecer as estruturas da retina e os sinais patológicos em imagens não analisadas anteriormente. Estas metodologias necessitam de um grande número de imagens do fundo do olho [14], sendo normalmente necessário uma nova etapa de treino para uma nova base de dados de imagens (normalmente obtidas com diferentes dispositivos de aquisição), limitando a sua generalização [15].

Recentemente, foram propostas abordagens baseadas em redes neuronais profundas, comummente designadas por aprendizagem profunda [16, 17, 18]. Estas metodologias conseguem realizar ambas as tarefas de aprendizagem das características para uma dada classificação ou deteção, e a classificação final. Uma vantagem dos métodos de aprendizagem profunda é que, com suficientes imagens do fundo do olho e poder computacional, facilmente se supera o desempenho dos métodos tradicionais de aprendizagem máquina [19]. Os principais desafios na aplicação de métodos de aprendizagem profunda são o número limitado de imagens de fundo do olho anotadas [20], e uma distribuição desequilibrada de imagens de diferentes classes por base de dados [21]. Estas abordagens demoram muito tempo a elaborar, requerem um hardware especializado, um grande número de imagens [22, 23, 24, 20], e necessitam de um novo treino para diferentes bases de dados. Estas limitações comprometem o seu desempenho em múltiplas bases de dados e a sua generalização em prática clínica [20, 25].

De acrescentar ainda que, as bases de dados de imagens do fundo do olho apresentam uma considerável ambiguidade na correta identificação do número, tamanho e anotação dos sinais patológicos. Isto deve-se sobretudo às heterogeneidades dos sinais patológicos. Estas dificuldades comprometem a eficiência das metodologias de aprendizagem máquina [24]. Além disso, a má qualidade das imagens e as características heterogéneas das estruturas da retina e dos sinais patológicos, geram um desequilíbrio de classes nas imagens disponíveis na etapa de aprendizagem [26, 21].

Apesar da quantidade e diversidade dos métodos propostos para realizar a análise automática de imagens do fundo do olho, esta área de investigação é muito ativa devido ao seu potencial impacto e à necessidade de alcançar resultados mais exatos [27].

Objetivos de investigação e descrição do problema

Esta tese visa o desenvolvimento de sistemas computacionais automatizados capazes de detetar e segmentar as estruturas anatómicas da retina, assim como os sinais patológicos associados às doenças mais comuns da retina. Em particular, estes algoritmos automatizados foram desenvolvidos sob as premissas de robustez e eficiência, para lidar com as dificuldades e complexidades inerentes a estas imagens.

Foram considerados quatro objetivos para a análise de imagens do fundo do olho. Estes são, a segmentação dos exsudados, a localização do disco ótico, a deteção da linha central dos vasos

retinianos e a segmentação da rede vascular, e finalmente a deteção dos microaneurismas. Foi escolhida a utilização de metodologias de análise de imagem espacial, baseadas numa estrutura de espaço de escalas em vez de metodologias baseadas na aprendizagem máquina. A razão para escolher metodologias baseadas numa estrutura de espaço de escalas [28, 29] deve-se ao elevado grau de heterogeneidade em tamanho, forma, cor e extensão, das estruturas anatómicas e sinais patológicas da retina [30]. Estas características tornam-nas ideais para serem detetadas e analisadas utilizando uma abordagem baseada em espaço de escalas.

A escala é normalmente utilizada num contexto de extração de regiões de candidatos [31, 32, 33] ou para obter características em multi-escala [34, 35, 36, 37, 6] Outros trabalhos utilizam a escala combinada com alguma outra técnica de processamento de imagem para extrair as estruturas da retina.

Os trabalhos de [27, 38, 39] utilizaram um conjunto de múltiplas escalas de filtros bidimensionais de Gabor para sobressair e segmentar os vasos retinianos. Um conjunto de filtros de múltiplas escalas Gaussianos foi utilizado em [40] para efetuar a deteção do disco ótico e da fóvea, e em [41] para efetuar a deteção de microaneurismas. Normalmente as escalas são definidas empiricamente, o que limita a aplicabilidade e generalização das características e descritores. Foram também propostas abordagens em espaço de escalas para extrair a rede vascular retiniana [39, 38]. No entanto, não existem estudos para a generalização de uma mesma abordagem em espaço de escalas para a deteção de várias estruturas retinianas.

Nesta tese foi feito um esforço no desenvolvimento de abordagens baseadas em espaço de escalas que sejam robustas e fiáveis na deteção de diferentes estruturas retinianas e sinais patológicos da retina. Além disso, a definição de um conjunto mais generalizável de escalas pode fornecer informação importante no desenvolvimento de novos descritores e características, para melhorar o desempenho dos métodos de aprendizagem.

Bases de dados e avaliação de desempenho

Nesta tese foram consideradas dez bases de dados de imagens do fundo do olho descritas no Anexo A. São estas a Stare (STructured Analysis of the REtina) [42], Drive (Digital Retinal Images for Vessel Extraction) [43], Diaretdb0 (Standard Diabetic Retinopathy Dataset Calibration level 0) [44], Diaretdbv1 (Standard Diabetic Retinopathy Dataset Calibration Level 1) [45], Messidor (Méthodes d'Evaluation de Systèmes de Segmentation et d'Indexation Dédiées à l'Ophtalmologie Rétinienne) [46], Latim (Laboratoire de Traitement de l'Information Medicale) [47], ROC (Retinopathy Online Challenge) [48], E-ophtha-MA [49], E-ophtha-EX [49] e HRF (High-Resolution Fundus) [50].

As métricas mais usadas para a avaliação do desempenho dos métodos descritos na literatura são a Sensibilidade, Especificidade e a Precisão [51, 52].

A Sensibilidade (definida na equação 3.1) é o rácio de pixeis ou regiões pertencentes a lesões, ou estruturas retinianas classificados corretamente. A Especificidade (definida na equação 3.2) é o rácio de pixeis ou regiões não pertencentes a lesões, ou estruturas retinianas classificados corretamente. A Precisão (definida na equação 3.3) é proporção das verdadeiras deteções considerando o número total de pixeis ou regiões pertencentes a lesões, ou estruturas retinianas. A Sensibilidade, Especificidade e Precisão podem ser calculados segundo um critério baseado na deteção pixeis, lesão ou imagem.

Outra métrica comum de desempenho é a curva *Receiver operating Characteristic*. Esta curva relaciona a Sensibilidade com a Especificidade. Com base neste gráfico determina-se a *area under the curve* (AUC) que varia no intervalo de [0,1], onde a unidade representa a situação ideal onde

todos os pixeis ou regiões foram corretamente identificados sem falsas deteções.
A avaliação de *desempenho* da deteção de microaneurismas é feita em termos de uma curva *free-response characteristic operating* (F_ROC) onde a Sensibilidade (verdadeiros positivos vs. o número de positivos) é comparada ao número de falsos positivos por imagem (FP/I) [48]. Esta métrica varia no intervalo [0,1] sendo a unidade o valor ótimo.

Principais contribuições

Em seguida é feita uma descrição das metodologias desenvolvidas segundo as estruturas anatómicas e sinais patológicos da retina estudados nesta tese.

Segmentação de exsudados retinianos

Para a segmentação de exsudados foram elaborados duas abordagens distintas.
A primeira abordagem baseia-se na curvatura em espaço de escalas. Primeiramente é efetuada a deteção de possíveis localizações de exsudados utilizando a curvatura extrínseca em espaço de escalas [53]. Em seguida é efetuada a segmentação de regiões utilizando a curvatura média em espaço de escalas. Apenas são considerados verdadeiros exsudados as regiões que também tenham sido detetadas no cálculo da curvatura extrínseca em conjunto com a resposta máxima local de um detetor de manchas em espaço de escalas. Todos os limiares são determinados de forma dinâmica. O método foi avaliado na base de dados Diaretdbv1 atingindo uma Sensibilidade de 97.07%, uma Especificidade de 99.90% e uma Precisão de 99.83% [54].
A segunda abordagem baseia-se na análise do mapa de distribuição de ruído. Inicialmente calcula-se uma imagem binária dos máximos locais da imagem original do fundo do olho. Os máximos locais seguem uma distribuição semelhante ao ruído presente nas imagens. No entanto, em algumas zonas observam-se padrões de distribuição diferentes. As zonas com maior dispersão de máximos pertencem a estruturas com valores de intensidade mais elevados (exsudados e disco ótico) ou estruturas com valores de intensidade mais baixos (rede vascular ou hemorragias). Em seguida, a imagem do fundo do olho é dividida em regiões concêntricas. Estas regiões seguem um padrão de iluminação típico destas imagens. Para cada uma destas regiões concêntricas consideram-se apenas os pixeis que têm uma intensidade acima de um determinado limiar. O cálculo do limiar considera apenas os pixeis locais dessa mesma região local sendo feito de forma adaptativa. Os pixeis selecionados permitem segmentar os exsudados. O método foi avaliado na base de dados Diaretdbv1 atingindo uma Sensibilidade de 97.49%, uma Especificidade de 99.95% e uma Precisão de 99.91% [54].
Ambas as abordagens fazem uma correta segmentação dos exsudados, além de lidarem eficazmente com a iluminação não-uniforme e a variação de contraste presente nas imagens do fundo do olho.

Localização do disco ótico

O método desenvolvido começa com uma nova técnica de melhoramento da rede vascular baseada num detetor de cantos modificado. Em seguida o resultado do melhoramento da rede vascular é combinada com operadores morfológicos, para detetar as quatro orientações principais dos vasos { 0^o, 45^o, 90^o, 135^o}. Estas quatro imagens têm toda a informação necessária para determinar uma localização inicial do disco ótico sendo combinadas para produzirem duas imagens, designadas de imagem vertical e imagem horizontal. As imagens verticais e horizontais permitem determinar a

coordenada vertical e horizontal do disco ótico. Estas imagens vertical e horizontal são divididas em regiões horizontais e verticais respetivamente. Para cada divisão é obtido o valor médio da intensidade. Repetindo o processo para regiões cada vez mais pequenas e somando as respostas obtém-se um campo de soma acumulativo para a imagem vertical e horizontal. Estes campos acumulativos têm regiões de mínimos que representam as coordenadas iniciais do disco ótico. A localização final do disco ótico é refinada através de um algoritmo de convergência usando as características de convergência vascular e valores de intensidade. O método proposto foi avaliado em oito bases de dados publicamente disponíveis, incluindo a Stare e a Drive. O disco ótico foi localizado corretamente em 1752 imagens de um total de 1767 (99,15%). O algoritmo revela ser fiável e eficiente, particularmente em imagens patológicas [55].

Deteção da rede vascular

Foram considerados dois objetivos distintos, nomeadamente a deteção da linha central dos vasos sanguíneos e a segmentação da rede vascular.

A deteção da linha central começa com a descrição de uma novo detetor de vasos sanguíneos baseado num detetor de cantos modificado. O resultado do detetor de vasos é processado com filtros binários em várias orientações, resultando no melhoramento das estruturas vasculares. As linhas centrais dos vasos principais são em seguida detetados através de um limiar dinâmico. Para lidar com bifurcações e cruzamentos não detetados, a imagem original do fundo do olho é processada com quatro operadores diferenciais. O resultado destes operadores diferenciais é adicionado às linhas centrais previamente calculadas, resultando na deteção das linhas centrais dos vasos sanguíneos. Para melhorar a robustez, o método foi implementado seguindo uma abordagem em espaço de escalas. O método foi avaliado nas bases de dados Drive e Stare. Na base de dados Drive obteve-se uma Sensibilidade, Especificidade, Precisão e uma AUC de 79.70%, 99.96%, 98.07% e 90.12% respetivamente [56].

A segmentação da rede vascular começa com um método de melhoramento da rede vascular através da resposta máxima acumulativa em espaço de escalas da curvatura principal. A segmentação é obtida calculando dois limiares através de uma técnica de histerese. O resultado da segmentação é aprimorado com a remoção de zonas ao redor do disco ótico pertencentes a uma atrofia peripapilar. Esta remoção é obtida com a implementação de um novo método baseado na magnitude do gradiente em espaço de escalas. O método foi avaliado nas bases de dados, Drive, Stare e HRF. Na base de dados Drive obteve-se uma Sensibilidade, Especificidade, Precisão e uma AUC de 70.44%, 99.97%, 94.16% e 83.98% respetivamente. Na base de dados Stare obteve-se uma Sensibilidade, Especificidade, Precisão e uma AUC de 75.08%, 99.96%, 94.23% e 86.45% respetivamente. Finalmente, na base de dados HRF foi conseguida uma Sensibilidade, Especificidade, Precisão e uma AUC de 80.30%, 99.96%, 94.81% e 88.85% respetivamente [57].

Deteção de microaneurismas retinianos

O método de deteção de microaneurismas começa com a extração de candidatos usando um método de melhoramento em espaço de escalas. Este melhoramento consegue-se por uma transformação do Laplaciano. Em seguida, a rede vascular é removida utilizando uma metodologia baseada na curvatura média em espaço de escalas. A rotulagem dos candidatos é feita com uma técnica em espaço de escalas de análise de vizinhança. Excetuando a etapa de pré-processamento, todo o método foi desenvolvido segundo uma metodologia em espaço de escalas. É também feita uma análise de quais são as melhores escalas individuais e a qual é a melhor combinação de escalas

para a deteção de microaneurismas. O método foi avaliado na deteção de microaneurismas em três bases de dados, nomeadamente na E-ophtha-MA, ROC e Latim com um F_{score} de 0.443, 0.227 e 0.207 respetivamente. Na deteção da retinopatia diabética alcançou-se uma AUC de 0.874 com uma Sensibilidade de 0.936. A metodologia apresentada mostrou ser competitiva quando comparado com outros métodos, em particular em imagens de alta resolução.
Usando o método desenvolvido para a deteção de microaneurismas procede-se também à avaliação da possibilidade da sua utilização visando detetar imagens do fundo do olho que revelem retinopatia diabética. A deteção da retinopatia diabética exibiu um desempenho semelhante a outros métodos assim como a especialistas humanos.

Conclusões

Esta tese descreve os estudos desenvolvidos sobre o desenvolvimento de metodologias computacionais para a análise de imagens do fundo do olho.
Foram desenvolvidos dois métodos diferentes para a segmentação dos exsudados. O primeiro método segmenta os exsudados utilizando uma metodologia baseada em espaço de escalas. O segundo método é baseado no mapa de distribuição de ruído. Ambos os métodos apresentam um bom desempenho, com resultados semelhantes a outros métodos publicados.
Foi desenvolvido um método baseado numa nova técnica denominada campos de soma acumulativa para a deteção do disco ótico. Este método foi testado em oito bases de dados com resultados semelhantes aos melhores métodos publicados anteriormente. Contudo, estes métodos foram avaliados num menor número de bases de dados [58, 59, 60, 61, 62].
A deteção da linha central dos vasos sanguíneos foi obtida por um detetor de cantos modificado combinado com filtros binários e limiares dinâmicos. Este método foi avaliado em duas bases de dados com resultados superiores aos melhores métodos já publicados.
A segmentação da rede vascular foi obtida com um novo método de melhoramento dos vasos sanguíneos desenvolvido em espaço de escalas. A avaliação do desempenho foi feita em três bases de dados com resultados competitivos.
Finalmente, o método de deteção de microaneurismas baseia-se numa estrutura espaço de escalas tanto para a extração de candidatos como para a rotulagem dos microaneurismas. A rotulagem utiliza uma abordagem baseada na vizinhança dos microaneurismas. Os resultados alcançados mostram que este método é muito competitivo quando comparado com os métodos atuais, particularmente em imagens de alta resolução.
Com base nas metodologias descritas nesta tese, conseguiram-se várias contribuições individuais. Entre elas, as mais destacadas são; um novo método de melhoramento das estruturas dos vasos, uma nova técnica designada campos de soma acumulativa para a deteção do disco ótico, um detetor de vasos sanguíneos baseado num detetor de cantos modificado em espaço de escalas, extração de exsudados utilizando o mapa de distribuição de ruído. Um novo procedimento para detetar e remover regiões de atrofia peripapilar em redor das margens do disco ótico, e uma análise de vizinhança para a deteção de microaneurismas apenas dependente da intensidade dos pixeis.
Um dos principais objetivos foi a criação de uma abordagem de alguma forma uniforme e comum para a deteção e segmentação das estruturas da retina e sinais patológicas. Em vários dos métodos descritos foi utilizada uma abordagem em espaço de escalas, nomeadamente para a segmentação dos exsudados, segmentação da rede vascular e deteção dos microaneurismas.
A investigação desenvolvida nesta tese levou a uma compreensão mais sólida de quais são as escalas mais adequadas (individuais ou combinadas) para realizar a deteção e segmentação de estruturas da retina e sinais patológicos. Este conhecimento pode ter um impacto importante no

desenvolvimento de novas técnicas, na deteção de características e descritores mais efetivos para métodos de aprendizagem máquinas [63, 24].
Tanto nas metodologias desenvolvidas para a segmentação dos exsudados e para o método de deteção do disco ótico, não foi realizada nenhuma etapa de pré-processamento para corrigir iluminação não uniforme ou normalizar o contraste. Esta abordagem difere de outras da literatura [64, 65], sendo relevante para a preservação da informação dos pixeis. Os resultados obtidos contradizem alguns métodos encontrados na literatura [66, 67] onde uma fase de pré-processamento é considerada fundamental para realizar a deteção e segmentação de estruturas e sinais patológicas da retina. A robustez dos métodos desenvolvidos foi verificada pela sua avaliação em várias bases de dados. Em particular, o método de deteção de disco ótico provou ser particularmente eficaz em imagens patológicas.
Visto que os MAs são a primeira manifestação oftalmológica de retinopatia diabética em imagens de fundus [68], a sua deteção permitiu a avaliação da sua deteção a partir de imagens do fundo do olho. A avaliação da efetividade deste método obteve um desempenho como outros métodos recentes assim como quando comparado com a capacidade de especialistas humanos detetarem esta patologia nestas imagens. Isto sugere o potencial para que este método seja utilizado como ferramenta de rastreio da retinopatia diabética.
Em resumo, embora a investigação recente para a análise automatizada de imagens de fundo do olho esteja centrada em abordagens de aprendizagem máquina, os métodos baseados em processamento de imagem e em particular baseados em metodologias de espaço de escalas podem fornecer contributos importantes para futuros desenvolvimentos na análise computacional automatizada. Em particular, o trabalho de investigação desenvolvido dá importantes perspetivas sobre novas técnicas, uma abordagem comum baseada em espaço de escalas, e informação sobre escalas que pode levar ao desenvolvimento futuro de melhores metodologias e algoritmos.

Trabalho futuro

A investigação desenvolvida durante esta tese mostrou a potencial aplicabilidade de abordagens em espaço de escalas na deteção automática de estruturas e de sinais patológicas da retina. Possíveis linhas futuras de investigação devem considerar sobretudo a aplicabilidade clínica.
Embora o método de deteção da retinopatia diabética exiba o potencial para ser utilizado como método de rastreio, é necessário primeiramente abordar alguns pontos. Em primeiro lugar, é necessário incluir outros sinais patológicas, como exsudados, hemorragias e neovascularização, numa possível ferramenta de rastreio da retinopatia diabética. Em segundo lugar, é necessário um maior conjunto de bases de dados para a validação do método, onde se considerem outros fatores como, diferentes grupos etários, etnia, e interação com outras doenças da retina.
Uma das principais desvantagens das bases de dados publicamente disponíveis é a falta de heterogeneidade étnica. É um facto bem conhecido que diferentes etnias têm diferentes alterações morfológicas e patológicas da retina nas mesmas doenças [2]. A maioria destas bases de dados são compostos por imagens que representam populações caucasianas e asiáticas. Uma exceção é a base de dados Hamilton Eye Institute Macular Edema Dataset (HEI-MED) [69] formada por uma mistura razoável de etnias. Esta falta de heterogeneidade de etnias resulta numa aplicabilidade limitada dos algoritmos automatizados a uma população mais vasta.
Além disso, um dos maiores problemas para a melhoria e generalização dos métodos automatizados advém da fiabilidade da anotação e da ambiguidade das lesões presentes nas bases de dados. Isto compromete o desenvolvimento de algoritmos robustos e eficazes que possam ser utilizados em cenários clínicos. A investigação futura deve incluir a revisão por pares das bases de dados

atuais e a criação de novas e melhores bases de dados anotados.

Organização da tese

Os trabalhos descritos nesta tese encontram-se organizados em oito capítulos. O primeiro capítulo fornece uma introdução dos temas abordados ao longo desta tese. O segundo capítulo descreve as estruturas e anatomia do olho humano. São também descrita de forma geral as principais doenças oculares e sistémicas com manifestações clínicas no olho. No terceiro capítulo é feita uma revisão da literatura sobre os métodos de análise computacional de imagens do fundo do olho. Em particular estes métodos abordam as temáticas da deteção do disco ótico e segmentação da rede vascular, e da segmentação de exsudados e deteção de microaneurismas. O quarto capítulo apresenta os métodos desenvolvidos para a segmentação dos exsudados da retina. No quinto capítulo é descrito o método para a deteção do disco ótico. O sexto capítulo aborda os métodos desenvolvidos para deteção da linha central dos vasos sanguíneos e segmentação da rede vascular. O capítulo sete apresenta o método utilizado para a deteção de microaneurismas, também usado na deteção da retinopatia diabética. Finalmente, o capítulo oito fornece descreve as conclusões finais com orientações futuras.

Abstract

Fundus images are one of the most important imaging examinations in modern ophthalmology because they are simple, inexpensive and, above all, non-invasive. Nowadays, the acquisition and storage of high-resolution fundus images is relatively easy and fast. Therefore, fundus imaging has become a fundamental investigation in retinal lesion detection, ocular health monitoring and screening programmes. Given the large volume and clinical complexity associated with these images, their analysis and interpretation by trained clinicians becomes a time-consuming task and is prone to human error. Therefore, there is a growing interest in developing automated approaches that are affordable and have high sensitivity and specificity. These automated approaches need to be robust if they are to be used in the general population to diagnose and track retinal diseases. To be effective, the automated systems must be able to recognize normal structures and distinguish them from pathological clinical manifestations.

The main objective of the research leading to this thesis was to develop automated systems capable of recognizing and segmenting retinal anatomical structures and retinal pathological clinical manifestations associated with the most common retinal diseases. In particular, these automated algorithms were developed on the premise of robustness and efficiency to deal with the difficulties and complexity inherent in these images. Four objectives were considered in the analysis of fundus images. Segmentation of exudates, localization of the optic disc, detection of the midline of blood vessels, segmentation of the vascular network and detection of microaneurysms.

In addition, we also evaluated the detection of diabetic retinopathy on fundus images using the microaneurysm detection method. An overview of the state of the art is presented to compare the performance of the developed approaches with the main methods described in the literature for each of the previously described objectives. To facilitate the comparison of methods, the state of the art has been divided into rule-based methods and machine learning-based methods.

In the research reported in this paper, rule-based methods based on image processing methods were preferred over machine learning-based methods. In particular, scale-space methods proved to be effective in achieving the set goals.

Two different approaches to exudate segmentation were developed. The first approach is based on scale-space curvature in combination with the local maximum of a scale-space blob detector and dynamic thresholds. The second approach is based on the analysis of the distribution function of the maximum values of the noise map in combination with morphological operators and adaptive thresholds. Both approaches perform a correct segmentation of the exudates and cope well with the uneven illumination and contrast variations in the fundus images.

Optic disc localization was achieved using a new technique called cumulative sum fields, which was combined with a vascular enhancement method. The algorithm proved to be reliable and efficient, especially for pathological images. The robustness of the method was tested on 8 datasets.

The detection of the midline of the blood vessels was achieved using a modified corner detector in combination with binary philtres and dynamic thresholding. Segmentation of the vascular network was achieved using a new scale-space blood vessels enhancement method. The developed methods have proven effective in detecting the midline of blood vessels and segmenting vascular networks.

The microaneurysm detection method relies on a scale-space microaneurysm detection and labelling system. A new approach based on the neighbourhood of the microaneurysms was used for labelling. Microaneurysm detection enabled the assessment of diabetic retinopathy detection. The microaneurysm detection method proved to be competitive with other methods, especially

with high-resolution images. Diabetic retinopathy detection with the developed microaneurysm detection method showed similar performance to other methods and human experts.
The results of this work show that it is possible to develop reliable and robust scale-space methods that can detect various anatomical structures and pathological features of the retina. Furthermore, the results obtained in this work show that although recent research has focused on machine learning methods, scale-space methods can achieve very competitive results and typically have greater independence from image acquisition. The methods developed in this work may also be relevant for the future definition of new descriptors and features that can significantly improve the results of automated methods.

Keywords

Retina; Fundus images; Fundus images analyses; Scale-space; Curvature; Segmentation; Features detection; Retinal diseases; Diabetic retinopathy; Exudates; Optic disc; Vascular network; Blood vessel centerline; Microaneurysms;

Contents

List of Algorithms

List of Figures

List of Tables

Acronyms

Acc	Accuracy
AM-FM	Amplitude-modulation frequency-modulation
AMD	Age-related macular degeneration
AUC	Area under the curve
BAMCA	Bi-level adaptive morphological component analysis
CLAHE	Contrast-limited adaptive histogram equalization
CNN	Regions with convolutional neural network
DOG	Difference of Gaussian
DoOG	Difference of offset Gaussian
DNN	Deep neural network
DR	Diabetic retinopathy
FDOG	First order derivative of Gaussian
FN	False negative
FP	False positive
GAN	Generative adversarial network
GMM	Gaussian mixture model
HLS	Hue luminance saturation
HSV	Hue saturation value
kNN	k-nearest neighbours
LoG	Laplacian of Gaussian
MF-FDOG	Matched-filter first order derivative of Gaussian
MM	Mathematical morphology
NN	Artificial neural network
OD	Optic disc
PCA	Principal component analysis
ResNet	Residual networks
RGB	Red Green Blue
Sen	Sensitivity
Spe	Specificity
SVM	Support vector machine
TP	True positive
TN	True negative

Chapter 1

Introduction

1.1 Overview

The aim of this thesis is to develop computer algorithms capable of recognizing retinal anatomical structures and pathological clinical manifestations of the retina. Two different anatomical structures of the retina are investigated, in particular the detection of the vascular network and the optic disc (OD). Segmentation of exudates and microaneurysms (MAs) are the pathological clinical manifestations of the retina studied. This chapter describes the motivation of the work, followed by a brief description of the main contributions and the organization of the work.

1.2 Motivation

The development of new image capture and processing technologies led to an expansion of possible applications. It is therefore not surprising that the medical field of ophthalmology has been intensively and extensively studied by the research community to develop automated computer-based analysis methods for retinal fundus images. With the ability to analyse large volumes of images in a short time, these systems are particularly useful in screening environments and in the detection and diagnosis of retinal diseases [1].

There are several ocular or systemic diseases that can manifest clinically in the retina. Based on their prevalence and frequency in the population, the most studied are diabetic retinopathy (DR), glaucoma, hypertensive retinopathy, and age-related macular degeneration (AMD). These diseases have multiple developmental stages characterized by distinct and specific retinal changes [2].

In addition, there are several difficulties that automated systems must overcome before they can be used in clinical practice. In normal fundus images, automatic analysis and recognition of retinal structures requires robust algorithms. In pathologic fundus images, anatomic structures and pathologic clinical manifestations exhibit great heterogeneity in size, color, extent, and shape cite-Tobin2008detectionofanatomicalstructures. Fundus images also show great variability within an image and between images. The variability within an image is mainly due to the presence of opacities, different reflectance, and retinal thickness. The variability between images is due to differences in acquisition hardware, illumination, acquisition protocol, and retinal pigmentation [4]. This variability in imaging, inherent in all fundus images, combined with the heterogeneity of clinical manifestations, makes the development of automated computational algorithms a complex and challenging task.

The preceding descriptions imply that detection and segmentation of clinical manifestations by automated systems must be robust, efficient, and accurate to achieve correct diagnosis or follow-up and to enable early detection and effective therapy. In recent years, a considerable number of papers have been published dealing with the automated detection of retinal structures and pathological clinical manifestations of the retina. A detailed review of the major contributions to selected cases of retinal analysis is described in Chapter 3. Some methods also address automatic detection

of retinal diseases such as DR [5, 6, 7], AMD [8, 9], glaucoma [10, 11], and hypertensive retinopathy [12]. Overall, these methods have been divided into rule-based methods and machine learning. More recently, deep-learning methods have also been proposed.

The first methods for detecting retinal structures and pathological clinical manifestations were rule-based approaches. Rule-based methods do not require a learning mechanism, although they need to be calibrated using examples. Many rule-based methods show very good results, but in general they have difficulty overcoming the complexity and heterogeneity of retinal features found in both normal and pathological fundus images [13]. Nevertheless, many important features and descriptors are based on rule-based approaches.

To overcome the difficulties of rule-based methods, approaches based on machine learning have also been proposed. These methods require a set of features or descriptors that are measured directly from the data. Then, in a learning phase, a particular classifier is trained based on these features or descriptors. After the learning phase, the classifier learns to recognize retinal structures and clinical manifestations from unseen images. Machine learning approaches depend on the discriminative power of the features and descriptors [13]. Moreover, these methods often require large datasets to train [14] and tend to require new training for different datasets (typically with different recording devices), which limits their generalization [15]. In addition, due to their heterogeneity, the available datasets have ambiguities in correctly identifying the number, size, and precise delineation of retinal lesions, which affects the efficiency of the methods [24].

More recently, deep neural network (DNN) approaches, commonly referred to as deep learning [16, 17, 18], have also been proposed. These methods perform the tasks of learning the features suitable for a particular classification or recognition, as well as the final classification. A main advantage of deep learning methods is that with enough data samples and computational power, they can outperform traditional machine learning methods [19]. The main challenges in applying deep learning approaches are the limited number of labeled data and the unbalanced distribution of data from different classes. These approaches are very time-consuming, require specialized hardware, need large datasets [22, 23, 24, 20], and also tend to require new training for each dataset. These limitations affect their performance on multiple datasets and generalization in clinical practice [20, 25]. In addition, poor image quality and heterogeneous features of retinal structures and pathological clinical manifestations create data imbalance in the learning process [26, 21].

Despite the amount and variety of methods proposed for automatic retinal image analysis, this area of research remains an active area of research today [**?**] due to the potential impact and the need for more accurate results.

To overcome some of the previously described limitations, a combination of deep Learning and domain knowledge was proposed by Orlando *et al.* [24]. It has been demonstrated that handcrafted features are better able to discriminate low-contrast lesions and thus improve the detection of pathological clinical manifestations [76, 24]. In addition, the application of new preprocessing techniques to overcome inter-image variability [63], as well as the use of discriminative features and more robust classifiers is needed to obtain more robust methods that outperform existing ones in terms of sensitivity and specificity [63, **?**].

The aim of this work is to develop automated systems capable of recognizing and segmenting the anatomical structures of the retina and detecting the pathological clinical manifestations of the most common retinal diseases. In particular, these automated algorithms have been developed under the premises of robustness and efficiency to deal with the difficulties and complexity of these images.

Four objectives were considered for the analysis of the ocular fundus image. These are segmentation of exudates, localization of OD, detection of the blood vessels centerline and segmentation of

the vascular network, and finally detection of MAs. We chose to use spatial image analysis methods based on a scale-space framework rather than machine learning methods. The reason for choosing a scale-space framework [28, 29] is the large heterogeneity in size, shape, color, and extent between retinal anatomic structures and retinal pathologic clinical manifestations [30]. These characteristics make them ideal for detection and analysis using a scale-space framework approach.
Scale information is typically used to extract candidate regions [31, 32, 33] or to obtain multiscale features [34, 35, 36, 37, 6]. Other work uses scale information in combination with a different processing technique to extract retinal structures. Work by [**?**, 38, 39] used a set of 2D Gabor filters at different scales to enhance and perform vessel segmentation. A set of Gaussian filters at different scales was described in [40] to perform OD and fovea detection, and in [41] to perform MAs detection. Usually, the scales are defined empirically, which limits the applicability and generalization of the extracted features and descriptors.
Scale-space approaches have also been proposed to extract retinal vessels [39, 38]. However, there are no studies generalizing the same scale-space framework for the detection of multi-retinal structures.
In this work, we attempted to develop robust and reliable scale-space approaches capable of detecting different anatomical structures and pathological clinical manifestations of the retina. In addition, the search for a generalized set of scales can provide important information for the development of new descriptors and features that can significantly improve the performance of automated methods.

1.3 Contributions

This section briefly describes the contributions resulting from the research work developed in the scope of this thesis. These can be summarised as follows,

- A fully automated method for the detection of the retinal exudates based on the scale-space extrinsic curvature analysis. A major difference of the proposed approach from other exudates detection methods is that no preprocessing of the images is done before they are processed. The results of the proposed algorithm are compared with other recent state-of-the-art methods, revealing a good performance. This study is described in chapter 4, which is based on an article published in the proceedings of the 7th International Symposium on Image and Signal Processing and Analysis [53].

- A fully automatic scale-space method for the segmentation of the retinal exudates. Contrary to common approaches, this is achieved without any preprocessing steps, but rather considering the curvature scale-space to perform the segmentation. The performance is evaluated in a publicly available dataset and results in very good performance. This study is described in chapter 4, which is based on an article published in the proceedings of the 9th IEEE International Symposium on Biomedical Imaging [54]

- A fully automatic method for the segmentation of retinal exudates based on the maxima noise map distribution function. Noise is an inherent characteristic associated with all digital images. The traditional approach is to perform some filtering mechanism to reduce it. In this work. noise is effectively used to segment the retinal exudates. Furthermore, contrary to the traditional approaches, no preprocessing is performed. The method proves to be resilient and robust to contrast variations and non-uniform illumination. The method performance is evaluated in a publicly available dataset, with results similar to other state-of-

the-art methods. This work is described in chapter 4, which is based on an article published in the proceedings of the 19th European Signal Processing Conference (EUSIPCO) [77].

- An automatic approach to detect the OD. This method presents a new vessel enhancement method and a new technique designated cumulative sum fields. The cumulative process increases the robustness of the detection in pathological fundus images and increases its adaptability to different fundus images resolution. Furthermore, no illumination equalization was performed as in other methods. The robustness of the method is demonstrated by its evaluation on 8 publicly available datasets. A better performance results when compared with other state-of-the-art methods. This work is described in chapter 5, which is based on a paper published in the journal of IEEE Journal of Biomedical and Health Informatics [55].

- An automatic method for the detection of the vessels centerline. This method uses a scale-space approach to increase the reliability and robustness. A new vessel detector is described based on a modified corner detector. The algorithm was evaluated in two publicly datasets and results in better performance when compared to other state-of-the-art methods. This work is described in chapter 6, which is based on an article published in the proceedings of the 22nd European Signal Processing Conference [56].

- A fully automatic method for the segmentation of the retinal vasculature. In this work, three new and important contributions are made. First, a new method capable of vessel enhancement is presented. Second, a new criterium to remove some false vessel caused by the proximity of bright regions is presented, avoiding the false vessels created by the presence of exudates or bright artefacts. Third, a new approach that discards the false vessel regions associated with peripapillary atrophy that usually appear in the border of the OD. The performance evaluation is made on three publicly available datasets. The method is described in chapter 6, which is based on an article published in the proceedings of Eurographics Workshop on Visual Computing for Biology and Medicine [57].

- A scale-space based method for the extraction of MAs candidates. The traditional approach involves selecting several features to characterize MAs. In this work, the MAs are detected using a new scale-space neighbourhood method using only the pixel intensity values between a region and its neighbourhood. The performance evaluation is made on a publicly available dataset. The neighbourhood analysis technique is described in chapter 7. This method is based on an article published in the proceedings of Ophthalmic Medical Image Analysis International Workshops [78].

- A new approach for the detection of MAs. Apart to the initial image preprocessing, the whole algorithm is developed using a scale-scale framework. The MAs detection is evaluated at each scale individually and also through a combination of scales. This provides important insights to understand which scales are better for MAs detection. Furthermore, the developed algorithm is also evaluated when used for DR detection. The proposed method proves to be competitive against other state-of-the-art approaches in both MAs and DR detection. The method is described in chapter 7, which consists in a paper submitted for publication in a journal.

1.4 Thesis organization

The research reported in this thesis is organized in eight chapters. Chapter 1 provides an introduction to the subjects considered in this thesis. Chapter 2 describes the eye structures and the pathological clinical manifestations associated with the most common ocular diseases. This chapter presents a brief description of the human eye anatomy and the main anatomical structures. An overview of the main ocular and systemic diseases with clinical manifestations is also described. Chapter 3 provides a general literature review on retinal image analysis methods for the detection and segmentation of the retinal structures (OD and retinal vasculature) and of the pathological clinical manifestations (exudates and MAs). Chapter 4 presents the research work developed in this thesis for the segmentation of the retinal exudates. In chapter 5 the method for the detection of the OD is described. Chapter 6 presents the developed methods for the detection of the vessels centerline and for the segmentation of the retinal vasculature. Chapter 7 describes the developed method for the detection of MAs and diabetic DR. Finally, chapter 8 provides some concluding remarks together with future directions.

Chapter 2

Eye and retina

2.1 Overview

This chapter presents a brief description of the human eye anatomy up to the neurosensory retina. A brief description of the imaging modalities available to analyse the eye is also presented. Last, an overview of the main ocular and systemic diseases with clinical manifestations is described.

2.2 Structure of the eye

The human eye is responsible for sensing the energy of visible light and transducing it into nerve impulses. This energy belongs to the visible light domain of the electromagnetic spectrum between 400 nm and 700 nm [2]. The eye is formed by three layers. The first layer is the outer fibrous layer of connective tissue that forms the cornea and the sclera. The second layer is the middle vascular composed of the iris, ciliary body, and choroid. The third layer is the inner neurosensory retina. [2]. The clinical visible part of the eye includes the cornea, the white sclera, the coloured iris, and the pupil [71]. These anatomical structures are presented in Figure 2.1. The light path through the eye consists of several steps. First, it passes through the cornea, followed by the aqueous humor and pupil. The pupil dilates or contracts, thus controlling the amount of light entering the eye. Next, the light passes through the crystalline lens, crossing the vitreous humor, and is finally focused on the fundus of the eye [71, 2].

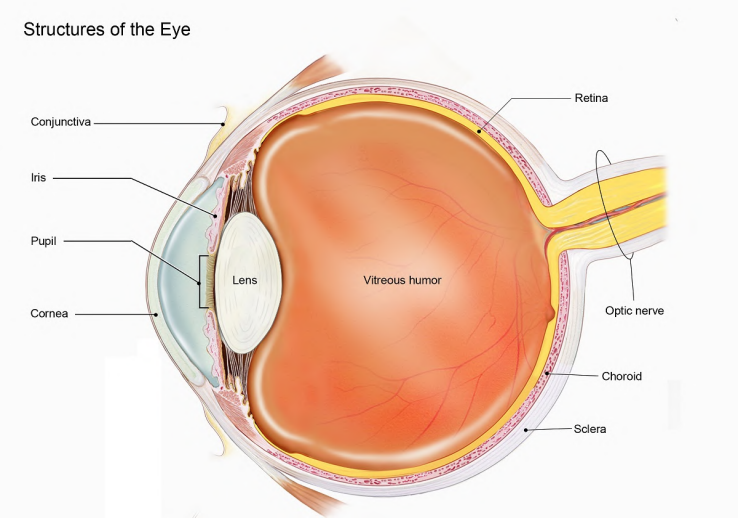


Figure 2.1: Schematic representation of the structures of the eye (Adapted from [2]).

The fundus of the eye is formed by the neurosensory retina, the retinal pigment epithelium, and the choroid. These layers of tissue extend to the periphery of the fundus, thus covering the vitreous body except for the anterior segment. A complete fundus image (see Figure 2.2) can be defined as the interior lining of the eyeball, including the retina, OD, the retinal vasculature, and the macula.

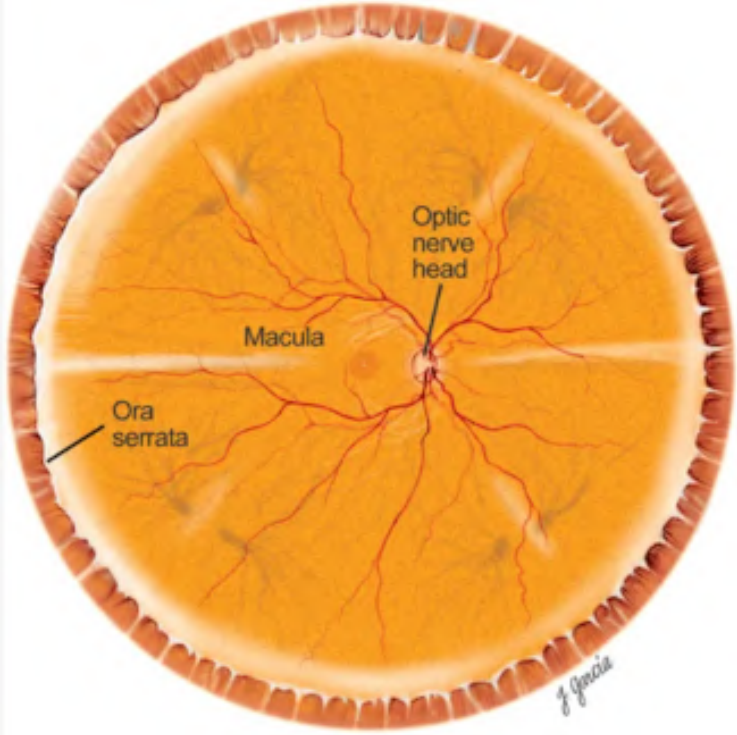


Figure 2.2: The anatomy of the right eye fundus (Adapted from [70]).

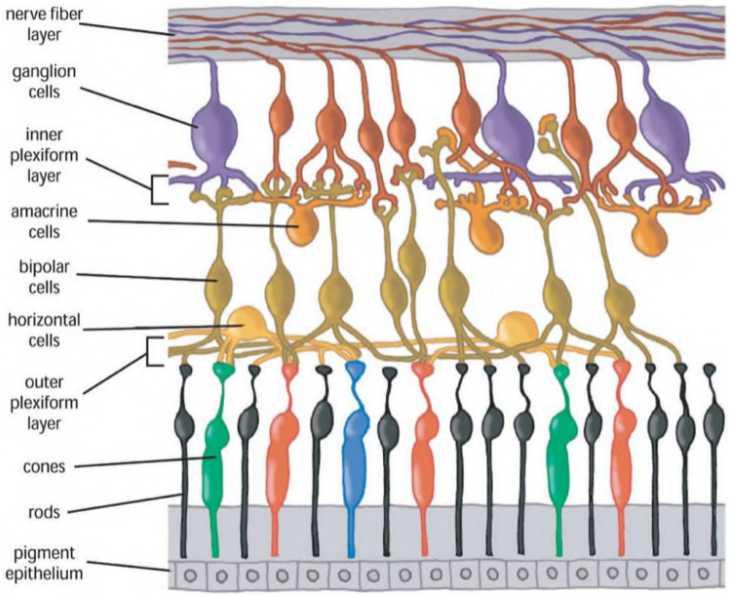


Figure 2.3: Schematic drawing of cellular layers of retina (Adapted from [71]).

The neurosensory retina, usually called the retina, is a multi-layered sensory tissue layer covering the largest part of the eye fundus. These layers contain the photoreceptors, bipolar and ganglion cells (see Figure 2.3). The retina contains millions of photoreceptors that translate the received light into electric impulses Bipolar and ganglion cells gather and carry the message to the OD, that it is transferred through the optic nerve until it reaches the brain. The brain then decodes and interprets the electrical message sent to it, resulting in the perceived vision.

In the external layer of the retina, there are two types of photoreceptors, namely, rods and cones. These are named after their shape. The cones have a conical shape, are active in bright illumi-

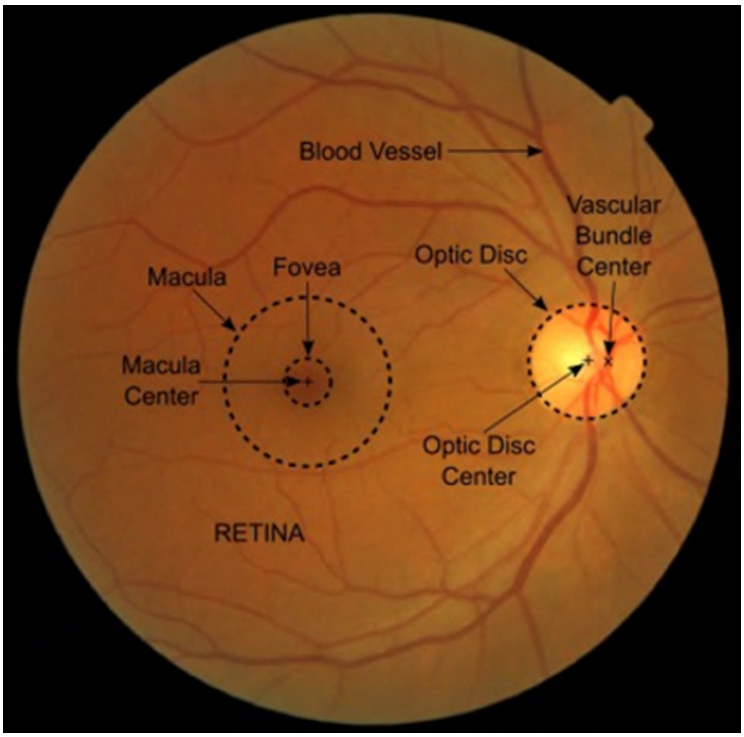


Figure 2.4: Fundus photograph with several annotated retinal landmarks. The most important are the fovea, OD and blood vessels (Adapted from [71]).

nation, and are responsible for colour vision. There are approximately 6 to 7 million cones, all located in the central region of the macula, designated by the fovea. There are three distinct types of cones, that are sensitive to different parts of the colour spectrum. Their combined response to the incoming light enables us to perceive colours. The rods, on the other hand, are not sensitive to colour [2]. The rods are thinner and longer than cones and are located in the peripheral regions of the retina. They are active in dim illumination conditions and are responsible for the black-and-white vision. There are about 125 million rods on the retina.

The light that enters the eye is translated to electrical impulses by the cones and rods. They are followed by the bipolar and ganglion cells that transmit the neuronal signals to the brain. The axons of the ganglion cells converge to a single region of the retina, forming the optic nerve and exiting the eye. The optic nerve is part of the central nervous system. The clinical visible part of the optic nerve in the retina is the optic nerve head or the OD. Since the OD is formed by the ganglion cells axons, it does not contain photoreceptors. This creates a physiological blind spot within the visual field. The blind spot is not perceptible because the brain *interprets* this visual field region with the surrounding information.

The OD is also the site of entry for the central retinal artery and the exit site for the central retinal vein. Clinically, these retinal vessels branch out from the OD to the retinal periphery, forming the vascular arches. The vascular arches surround the darker and central macular region of the retina. The darkest colour of the macula is due to a larger accumulation of pigment. The central part of the fovea has no vascularization and is then designated avascular. Furthermore, in the fovea, the density of cones is very high. This results in the fovea being the region of the retina with the highest visual acuity (See Figure 2.4) [71].

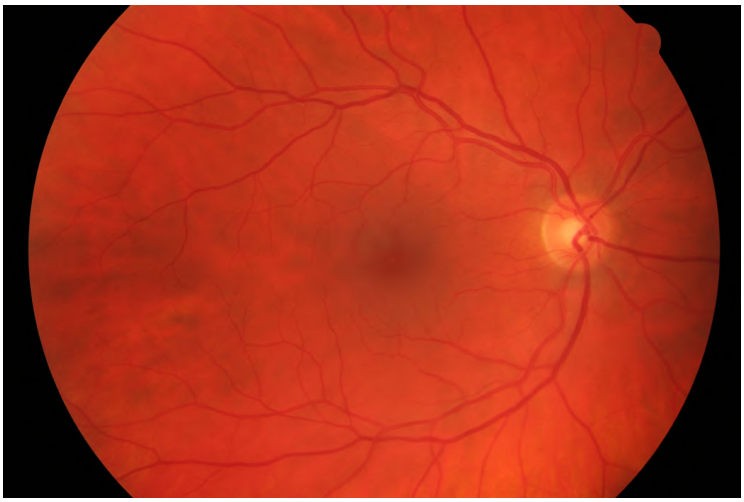


Figure 2.5: Example of a fundus photography (Adapted from [50]).

2.3 Imaging the eye

The retina is the only place in the human body where the blood circulation system and parts of the central nervous system (the retina and the optic nerve) [71] can be observed non-invasive and in-vivo [79]. Furthermore, several ocular systemic and cardiovascular diseases have clinical manifestations in the retina [79].
Several imaging modalities are available for imaging the fundus, such as digital colour fundus photography, stereo fundus photography, hyperspectral imaging, scanning laser ophthalmoscopy, adaptive optics scanning laser ophthalmoscopy, fluorescein angiography, and indocyanine angiography [71].
The choice of an imaging method usually depends upon the type of disease under study. Nevertheless, because it is non-invasive and widely available, digital colour fundus photography is today a mainstay of the clinical care of patients with retinal and systemic diseases as well as screening activities [71].
Colour fundus photography uses a retinal camera to photograph regions of the vitreous, retina, choroid, and optic nerve to document abnormalities related to disease processes affecting the eye or to follow the progress of the disease in response to therapy. An example of such a colour fundus photograph is shown in Figure 2.5. In colour fundus photographs, the images can be obtained in non-mydriatic (non-dilated pupil) or mydriatic conditions (dilated pupil). Mydriatric images provide better image quality since sufficient light passes through the pupil, but some patient discomfort may exist.
In this thesis. different publicly available datasets of digital colour fundus photographs (fundus images for short) are used. The datasets consist of fundus images captured using different protocols [46, 42, 48, 43, 44, 45, 49, 47, 50]. These datasets are described in the appendix A.

2.4 Retinal diseases

The eye is a complex organ composed of many types of tissues. These anatomical features make the eye susceptible to a wide variety of diseases, but also capable of providing insights on many body systems [71]. There are various diseases and age-related changes that originate either in the eye [25], the brain [80, 81], or the cardiovascular system that have clinical manifestations in the

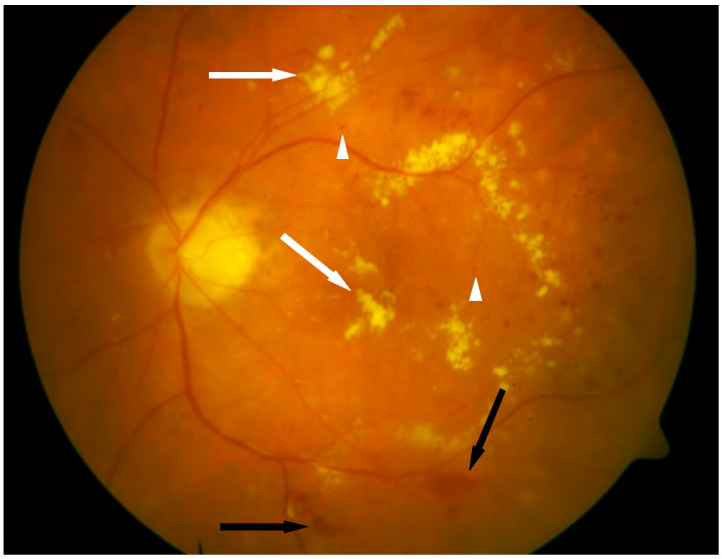


Figure 2.6: Fundus image with several annotated DR signs. The white arrows indicate exudates. The black arrows indicate haemorrhages. The white arrowheads indicate MAs (Adapted from [44]).

retina [82]. The most prevalent diseases are DR [83, 84], glaucoma [72, 85], AMD [25, 18] and hypertensive retinopathy [71]. Next, a brief description of these diseases is made, accompanied by fundus images examples of the typical retinal pathological clinical manifestations.

2.4.1 Diabetic retinopathy

The global prevalence of diabetes mellitus is predicted to increase in the coming decades, from an estimated 382 million in 2013 to 592 million by 2035 [86]. Diabetic retinopathy is a complication of diabetes mellitus, caused by high blood sugar levels (hyperglycemia) that damages the blood vessels in the retina [71]. Retinal changes are usually asymptomatic in their early stages and are only detected when they have evolved to a level that makes treatment difficult or impossible [87]. The hyperglycemia damages the retinal vessel walls [71], and as a result, microvascular lesions have been utilized as the major criteria for evaluating and classifying the DR grade [86]. Diabetes-induced changes also affect non-vascular cell types that play an important role in the development and progression of DR, albeit in unison with the vasculature [86].
Diabetic retinopathy is usually divided into two broad stages. The earlier stage is designated by nonproliferative DR, and the advanced stage is designated proliferative DR. The nonproliferative DR stage is classified based on clinical findings manifested by visible features, including microaneurysms (MAs), retinal haemorrhages, intraretinal microvascular abnormalities, and venous calibre changes [86]. At this stage, tiny vessels leak, making the retina swell. The swelling of the macula is designated macular oedema, and it is the most common cause of loss of vision in DR [88, 71]. Furthermore, there is also the presence of yellow flecks made up of lipid residues of serous leakage from damaged capillaries [89].
The proliferative DR stage is characterized by the hallmark feature of pre-retinal neovascularization [86]. This new, abnormal and superficial blood vessels lead to scarring and cell loss in the retina.
Figure 2.6 shows some examples of MAs, exudates, and haemorrhages.
The major issue with DR is that in most cases the patient does not perceive any vision loss at early stages. Furthermore, the rising prevalence of diabetes worldwide leads to the implementation of population-based DR screening programs that can perform automated retinal imaging detection

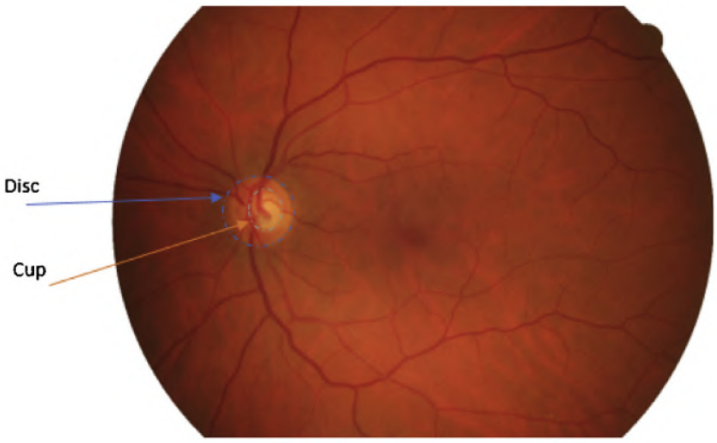


Figure 2.7: Fundus image with the annotated disc and cup regions (Adapted from [72]).

and analysis, rapidly and sensitively while minimizing inappropriate referrals to retina specialists [90].

2.4.2 Glaucoma

Glaucoma is the world's second prominent condition of irreversible vision loss after cataracts, accounting for 12% of annual cases of blindness. It is expected that the number of people between the ages of 40 to 80 affected by glaucoma will increase from 64.3 million to 80 million by 2020, and to 111.8 million by 2040 [72]. Glaucoma refers to a degeneration of the retinal ganglion cells and their axons, caused by a group of heterogeneous diseases [72], such as high intraocular pressure of the eye and damage to the OD [91]. In fundus images, the OD is divided into the disc and the cup, as shown in Figure 2.7.
The hallmark of glaucoma is the enlargement of the cup region with respect to disc [71]. This results in the gradual damage to the OD and resulting visual field loss [71]. In the early stages of the disease, patients do not have any symptoms of vision loss. As glaucoma progresses, the loss of peripheral vision occurs [91]. Early diagnosis and optimal treatment have been shown to minimize the risk of visual field loss. Early detection depends on manual observation by optometrists and ophthalmologists. This is costly and subjective, and as such prone to error. Automated tools that can automatically detect the OD, and quantify the disc cupping would create a positive impact on both people and the economy [72, 55].

2.4.3 Age-related macular degeneration

Age-related macular degeneration is responsible for approximately 9% of global blindness [18], particularly in the elderly. By 2040 it is estimated that the number of people affected globally by the disease will be 288 million [92]. Age-related macular degeneration arises from a complex combination of ageing, genetics, and environmental risk factors [18]. Based on the examination of fundus images, AMD is classified as wet or dry. Dry AMD leads to a gradual vision loss. Wet AMD, also designated by choroidal neovascularization, is the most visual aggressive and threatening form [71]. In the early stages, dry AMD patients are usually asymptomatic, although in fundus images it is possible to observe on the macula small yellow deposits of lipid acellular debris, named drusen (see Figure 2.8). Later signs of dry AMD, include retinal pigmentary changes that occur before the development of geographic atrophy. These patients usually have limitations in the vi-

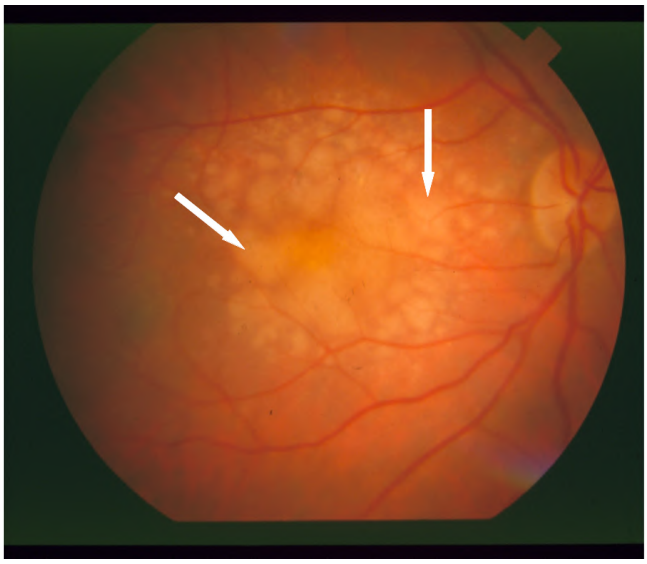


Figure 2.8: Colour retinal photograph with AMD. The white arrows indicate drusen (Adapted from the image *im0006* from the Stare dataset [42]).

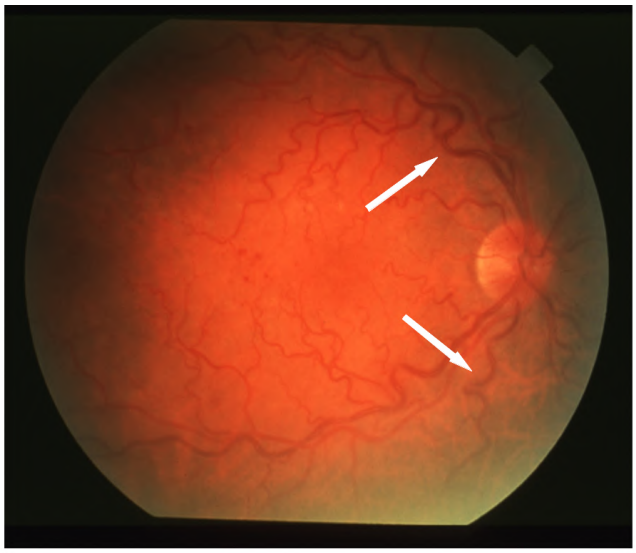


Figure 2.9: Colour retinal photograph with vascular abnormalities. The white arrows indicate examples of vessel tortuosity (Adapted from the image *im0033* from the Stare dataset [42]).

sion that happen under poor illumination. Wet AMD early stages are similar to dry AMD. In later stages, there is the appearance of exudative abnormalities and abnormal blood vessels under the macula, that damage its cells [25].

2.4.4 Hypertensive retinopathy

Cardiovascular diseases are the leading cause of death globally in the last 15 years [93]. Hypertension is the leading risk factor of cardiovascular disease, affecting both the structure and function of the retinal microsvasculature [71]. The retinal microvascular changes represent a manifestation of ongoing systemic microvascular damage and can be viewed directly and non-invasively in the retina [57] (see Figure 2.9).
These vascular changes include haemorrhages, microaneurysms, cotton wool spots, changes in retinal vascular calibre and changes from retinal vascular architecture (tortuosity, fractal dimen-

sion, and branching angle) [82, 94]. It was found that there is a relationship between retinal abnormalities and systemic vascular diseases, even though the blood pressure and other vascular risk factors are in control [95]. To detect these vascular changes, health care professionals explore and analyses the retinal vessels. This visual manual inspection is tedious, time-consuming, and prone to errors. Furthermore, the retinal vasculature is complex, with small blood vessels and several calibre variations. Hence, several automatic retinal vessels segmentation techniques have been proposed, to assist health care professionals in the timely identification of any vascular abnormality [96].

Chapter 3

State-of-the-art overview

The research work reported in this thesis considers the automatic detection of the retinal vasculature and OD, and the retinal pathological clinical manifestations, exudates, and MAs. Thus, it is appropriate to start with a literature review of the state-of-the-art methods that deal with the automatic detection of these retinal features.
The current state-of-the-art methods that detect exudates are presented in section 3.1. A review of the segmentation of the retinal vasculature is described in section 3.3. In section 3.2 several methods dealing with the localization of the OD are described. Finally, in section 3.4 a review of the methods dealing with MAs detection is presented.

3.1 Retinal exudates

Diabetic retinopathy is the main cause of blindness induced by diabetes. It is associated with the appearance of yellow spots in the retina, namely, hard and soft exudates. Hard exudates are yellow lipid sediments of serous leakage from damaged capillaries. Soft exudates (cotton-wool spots) are superficial retinal infarcts. They both appear as yellowish lesions, spread over the retina, with different sizes, shapes, and brightness [97]. The fuzzy appearance of soft exudates is the only visible difference from the hard exudates. The non-uniform illumination, contrast variation, and variable background among different images and patients are inherent to all retina fundus images. These characteristics make the exact detection and marking of exudates a challenging task [53]. Reliable detection and marking of these lesions make possible an early DR detection, providing invaluable help in screenings and follow-ups studies.
A considerable number of papers have been published in recent years dealing with the detection of retinal hard exudates. Depending on their methodology, these methods can be broadly divided into rule-based algorithms and machine learning based algorithms. Rule-based algorithms make use of image processing techniques like adaptive histogram equalization (CLAHE) [98], spatial and frequency filters [99], edge detectors [100], and adaptive thresholding [66]. Machine learning based methods follow a supervised, or in some cases unsupervised, approaches. Recently, deep learning models became the most popular supervised machine learning approach [19].
Most of the exudates detection and segmentation methodologies are evaluated on the publicly available Diaretdbv1 dataset [45]. For this dataset, no test/training subsets are provided. A detailed description is introduced in appendix A

3.1.1 Rule-based approaches

One of the first techniques employed in the detection and segmentation of exudates was mathematical morphology. Exudates have a very large grey level variation concerning the background, and this property was explored by Walter *et al.* [101]. The method starts localizing the OD through alternating sequential filtering [102] and watershed transformation [103] to find its contours. The exudate candidates are found by morphological dilatation and thresholding [104].

A morphological coarse-to-fine method was proposed by Welfer *et al.* [105]. First, a morphological contrast enhancement method [104] is applied to the L channel of the Luv color space. Next, a regional minima operator [106] is applied to remove brighter regions commonly associated with exudates or the OD. A reconstruction by dilatation operation using the contrast-enhanced image and the result of the regional minima operation segments the exudates from the fundus image.
Imani *et al.* [107] makes use of the morphological component analysis [108] to separate lesions from the normal retina structures. The final exudate segmentation is detected by dynamic thresholding and morphological operations.
Kaur *et al.* [109] starts by pre-processing the fundus images for uneven illumination normalization and local contrast enhancement. Next, a dynamic thresholding is applied to segment the exudates.
A more recent method, based on the inverse surface technique, was presented by Karkuzhali *et al.* [15]. First, a fuzzy c-means clustering [110] identifies high-intensity regions. The enhancement of the exudates regions is made with a set of Prewitt edge detectors [111] and a global Otsu thresholding [112]. The inverse surface adaptive thresholding [113] is used to detect the final exudates.
Wisaeng *et al.* [52] proposed a method that starts with an image normalization, contrast enhancement, noise removal, and the localization of the OD. Then, a coarse segmentation method based on the mean-shift filter [114] provides a set of exudates and non-exudate candidates. The final labelling of the exudates is performed by a mathematical morphology algorithm procedure.
Other relevant rule-based exudates segmentation methods that do not appear in the performance comparison reported in section 3.1.3 are described next. These methods are not reported in table 3.1.3 because either they do not reveal a competing performance with similar approaches, or they do not provide performance information that makes possible comparison with other methods.
A statistical atlas-based exudate segmentation method was proposed by Ali *et al.* [115]. The fundus image is first warped on the atlas coordinate, and then a distance map is obtained regarding the mean atlas image. This process leaves behind the exudate candidate's regions. Final segmentation is performed with a combination of Kirsch [116] and steerable filter responses [111]. Banerjee *et al.* [100] implemented a mean-shift and a normalized cut method [117] for hard exudates segmentation. First, a bottom-hat operation [104] was applied on the red channel of an RGB fundus image. The delimitation of the OD was obtained by using an active contour model [118]. As a post-processing step, a Canny edge operator [119] was used to retrieve the exudates boundaries and final segmentation. Kar *et al.* [120] presented a method for automatic detection of lesions for DR screening. Exudate candidates were enhanced by an optimum bandpass filter. Next, a maximization of the matched-filter response [121] and a Laplacian of Gaussian (LoG) [122] filter response was used to detect the exudates. Morphology-based post-processing is applied to exclude the falsely detected candidate regions.

3.1.2 Machine learning approaches

Several supervised classifiers were applied to the detection and segmentation of exudates in fundus images. The most common classifiers are support vector machine (SVM) [123], naive bayes [124], decision trees [125], random forests [126], ensemble of classifiers [127], k-nearest neighbours (kNN) [128] and artificial neural networks (NN) [129, 130].
A technique that uses Gabor filter banks [131] to extract the exudate candidate regions was proposed by Akram *et al.* [132]. After the extraction of several features for each candidate (area, compactness, mean intensity, mean hue, mean gradient magnitude, entropy, intensity), it applies a Bayesian classifier as a combination of Gaussian functions to label exudate and no-exudate re-

gions. The groundtruth labelled regions were divided into two equal datasets, i.e. training and testing data sets.

The same author proposed another exudate segmentation method [133]. First, they segmented the retinal vasculature using Gabor wavelets [134] and multilayered thresholding. After the vasculature segmentation, candidate exudate lesions were extracted using Gabor filtering. A set of features was extracted for each candidate based on lesion shape, intensity, and statistic descriptors. The extracted features were fed by an ensemble hybrid classifier to classify the candidates. The modelling of the classifiers was done at lesion level, using randomly selected 70% of data as training and the remaining 30% of data as testing. The performance values result from the average of 10 experiments.

An ensemble classifier of bootstrapped decision trees [135] for multiscale segmentation of exudates in fundus images was developed by Fraz *et al.* [136]. The candidate regions are extracted at fine levels using morphological reconstruction [104, 137], and at coarser levels using Gabor filtering. A 9D feature vector, characterizing each candidate region, was used with an ensemble classifier of bootstrapped decision trees to classify the candidate regions as exudate or non-exudate. The candidate region area, compactness, mean intensity, mean saturation, mean gradient magnitude, and entropy were used as features.

Zhang *et al.* [138] initially applies an image preprocessing to remove dark regions. Using contextual information and mathematical morphology, candidate regions were extracted. Thus, twenty-eight contextual features were used for the exudate regions classification through a random forest classifier. The used features were based on intensity, geometric, textural, hybrid, and contextual information. Testing was done on the E-ophtha-EX dataset with a leave-one-out cross-validation, while the rest is used to train the model.

A model-based approach based on Markov random field [139] was proposed by Harangi *et al.* [140] to extract the exudate candidates. These candidates are labelled as true or false using an SVM based upon certain features of the fundus image green channel. The classified features were based on pixel intensity, gradient magnitude, and morphological descriptors. For performance evaluation, the image datasets were divided into test (71%) and training sets.

Liu *et al.* [141] proposed a three-stage approach for segmentation of exudates. Initially, the major anatomical structures, i.e. blood vessels and the OD were segmented and removed. Next, a random forest classifies the exudate candidates. The candidate features were computed using the histogram of completed local binary patterns. The segmentation is improved using adaptive thresholding, and specific size and local contrast values. The data were divided into a test set (69%) and a training set for evaluation.

Kusakunniran *et al.* [142] proposed a two-stages method. The first stage is based on the multilayer perceptron [143] to identify the initial seeds. The exudates segmentation is finalized with an iterative graph cut [144] using clusters of the initial seeds. Features like mean and standard deviation of pixels values in the green, red, and luminance channels were used. The datasets were divided into 28% for training and 84% for testing.

Recently, Long *et al.* [145] developed and evaluated an automatic fundus image processing algorithm for exudates detection using dynamic thresholding and fuzzy c-means followed by SVM classification. After preprocessing and removal of the OD, exudate candidate regions are determined using dynamic threshold combined with global threshold based on fuzzy c-means. Next, eight texture features (based on pixel intensity, gradient values, and hue values) from the candidate region are fed into an SVM for automatic exudates classification. For training, a few small regions of each groundtruth image were manually selected from the E-ophtha-EX dataset as training samples. These selected regions have been divided into exudates regions and non-exudates

regions. 10-fold cross-validation was applied to evaluate the method. On each fold, The classifier was trained on 42 selected training images and tested on the remaining 5 images.

Based on [107], Javidi *et al.* [146] proposed an extension of the morphological component analysis designated by Bi-level adaptive morphological component analysis (BAMCA). For learning the exudate dictionaries in the proposed BAMCA framework, a K-singular value decomposition algorithm [147] is used. The learning dictionary is formed by randomly selecting 5% images from the total images. From each selected image, several patches that represent true exudates and non-exudates are considered.

The first deep learning method for exudates detection was proposed by Prentasic *et al.* [148]. The proposed network has 10 alternating convolutional and max-pooling layers [16]. After the last layer, a Softmax activation function [149] provides a probability-based classification of individual pixels i.e. exudate or non-exudate classes. Experimental validation was made with 3-fold cross-validation.

Khojasteh *et al.* [150] proposed a combination of a pre-trained residual network (ResNet-50) with an SVM to detect exudates. The convolutional neural network (CNN) and the first 49 layers of the ResNet-50 compute the features extraction. The later softmax layer was replaced with an SVM for final classification. A total of 67600 exudate patches and 23200 non-exudate patches were used for training. 10-fold cross-validation with 10 runs was used to evaluate the method.

Huang *et al.* [51] proposed a method based on superpixel multi-feature extraction [151] and patch-based deep CNN. First, the superpixels, regarded as candidates patches, enable the determination of 25 features. These are divided into pixel-level features and superpixel-level features. Pixel level features depend on the saturation values from HSV colour space, and pixel intensity values from the red, green, and blue channels. Superpixel level features are based on the mean intensity and standard deviation values for each superpixel. Next, these candidate patches are used to train a deep CNN to classify the exudate candidates. The Diaretdbv1 dataset (89 images) is used for performance evaluation, divided into 47 images for training and 42 for testing. Due to the limited number of fundus images, after multi-feature extraction, training sample patches are generated for each extracted feature. The patches are divided into two groups (background and exudates) with 2500 patches each. An 8-layer CNN network is trained on the patches.

Other relevant machine learning methods for exudates detection are described next. Osareh *et al.* [152] implemented a computational intelligence technique for exudate detection. After, normalization of illumination and contrast, the images were segmented using fuzzy c-means clustering. Next, features like, colour, edge strength, and texture were ranked by a genetic algorithm. High classification accuracy was confirmed by the use of the most optimal feature set found by a genetic algorithm. The selected optimal feature set was then fed into a multilayer neural network classifier [153]. A total of 300 images were used for performance evaluation, divided into a training set and a test set (150 images each). The training set is formed by an abnormal and normal image subset (75 each). Another method, proposed by Giancardo *et al.* [89] starts with a background normalization using a median filter and morphological reconstruction. An initial localization of exudates is performed with Kirsch filters [116]. The colour analysis and wavelet decomposition [154] were employed at the pixel level for feature extraction. For each set of pixels extracted, the following statistics were determined; median, mean, standard deviation, maximum and minimum, and combinations, resulting in a total of 48 features. The classifiers kNN, SVM, naive Bayes, and random forest were used to classify the exudates. the classifiers were trained and tested once for each image, each time holding out a different image from the training set which is then used for testing. Several combinations of classifiers/feature sets were performed, being SVM the best. Sánchez *et al.* [155] first applied colour normalization and contrast enhancement

of raw images through preprocessing operations. Then, a set of features were selected based on pixel colour information in RGB channels. Two training sets representative of non-exudate and exudate pixels are considered. To deal with intra-class variability between images, several training sets are automatically determined based on the edge responses. This avoids manual segmentation. The exudates classification was made with a Fisher's linear discriminant [156] analysis. Agurto *et al.* [157] employed a multiscale amplitude modulation frequency modulation (AM-FM) approach for differentiating normal and abnormal fundus images. First, optimal thresholding of instantaneous amplitude components extracted from multiple frequency scales, generate candidate exudate regions. For each candidate region, features, like colour, shape, and texture are used for classification. The classification is performed using partial least squares [158]. From a total of 652 images, 52 were selected to perform the training. Pereira *et al.* [159] used the ant colony optimization [160] method for exudates segmentation. The image is first filtered and normalized. Then an ant colony optimization performs a heuristic search and probabilistic growing, forming a segmentation of the image with candidate regions. Exudates were segmented using connected component analysis and a hard threshold.

3.1.3 Performance of the exudates segmentation algorithms

Three widely used performance measures, namely Sensitivity (Sen), Specificity (Spe), and Accuracy (Acc) are used for evaluation purposes [51, 52]. They are expressed as follows:

$$Sen = \frac{TP}{(TP + FN)} \quad (3.1)$$

$$Spe = \frac{TN}{(FP + TN)} \quad (3.2)$$

$$Acc = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (3.3)$$

where true positive (TP) is the correctly classified lesion regions, false positive (FP) is the non-lesion regions detected as lesion, true negative (TN) is the correctly classified non-lesion regions, and false negative (FN) is the lesion regions wrongly classified as non-lesion regions. The performance of the published methods are evaluated on an image based criteria or lesion based criteria. In the image-based criteria, an image is categorized as *normal* if it contains no lesion or *abnormal* when it contains at least one lesion. Due to the possible presence of noisy pixels, some authors only consider true exudates when the detected final region is larger than 30 pixels [66]. In the lesion based criteria considers the position and counts individual lesions. This requires the manually segmented groundtruth images created by clinical experts. This evaluation is done by counting the number of pixels which are correctly classified [138]. Other authors like Giancardo *et al.* [161] and Sánchez *et al.* [75] consider a component lesion validation. A connected component candidate region is considered a TP if, and only if, the ratio of the overlap between the candidate region and the groundtruth region is greater than or equal to 0.2. Thus, the connected components of the segmentation and the groundtruth can be classified as TP, FP, and FN [138]. For a more precise comparison between exudate segmentation results, the lesion base criteria is more appropriate. From a clinical point of view, especially for screening applications, an image-based evaluation is important as well [109]. Table 3.1 shows the performance of some mentioned methods. Rule-based methods (top part of the table) are separated from the machine learning methods (bottom

Table 3.1: Comparison of exudates segmentation methods.

Diaretdb1 Dataset						
Method	Sen(%)	Spe(%)	Acc(%)	Evaluation criteria	Techniques	Evaluation Method
Wisaeng *et al.* (2019) [52]	97.05	97.18	97.14	Lb	MM	All images
Kaur *et al.* (2018) [109]	71.34	88.56	85.34	Lb	DT	All images
Welfer *et al.* (2010) [105]	70.48	**98.84**	-	Lb	MM	All images
Karkuzhali *et al.* (2019) [15]	**98.87**	92.31	**99.67**	Ib	MM	-
Kaur *et al.* (2018) [109]	91.00	94.00	89.00	Ib	DT	All images
Khojasteh *et al.* (2019) [150]	99.00	96.00	98.20	Lb	ResNet+SVM	All images(Ex/Non-Ex patches)
Kusakunniran *et al.* (2018) [142]	89.10	**99.70**	96.20	Lb	Graph cut	All images(84% Test)
Fraz *et al.* (2017) [136]	92.42	81.25	87.72	Lb	Ensemble classifier	All images(71% Test)
Harangi *et al.* (2014) [140]	86.00	-	-	Lb	SVM	All images(71% Test)
Akram *et al.* (2012) [132]	**99.65**	98.35	**98.92**	Lb	Bayesian classifier	All images(50% Test)
Huang *et al.* (2021) [51]	98.21	91.38	98.01	Ib	CNN	All images(47% Test)
Long *et al.* (2019) [145]	97.50	97.80	97.70	Ib	FCM+SVM	All images(10-Fold crx val)*
Javidi *et al.* (2019) [146]	96.00	83.00	-	Ib	BAMCA	All images(95% Test)
Liu *et al.* (2017) [141]	83.00	75.00	79.00	Ib	Random forest	All images(69% Test)
Harangi *et al.* (2014) [140]	92.00	68.00	82.00	Ib	SVM	All images(71% Test)

Sen - Sensitivity; Spe - Specificity; Acc - Accuracy; Lb - Lesion-based; Ib - Image-based; MM - Mathematical morphology; DT - Dynamic thresholding; FCM - Fuzzy c-means; BAMCA - Bi-level adaptive morphological component analysis; crx val - Cross validation; Ex - Exudate; * - The cross validation was made in the E-Ophtha-EX dataset. Best values appear in bold.

part of the table).

The machine learning method proposed by Akram *et al.* [132] reports the highest sensitivity value (99.65%) and the highest accuracy value (98.92%). For the rule-based methods, the highest sensitivity and accuracy value was reported by Karkuzhali *et al.* [15] (98.87% and 99.67% respectively). Note, however, that no information is given about the evaluation criteria. Only two authors evaluated their methods in both lesion-based and image-based criteria, namely Kaur *et al.* [109] and Harangi *et al.* [140]. In both methods, the lesion-based criteria achieve lower performance values than the image-based criteria. This is important because in clinical practice it is also very important to detect the presence or absence of exudates, for screenings purposes. A precise segmentation is important to follow up the evolution of the disease. In general, several difficulties were reported by all the mentioned methods. Specifically, low contrast images, shading artefacts due to uneven illumination and retinal curvature, exudates heterogeneity in terms of size, shape, and magnitude, and resemblance of the exudates with other bright structures[15]. For the deep learning approaches, the major difficulty is the limited number of labelled data and the imbalanced distribution of data of different classes [150]. Moreover, deep learning based approaches also reveal a strong dependency of the data source, which typically requires retraining for every different dataset requiring more labelled data and leading to a problem of robustness and reliability of the models.

Between the rule-based methods and also for the descriptors definition of the machine learning based methods, it becomes relevant the importance that the scale might have for reliable detection of exudates [136, 157].

3.2 Optic disc

The OD is one of the main anatomical landmarks in a fundus image. Normally, it appears as a relatively circular yellowish disc, brighter than its surroundings. It is also the region where the retinal veins and arteries emerge and spread, covering the retina [162]. Although the OD main features and characteristics are relatively easy to describe, intensity and contrast inside the OD are variable. Furthermore, the OD can have a modified appearance by different types of retinal lesions or in pathological retinas. This makes the automatic localization of the OD a difficult task [55, 62, 163]. Many methods have been proposed for the detection of OD. Once again, they will be divided into rule-based methods and machine learning methods in the following.

To evaluate the OD detection algorithms, eight publicly available datasets are considered. These are the Stare [42], Drive [43], Diaretdb0 [44], Diaretdb1 [45], Messidor [46], ROC [48], E-ophtha-EX [49], and HRF [50] datasets. The Drive dataset is composed of 40 images, equally divided into a training subset (20 images) and a test subset (20 images). The ROC dataset is composed of 100 images separated into test/training subsets (50 images each). The HFR dataset is composed of 45 images equally divided between three categories(normal, DR, and glaucoma). No test/training subsets are provided. The Stare (20 images), Diaretdbv1 (89 images), Diaretdbv0 (130 images), and Messidor (1200 images) are not separated in test/training subsets. Only the HFR dataset provides the OD localization groundtruth. These datasets are described in detail in appendix A.

3.2.1 Rule-based methods

The rule-based methods are usually based on appearance-based approaches, the convergence of blood vessels approaches, and model-based approaches. Combinations of several approaches can also be found [164].

The first rule-based methods used *appearance-based* approaches. These methods rely on shape characteristics and the fact that the OD is usually the brightest region in a normal fundus image. Sinthanayothin *et al.*-[165] proposed a method that searches for rapid variation in intensity pixels values. The region with a high variance of the intensity of adjacent pixels is used for the OD location recognition.

Walter and Klein *at al.* [101] localized the OD in the luminance channel of the HLS colour space. First, a shade-correction algorithm is applied to remove slow image background variations. This is done by subtracting the luminance channel image with a filtered image of itself. The filtered image is computed by a larger median or mean filter, by homomorphic filtering [166] or by alternating sequential filters [102]. The final OD location is the pixel with the maximum intensity value.

More recently, Lu [167] proposed a circular transformation designed to capture both the circular shape of the OD and the image variations across the OD boundary. A recent improvement to this method was proposed by Reza [168] with the analysis of specific angles and intensity variations around the circular operator. Moreover, Lu and Lim [59] also proposed a line operator designed to capture circular brightness structures. This allows the evaluation of the image brightness variation along multiple line segments of specific orientations in each fundus image pixel. The orientation of the line segment with the minimum/maximum variation has a specific pattern used to locate the OD. Most of these methods result in effective OD localization in normal fundus images. However, they tend to exhibit difficulties in images with pathological clinical manifestations.

Convergence of blood vessels approaches are based on the fact that the major blood vessels of the retina emerge from the OD.

One of the earlier methods was proposed by Hoover *et al.* [169] with a fuzzy convergence approach, based on a vessel map voting type algorithm, that determines the convergence point near the OD. Initially, the retinal vasculature is segmented at different scales to reinforce the detection of different convergence points. A new algorithm called vessel fuzzy convergence was applied. This algorithm is a voting-type method that works on the spatial domain of the image. Each segmented vessel is modelled by a fuzzy segment, which contributes to a cumulative voting image. The output from the algorithm is a convergence image. The strongest point of convergence was identified as the OD location.

Mahfouz *et al.* [58] developed a technique where two projections of certain image feature maps to encode the x and y coordinates of the OD are obtained. The resulting 1D projections are then searched to determine the localization of the OD. Two feature maps are used. The first feature map

is constructed by computing the difference between the vertical and horizontal edge maps of the fundus image and dividing the result by the intensity map. The second feature map is defined by the summation of the vertical and horizontal edge maps and multiplying the result by the intensity map.

Following the work of Mahfouz *et al.* [58], Dietter *et al.* [61] estimated a preliminary OD location employing two sliding windows in the x and y direction. The retinal vessels are then segmented [170] and the vessel orientation and intensity are combined with the sliding windows approach to improving the OD location.

Welfer *et al.* [105] proposed a new adaptive method for the automatic detection and segmentation of the OD using mathematical morphology. The vascular tree is used as the reference to locate the OD in fundus images. A set of morphological operators like the regional minima and top-hat by opening are combined to segment the retinal vessels. A morphological skeletonization [104] is applied to the segmented vessels to detect the vessel arcades centerline. Finally, an iterative pruning method removes the vessel arcades until they converge in the OD localization.

Zhang and Zhao [171] use a vessel distribution and directional characteristics. A feature map combining three vessel distribution characteristics, i.e., local vessel density, compactness, and uniformity, finds possible candidates of the horizontal coordinate of the OD. Then, based on the global vessel direction characteristic, a general Hough transformation [172] is performed identifying the vertical coordinates of the OD. To reduce the computational cost of the proposed algorithm, the vertical range of possible vertical OD coordinates was limited to specific values.

The method proposed by Panda *et al.* [62] considers a new algorithm designated by vessel symmetry line. This algorithm determines the line that divides the retinal vessel tree into approximately two halves. The OD location is then located through an iterative convergence process, exploring the local vessel symmetry.

Model-based approaches make use of the vasculature information and the fact that the vessel tree has a shape similar to a parabola. This approach tends to be very effective and reliable, even in the presence of retinal diseases.

An example of a model-based approach is the geometrical model proposed by Foracchia *et al.* [162]. The proposed algorithm makes use of the fact that all the retinal vessels are originated from the OD. The general direction of retinal vessels at any given position in the image is described by a geometrical parametric model. Specifically, the retinal vessels are modelled as two parabolas, with their vertexes located in the OD position. Using experimental data samples of vessel centerline points and corresponding vessel directions, provided by a vessel identification procedure, the model parameters were determined using a simulated annealing optimization technique. The estimated values provide the coordinates of the OD location.

Lalonde *et al.* [173] presented a two stages OD detection algorithm. The first stage performs a pyramidal decomposition [174] for large-scale object tracking. The second stage computes the OD contour using a Hausdorff-based template matching [175] technique on the edge map. The edge map is determined upon a Rayleigh probabilistic modelling [176] of the noisy edge distribution on the fundus image green channel. Each stage generates a set of OD localization possibilities. The final OD localization is calculated by determining the candidate with the highest overall confidence between the two stages.

Youssif *et al.* [64] proposed to fit the vasculature orientations on a directional model to detect the OD. The method starts with a normalization of the luminosity and contrast using illumination and adaptive histogram equalization [177]. Next, the retinal vessels are segmented using a 1D Gaussian matched filter. Consequently, a vessel's direction map of the segmented retinal vessels is obtained. The segmented vessels are morphologically thinned [104] and filtered using local

intensity, to determine the OD center candidates. Next, the difference between the original image and the filtered image is performed, and the result is resized into four different sizes. At each OD candidate, it is calculated the difference between the resized images and the vessel's directions at the surrounding area of each of the OD candidates is measured. The minimum difference gives the final OD localization.
Yu *et al.* [178] first performs an image illumination correction using an average filtering procedure. The OD candidates are then determined by a template matching technique using circular binary kernels. Furthermore, a directional matched filtering procedure was also implemented in the green channel image. The OD candidate with the highest response in the directional matched filtering is selected as the OD localization.
Ramakanth *et al.* [60] proposed a method that uses an approximate nearest neighbour field algorithm, to find the correspondence between a chosen OD reference image and a given query image. The correspondence provides a distribution of patches in the query image that are closer to the patches in the reference image. The likelihood map obtained from the distribution of patches in the query image is used for the OD localization [60, 179].

Other methods use a *combination of several approaches*. Giachetti *et al.* [163] starts by performing a rough vessel segmentation by subtracting a top-hat filtered image from the green channel image, and thresholding the result using the Otsu method. The segmented vessels are then removed from the fundus image with an inpainting algorithm [180]. To obtain an approximate OD location, both vascular and brightness-related priors are used. These priors are based on a simple probabilistic approach that combines the output of two detectors related to radial symmetry [181] and vessel density.
A novel approach for the localization of the OD using 1D intensity profile analysis was presented by Kamble *et al.* [182]. The proposed method uses time and frequency domain information for the localization of the OD. For the horizontal OD coordinate detection, a peak-valley detection procedure searches for the pixel with the maximum intensity variation in the spatial domain, and discontinuities detection in the wavelet frequency domain [154]. The vertical coordinate is determined by a peak-valley detection procedure similar to the horizontal coordinate.
The method proposed by Uribe-Valencia [164] after image enhancement, image partitioning, and local thresholding enables the selection of the OD candidate's locations. An intensity profile analysis is performed at each OD candidate in terms of the number and the distance of signal peaks. A specific signal pattern was defined as a representative of a true OD based on the number and distance of signal peaks. The OD candidate with an intensity profile analysis more similar to the reference signal pattern is chosen as the final OD location.

3.2.2 Machine learning methods

Several machine learning methods were also proposed for OD localization. More recently, some deep learning methods have also been developed [183].

The method proposed by Abràmoff *et al.* [40] defines the OD localization as a regression problem. A kNN regression [128] is used to predict the distance in pixels to the object of interest at any given location in the image based on a set of features measured at that location. The proposed method combine cues measured directly in the fundus image with cues derived from a segmentation of the retinal vasculature to extract the features. The features include vessel width, number of vessels, vessel orientation, and vessel density, as well as pixel intensities information. A distance predic-

tion is made for certain image locations, and the point with the lowest predicted distance to the OD is selected as the OD location. For training, a private dataset of 1000 images was divided into a test and training subsets.

The method proposed by Roychowdhury *et al.* [184] first resizes the fundus image and performs morphological reconstruction using a circular structuring element. The resulting bright regions are classified as probable OD regions and non-OD regions using six features and a Gaussian mixture model classifier. The features are based on structural and pixel intensity values. The classified signalizes the probable OD region with the maximum Vessel-Sum and morphological solidity [104] is considered the best candidate region for the OD location. To perform the feature-based classification of the OD regions, 50 images from five public datasets (Drive, Diaretdbv1, Diaretdbv0, ChaseDB1 [185] and Messidor) are aggregated in a training subset. To determine the optimal preprocessing parameters, a preprocessing dataset, comprises the previously training subset with 10 images from the Stare dataset. The algorithm is tested on the remaining images from the six datasets.

More recently, deep learning approaches were also presented. The first deep learning approach was developed by Al-Bander *et al.* [183] a deep multiscale sequential CNN is designed and trained for the OD localization. The method was trained in the Kaggle [1] dataset and tested in the Messidor dataset.

Meyer *et al.* [186] proposed a pixel-wise regression task capable of simultaneously OD and fovea detection. The regressed quantity consists of the distance from the closest landmark of interest. This distance is detected by a fully-convolutional deep neural network. The model was trained and evaluated on the Messidor dataset (50% for training).

Finally, a CNN proposed by Bajwa *et al.* [187] is used for localizing the OD from fundus images. The proposed algorithm consists of three major modules, notably, a region proposal network, a CNN classifier, and a bounding box regression. The region proposal network generates several random rectangular object proposals and their confidence scores. These proposals are fed to a CNN that classifies it as OD or non-OD. Then bounding box regression is performed to fit the rectangle closely to the classified OD to provide a more precise OD location. The Messidor dataset was separated into training and a test subset with 100 and 1100 images respectively. A private and another publicly available dataset (IDRiD [188]) were also used. The training process adopted a 5-fold cross-validation approach.

3.2.3 Optic disc detection evaluation

The HRF dataset provides the OD region groundtruth. For the remaining datasets, a manual OD center label is made by a retinal expert. These datasets are described in Appendix A. The OD localization performance is usually made using an accuracy metric [162, 169]. Accuracy reflects the ability of an algorithm to correctly determine the OD center based on equation 3.3. In this evaluation scenario context, TP is a true positive OD localization, TN a true negative OD localization, FP a false positive OD localization, and TN a true negative OD localization [58].

[1] Kaggle Diabetic Retinopathy Detection, Accessed July 2022, www.kaggle.com/competitions/diabetic-retinopathy-detection/overview/description

Table 3.2: Optic disc localization methods comparison on several datasets.

Stare dataset					
Methods	Acc(%)	Fail	Evaluation Criteria	Techniques	Evaluation Method
Dietter *et al.* (2019) [61]	90.10	8	Boundary	Convergence	All images
Reza [168] (2018)	81.48	15	Boundary	Appearance	All images
Panda *et al.* (2017) [62]	97.53	2	60 pix	Convergence	All images
Kamble *et al.* (2017) [182]	95.07	4	Boundary	Combination	All images
Zhang and Zhao (2016) [171]	**98.77**	1	Boundary	Convergence	All images
Ramakanth *et al.* (2014) [60]	93.83	5	60 pix	Modelling	All images
Lu and Lim (2011) [59]	96.30	3	Boundary	Appearance	All images
Lu (2011) [167]	**98.77**	1	Boundary	Appearance	All images
Mahfouz *et al.* (2010) [58]	92.59	6	60 pix	Convergence	All images
Youssif *et al.* (2008) [64]	**98.77**	1	60 pix	Modelling	All images
Foracchia *et al.* (2004) [189]	97.53	2	60 pix	Modelling	All images
Hoover *et al.* (2003) [169]	88.89	9	60 pix	Convergence	All images
Roychowdhury *et al.* (2016) [184]	**98.77**	1	Boundary	GMM	All images*

Drive dataset					
Methods	Acc(%)	Fail	Evaluation Criteria	Techniques	Evaluation method
Uribe-Valencia *et al.* (2019) [164]	**100.00**	0	Overlapping	Combination	All images
Dietter *et al.* (2019) [61]	**100.00**	0	Boundary	Convergence	All images
Reza (2018) [168]	97.50	1	Boundary	Appearance	All images
Panda *et al.* (2017) [62]	**100.00**	0	60 pix	Convergence	All images
Kamble *et al.* (2017) [182]	**100.00**	0	Boundary	Combination	All images
Ramakanth *et al.* (2014) [60]	**100.00**	0	60 pix	Modelling	All images
Lu and Lim (2011) [59]	97.50	1	Boundary	Appearance	All images
Mahfouz *et al.* (2010) [58]	**100.00**	0	60 pix	Convergence	All images
Welfer *et al.* (2010) [105]	**100.00**	0	60 pix	Convergence	All images
Youssif *et al.* (2008) [64]	**100.00**	0	60 pix	Modelling	All images
Bajwa *et al.* (2019) [187]	97.50	1	Overlapping	RCNN	All images*
Roychowdhury *et al.* (2016) [184]	**100.00**	0	Boundary	GMM	All images*

Diaretdbo dataset					
Methods	Acc(%)	Fail	Evaluation Criteria	Techniques	Evaluation method
Uribe-Valencia *et al.* (2019) [164]	**100.00**	0	Overlapping	Combination	All images
Dietter *et al.* (2019) [61]	97.69	3	Boundary	Convergence	All images
Reza (2018) [168]	96.15	5	Boundary	Appearance	All images
Panda *et al.* (2017) [62]	96.92	4	60 pix	Convergence	All images
Kamble *et al.* (2017) [182]	97.69	3	Boundary	Combination	All images
Ramakanth *et al.* (2014) [60]	98.46	2	60 pix	Modelling	All images
Lu and Lim (2011) [59]	99.23	1	Boundary	Appearance	All images
Mahfouz *et al.* (2010) [58]	98.46	2	60 pix	Convergence	All images
Roychowdhury *et al.* (2016) [184]	**100.00**	0	Boundary	GMM	All images*

Diaretdb1 dataset					
Methods	Acc(%)	Fail	Evaluation Criteria	Techniques	Evaluation method
Uribe-Valencia *et al.* (2019) [164]	**100.00**	0	Overlapping	Combination	All images
Dietter *et al.* (2019) [61]	98.88	1	Boundary	Convergence	All images
Reza (2018) [168]	94.38	5	Boundary	Appearance	All images

Panda (2017)*et al.* [62]	**100.00**	0	60 pix	Convergence	All images
Kamble (2017)*et al.* [182]	97.75	2	Boundary	Combination	All images
Ramakanth *et al.* (2014) [60]	98.88	1	60 pix	Modelling	All images
Lu and Lim (2011) [59]	98.88	1	Boundary	Appearance	All images
Mahfouz *et al.* (2010) [58]	97.75	2	60 pix	Convergence	All images
Bajwa *et al.* (2019) [187]	**100.00**	0	Overlapping	RCNN	All images*

Messidor dataset

Methods	Acc(%)	Fail	Evaluation Criterion	Techniques	Evaluation Method
Dietter *et al.* (2019) [61]	98.00	24	Boundary	Convergence	All images
Giachetti *et al.* (2018) [163]	**99.83**	2	1R	Combination	All images
Panda *et al.* (2017) [62]	99.75	3	60 pix	Convergence	All images
Kamble *et al.* (2017) [182]	99.75	3	Boundary	Combination	All images
Ramakanth *et al.* (2014) [60]	99.42	7	60 pix	Modelling	All images
Yu *et al.* (2012) [178]	99.08	11	1R	Modelling	All images
Lu (2011) [167]	99.75	3	Boundary	Appearance	All images
Bajwa *et al.* (2019) [187]	99.17	10	Overlapping	RCNN	5-Fold cross val*
Meyer *et al.* (2018) [186]	98.94	-	1R	CNN	All images(50% Test)
Al-Bander *et al.* (2018) [183]	97.00	36	1R	CNN	All images*
Roychowdhury *et al.* (2016) [184]	**100.00**	0	Boundary	GMM	All images*

ROC dataset

Methods	Acc(%)	Fail	Evaluation Criterion	Techniques	Evaluation Method
Panda *et al.* (2017) [62]	**99.00**	1	60 pix	Convergence	All images

E-ophtha-EX dataset

Methods	Acc(%)	Fail	Evaluation Criteria	Techniques	Evaluation Method
Uribe-Valencia *et al.* (2019) [164]	98.78	1	Overlapping	Combination	All images
Panda *et al.* (2017) *et al.* [62]	**100.00**	0	60 pix	Convergence	All images

HRF dataset

Methods	Acc(%)	Fail	Evaluation Criteria		Evaluation Method
Dietter *et al.* (2019) [61]	**100.00**	0	Boundary	Convergence	All images
Panda *et al.* (2017) [62]	**100.00**	0	60 pix	Convergence	All images
Kamble *et al.* (2017) [182]	**100.00**	0	Boundary	Combination	All images

Acc - Accuracy; Fail - number of images failed by the method; CNN - convolutional neural network; RCNN - Regions with convolutional neural network; GMM - Gaussian mixture model; crx val - cross-validation * - Training is performed on other dataset. Best values appear in bold.

Table 3.2 shows some results in the literature for the OD localization performance for several datasets. For each dataset, the rule-based methods are in the top lines of the table 3.2 and the machine learning methods are in the bottom after the separating line. For the ROC, E-ophtha-EX, and HRF datasets, only rule-based methods are shown. Several criteria can be found in the literature to validate the OD localization. These appear in *Evaluation Criteria* column in Table 3.2. The *Boundary* evaluation criterion was defined by Hoover and Goldbaum [169]. In this case, the OD center localization is correct if it is inside the OD contour, including its borders. Some authors consider the OD localization correct if it is positioned within 60 pixels of the groundtruth center. This corresponds to the (*60 pix* evaluation criteria). This criterion can also be applied in images with no visible OD boundaries [162, 64]. Al-Bander *et al.* [183] and Meyer *et al.* [186] considers a

1R criterion, where R refers to the OD radius. The detected OD localization is considered correct if its distance to the groundtruth is less or equal to the R value in each image. This criterion is defined as 1R. Finally, Bajwa *et al.* [187] and Uribe-Valencia *et al.* considers a criterion based on the overlapping of the calculated and groundtruth OD regions. If this overlapping is larger than 50% then the OD is correctly detected. This criterion is defined as *Overlapping*. Although the anatomical and morphological characteristics of the OD are easily described, its detection from fundus images can be a challenging task, particularly in pathological fundus images. Hence, ideally, an OD localization method must be effective in different datasets. From Table 3.2 it is possible to observe that the majority of the proposed methods for OD localization are rule-based. Considering the applicability in multiple datasets, the most relevant rule-based methods are Panda *et al.* [62] and Uribe-valencia *et al.* [164]. For the machine learning methods, the most relevant are Roychowdhury *et al.* [184] and Bajwa *et al.* [187]. Not surprisingly, the lower values of accuracy are found in the Stare dataset, which is formed by pathological fundus images with different resolutions. This indicates the need for robust OD localization algorithms.

3.3 Retinal vasculature

Several systemic and non-systemic diseases have clinical manifestations through changes in the retinal blood vessels [190]. Furthermore, the eye is unique since it's the only region of the human body where the vascular network can be observed in-vivo [191]. The analysis of the retinal vasculature is fundamental in glaucoma, DR, hypertensive retinopathy, cardiovascular diseases, and other cerebral diseases as stroke [190, 192]. In DR, vascular changes include neovascularization and hemorrhages. Hypertension and atherosclerosis alter the diameter of retinal arteries and veins. Furthermore, thinning of the arteries and widening of the veins is associated with an increased risk of stroke and myocardial infarction [190].
The manual segmentation of the retina vessels is a very time-consuming task prone to errors [193]. Hence, the automatic quantification of the retinal vessels as well as the vascular features, such as the length, width, tortuosity, branching pattern, and arteries/veins differentiation can provide important insights to a proper diagnose and follow-up of the previously mentioned diseases. Some of these vessel properties can be measured by either vessel segmentation or vessel centerline detection methods. A brief review of both types of models is made in the following.
The majority of the retinal vessels segmentation methods are evaluated on three publicly available datasets, namely the Drive [43], Stare [42] and the HRF [50] datasets. The Drive dataset is composed of 40 images, equally divided into a training subset (20 images) and a test subset (20 images). Both subsets have the corresponding groundtruth images. The Stare dataset is composed of 20 images with the correspondent groundtruth, with no training/test subsets. The HFR dataset is composed of 45 images equally divided between three categories,(normal, DR and glaucoma). No training/test subsets are provided. These datasets are described in detail in Appendix A.

3.3.1 Vessel segmentation methods

The accurate segmentation of the retinal vessels is a challenging task for several reasons; the presence of lesions, exudates, or haemorrhages; other retinal structures, like the OD and fovea; a wide range of scales; a variable contrast between the vessels and retina background, and the central reflex in major arteries [193]. Many methods have been reported for the segmentation of retinal vessels. In this review, they are divided into two major groups, Rule-based methods and Machine learning methods [193, 194].

3.3.1.1 Rule-based methods

Rule-based methods use specific properties of the vessels structures in an algorithmic framework [194]. Typically, these methods approaches can be further divided into kernel-based techniques, vessel-tracking, mathematical morphology-based, multiscale, model-based, and region growing.

Kernel-based techniques (also known as matched-filter techniques) convolve a 1D kernel with the fundus image. The kernel is designed to model a retinal vessel with some unknown position and orientation. The kernel filter response is proportional to the presence of the feature. Kernel-based techniques are particularly effective when used in conjunction with additional processing techniques [195]. One of the earliest kernel-based methods was proposed by Chaudhuri *et al.* [196] where a Gaussian profile is used to estimate the shape of the cross-section of the retinal vessel. The kernel is applied to extract piecewise linear regions of retinal vessels. The template is rotated

at several orientations, and the highest matched-filter response is chosen for each pixel. A threshold is then applied to obtain the vessel map. Post-processing steps are used to trim and extract the vessel regions. Zhang *et al.* [197] improved the method proposed by Chaudhuri *et al.* [196] using double-sided thresholding to decrease the false-positives response to nonlinear boundaries.
Hoover *et al.* [198] proposed a method based on vessel's local and region-based attributes using a threshold probing method on a matched-filter response image. The pixels that are not categorized as vessels are reanalysed for further probing. Each pixel is classified using local and region-based properties. Matched-filter operations react not only to vessels but also to background boundaries, leading to an increase of false-positive responses.
Zhang *et al.* [32] proposed a matched-filter with the first-order derivative of Gaussian (FDOG) to extract the vessels, designated by MF-FDOG. The use of different scales makes it possible to extract both small and wider vessels. A logical OR operation merges the different scales of response. The MF-FDOG approach decreases the false positives created by the matched-filter operation and identifies several thin vessels overlooked by a single-scale matched-filter operation.
A novel matched-filter approach using Gumbel probability density function kernel [199] was proposed by Singh and Srivastava [200]. First, an image contrast enhancement was performed. The matched-filter response image was then thresholded by an entropy-based optimal approach. The final vessel segmentation was obtained by removing artifacts through length filtering.
Vessel tracking techniques track the vessels based on manual or automatic selection of some seed points, leading to the detection of the vessel midline steered by local data. One of the advantages of these approaches is that only pixels near the seed points are analyzed. However, complex situations like vessel crossings and vessel terminations are difficult to be analyzed with this approach.
Chutatape *et al.* [201] uses a combination of Gaussian and Kalman filters [202] for retinal detection. The second-order Gaussian matched-filtering is used to detect the vessel midline center point. The vessel tracking is performed starting from the OD boundary and using the defined midlines. The Kalman filter is used to detect the subsequent vessel segment position. Odstrcilik *et al.* [50] proposed a methodology based on multi-orientation matched-filtering. Five 1D filters are convolved with the fundus image at 12 different orientations. The resulting filtered images are then fused by selecting the local maximum response for each pixel. The fused image is then thresholded to obtain a binary map of the blood vessel tree. Morphological operations are used to remove noise artifacts or other retinal structures.
Mathematical morphology techniques filters the fundus image with a specific structuring element [104]. These morphological filters are designed to preserve image shapes similar to the structuring element, while they modify dissimilar shapes. The morphological approaches are usually successful for discarding both linear and nonlinear noise forms [96]. Usually, mathematical morphology-based methods exhibit difficulties in vessel crossings, and elongated and tortuous vessels [195].
Zana and Klein [203] proposed a mathematical morphology and curvature-based approach for the detection of vessel structures. Morphological filtering is first used in larger vessels. Next, a cross curvature evaluation distinguishes vessels from background pixels.
Mendonça and Campilho [191] proposed a three-stage approach; preprocessing, vessel centerline identification, and vessel extraction. The prepossessing stage normalizes the image background. Thin vessels are enhanced with a bank of line detection filters with different orientations. In the second stage, a set of Difference of Offset Gaussian (DoOG) filters extracts the vessel centerline. The process is complete with a region-growing method for vessel centerline reconnection. The vessel extraction stage uses a modified top-hat transform with different structuring elements to enhance vessels of variable widths. Next, a morphological reconstruction extracts the binary vessel

map.

A differential filtering and morphological processing approach for the extraction of the retinal vessels was proposed by Fraz *et al.* [204]. The first-order derivative of Gaussian is used to detect the centerlines. Then, a directional morphological top-hat operation is performed to determine the structure and direction map of the vessel. Next, a bit plane slicing extracts the vasculature map. The final binary vessel map is obtained with a combination of the vessel centerlines and the vasculature map. Imani *et al.* [87] proposed a combination of morphological component analysis and Morlet wave transform [134] for vessels segmentation. This method assumes that an individual signal is a linear mixture of several morphological significant elements. The morphological component analysis technique is used to distinguish between the vessel and non-vessel pixels using suitable transforms. The Morlet wave transform enhances the retinal vessel image, and an adaptive threshold produces the final binary image.

Multiscale based techniques perform vessel segmentation on different image scales. In multiscale approaches, a scale-space is first created from the original fundus image. Larger vessels are segmented from larger (low resolution) scales, and thin vessels are segmented from smaller (high resolution) scales. This approach is well suited to vessel segmentation due to the vessel variation diameter in fundus images.

Frangi *et al.* [73] proposed a vessel enhancement filtering process based on geometrical attributes. The filtering process enhances elongated structures. These vessels contrast enhancement is achieved by computing all eigenvalues of the Hessian matrix of the image. A multiscale approach is used to capture the width variation. A vesselness measure based on this method can be used to extract the vessels in fundus images. Based on the vesselness measure, Sofka and Stewart [205] proposed a method based on the likelihood ratio vesselness with matched-filter responses, confidence, and edge measures to extract the retinal vasculature.

Martinez-Perez *et al.* [39] presents a method based on scale-space analysis. The width, size, and orientation of the retinal vessels are calculated using two main geometrical features based on the magnitude of the gradient and the maximum principal curvature along the scale-space. The local maxima over scales of the gradient magnitude and the maximum principal curvature are used as features in a two-stage region growing scheme. In the first stage, the growth is constrained to regions of low gradient magnitude together with spatial information using the 8-neighboring pixels. In the second stage, the previous constraints are relaxed to allow the definition of the borders between regions.

Yu *et al.* [206] introduced a multiscale enhancement and second-order local entropy-based technique for the segmentation of the retinal vasculature. After preprocessing, the eigenvalues of the second derivatives at different scales are calculated. The second-order local entropy [207] threshold scheme, segments the vessel tree. Finally, a rule-based decision step computes the geometric shape variance between vessels and background to decrease the false-positive detection rate.

A multiscale line identification method proposed by Nguyen *et al.* [208] extracts the vasculature in fundus images. This approach depends on the variation of the length of a line detector to achieve a multiscale response. These line responses at different scales are linearly combined to obtain a final binary image.

Azzopardi and Petkov [209] developed the trainable COSFIRE filters for the detection of vascular junctions. The COSFIRE filters rely on contour measures that exist along with concentric circles of given radii around a stated ROI. Gabor filters enhance elongated structures and boundaries. The combined responses of the filters are aggregated by weighted geometric mean criteria, extracting the retinal vasculature.

Annunziata *et al.* [193] proposed a new retinal vasculature segmentation technique that uses an

inpainting filter, called neighborhood estimator. Bright lesions, like exudates, are first filled with the inpainting procedure to reduce false-positive vessel detection. After preprocessing steps [210] a multiscale Hessian eigenvalue approach is used for retinal vascular enhancement. Percentile-based thresholding is used to obtain the final binary image.

Khawaja *et al.* [211] proposed a directional multiscale line detector for the segmentation of retinal vessels. The focus of this approach was the smaller vessels. First, a direction-specific image array using a band-pass directional filter bank is performed on the green channel. Next, a multiscale line detection is applied, followed by a Coherence-Enhancing Diffusion, which segmented the directional vessels. The final step involves a directional binarization operation to achieve the vasculature segmentation.

Yang *et al.* [212] makes use of the Frangi based multiscale level sets to segment retinal vessels from fundus images. First, the vascular structures are enhanced using the Frangi filter [73] with local optimal scales being obtained simultaneously. The enhanced image and local optimal scales are considered as inputs of the proposed level-set models that perform the vessel segmentation.

Model-based techniques use algorithms that abstractly model the variability of a certain class of objects in the retina. These algorithms consider shape variations, where the shape is a flexible curve, that deforms to match the retinal vasculature. The curve deformation is usually based on energy minimization and curve evolution algorithms [194].

Lam and Yan [213] proposed a vessel extraction method based on the divergence of the vector fields. The vessel centerlines are detected using the gradient vector field. Next, the elongated vessels are extracted using the Laplacian operator of a pixel. The same author proposed a regularisation-based multi-concavity model capable of dealing with both normal and pathological fundus images [214]. The differentiable concavity measure is designed to deal with bright lesions. The line-shape concavity measure removes low contrast lesions. The locally normalized concavity measure is suggested to handle the randomly scattered noise in fundus images. These concavity measures are combined based on their geometrical properties, enabling the extraction of the vessels in fundus images.

Kovács and Hajdu [215] proposed a template matching and contour reconstruction for the retinal vessels extraction. The method uses kernels based on the generalized Gabor function to detect the centerlines of the vasculature. Then the intensity features of vessel shapes computed in training datasets are modeled.

Wang *et al.* [216] suggested a region-based active contour model, based on intensity and vesselness metrics as autonomous random variables. These variables are used to build a multi-feature Gaussian distribution fitting energy capable of retinal vasculature extraction.

A new fractional filter and an algorithm for retinal blood extraction were proposed by Shukla *et al.* [14]. The proposed fractional filter was designed considering a weighted fractional derivative and an exponential weight factor. The algorithm is based on the fractional filter and the eigenvalue maps of a local covariance matrix. The local covariance matrix corresponds to the local principal component analysis (PCA) [217] that is used to extract the blood vessels.

Region growing techniques collect pixels that have similar characteristics together to define a region. The effectiveness of these techniques depends on the seed locations and growth conditions. The blood vessels have a very wide gray-level distribution, making the region growing techniques highly dependent on the applied thresholds [96].

Palomera-Pérez *et al.* [218] described a parallel method for the vasculature segmentation of high-resolution fundus images. The described approach depends on data splitting, which allowed faster handling of those images. Data partitioning is made through horizontal partitioning for attribute extraction, and mixed (horizontal and vertical) partitioning for region growing.

Jiang *et al.* [219] proposed a region growing method that depends on spectrum information. This approach used the Fourier transform of vessel information to acquire its spectrum knowledge, according to which its initial feature direction will be detected. Next, a coupled boundary knowledge with initial feature direction records the vessel network middle points as the seed points. Finally, the region growing technique approach with a branch-based growth scheme extracts the vessels.
A retinal vasculature extraction based on level-set and region growing technique is proposed by Zhao *et al.* [220]. First, adaptive histogram equalization and a set of 1D Gabor filters are used to improve the blood vessels contrast. Next, an anisotropic diffusion filter is used to eliminate the uneven illumination. Finally, the region growing approach and a level set operation to identify the retinal vessels. Their results are superimposed to obtain the resulting binary image.
An automatic seed identification based on a novel binary Hausdorff symmetry and edge distance seeded region growing method for extraction of the retinal vasculature is proposed by Panda *et al.* [221]. The binary Hausdorff symmetry method directly computes a binary symmetry decision at every pixel, without image thresholding. The vessels are extracted by an edge distance seeded region growing approach.

3.3.1.2 Machine learning methods

As it is common, the machine learning methods initially require the *training process* where a retinal vessel segmentation model is created based on labelled data, followed by the *testing process* where the learned model is applied to unseen images. In the literature, the vessel segmentation methods are mainly divided into supervised approaches and unsupervised approaches. Supervised methods exploit some prior labelling information to decide whether a pixel belongs to a vessel or not. Unsupervised methods perform the vessel segmentation without any prior labelling knowledge. As supervised methods use classified data for the segmentation model definition, results in general in better performance than unsupervised methods. Some main techniques considered in the literature, are NN, fuzzy c-means, SVM, KNN, an ensemble of classifiers, and deep learning models.
Ricci and Perfetti [222] proposed an approach based on line operators and an SVM classifier to perform vessel segmentation. First, a line detector is convolved with the green channel of the fundus image. Additionally, two orthogonal line detectors extract a pixel feature array based on the pixel intensities. An SVM classifies the pixels as vessels or non-vessel. For the training, 20000 pixels were randomly selected in 20 images from both the Stare and Drive datasets. The classification was performed on the remaining images of each dataset using a cross-validation approach.
Roychowdhury *et al.* [223] proposed a three-step retinal vessel extraction algorithm. In the first step, a high-pass filtering and morphological reconstruction extract an initial binary vessel map. A total of 57 features are considered for each pixel in the binary vessel map. These features use information regarding the pixel neighbourhood and pixel gradient. In the second step, all residual pixels in the binary vessel map are classified using a Gaussian mixture-model classifier [224]. Finally, the major segments of the retinal vessels are combined with the categorized vessel pixels. The training was performed in the Drive training subset. The evaluation was made in the Drive test subset, and in Stare dataset using a leave-one-out approach.
A discriminatively trained fully connected conditional random field model for segmentation of retinal vessels was proposed by Orlando *et al.* [27]. This approach does not require human intervention for the training process. The retinal vasculature was modelled as an energy minimization problem in a conditional random field. Several filter responses (Gabor wavelets [134] and multiscale line detectors [208]) were used as vessels features. A structured output SVM was used in

the conditional random field learning process for retinal vessel extraction. The method was evaluated in the Drive, Stare, and HRF datasets. The Drive training subset was divided into two new subsets (training* and validation, containing 70% and 30% of the original training subset images) for the training stage. The evaluation was made in the Drive test subset. For the Stare dataset, a leave-one-out cross-validation approach was performed. For the HRF dataset, a training subset was made with 15 images. The remaining 30 images were used for testing.

A supervised deep neural network method was developed by Liskowski *et al.* [225] for the segmentation of the vessel map. The deep neural network is trained on a large sample of examples, previously pre-processed with a global contrast normalization and zero-phase whitening methods. The training samples were augmented using geometric transformations and gamma corrections. The test and training subsets of the Drive dataset were considered for the method validation. For the Stare dataset a leave-one-out approach validation was performed. From each training image, a sample of 20000 patches was selected for both datasets. For testing, all patches from each test image were extracted.

Other authors considered ensemble methods. Fraz *et al.* [185] proposed an approach using an ensemble classifier of bagged decision trees for extraction of retinal vessels. Decision trees work as a categorization model. The outputs of the classifiers are pooled using bootstrap aggregation. A 9D feature array is used, containing the vessel tree acquired from the orientation values of the gradient vector field, the morphological procedure, line strength computation, and Gabor filter's response. For the Drive dataset, 10000 random pixels were selected from the training subset and evaluated at the test subset. For the Stare dataset the classifier training was performed using 75000 pixels randomly extracted from each image. The evaluation was performed in the complete Stare dataset.

Wang *et al.* [226] proposed a feature and an ensemble learning-based technique, using a combination of a CNN and a random forest classifier. The CNN performs as a hierarchical feature extractor, and the ensemble random forest works as a classifier. The integration of the feature learning and the classifier procedures results in the automatic learning features from the raw fundus images. For the Drive and Stare datasets, the training of both feature extractor and predictor was conducted on 20 training images using 193325 patches and 192419 patches, respectively. About 10000 pixels per image were selected. Since the Stare dataset does not provide a testing subset, both the training and the evaluation were performed in the complete dataset.

A combination of hybrid context-aware features with a discriminative learning system was proposed by Cheng *et al.* [227]. The stroke width transform [228] and Webers local descriptors [229] extract the curvilinear structures. The responses of Gabor filter and a vesselness filter are used for vessel detection. Vessel attributes and filter-based attributes are used as an input for a random forest approach for pixel-level vessel extraction. For the Drive dataset, the classifier was trained by selecting 20000 pixels (10000 vessels/10000 non-vessel). The samples were randomly selected from each training image. The performance was evaluated in the test subset. For the STARE dataset, the training and testing were performed on the whole dataset using the same pixel criteria used for the Drive dataset. For the HRF dataset, three-fold cross-validation was used for the evaluation. Each fold is composed of 15 images, 5 from each of the three image categories. For each run, there were 30 training images and 15 testing ones. For each of the three training runs 10000 samples and 10000 non-vessel samples were randomly selected to train a decision tree.

Li *et al.* [37] proposed a deep learning method that considers the multiscale nature of the retinal vasculature. A fully attention-based network based on attention mechanisms that adaptively learn rich features representation and aggregate the multiscale information is proposed. First, the green channel image is extracted and preprocessed with a CLAHE contrast normalization procedure. Then a U-net network is combined with dual-direction attention modules. These modules are

attention maps produced by horizontal and vertical polling operations. The standard convolution operations are replaced by a selective kernel to obtain multiscale features of different receptive field sizes generated by soft attention. The Drive dataset training and test subsets were used for training and testing, respectively. For the Stare dataset, the original dataset was split in order and the first 10 images were used as a training subset, while the remaining 10 images were used as a testing subset. Each image of the training subsets was divided into patches. A random flipping for data augmentation was also used. The method performance evaluation was performed in the testing subsets.

Park *et al.* [230] presented a new conditional generative adversarial network-based on [231] with a *M* network structure. The generative adversarial network conducts retinal vessel segmentation by balancing losses through stacked deep fully convolutional networks. It consists of an M-generator with deep residual blocks and an M-discriminator with a deep network for more efficient training of the adversarial model. The scale-invariance of vessel segmentation of different sizes is supported by a multi-kernel pooling block, added between the stacked layers. The M-generator has down-sampling layers to extract features and up-sampling layers that perform the vessels segmentation from the extracted features. The M-discriminator has a deeper network similar to the down-sampling of the M-generator, but the final layer is constructed as a fully connected layer for decision-making. The Drive, Stare, and HRF datasets were used for performance evaluation. The Stare dataset was divided into test and training subsets of 10 images each, and the HRF dataset in a test (15 images) and training (30 images) subsets. All training subsets were augmented through transforming images using flipping and rotation.

3.3.2 Vessel centerline detection

Several methods were developed for vessel centerline detection. These methods aim the extraction the centerline or the skeletonization of the retinal vasculature. Vessel centerline is particularly useful in vessel morphology and anatomical properties, like vessel crossings, branching angles, and tortuosity. [232, 57]. One of the earliest vessel centerline detection methods was proposed by Sofka and Stewart [205], where a likelihood ratio combines matched-filter responses, confidence measures, and vessel boundary measures. The matched-filter responses are derived in scale-space to extract vessels of different widths. The vessel confidence measure is defined as a projection of a vector, formed from a normalized pixel neighbourhood, into a normalized ideal vessel profile. The vessel boundary measure and associated confidences are calculated at potential vessel boundaries. The combined responses form a 6D vector for each pixel. A training technique uses the 6D vector to produce a likelihood ratio that measures the “vesselness” for each pixel.

Morales *et al.* [233] proposed a method that determines significant points (bifurcations and crossovers) based on a morphological hit-or-miss transformation [104, 137]. Complex vessel intersections are determined by a post-processing stage. This stage focuses on the idea that the intersection of two vessels creates a close loop formed by the vessels. This effect can be used to differentiate bifurcation from a crossover.

Kromm *et al.* [17] proposed a capsule network [234] with an inception architecture [235]. The proposed inception capsule network is based on a shallow architecture that has fewer parameters and does not require data augmentation. The method was evaluated on the Drive dataset. For each image in the training, a subset were extracted with 80000 random patches. the testing was performed on all pixels of the test subset.

Xu *et al.*[232] proposed a deep semantics and multiscale cross-task aggregation networks. The combined network is featured by two sub-networks. The forepart is a deep semantics aggregation

sub-network that aggregates strong semantic information to produce more powerful features for vessel centerline detection. The tail part is a multiscale cross-task aggregation sub-network that explores complementary information to refine the results. For the Drive dataset, the method is trained in the training subset and tested on the test subset. For the Stare dataset, a 5-fold cross-validation is performed on the whole dataset.

3.3.3 Retinal vessels segmentation and centerline detection performance evaluation

The Drive and Stare datasets have two distinct manual groundtruth segmentation. The first human expert manual segmentation is considered the gold standard for testing and evaluation procedures [227]. A single manual groundtruth segmentation is provided for the HRF dataset. The vessels segmentation is usually evaluated based on Sensitivity (Sen), Specificity (Spe), and Accuracy (ACC) as defined in equations 3.1, 3.2 and 3.3. In the vessel segmentation context, accuracy is the proportion given by the ratio of the total number of correctly classified pixels (sum of true positives and true negatives) to the number of pixels (only pixels in the field-of-view are considered). Sensitivity measures the ability of the method to detect the vessel pixels. Specificity measures the ability to detect non-vessel pixels. where the true positive (TP) is the number of pixels identified as a vessel in both the ground truth and segmented image. True negative (TN) is the number of pixels classified as a non-vessel in the ground truth and the segmented image. False-negative (FN) is the number of pixels classified as non-vessel in the segmented image, but as a vessel pixel in the ground truth image. False-positive (FP) is the number of pixels marked as a vessel in the segmented image but non-vessel in the ground truth image. Some methods also report the area under the ROC (AUC). The ROC curve is a plot of the true-positive fraction versus the false-positive fraction for several cut-off values of a given algorithm parameter. Table 3.3 compares some of the previously described vessel segmentation methods.

For the vessel centerline evaluation, the same metrics defined in equations 3.1, 3.2 and 3.3 as well as the AUC are used. The groundtruth images are produced by a skeletonization process of the original groundtruth images at each dataset [232]. Table 3.4 compares some of the previously described vessel centerline detection methods. The reported performance of machine learning methods is typically better than rule-based methods. However, either do not use the full datasets or use a cross-validation methodology that typically favours the performance measures when compared with those that use the complete dataset. The highest AUC is reported by Li *et al.* [37] for the Drive and Stare dataset (98.95% and 99.24% respectively). For the HRF dataset, only the rule-based method proposed by Odstrcilik *et al.* [50] reported the AUC value (96.78%). It is worth mentioning that on datasets without a training set (Stare and HFR) the machine learning methods report average values of partial evaluations since the larger part of the tested dataset is used for training. This makes the comparison between methods difficult or even impossible. Although kernel-based approaches have been extensively proposed, they cannot handle vessel segmentation in fundus images with pathological clinical manifestations. The main advantage of the vessel tracking approaches is the effective detection of vessel crossings and bifurcations. The machine learning approaches, in particular, deep learning approaches, achieve more accuracy. These methods however require larger image samples and are computationally more expensive.

Table 3.3: Comparison of retinal vessel segmentation detection methods.

Drive Dataset						
Method	Acc(%)	Sen(%)	Spe(%)	AUC(%)	Techniques	Evaluation Method
Shukla *et al.* (2020) [14]	94.76	70.15	**98.36**	-	Model-based	All images
Khawaja *et al.* (2019) [211]	**95.53**	**80.43**	97.30	-	Multiscale	All images
Singh *et al.* (2016) [200]	95.22	75.94	-	92.87	Kernel-based	All images
Panda *et al.* (2015) [221]	95.39	73.37	97.52	-	Region growing	All images
Azzopardi and Petkov (2015) [209]	94.42	76.55	97.04	**96.14**	Multiscale	All images
Zhao *et al.* (2014) [220]	94.77	73.54	97.89	-	Region growing	All images
Odstrcilik *et al.* (2013) [50]	93.40	70.60	96.93	95.19	Vessel tracking	All images
Fraz *et al.* (2011) [204]	94.30	71.52	97.68	-	MM	All images
Mendonça and Campilho (2006) [191]	94.52	73.44	97.64	-	MM	All images
Martinez-Perez *et al.* (2007) [39]	93.44	72.46	96.55	-	Multiscale	All images
Li *et al.* (2020) [37]	**97.69**	81.45	**98.83**	**98.95**	U-Net	All images
Orlando *et al.* (2016) [27]	-	78.97	96.84	-	SVM	All images
Wang *et al.* (2015) [226]	97.67	**81.73**	97.33	94.75	CNN + RF	All images
Roychowdhury *et al.* (2014) [223]	95.20	72.50	98.30	96.20	GMM	All images
Fraz *et al.* (2012) [185]	94.80	74.06	98.07	97.47	Ensemble classifier	All images
Cheng *et al.* (2014) [227]	94.74	72.52	97.98	96.48	RF	All images
Ricci and Perfetti *et al.* (2007) [222]	95.63	-	-	95.58	SVM	All images

Stare Dataset						
Method	Acc(%)	Sen(%)	Spe(%)	AUC(%)	Techniques	Evaluation Method
Shukla *et al.* (2020) [14]	**95.73**	70.17	**98.63**	-	Model-based	All images
Khawaja *et al.* (2019) [211]	95.45	80.11	96.94	-	Multiscale	All images
Singh *et al.* (2016) [200]	92.70	79.39	-	91.40	Kernel-based	All images
Annunziata *et al.* (2016) [193]	95.62	71.28	98.36	-	Multiscale	All images
Panda *et al.* (2015) [221]	94.24	**84.03**	95.47	-	Region growing	All images
Azzopardi and Petkov (2015) [209]	94.97	77.16	97.01	95.63	Multiscale	All images
Zhao *et al.* (2014) [220]	95.09	71.87	97.67	-	Region growing	All images
Odstrcilik *et al.* (2013) [50]	93.41	78.47	95.12	**95.69**	Vessel tracking	All images
Fraz *et al.* (2011) [204]	94.42	73.11	96.80	-	MM	All images
Martinez-Perez *et al.* (2007) [39]	94.10	75.06	95.69	-	Multiscale	All images
Mendonça and Campilho (2006) [191]	94.40	69.96	97.30	-	MM	All images
Li *et al.* (2020) [37]	97.97	**85.05**	98.89	**99.24**	U-Net	All images(50%Test)
Orlando *et al.* (2016) [27]	-	76.80	97.38	-	SVM	crx val(leave-one-out)
Wang *et al.* (2015) [226]	**98.13**	81.04	97.91	97.51	CNN+RF	All images(pix for training)
Roychowdhury *et al.* (2014) [223]	95.10	77.20	97.30	96.90	GMM	All images(training in Drive Dataset)
Cheng *et al.* (2014) [227]	96.33	78.13	98.43	98.44	RF	Leave-one-out
Fraz *et al.* (2012) [185]	95.34	75.48	97.63	97.68	Ensemble classifier	All images(pix for training)
Ricci and Perfetti *et al.* (2007) [222]	95.84	-	-	96.02	SVM	All images(pix for training)

HFR Dataset						
Method	Acc(%)	Sen(%)	Spe(%)	AUC(%)	Techniques	Evaluation Method
Annunziata *et al.* (2016) [193]	**95.81**	71.28	**98.36**	-	Multiscale	All images
Panda *et al.* (2016) [221]	94.20	**81.59**	95.25	-	Region growing	All images
Odstrcilik *et al.* (2013) [50]	94.94	77.41	96.69	**96.78**	Vessel tracking	All images
Fraz *et al.* (2012) [185]	94.80	74.06	98.07	97.47	MM	All images
Park *et al.*(2020) [230]	**97.00**	69.48	**99.31**	-	GAN	All images(30%Test)
Orlando *et al.* (2016) [27]	-	78.74	95.84	-	SVM	crx val(Leave-one-out)
Cheng *et al.* (2014) [227]	96.47	**78.89**	98.65	-	RF	All images(pix for training)

Sen - Sensitivity; Spe - Specificity; Acc - Accuracy; MM - Mathematical morphology; GMM - Gaussian mixture-model; RF - Random forest; GAN - Generative adversarial network; crx val - Cross validation. Best values appear in bold.

Table 3.4: Comparison of retinal vessel centerline detection methods.

Drive Dataset						
Method	Acc(%)	Sen(%)	Spe(%)	AUC(%)	Techniques	Evaluation Method
Kromm *et al.* (2020) [17]	90.70	**75.20**	84.20	88.20	Inception capsule network	All images
Xu *et al.* (2020) [227]	**95.37**	74.22	**96.37**	**96.12**	Deep semantics+AN	All images

Stare Dataset						
Method	Acc(%)	Sen(%)	Spe(%)	AUC	Techniques	Evaluation Method
Xu *et al.* (2020) [227]	**96.89**	**77.30**	**97.45**	**97.81**	Deep semantics+AN	5-Fold crx val

Sen - Sensitivity; Spe - Specificity; Acc - Accuracy; AN - Aggregation networks; crx val - Cross validation. Best values appear in bold.

3.4 Microaneurysms detection

Microaneurysms appear in fundus images as small round dark dots, smaller than the thinnest visible vessels. Variations of size, shape, and illumination between fundus images, make the automatic detection of MAs a challenging task [48]. MAs detection is usually divided into three stages [31]. First, image preprocessing improves the image quality by removing non-uniform illumination and noise reduction [26]. The second stage aims at the extraction of as many MA candidates as possible [36]. The last stage classifies the MA candidates as true or false using some classification method [236, 31, 26]. Most of the proposed methods are based on a machine learning approach, but some rule-based methods, that do not require any learning mechanism, have also been proposed [237, 238, 239].
The datasets used to evaluate the MAs detection methods are the ROC [48], Latim [47], and E-ophtha-MA [49] datasets. Each dataset has the corresponding groundtruth. The ROC dataset is divided into test and training datasets (50 images each). However, the ROC test groundtruth was never released and was used in the past for contests purposes. Hence, in recently developed methods, only the ROC training dataset is used for performance evaluation. These datasets are introduced in appendix A.

3.4.1 Rule-based methods

The first method developed for MAs detection in fundus images was a hit-and-miss transform based on mathematical morphology, proposed by Øien and Osnes [240]. However, simple morphological methods are sensitive to changes in the size and shape of the structuring element [241]. Zhang *et al.* [32] proposed a multiscale Gaussian correlation filtering with dynamic thresholding for the extraction of the MA candidates. For each MA candidate, a set of features are extracted. A discrimination table with the minimum and maximum values for each feature is defined for MA candidate labelling as true or false.
Quellec *et al.* [47] proposed a template matching in sub-bands of wavelet transformed images to find the MA candidates. To improve the method performance, a search was made for the best-adapted wavelet within a lifting scheme framework [242]. The optimization process was based on a genetic algorithm followed by Powell's direction set descent method [243].
Antal and Hajdu [5] proposed an ensemble framework based on a combination of preprocessing methods and candidate detectors. A total of five preprocessing methods and five MA candidate extractors were evaluated. Simulated annealing was used as a search algorithm to select the best preprocessing and MA candidate extractors combinations. The sum of the several combinations gives a confidence value for each MA candidate to be true or false.

3.4.2 Machine learning methods

One of the first machine learning-based methods for MA detection was proposed by Niemeijer *et al.* [34]. First, the retinal vessels are removed employing a top-hat transform. Next, the image is enhanced with a 1D matched filter and thresholded to extract MA candidates. Finally, a kNN classifies the MA candidates using shape and intensity features. A dataset of 40 images was used to train the kNN classifier in the first candidate extraction step. A second set of 100 images was used to train and test the complete system.
Ram *et al.* [6] formulated the problem of MAs detection as a problem of target detection from clutter. In this context, the target is the MAs and the clutter is noise and other spurious regions. The MA candidates are extracted using morphological bottom-hat and opening operations [104]. Next,

two successive rejection-based strategies progressively reduce the number of clutter responses. These stages were designed to reject specific classes of clutter while passing the majority of true MAs, using a set of shape-based features. In the final stage, the degree of similarity of each MA candidate is compared with a true MA profile, and a score is assigned. This stage uses a composite set of features based on morphological and appearance-based information, and a Bayesian classifier. The final MAs are obtained by applying a threshold on the similarity score. For training, the ROC (100 images) and the Diaretdbv1 (89 images) datasets were used. The Diaretdbv1 dataset was divided into test and training subsets (68/21 images, respectively). using a 8-fold cross-validation was used for the method validation.

One of the main difficulties in template matching approaches is the design of suitable filters [241]. Furthermore, in multiscale frameworks, if scales are chosen empirically or are scarce for MAs characterization, the ability to detect all possible MAs becomes limited [26]. To overcome some of these difficulties, Adal *et al.* [36] propose a method where MAs detection is modeled as finding interest regions or blobs from an image with an automatic local-scale selection technique. Several scale-adapted region descriptors (Hessian [244], SURF [245], and Radon [239]) are introduced to describe these blob regions. A semi-supervised based learning approach [246], that requires a few manually annotated learning examples, was also proposed to train a classifier to detect true MAs. The training was made in a private dataset hosted by University of Tennessee Health Science Center (UTHSC) of fundus images [36]. A self-training and co-training semi-supervised learning approaches were considered for training using the ROC training dataset. The testing stage was made in the ROC test dataset.

Sparse representation-based methods were also proposed. The method proposed by Zhang *et al.* [241] is an improvement to their original work [32], with the inclusion of a dictionary learning with a sparse representation classifier. First, the MA candidates are detected with multiscale Gaussian filtering. Next, two sub-dictionaries for the MA candidates and non-MA candidates were found. The MA candidates previously detected are separated in true MAs and false MAs to compute the two sub-dictionaries. The ROC training dataset was further divided into training (80%) and test (20%) subsets. The final evaluation was performed in the remaining images of the test subset and in the ROC test dataset. Sparse methods depend on the original grey-scale features dictionary. Since fundus images exhibit inter and intra-variability in contrast and luminosity, using a single greyscale feature may affect the MA detection performance [23].

Pereira *et al.* [247] proposed a multi-agent approach, where an environment is designed in which the agents are situated and interact [248]. After, shade correction, a Gaussian matching filter is applied to enhance small and dark structures. Next, a Kirsch filtering [116] is performed to obtain edges of a two-pixel thickness. This enables the MAs model detection process, since MAs have specific gradient patterns. In the final stage, each pixel is considered an active agent. Based on the Kirsch edges information and the intensity profile in several directions, the agents interact with the surrounding environment, either becoming active or inactive. The agent validates its position as a true lesion using shape and intensity-based features analysis. The remaining pixels establish the detected MAs.

A statistical approach based on a mixture model-based clustering was proposed by Sanchéz *et al.* [237]. Features like colour, shape, and texture are extracted and fed to a three-class Gaussian mixture model. The classification is made on the assumption that each pixel belongs to one of the three classes: background, foreground (vessels, lesions, and OD), an outlier. The ROC training and test datasets were used for the training and evaluation stage of the algorithm.

Lazar *et al.* [68] proposed a method based on the analysis of the intensity profile using different scan lines at different directions. For each direction, the scan lines are translated horizontally and

vertically to cover all pixels of the image. Each 1D profile is binarized using adaptive thresholding selecting the MA candidates. Statistical measures of these profiles define MAs candidate features used by a na"ive Bayes classifier to perform MAs classification. Others classifiers (kNN and SVM) were considered. To determine the best classifier, the ROC training dataset was used both for training and testing. The ROC training dataset was used for method training, while the ROC test dataset was used for the method evaluation.

An approach based in [68] was used by Wu *et al.* [22]. After a similar feature detection algorithm, and the inclusion of new features based on the MAs surroundings, a total of 27 local and profile features were used to perform classification using AdaBoost and kNN. The E-ophtha-MA composed of 148 images was divided into training and test subsets of 74 images each. The ROC training dataset was divided into two new subsets for training and testing using 10-fold cross-validation.

Du *et al.* performs MA candidates extraction exploiting the regions where MAs may exist using a morphological local minimum region operator and local block filtering [249]. Thus, irrelevant background regions may be discarded. Then, multiple features based on the candidate profile, intensity, edge values, and saliency analysis are extracted to train the MA classifier. To distinguish MA from vascular regions, a series of descriptors based on the cross-section profile of the MAs are defined. An under-sampling boosting-based classifier (RUSBoost) is trained to classify the MA candidates. To evaluate the method's performance, 10-fold cross-validation was used in both the E-ophtha-MA and ROC training dataset.

Wang *et al.* [250] analysed the cross-section profiles by singular spectrum analysis. Each profile is correlated with a typical MA profile and described by a correlation coefficient. The correlation coefficient between each processed profile and a typical MA profile is measured and used as a scale factor to adjust the shape of the candidate profile. This process increases the difference between true MA and non-MA candidate profiles. A set of statistical features based on the intensity profiles are extracted and classified with a kNN. The ROC training dataset was divided in half for training and testing. The performance evaluation values are reported based on the full ROC training dataset.

Dai *et al.* [35] selects MA candidates by gradient vector analysis. The method starts with MA candidates localization based on the second order directional derivatives in different directions. Several features like contrast, intensity, edge values, texture, and region descriptors are used to classify MAs using a class-imbalance classifier (RUSBoost) [251]. To evaluate the method, the ROC training dataset was divided into two equal subsets for (training and test). 2-fold cross-validation was performed.

A new set of shape features were used by Seoud *et al.* [7] to extract MA candidates. The called dynamic shape features, do not require a precise segmentation to be classified. These features represent the evolution of the shape during an image flooding process, allowing to discriminate between lesions and retinal vessels. Random forests were used for MAs classification. An image-based evaluation was made in the Messidor dataset.

Recently, Melo *et al.* [63] proposed a sliding band filter for MAs enhancement, aiming at the extraction of MA candidates. Next, a combination of filter responses and features based on shape, intensity, color and contrast, is used by an ensemble of classifiers (RUSBoosted trees) for MAs detection. For training, the ROC training dataset was used. The RUSBoost algorithm applies random undersampling to create class-balanced training subsets from the initial imbalanced training set. The method was evaluated at the lesion level in the ROC training and E-ophtha-MA datasets. Furthermore, it is the only method that presents evaluation results considering different evaluation methods. In both datasets, the results of the 10-fold cross-validation and the 10-by-10 repeated cross-validation were selected. The advances of machine learning methods, as well as the

availability of publicly available large datasets of high-resolution images with the corresponding groundtruth, led to the proposal of deep learning methods [63].
After image preprocessing and MA candidates detection, Budak *et al.* [252] uses a deep CNN for MA final detection. Reinforcement sample learning was used to train the system. A set of patches (MA and non-Ma) were selected at the ROC training dataset to perform training and validation of the method using a holdout validation.
More recently, a two-stage CNN process was proposed by Eftekhari *et al.* [26]. An initial CNN extracts a MAs candidate probability map. A second CNN classifies the final MAs. The training and evaluation were performed in the ROC training dataset and the E-ophtha-MA dataset using 10-fold cross-validation.
Chudzik *et al.* [253] proposed a full CNN with batch normalization layers and Dice loss functions. 2-fold cross-validation was implemented for training and testing the algorithm in the ROC training and the E-ophtha-MA datasets.
Deep learning methods do not require hand-crafted features since theoretically, they can learn the most discriminative features for MAs detection [253]. However, they are time-consuming and require a large amount of training data [22, 24], not currently available. Poor image quality, differences in the size of MAs, the closeness of some MAs to vessels, also generate an imbalance of data in the learning process [26]. To improve the performance due to class imbalance, a two-stage deep learning framework was proposed by Bria *et al.* [254]. First, a deep cascade of long sequences of decision trees reduces the number of background samples. The remaining samples are fed to a CNN to obtain the final MAs. Lesion labeling requires the intervention of experienced experts to identify accurately each pathology limiting its performance [24]. To overcome these limitations, a combination of deep learning and domain knowledge was proposed by Orlando *et al.* [24]. It was verified that hand-crafted features are better in discriminating low contrast lesions, hence improving the MAs detection.

3.4.3 Microaneurysms detection performance evaluation

The earliest clinical sign of DR is the appearance of MAs in the retina [68]. Hence, both a lesion-level [48] (MAs detection) or image-level evaluation [75] (DR detection) can be performed.

3.4.3.1 Lesion-level performance evaluation

The MAs detection evaluation is made using the metric F_{Score} [48]. This metric is defined as the average sensitivity (True positive (TP) per number of Positives) at seven predefined FP/I (False positives per image) values, notably, $\{1/8, 1/4, 1/2, 1, 2, 4, 8\}$. In addition, a free-response operating characteristic curve (F_{ROC}), where the several lesion sensitivities are plotted against the corresponding FP/I is also used [48]. Thus, the performance of the MAs detection methodologies can also be measured through the area under the F_{ROC} curve, defined as F_{AUC}. Table 3.5 shows the performance of some MAs detection methods. For the E-Ophtha-MA and Latim datasets, only machine learning methods are reported. For the ROC Training and Test datasets, rule-based methods appear on top and machine learning methods appear on the bottom. The machine learning methods have a higher performance than rule-based methods. However, rule-based methods performance use the full datasets, while machine learning approaches, report performance average values on partial datasets, typically using a 10-fold cross-validation. This makes the comparison between these two approaches difficult. The only exception is in the ROC Test dataset, where the evaluation of the methods was made using the full dataset. When this dataset has used, the dif-

Table 3.5: MA detection comparison on several datasets.

E-Ophtha-MA Dataset

Methods	Sensitivity against FP/I							F_{Score}	F_{AUC}	Techniques	Evaluation Method
	1/8	1/4	1/2	1	2	4	8				
Du *et al.* (2020) [20]	0.227	0.309	0.407	0.488	0.622	0.739	0.820	0.516	-	RUSBoost	10-Fold crx val
Melo *et al.* I (2020) [63]	0.209	0.284	0.365	0.449	0.551	0.551	0.551	0.423±0.003	0.515	Classifiers ensemble	10-by-10 Repeated crx val(10%)
Melo *et al.* II (2020) [63]	0.145	0.197	0.303	0.404	0.499	0.598	0.598	0.392	0.525	Classifiers ensemble	10-Fold crx val(10% Test)
Eftekhari *et al.* (2019) [26]	0.091	0.258	0.401	0.534	0.579	0.667	0.771	0.471	0.637	CNN	10-Fold crx val(10% Test)
Chudzik *et al.* (2018) [253]	0.185	0.313	0.465	**0.604**	**0.716**	**0.801**	**0.849**	**0.562**±0.233	**0.640**	CNN	2-Fold crx val(50% Test)
Dashtbozorg *et al.* (2017) [31]	**0.358**	**0.417**	**0.471**	0.522	0.558	0.605	0.638	0.510	0.575	RUSBoost	10-Fold crx val(10% Test)
Wu *et al.* (2017) [22]	0.063	0.117	0.172	0.245	0.323	0.417	0.573	0.273	0.386	kNN	All images(50% Test)

ROC Training Dataset

Methods	Sensitivity against FP/I							F_{Score}	F_{AUC}	Techniques	Evaluation Method
	1/8	1/4	1/2	1	2	4	8				
OKMedical I (2010) [32]	**0.060**	0.096	0.138	0.189	0.267	0.311	0.345	0.201	-	-	All images
Du *et al.* (2020) [20]	0.155	0.166	0.197	0.260	0.331	0.436	0.504	0.293	-	RUSBoost	10-Fold crx val
Melo *et al.* I (2020) [63]	0.073	0.090	0.115	0.149	0.204	0.262	0.327	0.174±0.005	0.240	Classifiers ensemble	10-by-10 Repeated crx val(10%)
Melo *et al.* II (2020) [63]	0.077	0.092	0.113	0.149	0.205	0.283	0.348	0.181	0.254	Classifiers ensemble	10-Fold crx val(10% Test)
Eftekhari *et al.* (2019) [26]	0.047	0.173	0.351	**0.552**	**0.613**	**0.722**	**0.769**	**0.461**	**0.660**	CNN	10-Fold crx val(10% Test)
Chudzik *et al.* (2018) [253]	0.039	0.141	0.174	0.243	0.306	0.385	0.193	0.298±0.116	-	CNN	2-Fold crx val(50% Test)
Wang *et al.* (2017) [250]	**0.273**	**0.379**	**0.398**	**0.481**	**0.545**	0.576	0.598	**0.464**	**0.548**	kNN	All images(50% Test)
Budak *et al.* (2017) [252]	0.039	0.061	0.121	0.220	0.338	0.372	0.394	0.221	-	DCNN	All images(MA/Non-MA patches)
Dashtbozorg *et al.* (2017) [31]	**0.435**	**0.443**	**0.454**	0.476	0.481	0.495	0.506	**0.471**	0.484	RUSBoost	10-Fold crx val(10% Test)
Wu *et al.* (2017) [22]	0.037	0.056	0.103	0.206	0.295	0.339	0.376	0.202	0.302	kNN	10-Fold crx val(10% Test)
Dai *et al.* (2016) [35]	0.219	0.257	0.338	0.429	0.528	0.598	**0.662**	0.433	0.553	RUSBoost	2-fold crx val (50% test)
Lazar *et al.* (2013) [68]	0.037	0.055	0.103	0.162	0.196	0.223	0.285	0.152	-	Naive Bayes	All images(MA/Non-MA patches)
Inoue *et al.* (2013) [255]	0.088	0.132	0.175	0.201	0.291	0.343	0.409	0.234	-	Neural network	All images (50% Test)
OKMedical II (2012) [241]	0.046	0.064	0.100	0.127	0.150	0.197	0.289	0.139	-	SRC	All images(20% Test)
IRIA Group* (2011) [6]	0.041	**0.160**	**0.192**	**0.242**	**0.321**	**0.397**	**0.493**	**0.264**	**0.368**	kNN	8-Fold crx val(10% Test)
ISMV† (2010) [239]	0.131†	0.1494†	0.201†	0.251†	0.289†	0.345†	0.432†	0.257	-	SVM	All images
Fujita Lab (2009) [256]	0.050	0.068	0.106	0.146	0.195	0.252	0.323	0.163	-	Neural network	All images(25% Test)

ROC Test Dataset

Methods	Sensitivity against FP/I							F_{Score}	F_{AUC}	Techniques	Evaluation Method
	1/8	1/4	1/2	1	2	4	8				
DRSCREEN (2012) [5]	0.173	**0.275**	**0.380**	**0.444**	**0.526**	***0.599***	**0.643**	**0.434**	**0.551**	-	All images
OkMedical I (2010) [32]	**0.198**	0.265	0.315	0.356	0.394	0.466	0.501	0.357	0.430	-	All images
Latim (2008) [47]	0.166	0.230	0.318	0.385	0.434	0.534	0.598	0.381	0.489	-	All images
Lazar *et al.* (2014) [68]	0.251	0.312	0.350	0.417	0.472	0.542	0.615	0.423	0.510	Na"ive Bayes	All images
Adal *et al.* (2014) [36]	0.204	0.255	0.297	0.364	0.417	0.478	0.532	0.364	0.446	SVM,kNN, RF, Bayesian	All images
Pereira *et al.* (2014) [247]	0.053	0.083	0.135	0.187	0.276	0.407	0.540	0.240	0.366	Multi-agent	All images
OKmedical II (2012) [241]	0.175	0.242	0.297	0.370	0.437	0.493	0.569	0.369	0.465	SRC	All images
ISMV (2010) [239]	0.217	0.270	0.366	0.407	0.440	0.459	0.468	0.375	0.435	SVM	All images
GIB Valladolid (2009)[237]	0.190	0.216	0.254	0.300	0.364	0.411	0.519	0.322	-	LR	All images
Fujita Lab (2009) [256]	0.181	0.224	0.259	0.289	0.347	0.402	0.466	0.310	0.378	Neural network	All images
Niemeijer *et al.* (2005) [34]	0.243	0.297	0.336	0.397	0.454	0.498	0.542	0.395	0.469	kNN	All images

Latim Dataset

Methods	Sensitivity against FP/I							F_{Score}	F_{AUC}	Techniques	Evaluation Method
	1/8	1/4	1/2	1	2	4	8				
Pereira *et al.* (2014) [247]	-	0.050	0.077	0.130	0.215	0.308	0.449	0.205	-	Multi-agent	All images

† - Values are from [239]; crx val - Cross validation; CNN - Convolutional neural networks; RF - Random forest; kNN -K-nearest neighbours; NN - Neural networks; DCNN - Deep convolutional neural networks; SRC - sparse representation classifier; LDA - linear discriminative analysis; LR - logistic regression. Best values appear in bold.

ference between the best performing rule-based and machine learning approaches was not a very significant (DRSCREEN [5] and Wang *et al.* [250] with a F_{Score} of 0.434 and 0.464 respectively).

3.4.3.2 Image-level performance evaluation

The Messidor [46] dataset is used to evaluate the DR detection [75]. Each fundus image has a corresponding DR grade, namely, R0, R1, R2 and R3. R0 correspond to the absence of DR. R1 and R2 correspond to mild and severe cases of non-proliferative DR. R3 stands for proliferative, DR. The grading depends on the number of MAs, retinal haemorrhages and neovascularization as described in Table 3.6 [75]. Thus, considering the detection of MAs, it is possible to evaluate the performance of the methodologies in screening procedures (R0 vs. (R1,R2,R3)) [75].

To compare the DR detection between methods, it is used the area under the curve AUC and the Sensitivity at an operating point of 50% of Specificity [75, 24]. Sensitivity reflects the ability of an algorithm to detect DR fundus images based on equation 3.1. Specificity is the ability to detect non-DR fundus images using equation 3.2. TP is the number of true positive images, TN of true

Table 3.6: DR grading scheme for fundus image [75].

DR grade	Description
R0	$\#MA = 0$ and $\#H = 0$
R1	$(0 < \#MA \leq 5$ and $\#H = 0)$
R2	$(5 < \#MA < 15$ and $0 < \#H < 5)$ and $(\#NV = 0)$
R3	$(\#MA \geq 15$ or $\#H \geq 5)$ or $(\#NV > 0)$

$\#MA$, $\#H$, $\#NV$ - Number of MAs, haemorrhages and Neovascularization respectively.

Table 3.7: DR classification considering the AUC and Sensitivity (defined at a Specificity of 50%) values for the Messidor dataset. Expert A and B represent the human performance for two retinal experts [75].

Method	AUC	Sensitivity
Giancardo *et al.* [74]	0.854	0.910
Sanchéz *et al.* [75]	0.876	0.922
Orlando *et al.* [24]	0.893	0.911
Seoud *et al.* [7]	0.899	0.939
Expert A [75]	0.922	0.945
Expert B [75]	0.865	0.912

negative fundus images, FP of false positive, and FN of false negative. An image classification only depends on the detection or not of MAs, and it is independent of the performance of the MAs detection in each image. Table 7.3 shows some of the results in the literature for the DR detection using the Messidor dataset.

Chapter 4

Retinal exudates segmentation

Retinal exudates are one of the most prevalent lesion signs in early DR. Hence, their detection and segmentation are of primary importance in early DR diagnose and follow-ups. Exudates are lipid residues of serous leakage from damaged capillaries, appearing like yellowish lesions. They can be spread over the retina, and have different sizes, shapes, and brightness [257].

The non-uniform illumination, contrast variation, and variable background are inherent to all fundus images. These typical differences are one of the main obstacles to good exudates detection. Shade correction and non-linear point transformation are approaches to overcome non-uniform illumination. However, they need to deal with the difficult discrimination between the variations due to features or illumination [4]. Techniques used in contrast enhancement like histogram equalization or local contrast enhancement face problems like information loss and noise increment. Furthermore, there are no specific metrics to quantify contrast enhancement results, being the final result judged by visual inspection [4]. These characteristics are one of the main obstacles in the accurate detection and segmentation of exudates.

In this chapter, it is presented two different methods for exudates segmentation. The main motivation was the development of a fully automatic exudate segmentation method that does not require any preprocessing steps.

The first developed method is based on a scale-space extrinsic curvature approach, described in the paper, *"Scale-space Curvature Detection of Retinal Exudates with a dynamic threshold"* [53] and further developed in *"Curvature Detection And Segmentation of Retinal Images"* [54]. True exudates are detected by a blob response and dynamic thresholding. The use of the scale-space curvature provides a good uniformization of the retinal structures, with no need for contrast variation and non-uniform illumination correction.

The second developed exudates segmentation method uses the noise present in every retinal image as a framework for exudates segmentation. It is based on the paper *"Exudates Dynamic Detection in Retinal Fundus Images Based on the Noise Map Distribution"* [77]. Noise is an unwanted characteristic inherent to all retina fundus images, resulting from stochastic fluctuations. The normal approach is to remove or reduce it [258]. It does not come associated with techniques specially designed to perform the segmentation of exudates. The proposed exudates segmentation method is based on the computation of the noise map distribution and the use of morphological operators. A binary image called Maxima Noise Distribution Function ($I_{ndf}(\mathbf{x})$) is used to perform the detection of the exudates using concentric binary masks and adaptive threshold. The segmentation of exudates is achieved without any preprocessing. That shows the robustness of the method for exudates segmentation on images with non-uniform illumination.

Section 4.1 describes the *Scale-space curvature exudates segmentation* method, and section 4.2 the *Noise based exudates segmentation* method.

Section 4.3 presents some resulting exudates segmentation using the publicly available Diaretdb1 dataset. Furthermore, a comparison is made with recent works. Finally, section 4.4 summarizes the concluding remarks.

4.1 Scale-space curvature exudates segmentation

The scale-space curvature exudates segmentation method consists of the five main processing stages, described in the next subsections:

- Image scale-space computation
- Curvature extremes computation
- Dynamic threshold of the curvature extremes
- Curvature extremes tracking through different scales
- Exudates segmentation

Appendix B defines the used scale-space. The number of scales N_s was set to 6.

4.1.1 Curvature extremes computation

According to Sivaswamy *et al.* [259] the curvature at a given location of an image is a measure of the rate of the bending change in the surface along a particular direction. Hence, let $y = f(x)$ be a 1D function, and the tangent at the point P of function makes an angle θ with the x-axis as represented in Figure 4.1.

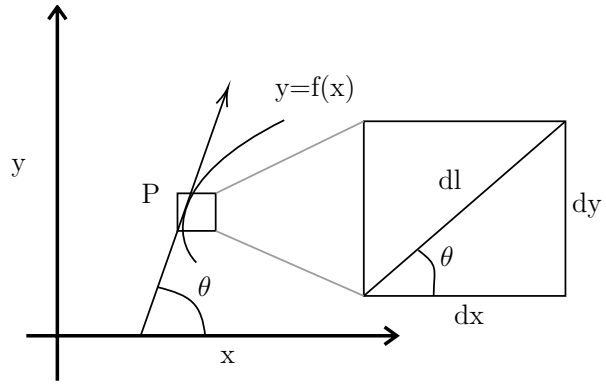


Figure 4.1: Curvature definition.

Defining dl as the differential arc-length at the point P, then the extrinsic curvature of the function $f(x)$ at this point is given by,

$$k(x) = \frac{d\theta}{dl} = \frac{d\theta}{\sqrt{dx^2 + dy^2}} = \frac{\frac{d\theta}{dx}}{\sqrt{1 + \frac{dy^2}{dx^2}}}. \quad (4.1)$$

Considering only the numerator of $k(x)$ in equation 4.1, the surface tangent derivative (STD) is given by [259],

$$\gamma(x) = \frac{d\theta}{dx} = \frac{d}{dx} \arctan(\frac{dy}{dx}). \quad (4.2)$$

Although $\gamma(x)$ is not the true curvature measure, it also provides a measure of the rate of change of θ along a particular direction, and can be easily adapted in the exudate locations. Hence, for any given scale t_n, $\gamma(x)$ is calculated in $L(\mathbf{x}; t_n)$ in four specific directions $\alpha \in \{45°, 90°, 135°, 180°\}$, designated by $S_\alpha(\mathbf{x}; t_n)$. Let tr_n be the curvature extremes threshold for the scale t_n, and $N_3(\mathbf{x})$

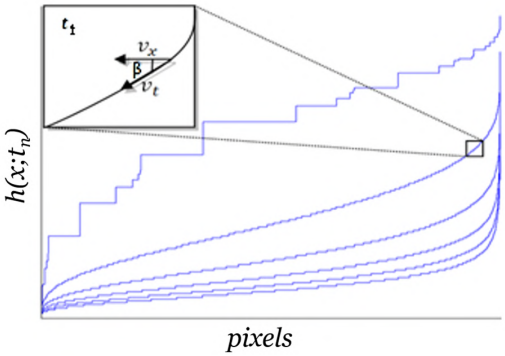


Figure 4.2: Representation of $h(x;t_n)$ for all scales t_n in ascending order. The top outline represents the curvatures of the original image. The following outlines show the curvatures in the successive scales. The upper left image represents the calculation of the threshold for the scale t_o.

the 3×3 neighbourhood at any location x. For each pixel location x, the procedure to compute the scale-space extremes is described in the Algorithm 1.

$M(\mathbf{x};t_n)$ is a binary image matrix that represents the thresholded scale-space curvature extremes locations. The threshold tr_n for each scale t_n is dynamically computed, as explained in the following.

4.1.2 Dynamic threshold of the curvatures extreme

The performance of the Algorithm 1 is strongly dependent on the threshold value tr_n. If tr_n is too small, all the curvature extremes values belonging to the exudates are detected, but other retinal features are also detected. If tr_n is too large, several curvature extremes values belonging to exudates may not be detected. A possible approach is to define tr_n empirically, but in that case, the method will not be an adaptive detection process. In this work, the threshold tr_n is dynamically computed at any scale t_n.

The values $|S_{max}|$ computed in the step 2 of the Algorithm 1 can be saved in ascending order in a vector $h(x)$. If this operation is performed for all the scales it is possible to define $h(x;t_n)$, that represents the curvature extremes values for any scale t_n.

In Figure 4.2 the outlines of $h(x;t_n)$ are represented for each scale t_n. These outlines are present in all ocular fundus images, either if exudates are present or not. This is guaranteed by the presence of the OD that is similar to the exudates intensity, resulting in similar curvature extremes values. An analysis of these outlines reveals that there is a particular value where $h(x,t_n)$ bends more sharply, for any scale t_n. This value divides the features with high curvature extremes, like exudates and the OD, from others features with low curvature extreme values. This value is hence ideal to be chosen as the threshold value tr_n. Considering Figure 4.2, for the scale t_1 the threshold tr_1 of $h(x;t_1)$ is computed as follows. Consider θ the angle between the tangent vector at x, v_t with the x-axis. The selected value of tr_1 is the point where the angle θ is $\frac{\pi}{4}$. This process is repeated for every scale t_n, using $h(x;t_n)$, resulting in a set of threshold values tr_n, for each scale. Only curvature extremes values larger than the computed threshold tr_n at each scale are considered for the exudate detection.

Algorithm 1: Curvature extremes computation algorithm

Requirement: The 3×3 neighbourhood at pixel location **x**, denoted $N_3(\mathbf{x})$
Input: $S_\alpha(\mathbf{x}; t_n)$
Output: $M(\mathbf{x}; t_n)$

for *every pixel location* **x** **do**
  for *each scale* t_n **do**
    Evaluate $|S_{max}| = max\{S_\alpha(\mathbf{x}, ; t_n)\}$, and the corresponding maximum orientation α_{max}

$$M(\mathbf{x}; t_n) = \begin{cases} 1, \textbf{ If } \begin{cases} |S_{max}| > tr_n \wedge \\ S_{\alpha_{max}}(\mathbf{x}; t_n) < 0 \wedge \\ |S_{max})| \geq |S_{\alpha_{max}}(N_3(\mathbf{x}); t_n)| \end{cases} \textbf{ end for} \\ 0, \textbf{ Otherwise} \end{cases}$$

  end for

4.1.3 Curvature extremes tracking through different scales

The next step defines the final exudate location using the detected thresholded curvature extremes. Garg *et al.* [260] performed an OR operation on the binary matrices $M(\mathbf{x}; t_n)$ at every location $\mathbf{x}$ at all scales t_n. This procedure results in multiple detections of the same feature at different scales because curvature extremes tend to change their location with scale. Furthermore, there is the possibility that two different extremes, with close locations, may merge on larger scales. In this work, a pixel tracking method along with the scale-space matrices $M(\mathbf{x}; t_n)$ is proposed in order to achieve a more precise location of the exudates in the scale t_1. The scale t_1 is chosen to select the final locations, since it is the closest to the initial image, providing a more accurate marking.

A curvature extreme is only considered as being representative of an exudate when the following rules are valid: 1) A pixel marked in a scale t_n must also appear in the lower scales. 2) Between adjacent scales, a displacement of pixels can occur. It is considered that these pixels displacements are inside the adjacent 3×3 neighbourhood. 3) Starting from the highest scale to the lowest scale, only the pixels that appear marked in all scales are considered. 4) The feature locations in scale t_1 are the true feature location. The coarse-to-fine pixel tracking algorithm describing the exudates detection process is described in Algorithm 2.

Algorithm 2: Coarse-to-fine pixel tracking algorithm.

Requirement: The 3×3 neighbourhood at pixel location $\mathbf{x}$, denoted $N_3(\mathbf{x})$
Input: $M(\mathbf{x}; t_n)$
Output: $M_f(\mathbf{x}; t_n)$

for *every pixel location* $\mathbf{x} \in M(\mathbf{x}; t_{N_s})$ **do** `// starts at scale` t_{N_s}
  if $M(\mathbf{x}, t_{N_s}) == 1$ **then**
    for $n = (N_s - 1)$ ***to 1*** **do**
      if $M(N_3(\mathbf{x}), t_n) == 1$ **then**
        The location $\mathbf{x}$ where the condition is verified, is used in the previous scale.
        if $n == 1$ **then**
          $M_f(\mathbf{x}) = 1$
          go to step **1**
        end
      else
        go to step **1**
      end
    end for
  end
end for

After the pixel tracking is performed, a binary image result, representing all the true exudate locations, called $M_f(\mathbf{x})$. Figure 4.3 represents examples of the final detection and markings of true exudate locations using the scale-space curvature extreme.

4.1.4 Exudates segmentation

As defined in section B.1 the Hessian is the Hessian is a matrix of the second order partial derivatives that describe the local curvature of a function [261]. The eigenvalues of the Hessian matrix defined by $\lambda_1(\mathbf{x})$ and $\lambda_2(\mathbf{x})$ represent the minimal and maximal curvatures at each location $\mathbf{x} \in L(\mathbf{x}; t_n)$ [261]. The mean curvature at the location $(\mathbf{x})$ is the average of $\lambda_1(\mathbf{x})$ and $\lambda_2(\mathbf{x})$,

$$MK(\mathbf{x}; t_n) = \frac{\lambda_1(\mathbf{x}; t_n) + \lambda_2(\mathbf{x}; t_n)}{2}. \quad (4.3)$$

Figure 4.4 shows the comparison between an intensity profile and the mean curvature profiles at several scales, considering the same row. In the intensity profile, the effects of non-uniform illumination and contrast variation are well observed. However, in the mean curvature profiles, these effects almost disappear, revealing $MK(\mathbf{x}; t_n)$ uniformization along with the scales. It is also visible the presence of noise effects or small artefacts in the lower scales of the mean curvature profiles. Furthermore, Figure 4.4 b) shows that exudates belong to dark regions and that they can be selected with an appropriate threshold. Hence, the curvature representation is used for image segmentation instead of the intensity. In order to avoid the noise effects present in the lower scales, and the excessive blurring in larger scales, the segmentation is performed in a middle scale, t_s. After a normalization between 0 and 1, followed by a histogram equalization of the $MK(\mathbf{x}; t_s)$ image, $MK^h(\mathbf{x}; t_s)$, is computed. An initial segmentation is achieved by selecting all the pixels in $MK^h(\mathbf{x}; t_s)$ with values equal to zero, given by,

$$E(\mathbf{x}) = \begin{cases} 1, & \text{if } MK^h(\mathbf{x}; t_s) = 0. \\ 0, & \text{otherwise} \end{cases} \quad (4.4)$$

$E(\mathbf{x})$ is a binary image where white patches represent exudate candidates. This segmentation is

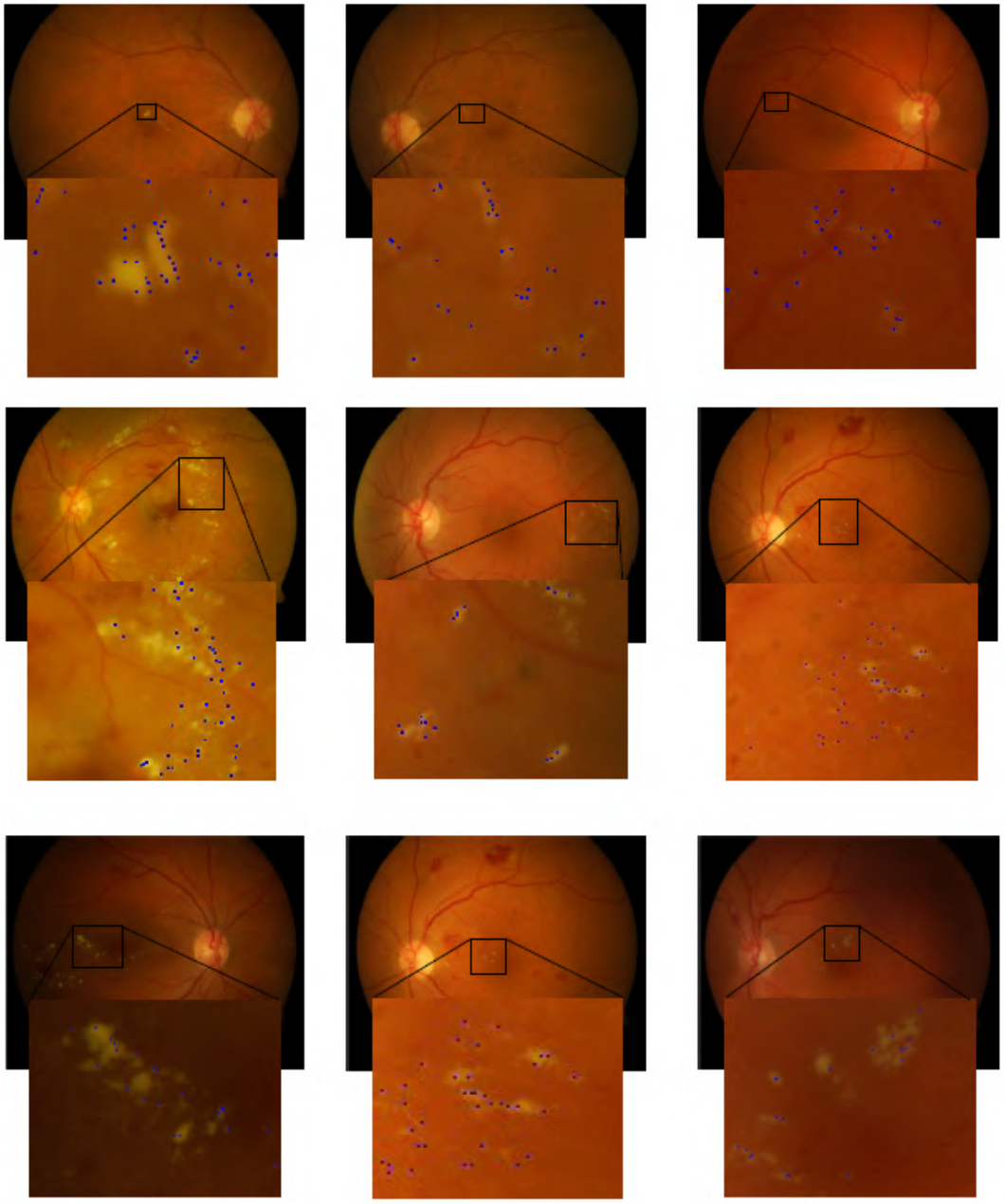
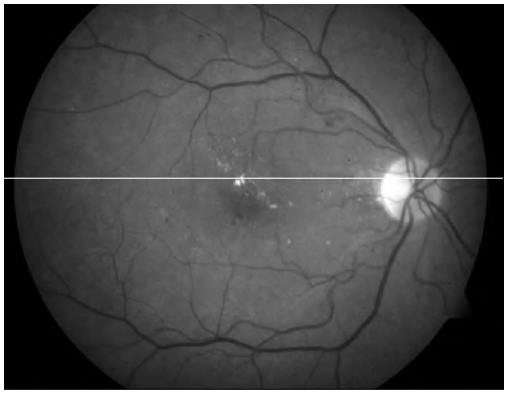


Figure 4.3: Example of extremes detected with the scale-space curvature method.

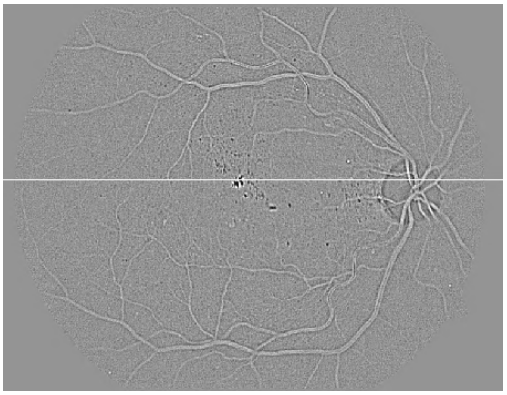
improved, considering only the patches where at least one point of the $M_f(\mathbf{x})$ image is present. After this step, because there are some regions with high curvature values that do not belong to exudates, few artefacts remain in the segmented image. Since exudates have a hill resemblance, they must have a high response to a blob detector and can be selected using an appropriate threshold c. For the scale t_s, a blob detector can be defined as [261],

$$B(\mathbf{x}; t_s) = L_{xx}(\mathbf{x}; t_s)L_{yy}(\mathbf{x}; t_s) - L^2_{xy}(\mathbf{x}; t_s). \quad (4.5)$$

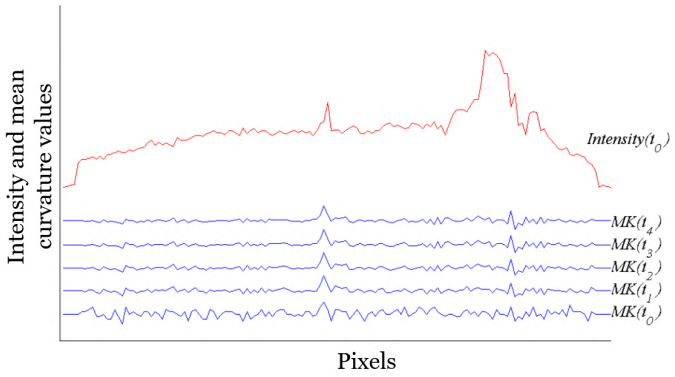
For each patch in $E(\mathbf{x})$, the local maxima blob response is computed using equation 4.5. These maxima responses are saved in a vector $\mathbf{V}(x)$. It was found, after extensive testing, that the median of $\mathbf{V}(x)$ is a good value for the threshold c. The final segmentation is achieved, selecting only the patches in $E(\mathbf{x})$ that have a local maxima blob response in $B(\mathbf{x}; t_s)$ larger than c. Figure 4.5 shows



(a) Green channel



(b) Mean curvature image $MK(\mathbf{x}; t_4)$



(c) Intensity (red) and mean curvature (blue) variation for growing scales of the marked row.

Figure 4.4: Mean curvature scale-space uniformization.

an example of the scale-space curvature exudates segmentation method. Results of the developed method are described in section 4.3.

4.2 Noise based exudates segmentation

The developed noise-based exudates segmentation method consists of the two main processing stages, described in the next subsections:

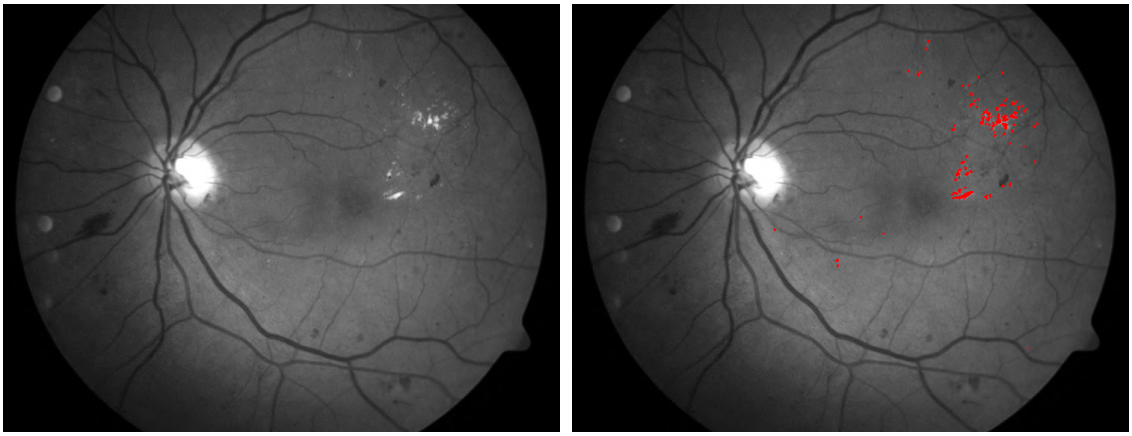
- Local maxima noise distribution computation
- Exudates segmentation

4.2.1 Local maxima noise distribution computation

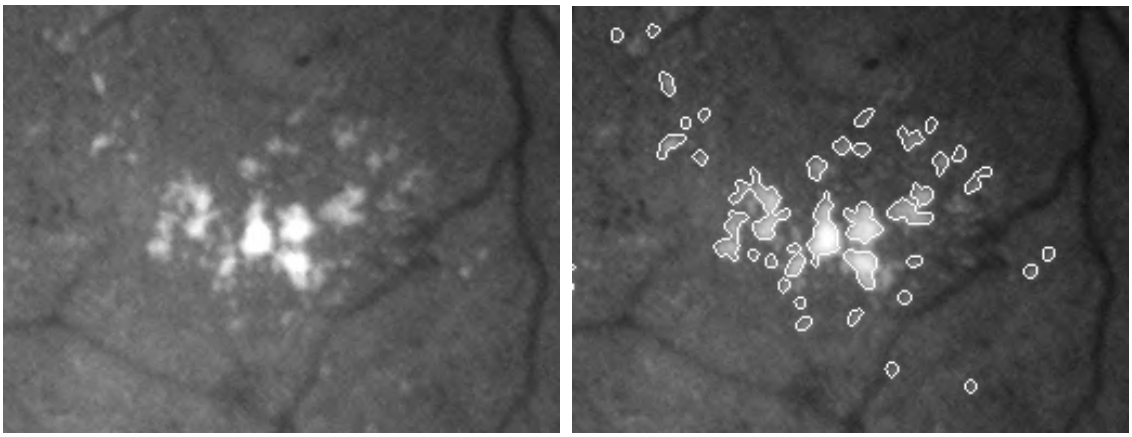
For the exudates detection, we make use of the green component of the RGB image $I_G(\mathbf{x})$, since it is the component that offers larger contrast [258]. $I_G(\mathbf{x})$ may be interpreted as the composition,

$$I_G(\mathbf{x}) = I_F(\mathbf{x}) + I_B(\mathbf{x}) \quad (4.6)$$

where $I_F(\mathbf{x})$ is the foreground image and $I_B(\mathbf{x})$ is the background image [189]. Features like the OD, vasculature branches, and any visible lesion belong to the $I_F(\mathbf{x})$ image. The background image



(a) Green channel (b) Segmented image



(c) Detailed green channel (d) Detailed segmented image

Figure 4.5: Scale-space curvature exudates segmentation result.

$I_B(\mathbf{x})$, contains all the illumination variation due to the acquisition transformation function, and it is the ideal image of the retina without any feature [4]]. $I_B(\mathbf{x})$ can be statistically modelled as,

$$I_F(\mathbf{x}) = \mathcal{N}(\mu, \sigma), \quad (4.7)$$

i.e., as a white and random field with a normal distribution and a mean value μ, that represents the ideal uniform luminosity value, and a standard deviation σ, that represents the natural variability of retinal fundus pigmentation. Furthermore, sub-regions in $I_B(\mathbf{x})$ also obey a normal distribution [162]. To separate exudates from the background and low-intensity features, the $I_G(\mathbf{x})$ image is considered as a topographic map, where peaks and valleys correspond respectively to high and low-intensity pixels. Since exudates are regions with high-intensity pixels, a local maximum detection technique is employed. A 5×5 neighbourhood $N_5(\mathbf{x})$, centred in an arbitrary location, $\mathbf{x}$ is considered. The size of the neighbourhood was selected based on experimenting tests. For all the locations $\mathbf{x}$, the local maxima of the image $I_G(\mathbf{x})$ are determined with the following condition,

$$I_{ndf}(\mathbf{x}) = \begin{cases} 1, & \text{if} I_G(\mathbf{x}) \geq I_G(N_5(\mathbf{x})) \\ 0, & \text{otherwise.} \end{cases} \quad (4.8)$$

where $I_{ndf}(\mathbf{x})$ is a binary image that represents the local maxima of the noise map distribution of

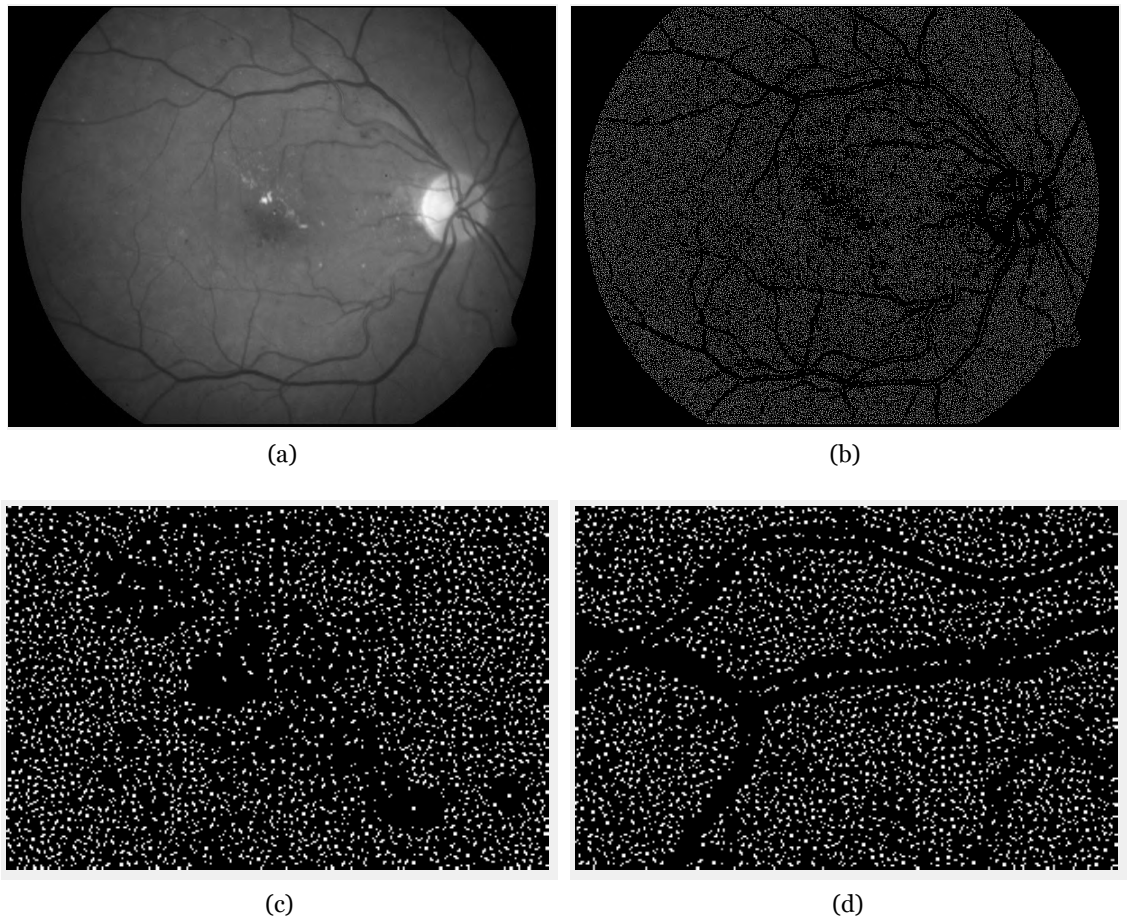
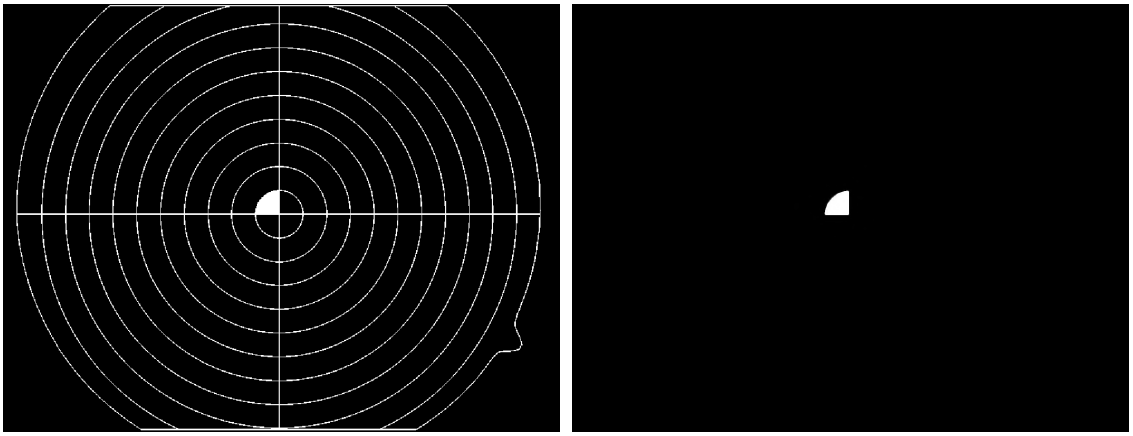


Figure 4.6: Upper images: a) represents a fundus image and b) $I_{ndf}(\mathbf{x})$. Lower images: c) exudates and d) vascular samples of $I_{ndf}(\mathbf{x})$.

the image $I_G(\mathbf{x})$. Equation 4.8 can produce the detection of pixels or regions of maxima [262]. The regions of maxima are replaced by the pixel with the higher intensity value in each region (See example Figure 4.6).

Although noise affects the entire image $I_G(\mathbf{x})$, certain patterns in the noise distribution can be noticed by analysing $I_{ndf}(\mathbf{x})$. There are areas with a denser and uniform distribution of points and areas with a sparse and irregular distribution. The noise effect is more noticeable in the background regions of $I_G(\mathbf{x})$, where the pixels have a small variation of the colour components between neighbour pixels. Because of that, Equation 4.8 produces a dense and uniform distribution of points in those background regions. In opposition, features like exudates and vascular branches represent regions of $I_G(x, y)$ where the pixels intensities values vary more abruptly and the noise effects are not so noticeable. Features will then belong to sparse and irregular regions in $I_{ndf}(\mathbf{x})$. Equation 4.8 only searches for local maximum, meaning that only peaks will be selected and marked, in opposition to the valleys points. Features with low-intensity values like vascular branches and haemorrhages will not result in valley points marked in $I_{ndf}(\mathbf{x})$. Exceptions can occur in arteries where the central zones can have brighter pixels (Figure 4.6 d)). Features with high-intensity pixel values like exudates belong to sparse and irregular regions. The effect of noise is not so noticeable due to the intensity growing towards the exudate center. Hence, in such areas, only one or a few pixels are marked (Figure 4.6 c)). These characteristics enable the separation of exudates from the background, resulting in their detection and segmentation. The OD is a feature that shares



(a) (b)

Figure 4.7: Example of the binary mask. a) Binary mask $D(\mathbf{x})$ and b) region $D_1(\mathbf{x})$.

some common characteristics, and it is manually removed from all the images before exudates segmentation. Although resulting from a simple condition (equation 4.8), $I_{ndf}(\mathbf{x})$ is very resilient to contrast changes, and non-uniform illumination, retaining the limit pixels, shape characteristics, and the peak pixels of the exudates, independently of their location.

4.2.2 Exudates segmentation

To perform the exudates segmentation, a region-processing solution is adopted. In each region, analysis of the intensity pixels of $I_G(\mathbf{x})$ that are also marked in $I_{ndf}(\mathbf{x})$, is performed using dynamic thresholding. In the usual approach, square regions with different sizes are considered [263]. In this work, the regions are defined following a different approach. The shapes of the regions are chosen to take into account the oblate cup shape of the retina and that the retinal image acquisition has approximately a spherical illumination distribution pattern. Anatomically, the retina can be modelled as an oblate cup shape, where the macula is the pole. Most of the retina images are acquired with the macula in the center or near the center since the patient is instructed to look towards a fixation central point. This ensures the alignment of the macula with the fixation point of the fundus camera. Hence, it is assured that the macula will be in a central region of the retinal image. Although it can be affected by defocusing, opacities, and acquisition with different fundus cameras, the approximately spherical illumination distribution pattern, result in lighter central areas than the peripheral areas. Consequently, different levels of contrast appear across the retina image. This must be considered for the definition of the segmentation regions. A black-and-white mask $BW(\mathbf{x})$, where white represents retina information and black does not have any information, is defined for the retina image. The dimensions of $BW(\mathbf{x})$ define the limits of the regions. These limits are dynamically computed, according to a binary mask $D(\mathbf{x})$ centred in $BW(\mathbf{x})$ and formed by concentric quarter rings, equally spaced, as shown in Figure 4.7.

Each concentric ring delimits the regions where the segmentations will be performed. Figure 4.7 b) shows the first binary region, $D_1(\mathbf{x})$. N_D is the total number of regions $D_i(\mathbf{x})$. Another advantage with this region model, is that although $D(\mathbf{x})$ may not be centred with the macula, it will be very close to it. Since the central regions in $D(\mathbf{x})$ are smaller than the peripheral ones, a more precise macular analysis is performed. For each mask $D_i(\mathbf{x})$, ($i = 1, ..., N_D$), $M_i(\mathbf{x})$ is defined as,

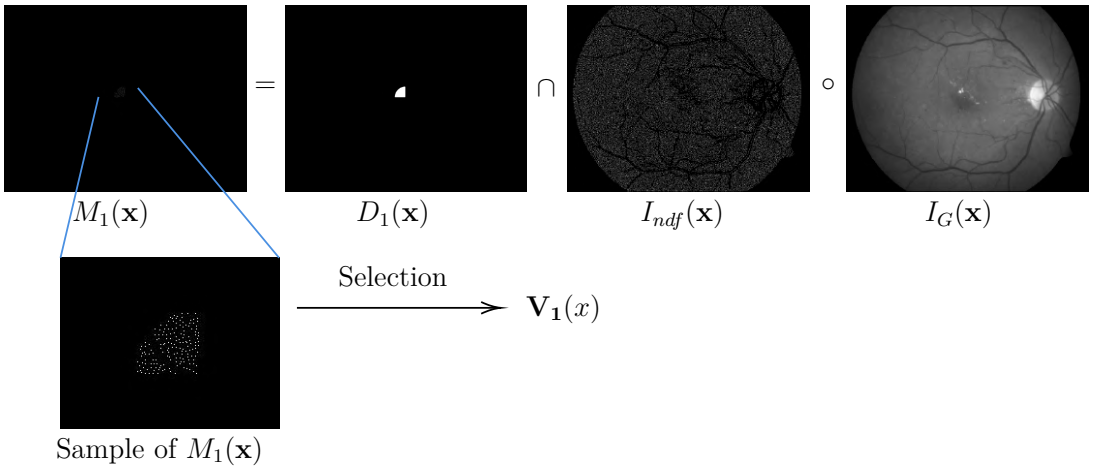


Figure 4.8: Example of $\mathbf{V_1}(x)$ computation.

$$M_i(\mathbf{x}) = (D_i(\mathbf{x}) \cap I_{ndf}(\mathbf{x})) \circ I_G(\mathbf{x}), (i = 1, ..., N_D). \quad (4.9)$$

where $\cap$ is the Hadamard product. The result is a set of N_D images $M_i(\mathbf{x})$, with the intensity pixels selected, Equation 4.9. After selecting these intensity pixels for each $M_i(\mathbf{x})$, it is possible to construct a set of N_D vectors $\mathbf{V_i}(x)$, where x are the select intensity pixels of $M_i(\mathbf{x})$. $\mathbf{V_i}(x)$ are the sampling of the ordered intensities pixels belonging to $I_G(\mathbf{x})$ and to $I_ndf(\mathbf{x})$ for each region $M_i(\mathbf{x})$. The process is illustrated in Figure. 4.8 for $\mathbf{V_1}(x)$.

The pixels in $\mathbf{V_i}(x)$ can represent two different cases; the pure background pixels or other feature with high-intensity pixels that includes exudate regions, respectively. The first case is selected in regions where no features exist. In that case, they can be statistically modelled as a normal distribution, according to Equation 4.8. In the second case, they cannot be approximated by a normal distribution. However, in both cases $\mathbf{V_i}(x)$ values plotted in ascending order, result in similar plots to the ones represented in Figure 4.9.

The plots show the existence of a point where the curves $\mathbf{V_i}(x)$ bend more sharply. This point is called the critical intensity value c_i. Its computation is explained in the next section. Every vector $\mathbf{V_i}(x)$ has a corresponding critical intensity value, c_i that is used to segment the bright areas from the background in each region $D_i(\mathbf{x})$. To perform the exudates segmentation, we select all the points in each region mask $D_i(\mathbf{x})$ that simultaneously belong to $I_G(\mathbf{x})$ with an intensity value larger than c_i. For every location x belonging to the $I_G(\mathbf{x})$ image, and the region $D_i(\mathbf{x})$, it is computed:

$$S_i(\mathbf{x}) = \begin{cases} 1, & \text{if } I_G(\mathbf{x})_{D_i} > c_i \\ 0, & \text{otherwise.} \end{cases}, (\forall \mathbf{x} \in D_i(\mathbf{x})) \quad (4.10)$$

The process in equation 4.10 is repeated for each region $D_i(\mathbf{x}), i = 1, ..., N_D$. The resulting segmentation image $S(\mathbf{x})$, is given by,

$$S(\mathbf{x}) = \cup_i^{N_D} S_i(\mathbf{x}) \quad (4.11)$$

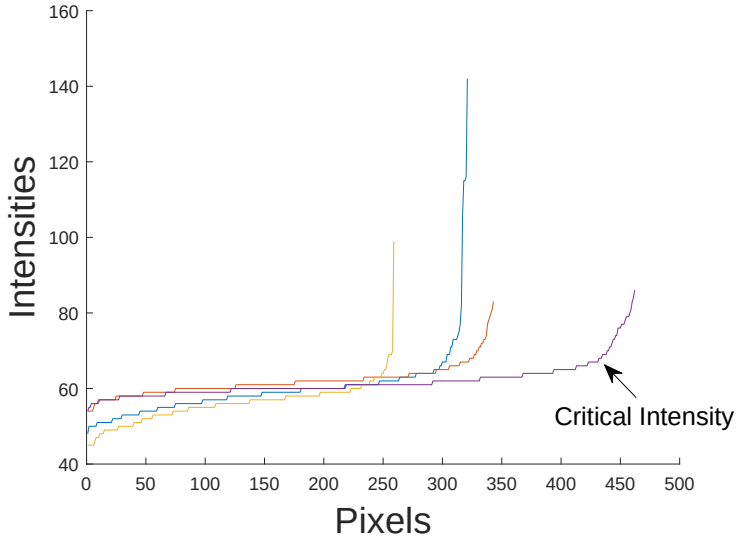


Figure 4.9: Plotted intensities of some ascending order of some $\mathbf{V_i}(x)$.

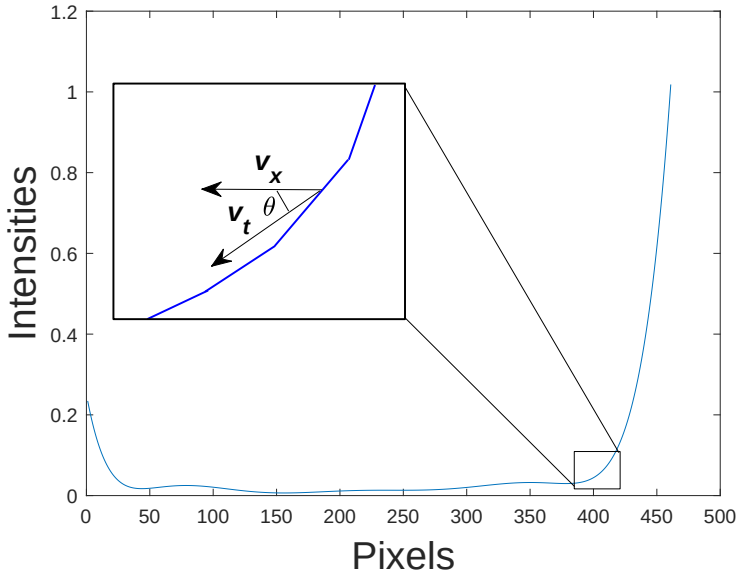
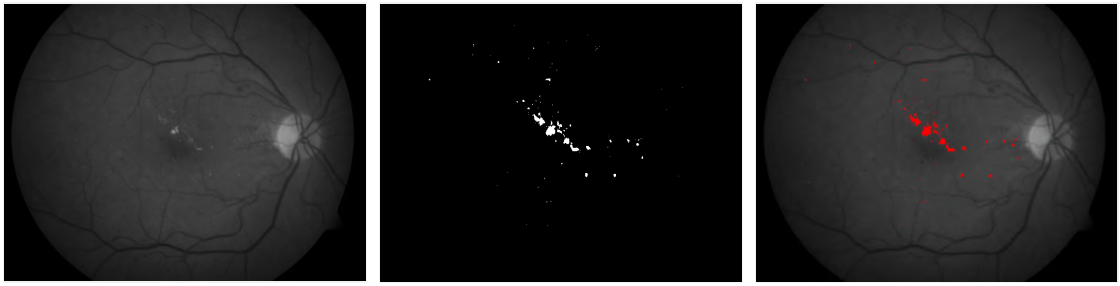


Figure 4.10: Critical intensity calculation.

The method reveals to be very effective in regions where exudates exist. The problem arises in regions without exudates. The calculation of c_i in these regions causes the appearance of artefacts. To perform the selection between artefacts and exudates, an algorithm based on the noise distribution $I_{ndf}(\mathbf{x})$ is presented in section 4.2.2.2.

4.2.2.1 Critical intensity determination

The critical intensity c_i will strongly influence the performance of the method. If it is too large, points that may belong to exudates may not be selected, resulting in under-detection. If it is too



(a) Green channel (b) Exudates segmentation result (c) Overlay of segmentation with original image

Figure 4.11: Noise based exudates segmentation result given by $S_{noise}(\mathbf{x})$.

small, it will result in over-detection. To determine c_i, $\mathbf{F_i}(x)$ was computed according to,

$$\mathbf{F_i}(x) = \frac{d\mathbf{V_i}(x)}{dx}. \quad (4.12)$$

Figure 4.10 represents $\mathbf{F_i}(x)$ for any given region $D_i(\mathbf{x})$. The derivative results in a stronger bending of, $\mathbf{F_i}(x)$ leading to an appropriate computation of the threshold value. For that, an angle θ between the tangent vector in x, $\mathbf{v_t}$, and the horizontal vector $\mathbf{v_x}$, is defined as shown in Figure 4.10. The critical intensity will be the point in which the angle θ, between the vectors $\mathbf{v_x}$ and, $\mathbf{v_t}$ reaches 45^o. Its determination is made starting from the last position in $\mathbf{F_i}(x)$ and computing the angle θ for every point x until it reaches 45^o. This point is selected as the critical intensity, c_i used in equation 4.10.

4.2.2.2 Artifacts removal

The described method is likely to result in the appearance of artefacts in $S(\mathbf{x})$ for regions without exudates. These artefacts are removed through a selection process based on the pixels dispersion of $I_{ndf}(\mathbf{x})$, and the segmented regions in $S(\mathbf{x})$. Algorithm 3 labels the regions in $S(\mathbf{x})$ as exudates or artefacts. First, $S(\mathbf{x})$ is interpreted as a binary image set according to,

$$S(\mathbf{x}) = \bigcup_{i=1}^{k_E} S_i(\mathbf{x}). \quad (4.13)$$

where k_E is the number of regions in $S(\mathbf{x})$. The algorithm then analyses the pixels of each region $S_i(\mathbf{x})$ and their position in $I_{ndf}(\mathbf{x})$. $S_i(\mathbf{x})$ pixels located in the proximity of high density regions are deleted. The value Sup_{S_i} represents the percentage of suppressed pixels that each $S_i(\mathbf{x})$ region suffers after the selection process is complete. $S_i(\mathbf{x})$ regions that suffer a pixels suppression larger than a given percentage value, are considered artefacts and removed. The best performance was achieved for a value of Sup_{S_i} =12% (similar Sup_{S_i} values result in identical results).

The region in $S_i(\mathbf{x})$ that represents a true exudate, belongs to regions of $I_{ndf}(\mathbf{x})$ with a low pixels density, resulting in almost no pixel suppression. Hence, those areas are preserved as segmented areas, resulting in the final exudates image $S_{noise}(\mathbf{x})$. Figure 4.11 shows an example of the final segmentation result.

Table 4.1: Scale-space curvature method parameters.

Variable	Simulation set	Best performance
N_s - Number of scales	$\{0, 1, ..., 16\}$	6
t_s - Scale to perform segmentation	$s \in \{0, 1, ..., 16\}$	4
c - Threshold	$mean, median, OTSU$	$median$

Algorithm 3: Artefacts removing algorithm.

Requirement: A binary circular mask $C_M(\mathbf{x})$ with radius=4 and size 9×9 and the 9×9 neighbourhood at location $(\mathbf{x})$ denoted by $N_9(\mathbf{x})$
Input: $I_{ndf}(\mathbf{x})$ and $S(\mathbf{x})$
Output: $S_{noise}(\mathbf{x})$

set $S(\mathbf{x}) = \bigcup_{i=1}^{k_E} S_i(\mathbf{x})$
for $i = 1$ ***to*** k_E **do**
  Compute $N_i = Card(S_i(\mathbf{x}))$ `// initial number of pixels of` $S_i(\mathbf{x})$
  for *every white pixel* $\mathbf{p} \in S_i(\mathbf{x})$ **do**
    $k = Card(I_{ndf}(N_9(\mathbf{p})) \cap C_M(\mathbf{p}))$
    if $k < 2$ **then**
      $S_i(\mathbf{p}) = 0$
    end
  end for
  Compute $N_f = Card(S_i(\mathbf{x}))$ `// final number of pixels of` $S_i(\mathbf{x})$
  if $\frac{100 \times N_f}{N_i} > Sup_{S_i}$ **then**
    remove $S_i(\mathbf{x})$ from $S(\mathbf{x})$
  end
end for
$S_{noise}(\mathbf{x}) \Leftarrow S(\mathbf{x})$

4.3 Results

The Diaretdbv1 dataset with 89 colour retinal fundus images and resolution of 1500×1152 pixels was used for testing and performance evaluation (46 abnormal/43 normal). Groundtruth images are provided for all the fundus images in the dataset [45]. These groundtruth images are the result of several markings made by experts for each retinal fundus image. The groundtruth areas are marked with a confidence level that depends on the number of experts that marked it as exudates or not. A higher number of expert markings leads to a higher level of confidence. The confidence level ranges from 0.0 to 1.0. The OD was manually removed from all the images. The exudates segmentation evaluation of the scale-space curvature and noise-based methods are described next.

4.3.1 Scale-space curvature exudates segmentation evaluation

The scale-space curvature exudates segmentation method depends on several parameters. Table 4.1 shows these parameters, the corresponding simulation set, and the values that provide the best segmentation performance.
The reliability of the developed method is based on the selection of the threshold values tr_n, that separate the curvatures extremes that are considered as valid exudate detections from the rejected ones. To study how the variation of these thresholds can influence the performance of the segmentation, they have been grown in steps. For each step, the number of extremes was grown 5% in each scale by selecting a new threshold value. The performance of the proposed method was evaluated

on a lesion basis for each image. A pixel is considered a true positive (TP) if it is contained in the groundtruth labelled regions, a false negative (FN) if there is no pixel marking in the groundtruth labelled region, a false positive (FP) if the pixel is marked in a groundtruth region where no label exists, and a true negative (TN) if no marking appears in a non labelled region in the groundtruth image. For each confidence level and for each step in the threshold value, the global sensitivity is computed for each image as SE=TP/(TP+FN), and then averaged for the 89 images. The global specificity (SP) is computed similarly by employing SP=TN/(TN+FP). Figure 4.12 represents several Receiver Operating Characteristic (ROC) curves. The best result is achieved for a confidence level of 1.0. However, when compared with the ROC curve of a confidence level of 0.8 there is no meaningful difference. Hence, the best sensitivity and specificity are respectively, 97.07% and 99.90% for a confidence level of 0.8.

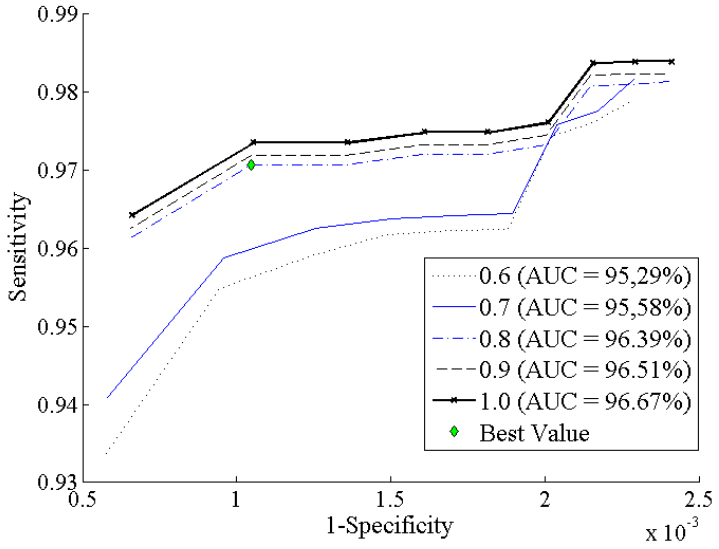


Figure 4.12: ROC curve for exudates segmentation. Each curve represents a different confidence level with the respective Area Under the Curve (AUC).

4.3.2 Noise based exudates segmentation evaluation

The performance of the noise-based exudates segmentation method was also evaluated on a lesion basis for each image. This evaluation results in the computation of the Sensitivity, Specificity, and Accuracy as previously described. To evaluate this method, a set of 28 training images defined in Diaretdbv1 (19 with exudates and 9 normal or with other types of a lesion) were used for calibration and selection of the parameters. Table 4.2 shows the parameters of these methods, the corresponding simulation set, and the values that provide the best segmentation performance.

Table 4.2: Noise based method parameters.

Variable	Simulation set	Best performance
N_{ring} - Number of rings to create $D(\mathbf{x})$	$\{2, ..., 15\}$	9
Sup - Pixels suppression factor	$\{8, ..., 30\}$	12
I_{c_i} - Incremental interval for each c_i	$\{-10, ..., 10\}$	0
k - constant	$\{1, ..., 4\}$	2

Using the parameter values of Table 4.2, the performance evaluation of the noise base exudates segmentation method was done using the evaluation set of the 61 remaining images defined in

Diaretdbv1 (28 with exudates and 33 normal or with other lesions). The best sensitivity, specificity, and accuracy are 97.49%, 99.95%, and 99.91% respectively with a confidence level of 0.7. The ROC curve for our algorithm was achieved considering the best parameters values and different confidence levels, showing a normalized area under the curve of 96.89% (Figure 4.13). Moreover, the images with no exudates result in a mean error of 0.004% of marked regions. This is a critical situation because images with no exudates might create spurious segmented regions, due to the dynamic thresholding.

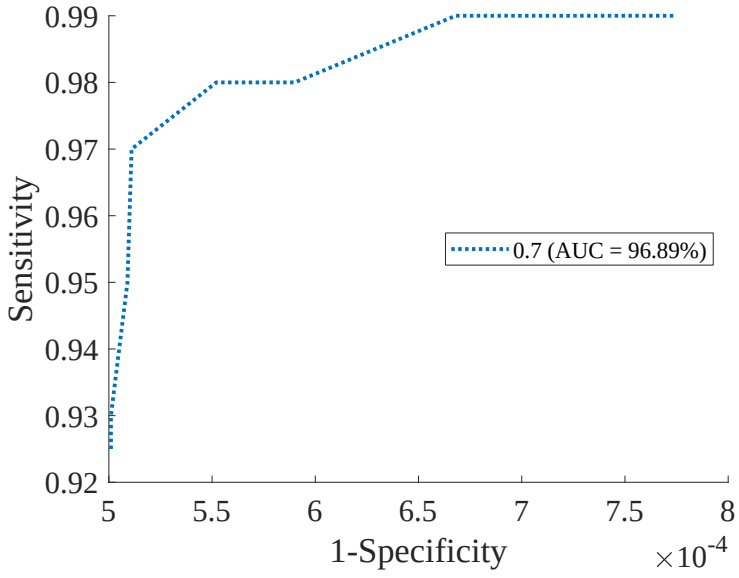


Figure 4.13: ROC curve for the noise based exudates segmentation method.

4.3.3 Exudates segmentation methods comparison

Figure 4.14 and Figure. 4.15 show a visual comparison between the two exudates segmentation methods described previously, in fundus images with and without exudates, respectively. Table 4.3 shows the results achieved with the two proposed exudates segmentation methods(curvature-based - Proposed **Method A**, and noise-based - Proposed **Method B**), and a comparison with some recent published works.

Table 4.3 is divided in rule-based methods (Top) and machine learning methods (Bottom). For the Diaretdbv1 dataset, the **Method A** and **Method B** for exudates segmentation has a sensitivity of 97.01% and 97.49%, a specificity of 99.90% and 99.95%, and an accuracy of 99.83% and 99.91% respectively.

In rule-based methods, except for Karkuzhali *et al.* [15] the **Method A** and **Method B** have the highest sensitivity. Karkuzhali *et al.* [15] has a higher sensitivity (98.87%) with the cost of a low specificity (92.31%). When an analysis is made between Lesion-based (Lb) criteria and Image-based (Ib) criteria methods (for both rule-based and machine learning methods), both **Method A** and **Method B** exhibits the highest sensitivity and accuracy.

The good performance of the proposed methods on fundus images with different characteristics of the Diaretdbv1 dataset demonstrates the robustness of the developed methods to illumination and contrast variations. This is further demonstrated in Figure 4.14, whereby visual inspection can be observed in a very accurate detection of the retinal exudates. Furthermore, Figure 4.15 shows that in fundus images with no exudates, both the proposed methods detect only small and spurious regions.

Table 4.3: Comparison of exudates segmentation methods.

Diaretdb1 Dataset (81)						
Method	Sen(%)	Spe(%)	Acc(%)	Evaluation	Method	Classifier
Method A (2011)	97.07	99.90	99.83	Lb	Scale-space curvature	Threshold
Method B (2011)	97.49	**99.95**	**99.91**	Lb	Noise distribution	DT
Karkuzhali *et al.* (2019) [15]	**98.87**	92.31	99.67	-	FCM, Prewitt edges	Inverse Surface
Wisaeng *et al.* (2019) [52]	97.05	97.18	97.14	Lb	Mean-Shift	MM
Kaur *et al.* (2018) [109]	71.34	88.56	85.34	Lb	MF-FDOG,Hough transform	DT
Kaur *et al.* (2018) [109]	91.00	94.00	89.00	Ib	MF-FDOG,Hough transform	DT
Huang *et al.* (2021) [51]	98.21	91.38	98.01	Ib	Superpixel	CNN
Long *et al.* (2019) [145]	97.50	**97.80**	97.70	Ib	FCM + DT	SVM
Khojasteh *et al.* (2019) [150]	**99.00**	96.00	**98.20**	-	-	ResNet-50 + SVM
Javidi *et al.* (2019) [146]	96.00	83.00	-	Ib	BAMCA	k-SVD

Sen - Sensitivity; Spe - Specificity; Acc - Accuracy; Lb - Lesion-based; Ib - Image-based; DT - Dynamic Thresholding; MM - Mathematical Morphology; GAN - Generative Adversarial Network; MU-net - Ensemble convolutional neural network; ResNet-50 - Pre-trained Residual networks;BAMCA - Bi-level adaptive morphological component analysis; Best values appear in bold

4.4 Concluding remarks

In this chapter, two different methods for the automatic segmentation of retinal exudates were presented. Both the methods perform a correct segmentation of the exudates. Furthermore, both methods effectively deal with non-uniform illumination and contrast variation in fundus images. This is achieved without any preprocessing steps, but rather considering the curvature space or the noise distribution when performing the segmentation.

The developed methods exhibit good performance, achieving results in line with the most recent techniques. The scale-space curvature-based method reliability is provided by its dynamic structure, based on the scale-space analysis and dynamic thresholding. The noise-based method uses the selection of the critical intensity locally, resulting in a dynamic thresholding computation. The concentric regions dynamically defined to every fundus image proved to be very effective in the segmentation task.

It is worth mentioning that the groundtruth regions are not accurate. In some cases, they join two different exudates in one region [264]. In those cases, the validation of the detection results is impossible, because in case only one of those exudates is detected, there is no possibility of quantifying that error. Although uncommon, that situation can produce a slight decrease in the Sensitivity results. Moreover, this is a common situation to other results obtained with this dataset. This situation was tried to be alleviated in [264], without accurate results. Because of this, the evaluation of the segmentation accuracy is an important issue that requires future consideration.

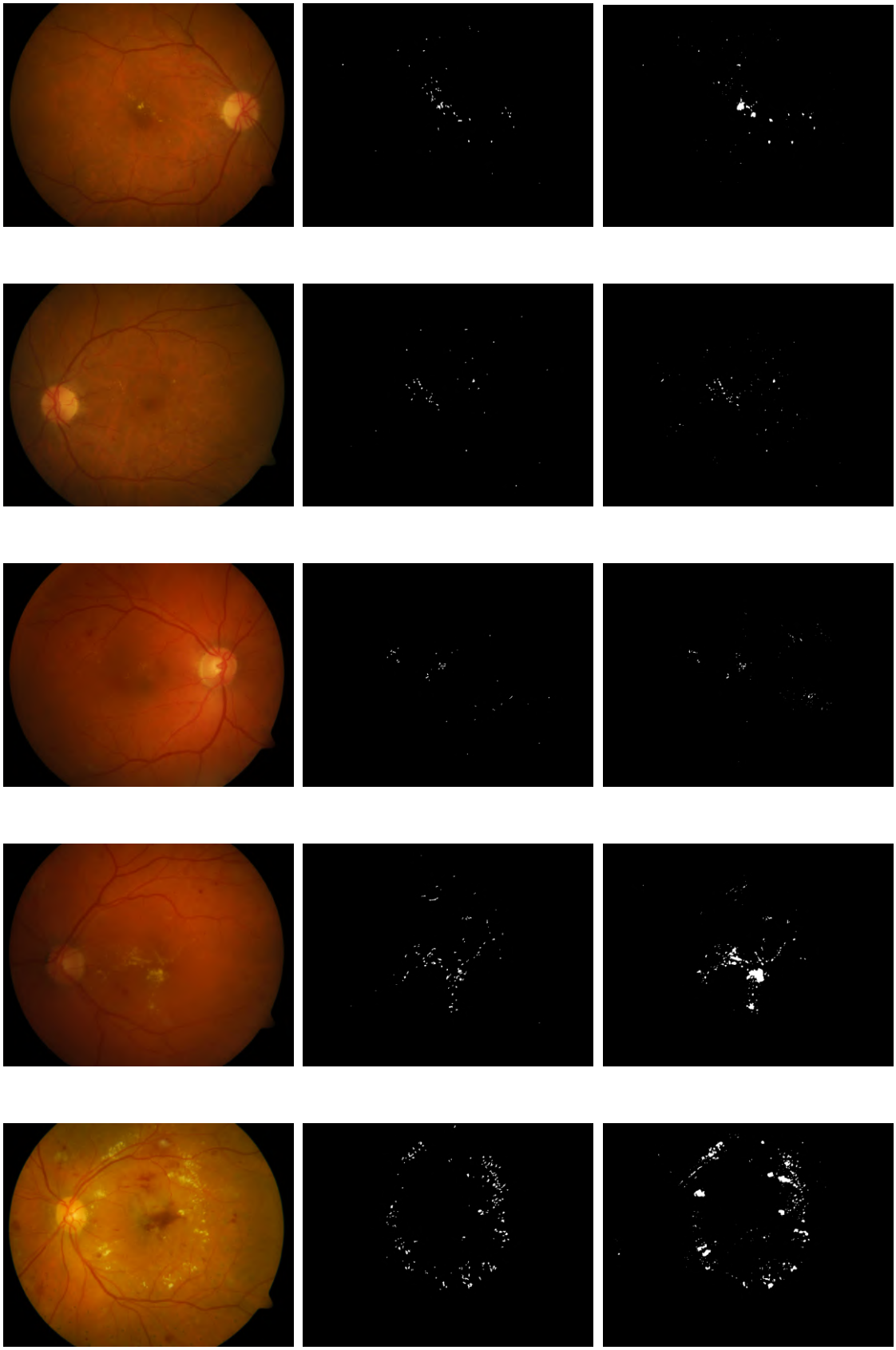


Figure 4.14: Exudates segmentation results in fundus images with exudates. The first column represents some fundus images from the Diaretdbv1 dataset. The middle column shows some segmentation results S_{curv} described in section 4.1, The last column shows some segmentation results S_{noise} described in section 4.2.

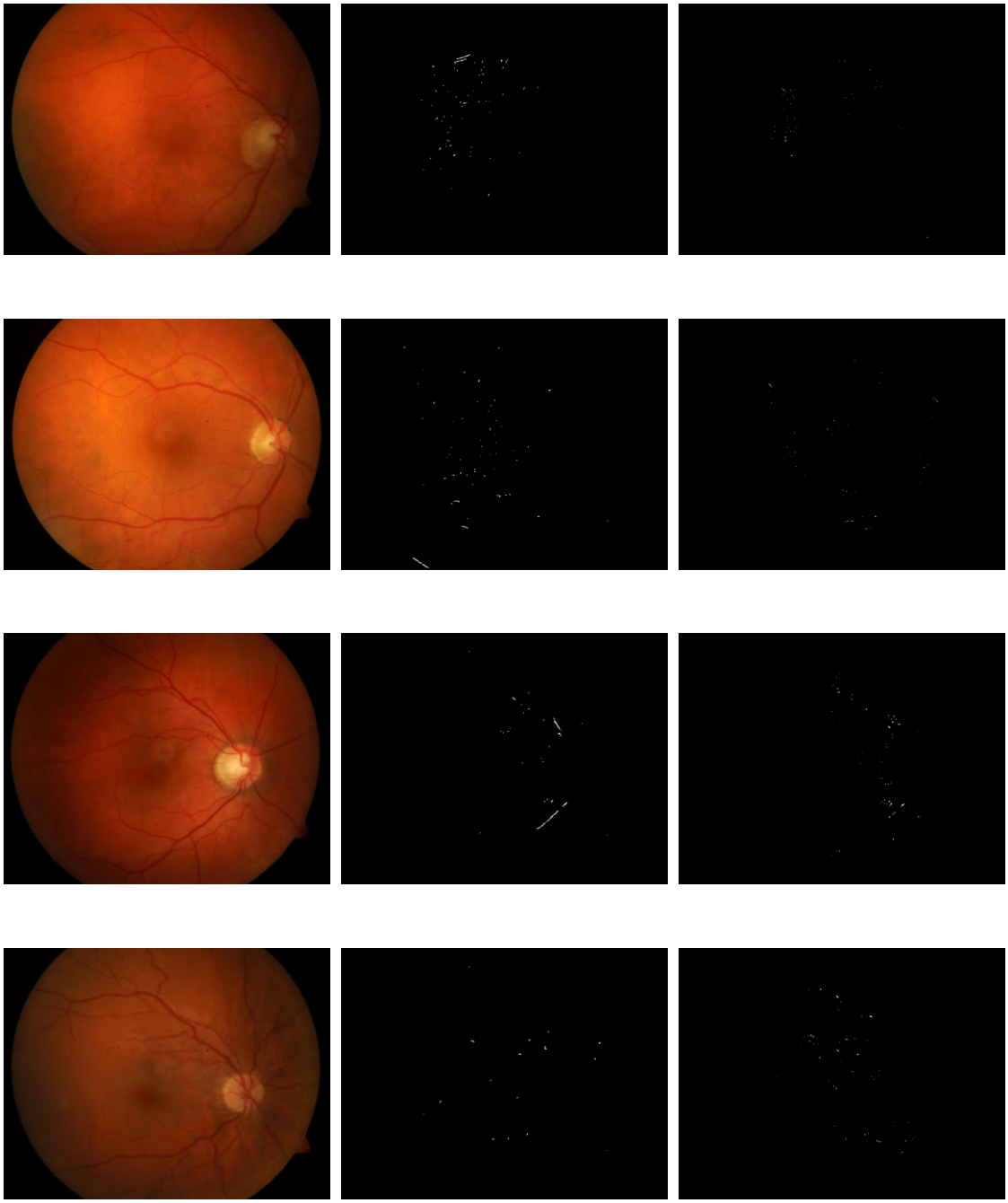


Figure 4.15: Exudates segmentation results in fundus images with no exudates. The first column represents some fundus images from the Diaretdbv1 dataset. The middle column shows some segmentation results S_{curv} described in section 4.1, The last column shows some segmentation results S_{noise} described in section 4.2.

Chapter 5

Optic disc localization

The OD is one of the most important retinal landmarks to be detected. In normal conditions, it appears as a nearly circular yellowish disk, brighter than its surroundings. It is also the region where the retinal veins and arteries emerge and spread, covering the retina [162].
Significant changes in the shape, colour, and depth provide quantitative metrics information for the detection of pathologies associated with the OD, e.g., glaucoma, neovascularization on the disc, papilledema, and assessment of white lesions related with DR [178, 265].
Moreover, the automatic localization of the OD provides invaluable help in the screening of DR and macular degeneration, by discarding the OD as a potential confounder, relative to retinal exudates and other bright lesions (see Figure 5.1). In addition, the OD provides a reference point for the detection of other important retinal landmarks like the fovea and retinal vasculature [162, 64].
Although the OD main features and characteristics are relatively easy to describe, its automatic localization can be a difficult task. Furthermore, this localization is particularly difficult in pathological retinas or in cases where the OD exhibits an altered appearance.
In this chapter, an automatic method for the OD localization in fundus images is described, based on the paper *Optic Disc localization in Retinal Images based on Cumulative Sum Fields* [55]. The main motivation was the development of an automatic method for the localization of the OD that is simultaneously effective in multiple datasets, and in fundus images with large pathological clinical manifestations.
The algorithm begins with a new vessel enhancement method based on a modified corner detector. Subsequently, a weighted version of the vessel enhancement is combined with morphological operators, to detect the four main vessels orientations $\{0^o, 45^o, 90^o, 135^o\}$. These four image functions have all the necessary information to determine an initial OD localization, resulting in two images that are respectively divided along the vertical or horizontal orientations with different division sizes. Each division is averaged creating a 2D step function, and a cumulative sum of the different sizes of step functions is calculated in the vertical and horizontal orientations, resulting in an initial OD position. The final OD localization is determined by a vessel convergence algorithm using its two most relevant features; high vasculature convergence and high-intensity values. The proposed method was evaluated on eight publicly-available datasets, including the Stare and Drive datasets. The OD was localized correctly in 1752 out of the 1767 fundus images (99.15%) with an average computation time of 18.34 seconds per fundus image.
This chapter starts by describing the OD detection method in section 5.1. Afterward, section 5.2 presents the results, where a comparison with other relevant methods is also included. Finally, some concluding remarks are drawn in section 5.4.

5.1 Proposed method

The proposed method comprises several steps. First, the green channel component $I_G(\mathbf{x})$ of each RGB fundus image is selected [4]. If necessary, a resize of $I_G(\mathbf{x})$ is performed, since the method is tested in different images from several datasets, and there is a need to maintain the coherence

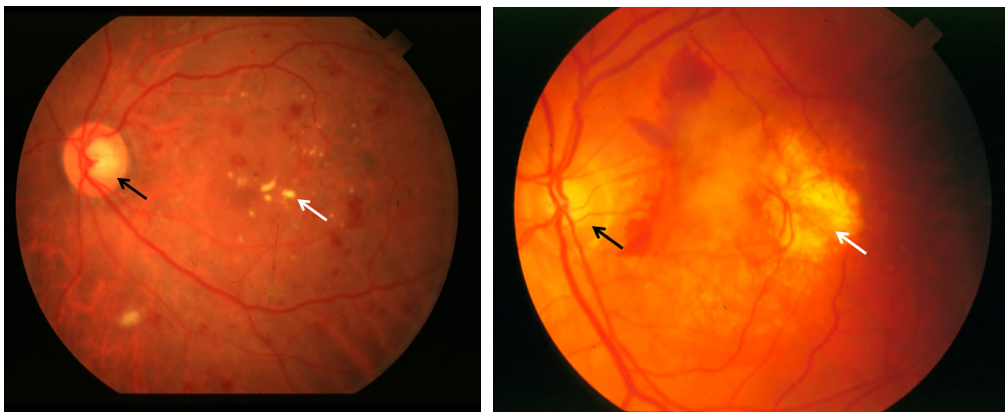
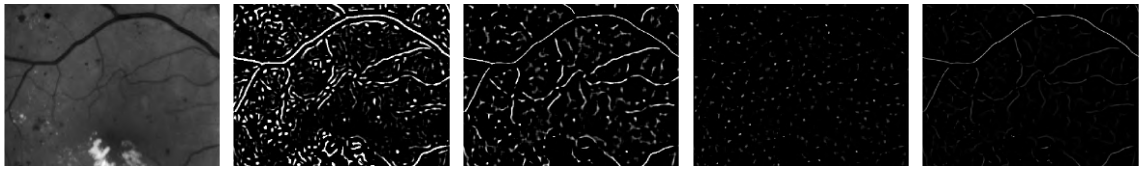


Figure 5.1: Fundus images examples. Black arrows denote the OD. White arrows denote exudative lesions.



(a) (b) (c) (d) (e)

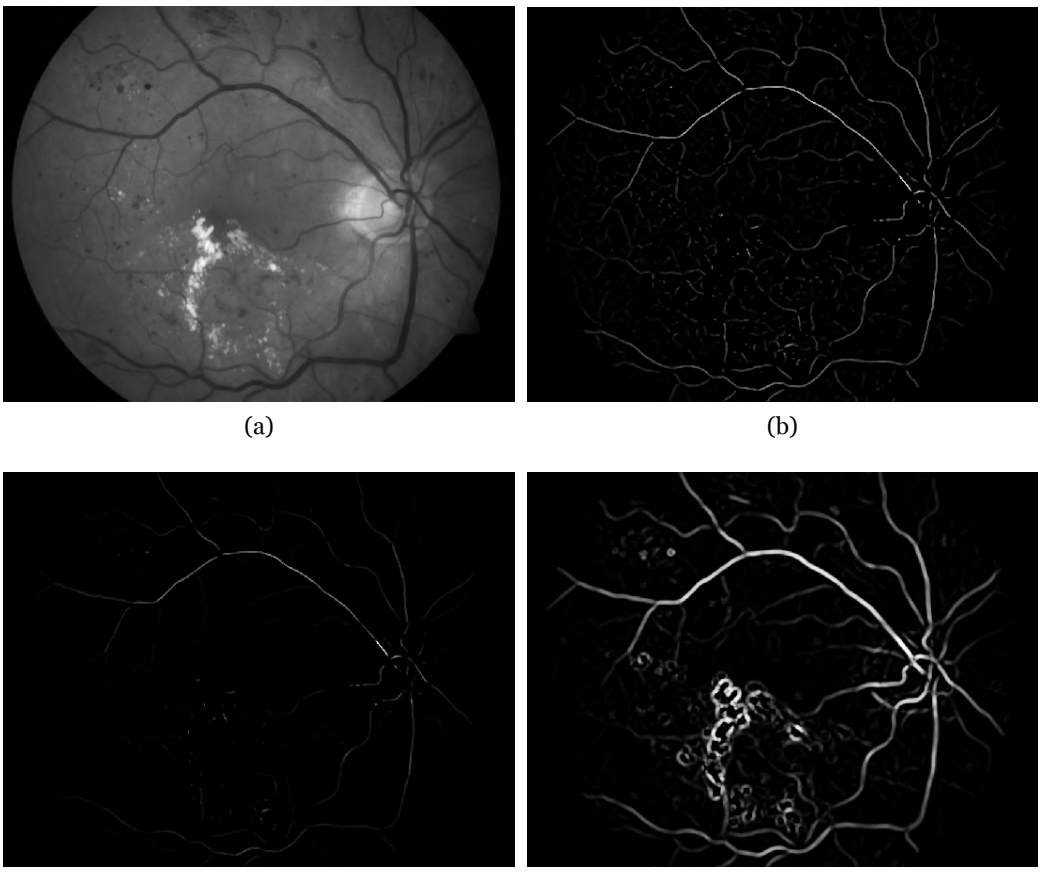
Figure 5.2: Vessel enhancement. a) Detail of a fundus image. b) The corner detector $K(\mathbf{x})$. c) The vessel detector $\tilde{K(\mathbf{x})}$. d) The second term in equation (B.4). e) The vessel enhancement $V(\mathbf{x})$.

of the parameters that govern the algorithm. If $I_G(\mathbf{x})$ has either a width or length larger than 900 pixels, it is proportionally scaled using a bicubic interpolation in such a manner that the larger dimension is 900 pixels. Next, a binary mask $BW(\mathbf{x})$ is created based on the method proposed by Frank ter Haar [65]. An empirical threshold t =28 is applied to the red channel component $I_R(\mathbf{x})$, followed by the morphological operations closing and erosion using a "disk" structuring element of size 5 pixels. Only the $BW(\mathbf{x})$ white pixels are considered in further calculations. In the following, a vessel enhancement method is described.

5.1.1 Vessel enhancement

Ideally, a good vessel enhancement method must be fast and minimize the detection of non-vessel structures or pathological clinical manifestations [205, 73]. The influence of noise and the presence of bright central reflections in arteries must also be considered [266]. The proposed vessel enhancement is specifically constructed to deal with all the previous mentioned aspects, being a modified version of the corner detector developed by Wang and Brady [267]. First, we create a filtered image $L(\mathbf{x})$ by blurring $I_G(\mathbf{x})$ with a normalized Gaussian filter with a $\sigma = 4$ and a kernel size of 13×13, reducing noise and eliminating the arterial central bright regions. The initial corner detector $K(\mathbf{x})$ proposed in [267] is given by,

$$K(\mathbf{x}) = (\nabla^2 L(\mathbf{x}))^2 - c|\nabla L(\mathbf{x})|^2, \quad (5.1)$$



(a) (b)

(c) (d)

Figure 5.3: Vessel enhancement examples. a) Fundus image. b) Result of the proposed vessel enhancement $V(\mathbf{x})$. c) Result of the proposed weighted vessel enhancement $\Gamma(\mathbf{x})$. d) Result of the vessel enhancement proposed by Frangi *et al.* [73]. All images are normalized.

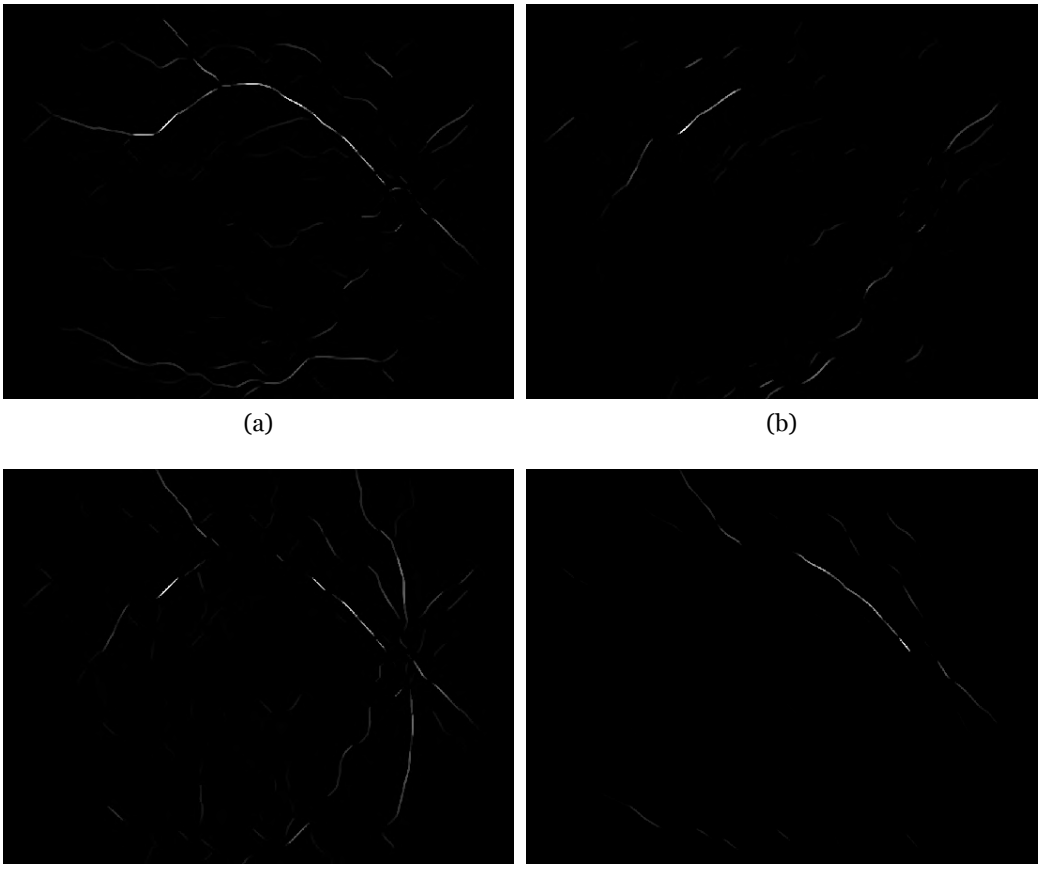
where $\nabla^2 L(\mathbf{x})$ is the Laplacian and $|\nabla L(\mathbf{x})|$ is the absolute value of the gradient of the image $L(\mathbf{x})$, considered as an intensity surface. $K(\mathbf{x})$ enhances regions where a rapid change in the edge direction occurs. The parameter c defines how edge-phobic is $K(\mathbf{x})$, being set to $c = 1$ in our implementation. To enable a stronger response of the vessel structures, the square of the first term in equation 5.1 is removed, resulting in,

$$\tilde{K(\mathbf{x})} = \nabla^2 L(\mathbf{x}) - c|\nabla L(\mathbf{x})|^2. \quad (5.2)$$

The vessel enhancement equation is then given by,

$$V(\mathbf{x}) = \tilde{K(\mathbf{x})} - \min_{\forall \mathbf{x} \in BW(\mathbf{x})} (\lambda_1(\mathbf{x}), \tilde{K(\mathbf{x})}). \quad (5.3)$$

$\lambda_1(\mathbf{x})$ is the minimum principal curvature as defined in chapter B, and enhances dark blood blobs. Hence, for the vessels localizations, $\lambda_1(\mathbf{x})$ becomes smaller than $\tilde{K(\mathbf{x})}$ and a large value of $V(\mathbf{x})$ is observed. Moreover, regions like blood blobs or microaneurysms close to the vessels, resulting in the most responsive locations, the minimum of the blob or microaneurysm, become less relevant in the final vessel detection. The process can be visualized in Figure 5.2. The vessel enhancement described in equation B.4 allows the preservation of elongated structures and good discrimination between a vessel and non-vessel structures. The proposed vessel enhancement $V(\mathbf{x})$, and the



(a) (b)

(c) (d)

Figure 5.4: The four main vessels orientations for the single fundus image in Figure 5.3. a) $\gamma(\mathbf{x})^{0^o}$, b) $\gamma(\mathbf{x})^{45^o}$, c) $\gamma(\mathbf{x})^{90^o}$, and d) $\gamma(\mathbf{x})^{135^o}$. The images were dilated with a small structuring element to allow a better visualization.

Frangi vessel enhancement [73] also based in the Hessian matrix eigenvalues, are shown in Figure 5.3. Both $V(\mathbf{x})$ and Frangi vessel enhancement was calculated at the same scale. From the comparison of Figure 5.3 (b) and Figure 5.3 (d) it can be observed that the vessel enhancement method given by $V(\mathbf{x})$ provides better discrimination between vessels and non-vessels structures, particularly when bright lesions are present. Nevertheless, in pathological images $V(\mathbf{x})$ can have two disadvantages. Firstly, although blood blobs and microaneurysms become less relevant, they can still be very intense. Secondly, $V(\mathbf{x})$ may enhance regions that appear similar to vessels, created by the depressions between bright regions, e.g., exudates. Since blood blobs and microaneurysms are transformed in their most responsive point, their influence can be reduced as explained in the next subsection. The second disadvantage can be overcome with the inclusion of a proper weight parameter. Also defined in B, the maximum curvature $\lambda_2(\mathbf{x})$ offers a stronger response in more thick vessels, favoring the ones belonging to the main arcades, and the ones crossing the OD. A weighted version of $V(\mathbf{x})$, using $\lambda_2(\mathbf{x})$ as a weight parameter can be defined. Hence, for any point $(\mathbf{x})$ the following equation was defined,

$$\Gamma(\mathbf{x}) = \begin{cases} \lambda_2(\mathbf{x})V(\mathbf{x}), & \text{if } \lambda_2(\mathbf{x}) > 0 \\ 0, & \text{otherwise} \end{cases} . \quad (5.4)$$

The weighted vessel enhancement $\Gamma(\mathbf{x})$ minimizes the impact of false vessels and will play a fundamental role in the OD localization (see an example in Figure 5.3).

5.1.2 The four main vessels orientations detection

The retinal vessels emerge from the OD vertically and branch out horizontally [58]. Although other regions in the retina may have vertical vessels (e.g. around the macula), these are thinner than the ones around the OD. Between the vertical and horizontal orientations, the vessels assume mainly oblique (approximately 45^o and 135^o) orientations [64]. Morphological operators provide an efficient extraction methodology of the four main vessels orientations $\alpha = \{0^o, 45^o, 90^o, 135^o\}$ [191]. Since we intend to enhance linear structures, a logical choice for a structuring element is a "line" with a variable length and a variable angle covering both the short and long vessels. The morphological opening of $\Gamma(\mathbf{x})$ by a "line" structuring element S with an angle θ and a length l, can be defined as,

$$\gamma(\mathbf{x}) = \Gamma(\mathbf{x}) \circ S_l^\theta. \quad (5.5)$$

The opening operation is performed in a set of θ values, and the maxima are detected for each l value. This is done defining the set values $\theta_i = \{\theta_1, ..., \theta_n\}$, $i = 1...n$, and a fixed set $l = \{5, 10, 15, 20, 25\}$ of pixels. The orientation vessel enhancement in the α orientation is defined as the summation of the maxima obtained for each l value, with,

$$\gamma^\alpha(\mathbf{x}) = \sum_{k=1}^{5} \max_{\forall \theta_i} (\Gamma(\mathbf{x}) \circ S_{5k}^{\theta_i}). \quad (5.6)$$

For each $\alpha = \{0^o, 45^o, 90^o, 135^o\}$, the following θ values are used specifically,

$$\begin{array}{ll} \gamma^{0^o}(\mathbf{x}), & \theta = \{0^o, 15^o, 30^o, 45^o, 135^o, 150^o, 165^o\} \\ \gamma^{45^o}(\mathbf{x}), & \theta = \{30^o, 40^o, 50^o, 60^o\} \\ \gamma^{90^o}(\mathbf{x}), & \theta = \{45^o, 60^o, 75^o, 90^o, 105^o, 120^o, 135^o\} \\ \gamma^{135^o}(\mathbf{x}), & \theta = \{120^o, 130^o, 140^o, 150^o\} \end{array} \quad (5.7)$$

Note that by defining a minimum length value of 5 pixels, we avoid lesions like the microaneurysms and blood spots mentioned in the previous subsection. Furthermore, to assimilate the vessels tortuosity that can exist in some fundus images, a smaller angle increment in the oblique orientations is used. These values of θ and l were revealed to be the best trade-off between accurate detection and computation time after extensive testing. Examples of results for equations 5.6 and 5.7 are shown in Figure 5.4.

5.1.3 Initial optic disc detection

The orientation vessel enhancement functions $\gamma^\alpha(\mathbf{x})$, provide all the necessary information for the determination of an initial OD position, designated by (p_x, p_y). To calculate (p_x, p_y), the horizontal p_x and vertical p_y coordinates are determined separately. Two specific functions, one for each coordinate respectively are created according to the following equations,

$$f_1(\mathbf{x}) = \gamma^{0^o}(\mathbf{x}) - \gamma^{90^o}(\mathbf{x}), \quad (5.8)$$

$$f_2(\mathbf{x}) = \gamma^{90^o}(\mathbf{x}) + \gamma^{45^o}(\mathbf{x}) + \gamma^{135^o}(\mathbf{x}). \quad (5.9)$$

Although the vessels that emerge from the OD have a dominant vertical orientation, in some cases they can exhibit some variation. The same situation can be found when the vessels branch hori-

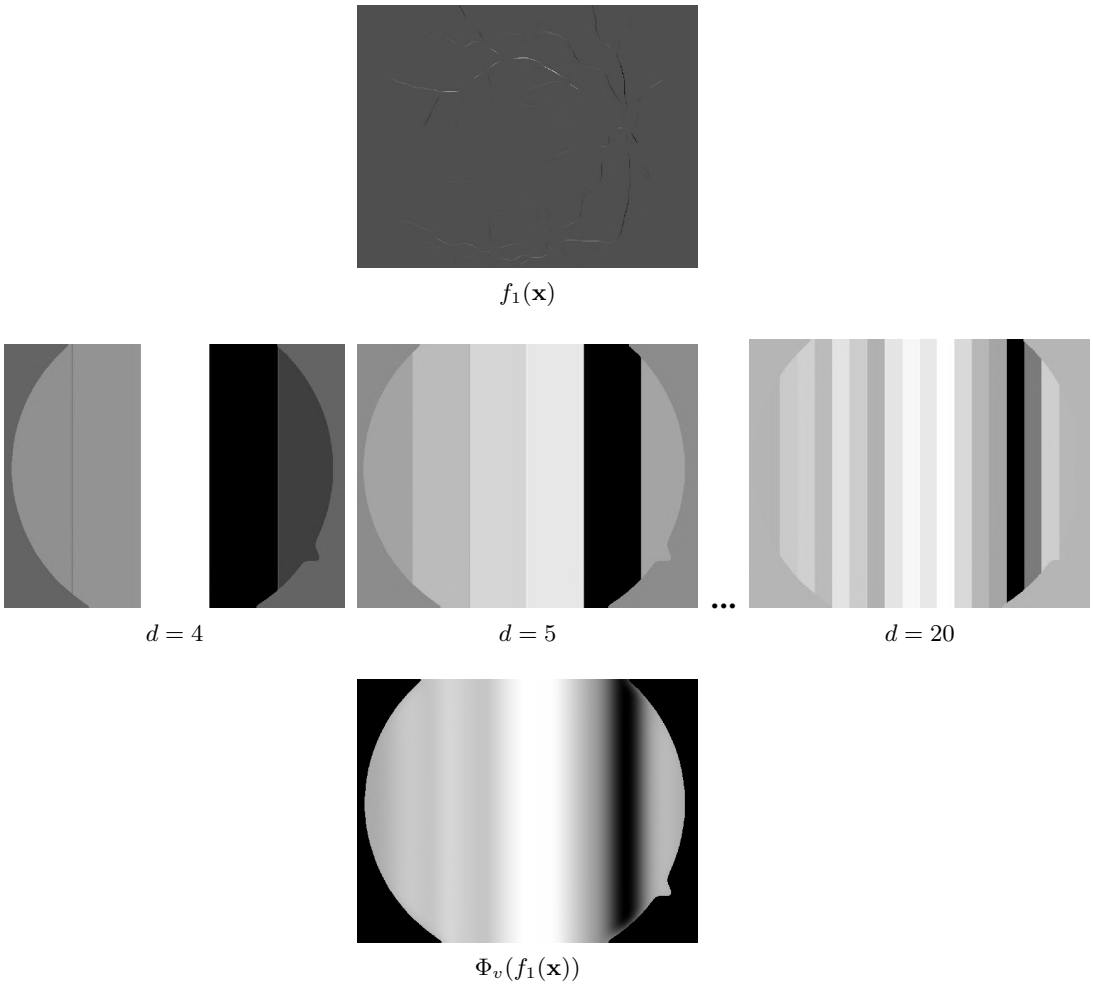


Figure 5.5: Computation of the vertical field $\Phi_v(f_1(\mathbf{x}))$. The first row shows the result of equation 5.8. The middle row represents 2D step functions of the successive subdivisions given by equation 5.10. The bottom row represents the filtered result of equation 5.13.

zontally, where they also exhibit some variation. Hence, to improve the method reliability, both vertical and horizontal orientations also consider oblique orientations, as can be seen in equation 5.7. Furthermore, the vertical vessels have a larger width in the OD region than in other regions of the retina, resulting in higher $\Gamma(\mathbf{x})$ values. The difference in equation 5.8 ensures a larger difference between the horizontal and vertical vessels values, allowing the determination of the OD horizontal coordinate p_x. In the OD horizontal coordinate, p_x, the difference of equation 5.8 is higher because vessels have dominant vertical directions. In order to determine the OD vertical coordinate p_y, both vertical and oblique orientations must be considered, leading to the summation in equation 5.9. The vertical coordinate uses the fact that in the proximity of the OD there is a high confluence of oblique and vertical vessels, creating the superior and inferior vessels arcades. These two vessels arcades create two maxima in $f_2(\mathbf{x})$. Furthermore, between the two vessels arcades, the vertical and oblique vessels almost disappear, creating a minimum in $f_2(\mathbf{x})$ close to the OD location. Since these coordinates are determined separately, the description of our method is continued using $f_1(\mathbf{x})$ and the coordinate p_x as an example. This is easily extended to $f_2(\mathbf{x})$ and the coordinate p_y.

The function $f_1(\mathbf{x})$ with domain $[1, N] \times [1, M]$, where N and M are, respectively, the number of rows and columns (first row in Figure 5.5), is d-times subdivided along the vertical direction, creating $r_i (i = 1, .., d+1)$, equally vertical disjoint regions. The following 2D step function, $\phi_v^d(f_1(\mathbf{x}))$,

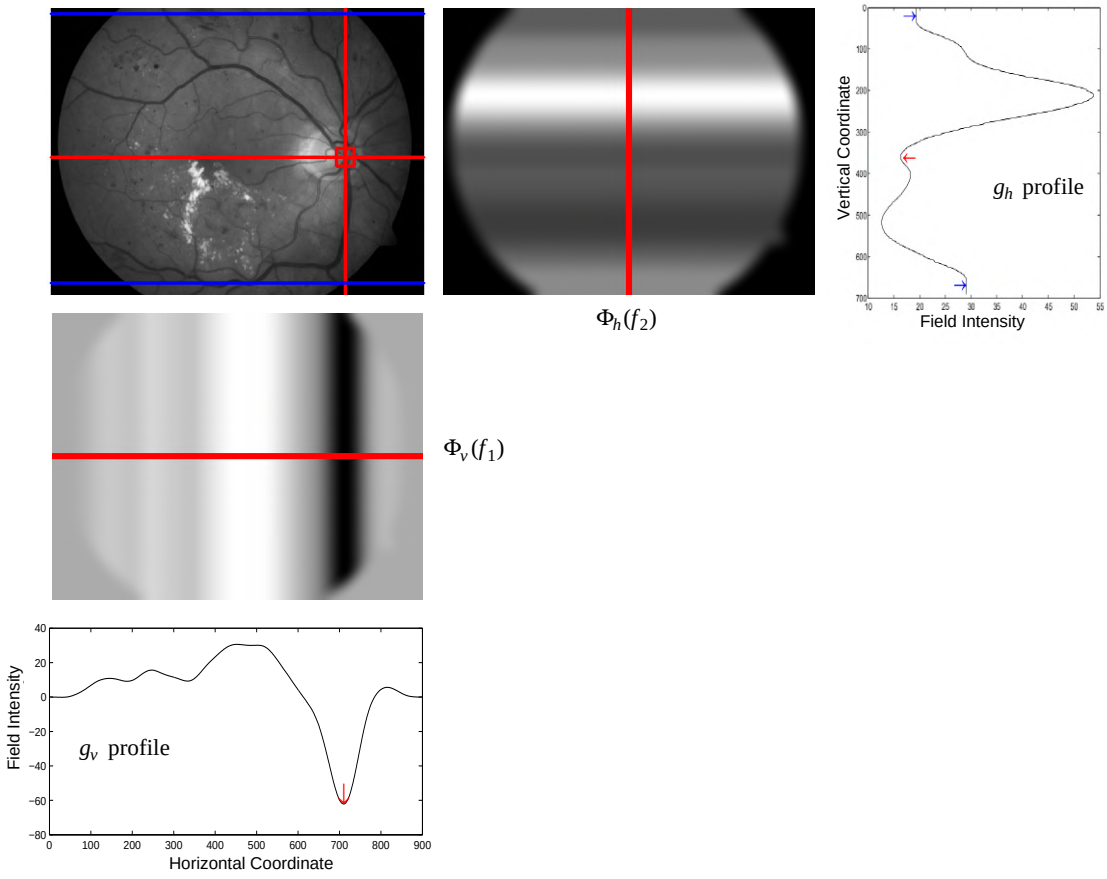


Figure 5.6: Determination of the initial OD localization (p_x, p_y). The horizontal blue lines in the $I_G(\mathbf{x})$ image limit the area where the OD can be found. The vertical and horizontal red lines in the $I_G(\mathbf{x})$ image represents minima from each profile. The perpendicular red lines, in each horizontal and vertical fields, select the points where the g_v and g_h profiles are determined. The red square defines the point calculated as (p_x, p_y).

is defined as,

$$\phi_v^d(f_1(\mathbf{x})) = \bigcup_{i=1}^{d+1} \frac{\sum\sum_{\forall \mathbf{x} \in (r_i \wedge BW(\mathbf{x}))} f_1(\mathbf{x})}{\sum\sum_{\forall \mathbf{x} \in r_i} BW(\mathbf{x})}, \quad (5.10)$$

where $BW(\mathbf{x})$ is the binary mask previously defined, d is the number of divisions, and v stands for vertical, indicating the orientation of the divisions (middle row in Figure 5.5). $\phi_v^d(f_1(\mathbf{x}))$ performs an average of $f_1(\mathbf{x})$ at each vertical region r_i. Next, $f_1(\mathbf{x})$ is successively subdivided in smaller intervals, and the cumulative sum of the corresponding 2D step functions $\phi_v^d(f_1(\mathbf{x}))$, is performed accordingly with the following equation,

$$\Phi_v(f_1(\mathbf{x})) = \sum_{d=\varepsilon}^{\xi} \phi_v^d(f_1(\mathbf{x})). \quad (5.11)$$

The values ξ and ε are respectively, the maximum and the minimum number of divisions. To suppress the abrupt changes resulting from the cumulative sums, several Gaussian filters were tested. The Gaussian filter with a σ =15 and a kernel size of 51×51 proved to be the most effective (see example in the last row of Figure 5.5). $\Phi_v(f_1(\mathbf{x}))$ represents an averaged version of $f_1(\mathbf{x})$

along the vertical direction. Theoretically, the minimum and maximum values that d can assume are 0 and $M-1$, respectively. However, in a small fundus image (700×600), the OD typically has a diameter of 80 pixels. Considering that the OD diameter can be up to double this value ($\sim \frac{1}{5}$ of 700), it is logical to assume a minimum value ε =4, creating 5 initial disjoint regions. The maximum value ξ, is defined according to the image dimensions as, $\xi = \text{round}((\max(N, M))/\mu)$, where $\text{round}(.)$ represents the rounding function and μ is set according to the orientation of the divisions.

Analogously, we define the average of $f_2(\mathbf{x})$ along the horizontal direction as,

$$\Phi_h(f_2(\mathbf{x})) = \sum_{d=\varepsilon}^{\xi} \phi_h^d(f_2(\mathbf{x})). \quad (5.12)$$

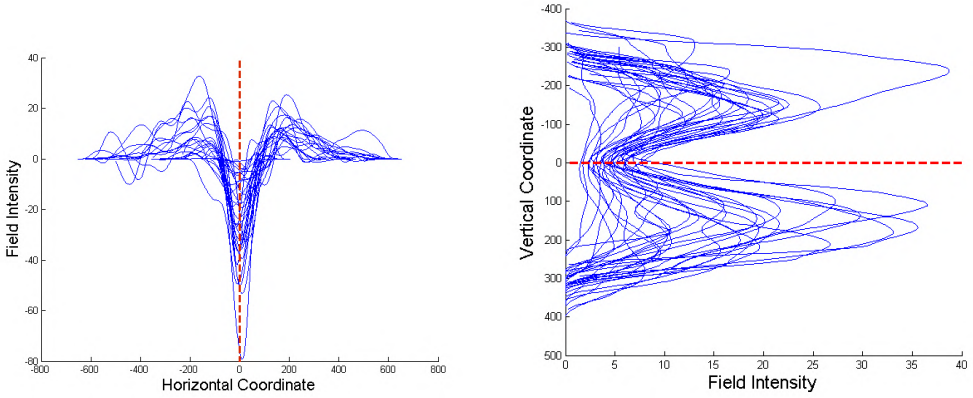
The construction of equations 5.11 and 5.12, with the corresponding upper and lower limits are subsequently,

$$\Phi_v(f_1(\mathbf{x})) = \sum_{d=4}^{\xi} \phi_v^d(f_1(\mathbf{x})),\ \mu = 70, \quad (5.13)$$

$$\Phi_h(f_2(\mathbf{x})) = \sum_{d=4}^{\xi} \phi_h^d(f_2(\mathbf{x})),\ \mu = 45. \quad (5.14)$$

Equations 5.13 and 5.14 can be interpreted as cumulative sum fields, indicating the presence or absence of specific vessels orientations on $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ (defined by the equations 5.8 and 5.9). The values of μ were defined empirically after extensive testing.

The use of several division values, d, intend to provide extra adaptability to different image resolutions and OD sizes, compensating the d values where the response on $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ fails. Hence, the cumulative sums combine the information that results from each d value division leading to a larger value on the OD region, reinforcing its position both in the horizontal and vertical orientations. Moreover, if the OD region is unfocused, a single value of d might lead to a larger response of $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ in other regions that has thinner vessels but are focused. This situation is also alleviated with the cumulative sums. By using the cumulative sums, the vertical and horizontal position of the OD becomes reinforced. A simplified analysis of the cumulative sums fields is achieved by tracing the respective profile g_v using the middle row of $\Phi_v(f_1(\mathbf{x}))$, and the profile g_h using the middle column of $\Phi_h(f_2(\mathbf{x}))$. The minima of each profile are identified by the downward red arrows and the maxima by the upward blue arrows (see Figure 5.6). These profiles are very similar for all the fundus images (see Figure 5.7 where profiles obtained from different fundus images are represented), being possible to derive several important properties. From Figure 5.7 it is possible to observe that both the horizontal p_x and the vertical p_y coordinates must be defined by a minimum in each profile. In particular, Figure 5.7 a) suggests that the horizontal coordinate p_x is defined by the absolute minimum of g_v. This is explained by $f_1(\mathbf{x})$ definition in equation 5.8 that results in a larger difference between the vertical and horizontal vessels orientations, leading to the lower region in equation 5.13 and the absolute minimum of g_v. Thus, p_x is set based on the absolute minimum of g_v. For the vertical coordinate, we can observe in Figure 5.7 b) that p_y is defined by the minimum between the two major maxima of the profile g_h. This is explained by $f_2(\mathbf{x})$ definition in equation 5.9 that uses the confluence of the vertical and oblique vessels to the OD region, and the fact that the OD is localized between the vessels arcades. However, in some fundus images there can be a vein occlusion, defocus or an untypical vessels disposition that di-



a) g_v profiles b) g_h profiles

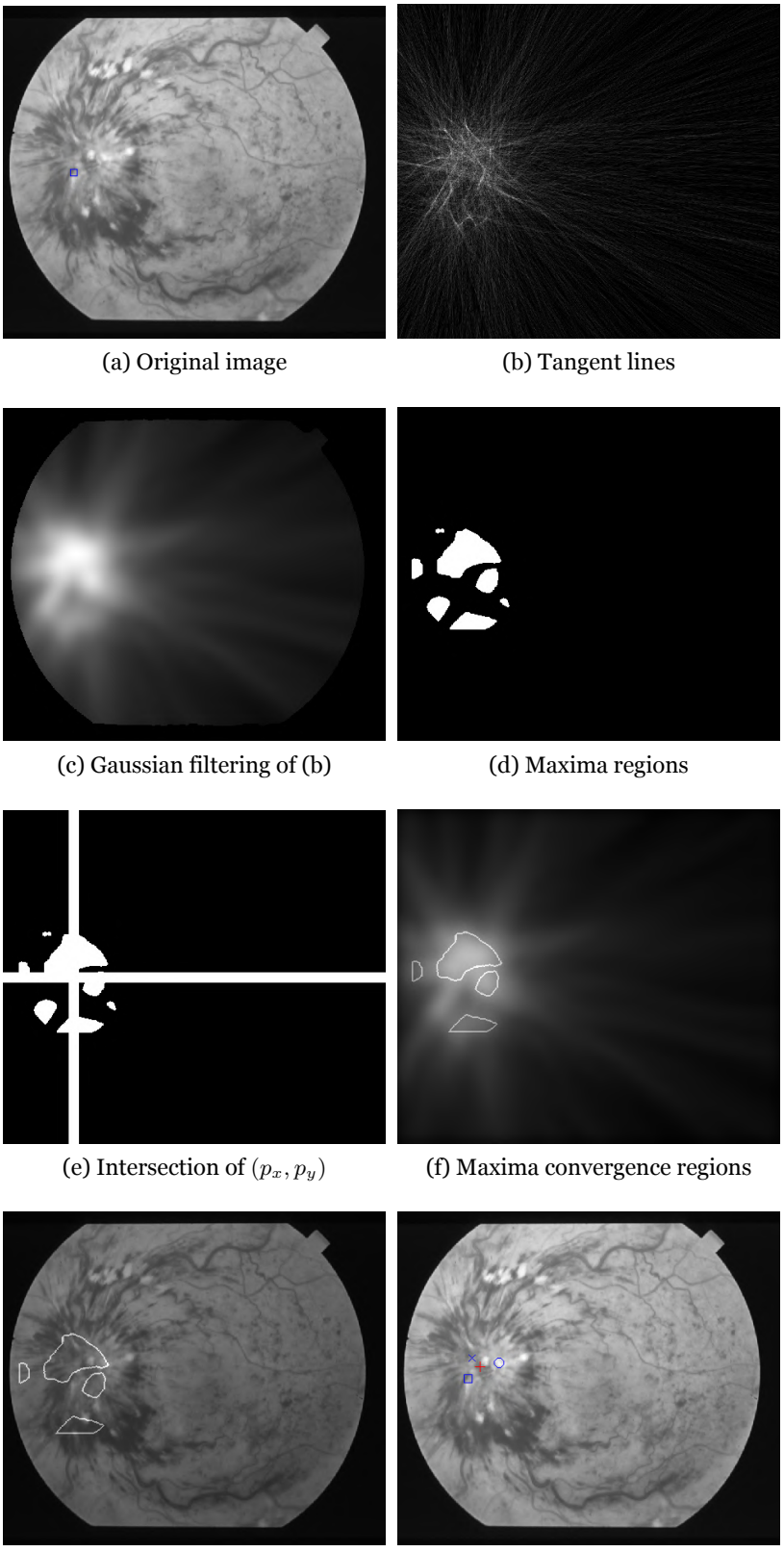
Figure 5.7: Superposition of 30 random g_v and g_h profiles. The profiles were shifted, localizing the OD in the origin (indicated by the dotted red line).

minishes the presence of one of the vessels arcades. These situations can create other minima and maxima in g_h. Hence, despite p_y being defined by a minimum of g_h, it might happen that it is not the absolute minimum. To determine p_y the two most displaced maxima, of g_h are initially considered (represented by the horizontal blue lines in $I_G(\mathbf{x})$ the image in Figure 5.6). If none of the problems described above affect the image, these blue lines correspond to the two major maxima of g_h. p_y is defined as the minimum that exhibits the larger variation of g_h among its neighbouring maxima. In other words, if a given minimum limited by the two most displaced maxima of g_h is defined as p_{min}, and its neighbouring maxima are defined as p_{max_1} and p_{max_2}. p_y is the minimum with the larger $(|p_{min} - p_{max_1}| + |p_{min} - p_{max_2}|)$ value. The extremes localization computation uses a window of 61 consecutive values of the considered profile. If the profile value is smaller than the extremes of the window, then a minimum exists inside that window. The smaller value of the profile inside the window is considered as the minimum location. A similar strategy is used for maxima localization. The point selected as (p_x, p_y) is marked in the $I_G(\mathbf{x})$ image as the red square (Figure 5.6 top left image).

In the large majority of the fundus images, (p_x, p_y) is the correct localization of the OD, but in some cases, it may be slightly dislocated and outside the OD region. To overcome these situations, a vessel convergence algorithm is described in the next section. It is derived to pinpoint the exact localization of the OD. The previously described approach has some similarities with the followed by Mahfouz *et al.* [58]. However, our method uses a different vessel detector that is more resilient than the edge detector of [58] to pathological clinical manifestations. This work also uses a set of image divisions with different sizes instead of the sliding window in [58], resulting in improved reliability to different image resolutions and OD sizes. Furthermore, this work uses the vessel's oblique orientations to improve the vertical coordinate detection.

5.1.4 Final localization of the optic disc

The final step of the developed method consists of the determination of the final OD localization. To achieve this, a technique based on the two most distinct features of the OD was developed: High convergence of vessels and high-intensity values. First, vessels convergence regions are defined. Subsequently, the maximum point of the vessels convergence, designated by (c_x, c_y), is computed inside those regions. Moreover, the maximum point of intensity inside the defined ves-



(a) Original image (b) Tangent lines

(c) Gaussian filtering of (b) (d) Maxima regions

(e) Intersection of (p_x, p_y) (f) Maxima convergence regions

(g) Maxima intensity regions (h) Final OD localization

Figure 5.8: A scheme of the final OD localization calculation. The blue square (□) in the first figure represent the (p_x, p_y) point. The blue cross (×) and circle (○), are respectively the (c_x, c_y) and (b_x, b_y) points. The red cross (+) is the final OD localization.

sels convergence regions, designated by (b_x, b_y), is also computed. The aim is to lead the initial OD localization (p_x, p_y) to a nearby localization with a higher vessel convergence (c_x, c_y) and a higher intensity (b_x, b_y).

In particular, the point (c_x, c_y) is obtained tracing a tangent line at each vessel pixel in $\Gamma(\mathbf{x})$, followed by the determination of the point where maxima intersection occurs. To evaluate the orientation at each vessel pixel, a structure tensor was used. As referred by van Vliet *et al.* [268], and Breuß *et al.* [269], the structure tensor yields an excellent characterization of the local dimensionality, and of the corresponding orientation for simple neighbourhoods. Simple neighbourhoods exhibit a single orientation, like in vessels features, making it a robust estimation of the orientation information. The structure tensor can be defined as,

$$T(\mathbf{x}) = \begin{bmatrix} \overline{\Gamma_x^2(\mathbf{x})} & \overline{\Gamma_x(\mathbf{x})\Gamma_y(\mathbf{x})} \\ \overline{\Gamma_x(\mathbf{x})\Gamma_y(\mathbf{x})} & \overline{\Gamma_y^2(\mathbf{x})} \end{bmatrix}, \quad (5.15)$$

where the overhead denotes a local Gaussian filtering with a σ =0.8 and a kernel size of 5×5 [270]. The differential operators $\Gamma_x(\mathbf{x})$ and $\Gamma_y(\mathbf{x})$, were calculated as in [271], considering the $\Gamma(\mathbf{x})$ defined in equation (5.4) improving the orientations calculation. The parameter β corresponds to the orientation of the vessel pixels, and it is computed by,

$$\beta = \frac{1}{2}\arctan\left(\frac{\overline{\Gamma_x^2(\mathbf{x})} - \overline{\Gamma_{y(\mathbf{x})}^2}}{2\overline{\Gamma_x(\mathbf{x})\Gamma_y(\mathbf{x})}}\right). \quad (5.16)$$

Before tracing the tangential lines, a morphological thinning was performed followed by a threshold of $\Gamma(\mathbf{x})$ to select the most relevant pixels. This threshold was set experimentally on $\frac{0.2}{\lambda_2(\mathbf{x})}$. Furthermore, only the pixels that lie inside a circular region centered in (p_x, p_y) with a diameter of δ pixels are considered. The value of δ depends on the initial resolution of the image, and its calculation is explained in subsection 5.1.4. The tangential lines are exemplified in Figure 5.8(b). The image with the tangential lines was filtered by a Gaussian (defined experimentally with a σ =15 and a kernel size of 51×51). Figure 5.8(c) represents the result of the Gaussian filtering, considering only the region defined by the binary mask $BW(\mathbf{x})$. In this image, the high intensity regions represent a high vessel convergence. To select the most relevant convergence regions, the maxima were selected using,

$$f_{xx}(\mathbf{x})f_{yy}(\mathbf{x}) - f_{xy}^2(\mathbf{x}) > 0 \wedge f_{xx}(\mathbf{x}) < 0, \quad (5.17)$$

where $f(\mathbf{x})$ is a generic 2D function, and the second-order partial derivatives were evaluated using a larger mask $[-1\ v_1\ v_2\ \ldots\ v_n\ 1]$ with $v_i = 0, \forall i$ and the corresponding transpose. n was experimentally set equal to 31 ensuring the detection of larger regions. Figure 5.8 (d) shows an example of the application of these masks. This way has avoided the use of a simple threshold that will not adapt to different OD dimensions. Next, the point (p_x, p_y) was prolonged in the vertical and horizontal orientations, followed by a dilation with a "disk" structuring element with a size of 10 pixels (see Figure 5.8 (e)). Then, only the regions that intersect these dilated lines are considered. In each of these regions the point with the highest convergence, (c_x, c_y), and the point with the highest intensity in $I_G(\mathbf{x})$, (b_x, b_y) are selected (Figure 5.8 (f) and 5.8 (g)). The final OD position is given by the average of (p_x, p_y), (c_x, c_y), and (b_x, b_y) (represented by the red cross in the example of Figure 5.8(h)).

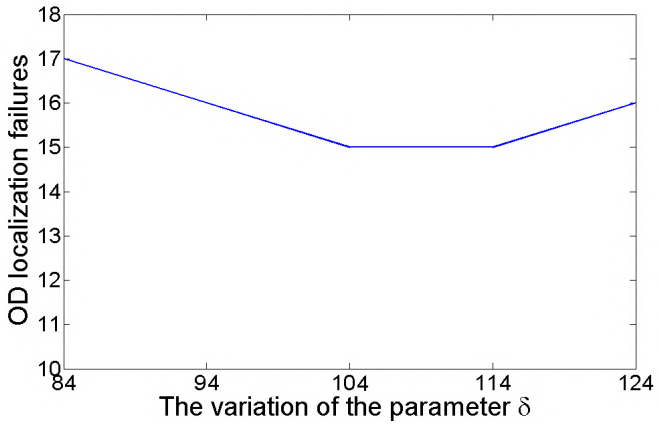


Figure 5.9: Variation of the OD localization failure with the δ parameter.

5.2 Results

An automatic system for the localization of the OD has been presented. To evaluate the proposed algorithm, eight publicly available datasets were used. In this section, the quantitative and qualitative results are presented. The public datasets are first described. Then the performance of the proposed method is presented and discussed.

5.2.1 Datasets and evaluation criteria

The eight public datasets of fundus images used for testing were: 1) the Stare dataset [169] with 31 images of healthy retinas and 50 images of pathological retinas, which are widely used for benchmarking of the OD localization in the literature. 2) The Drive dataset [43] is composed of 40 fundus images, and was created for retinal blood vessel segmentation benchmarking. 3) The Diaretdbv0 [44] and Diaretdbv1 [45], which are composed of 130 and 89 fundus images respectively, were created for DR detection benchmarking. 4) The Messidor [167] with 1200 fundus images was created for DR studies. 5) The sixth dataset is ROC [48], composed of 50 images and was created for microaneurysm detection benchmarking. 6) The E-ophtha-EX dataset [49] was particularly designed for scientific research in DR and is composed of 47 images with exudates and 35 images with no lesion. Finally, the High-Resolution Fundus (HRF) Image dataset [50] was also used, composed of 45 images and designed for scientific research in DR and glaucoma. These datasets are described in further detail in section A

The OD center was manually labelled by a retinal expert in all retina images of all datasets, except in the HRF dataset, creating the corresponding OD center groundtruth. For the HRF dataset, the provided OD center groundtruth was used. The calculated OD localization is considered correct if it is inside the OD region, including its borders. If no OD region is clearly visible, the OD localization was inferred as being correct if it was positioned within 60 pixels of the manually identified center, as proposed in [162, 64, 169]. Each image annotation was analysed and verified by an independent health care professional (Medical Doctor), with extensive experience in fundus image analysis.

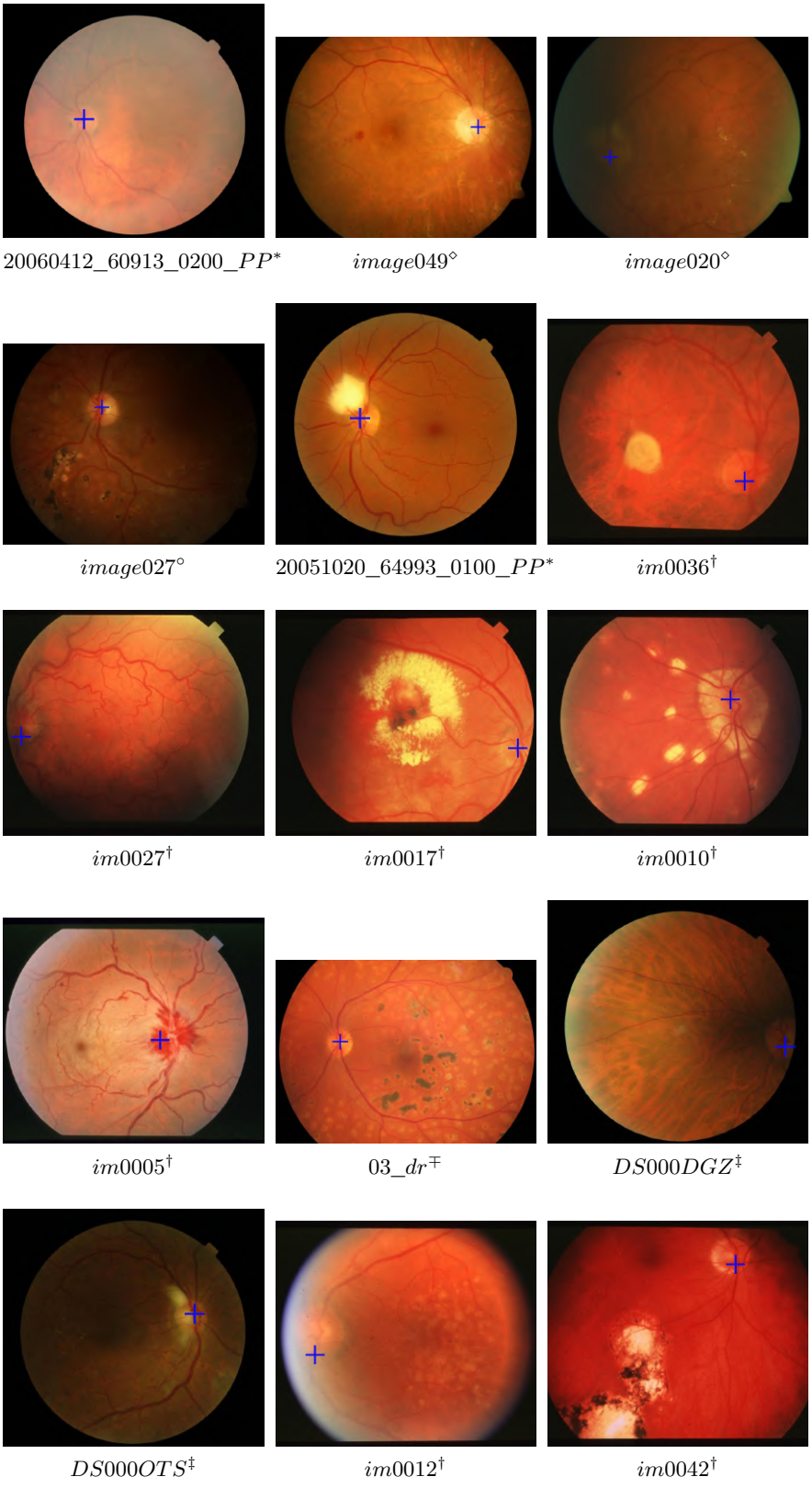
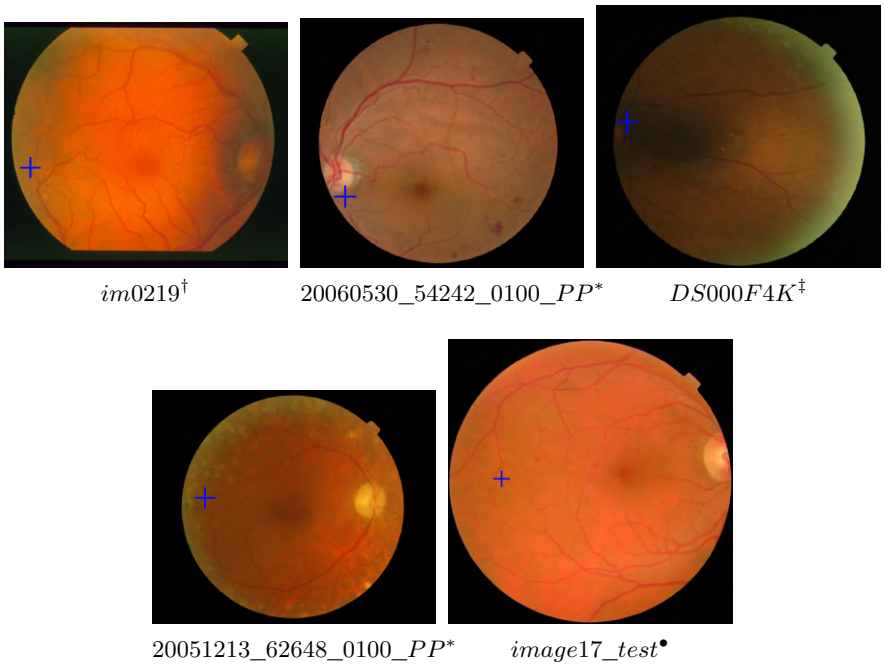


Figure 5.10: OD localization examples. The blue (+) represents the estimated OD center. Each image has the corresponding name and dataset (* Messidor, † Stare, ‡ E-ophtha-EX, ◇ Diaretdbv0, ° Diaretdbv1, ∓ HRF).

Table 5.1: Number of failed OD localizations using the proposed method with and without the final localization step of subsection 5.1.4.

	Stare	**Drive**	**Diaretdvbo**	**Diaretdvb1**	**Messidor**	**ROC**	**E-ophtha-EX**	**HRF**
Number of images	81	40	130	89	1200	100	82	45
Without step 5.1.4	6	2	5	3	38	6	4	1
With step 5.1.4	1	0	2	1	9	1	1	0



$im0219^{\dagger}$ $20060530\_54242\_0100\_PP^{*}$ $DS000F4K^{\ddagger}$

$20051213\_62648\_0100\_PP^{*}$ $image17\_test^{\bullet}$

Figure 5.11: Examples of OD localization failures. The blue (+) represents the estimated OD center. Each image has the corresponding name and dataset († Stare, * Messidor, ‡ E-ophtha-EX, $^{\bullet}$ ROC).

5.2.2 Optic disc localization results

A Matlab algorithm of the proposed method was implemented on a laptop with 2 GHz Intel Core i7 and 6 GB of RAM. Quantitative results for the proposed method without the final localization step of subsection 5.1.4 are shown in Table 5.1. It is evident that the inclusion of this last step considerably increases the performance of the proposed algorithm. As referred in subsection 5.1.4, the diameter value δ defines the size of the search area in which the OD can be localized and needs to be automatically defined. Figure 5.9 represents the variation of the number of failed detections when $\delta \in \{84, 94, 104, 114, 124\}$. After some tests, and based on the results of Figure 5.9, it was verified that δ depends on the initial resolution of the image. Hence, δ was set to 114 in case of the $min(width, height)$ $\leq$600 pixels or set to 104 otherwise. The $min(width, height)$ is the minimum value between the initial $width$ and $height$ of the image before the resizing process. Nevertheless, based on Figure 5.9 it is possible to observe that small variations of the δ do not change the results significantly. Qualitative results are shown in Figure 5.10 and Quantitative results are shown in Table 5.2, where a comprehensive and detailed review is made relative to the most used datasets in the literature, in terms of accuracy.

An OD localization accuracy of 99.15% is obtained for the 1767 fundus images within the eight publicly datasets. In particular for the Stare and Drive datasets, an accuracy of 98.77% (80/81), and 100.00% (40/40) was obtained respectively, which equals the best results of other state-of-the-art methods, ([64, 167, 184] in the Stare dataset and [64, 58, 60, 61, 184] in the Drive dataset).

Table 5.2: Optic disc localization methods comparison on several datasets.

Stare dataset				
Methods	Acc(%)	Fail	Evaluation Criteria	Evaluation Method
Proposed method (2016) [55]	**98.77**	1	60 pix	All images
Dietter *et al.* (2019) [61]	90.10	8	Boundary	All images
Ramakanth *et al.* (2014) [60]	93.83	5	60 pix	All images
Lu and Lim (2011) [59]	96.30	3	Boundary	All images
Lu (2011) [167]	**98.77**	1	Boundary	All images
Youssif *et al.* (2008) [64]	**98.77**	1	60 pix	All images
Roychowdhury *et al.* (2016) [184]	**98.77**	1	Boundary	All images*

Drive dataset				
Methods	Acc(%)	Fail	Evaluation Criteria	Evaluation method
Proposed method (2016) [55]	**100.00**	0	60 pix	All images
Dietter *et al.* (2019) [61]	**100.00**	0	Boundary	All images
Ramakanth *et al.* (2014) [60]	**100.00**	0	60 pix	All images
Lu and Lim (2011) [59]	97.50	1	Boundary	All images
Welfer *et al.* (2010) [105]	**100.00**	0	60 pix	All images
Youssif *et al.* (2008) [64]	**100.00**	0	60 pix	All images
Bajwa *et al.* (2019) [187]	97.50	1	Overlapping	All images*
Roychowdhury *et al.* (2016) [184]	**100.00**	0	Boundary	All images*

Diaretdbo dataset				
Methods	Acc(%)	Fail	Evaluation Criteria	Evaluation method
Proposed method (2016) [55]	98.46	2	60 pix	All images
Dietter *et al.* (2019) [61]	97.80	3	Boundary	All images
Ramakanth *et al.* (2014) [60]	98.46	2	60 pix	All images
Lu and Lim (2011) [59]	99.23	1	Boundary	All images
Roychowdhury *et al.* (2016) [184]	**100.00**	0	Boundary	All images*

Diaretdb1 dataset				
Methods	Acc(%)	Fail	Evaluation Criteria	Evaluation method
Proposed method (2016) [55]	98.88	1	60 pix	All images
Dietter *et al.* (2019) [61]	98.88	1	Boundary	All images
Ramakanth *et al.* (2014) [60]	98.88	1	60 pix	All images
Lu and Lim (2011) [59]	98.88	1	Boundary	All images
Mahfouz *et al.* (2010) [58]	97.75	2	60 pix	All images
Bajwa *et al.* (2019) [187]	**100.00**	0	Overlapping	All images*

Messidor dataset				
Methods	Acc(%)	Fail	Evaluation Criteria	Evaluation Method
Proposed method (2016) [55]	99.25	9	60 pix	All images
Giachetti *et al.* (2018) [163]	**99.83**	2	1R	All images
Ramakanth *et al.* (2014) [60]	99.42	7	60 pix	All images
Lu (2011) [167]	99.75	3	Boundary	All images
Bajwa *et al.* (2019) [187]	99.17	10	Overlapping	All images*
Roychowdhury *et al.* (2016) [184]	**100.00**	0	Boundary	All images*

ROC dataset				
Methods	Acc(%)	Fail	Evaluation Criteria	Evaluation Method
Proposed method (2016) [55]	**99.00**	1	60 pix	All images
Panda *et al.* (2017) [62]	**99.00**	1	60 pix	All images

E-ophtha-EX dataset				
Methods	Acc(%)	Fail	Evaluation Criteria	Evaluation Method
Proposed method (2016) [55]	98.78	1	60 pix	All images
Panda *et al.* (2017) *et al.* [62]	**100.00**	0	60 pix	All images

HRF dataset				
Methods	Acc(%)	Fail	Evaluation Criteria	Evaluation Method
Proposed method (2016) [55]	**100.00**	0	60 pix	All images
Dietter *et al.* (2019) [61]	**100.00**	0	Boundary	All images

Acc - Accuracy; Fail - number of images failed by the method; crx val - cross-validation * - Training is performed on other dataset. Best values appear in bold.

For the Diaretdbv0, an accuracy of 98.46% (128/130) is obtained. To the best of our knowledge,

this result is only exceeded in [59] and [184] with an accuracy of 99.23% and 100.00% respectively. For the Diaretdbv1, an accuracy of 98.88% (88/89) was achieved, being surpassed by [184] with an accuracy of 100.00%. In the Messidor dataset, an accuracy of 99.25% (1189/1200) was achieved, being slightly smaller than the reported on [184] with 100.00% accuracy. An accuracy of 99.00% (49/50), 98.78% (81/82), and 100% (45/45) were achieved respectively for ROC, E-ophtha-EX and HRF datasets. In these three datasets, only rule-based methods have been reported. Other relevant results were achieved in [40] with an accuracy of 99.90% (999/1000), and in [173] with 100.00% (40/40). However, these results were achieved in private datasets, and no comparison was possible. In total, our method failed in 15 fundus images, (see examples in the first row of Figure 5.11).

5.3 Discussion and conclusions

A new algorithm for the localization of the OD in fundus images was developed and presented in this paper. The algorithm reveals to be reliable and efficient. The robustness of the proposed technique is guaranteed by evaluating the method in eight publicly-available datasets. Experiments revealed an OD localization accuracy of 99.15% which is larger than other state-of-the-art methods. Although the used implementation is not optimized, the algorithm already provided an efficient computation, resulting in an average computation time of 18.34 seconds per image.

The defined model is based on two new techniques: 1) The cumulative sums of successive subdivisions described in section II-C, and 2) the vessel enhancement described in section II-A. The first has two major advantages. First, it is able to enhance the profiles g_v and g_h maxima and minima along a specific direction. It also fits well to different images with different resolutions because it considers several regions with different sizes.

The main cause of error resulted from an erroneous determination of g_v and g_h profiles minima, either in the horizontal or in the vertical coordinate. This occurs in images where the vessels belonging to the main arcades have low contrast or are not present, as can be seen in the first and second image in Figure 5.11. Another source of error occurs when the vertical vessels that cross the OD are either not present or exhibit a low contrast when compared with other vertical vessels in other regions of the retina, leading to an inaccurate determination of the horizontal coordinate (see images three, four and five in Figure 5.11). This suggests that a method capable of a better selection and discrimination of the correct profiles minima could reduce the small number of localization failures.

The developed method depends on several parameters. Apart from the δ parameter previously analysed, all other parameters were chosen based on the selection that results in the best performance, although the method reveals a low sensibility to the variation of these parameters. It is important to emphasize that the same algorithm with the same parameter values was used in all the images of all datasets, ensuring high reliability. This paper also presents a new vessel enhancement method based on a well-known corner detector [267]. The proposed vessel enhancement proves to be effective in the preservation of elongated structures, and in the discrimination between vessel and non-vessel structures. Furthermore, the vessel enhancement method does not provide segmentation of the retinal vasculature. However, it can be used as a preprocessing step for the segmentation of this type of retinal structure. Overall, and considering numerous datasets and images tested, the proposed method represents a significant improvement of the OD automated localization, surpassing other state-of-the-art methods. Another relevant aspect of our method is that it allows a considerable illumination variance, since no illumination equalization was per-

formed as in other methods like the described in [64] and [65]. Finally, and as a final comment, the reliability and computational efficiency of the proposed method allow the creation of an effective tool that can easily be incorporated into clinical practice.

5.4 Concluding remarks

In this chapter, an efficient and reliable method for OD localization in fundus images is proposed. Several important contributions are made. First, a new vessel enhancement is proposed. Second, a new technique designated by cumulative sum fields is introduced. Third, the proposed method was tested in many publicly-available datasets, achieving an improved performance to other state-of-the-art methods. Fourth, a detailed performance comparison is made with many several methods. The developed method was tested with 8 different datasets with results comparable to other state-of-the-art methods but evaluated on a smaller number of datasets [58, 59, 60, 61, 62]. No illumination equalization was performed as in other methods like the described in [64] and [65].

Chapter 6

Retinal vasculature detection

Several systemic and non-systemic diseases manifest through changes and modifications in the retinal blood vessels [190]. Furthermore, the eye is unique since it is the only region in the human body where the vascular network can be observed in-vivo [191]. The analysis of the retinal vasculature is fundamental in glaucoma, DR, hypertensive retinopathy, and myocardial infarct, from cardiovascular diseases, and other cerebral diseases like stroke [190]. The hallmark of glaucoma progression is the cupping of the OD [190]. Since this region is heavily crossed by the retinal vessels, their detection and removal can assist in glaucoma diagnose. One of the first manifestations of DR is the tiny capillary dilations known as microaneurysms [210]. In later stages, other vascular changes include neovascularisations and haemorrhages. Hypertension and atherosclerosis cause changes in the ratio between the diameter of retinal arteries and veins, known as the A/V ratio. A decrease in the A/V ratio, i.e., thinning of the arteries and widening of the veins, is associated with an increased risk of a stroke and myocardial infarction [190]. In addition, systemic vascular disease can cause arterial and venous occlusions, and also central and branch venous occlusions [190].

The manual segmentation of the retinal vessels is a very time-consuming task and prone to errors [193]. Hence, the automatic quantification of the retinal vessels as well the vascular features, such as the length, width, tortuosity, branching pattern, A/V ratio, and arteries/veins differentiation can provide important insights to a proper diagnose and management of the previously mentioned diseases. Moreover, the retinal vasculature can also be used to detect other retinal structures like the fovea and the OD, as well as to perform the registration of fundus images [210]. The accurate segmentation of the retinal vessels is a challenging task for several reasons, like the presence of lesions, exudates, and haemorrhages; other structures like the OD and fovea; a wide range of scales, covering both thin and wider vessels; variable contrast between the vessels and retinal background, and the central reflex in major arteries [193]. In this chapter, it is presented a vessel centerline detection method and a vessel segmentation method.

The vessel centerline detection method is described in the paper *"Vessel Centerline Detection in Retinal Images Based on a Corner Detector and Dynamic Thresholding"* [56]. The algorithm begins with a new vessel detector description method based on a modified corner detector. Next, the vessel detector image is filtered with a set of binary rotating filters, resulting in enhanced vessels structures. The main vessels can be selected with a dynamic thresholding approach. To deal with vessels bifurcations and crossovers, the initial fundus image is processed with a set of four directional differential operators. The resulting directional images are then combined with the detected vessels, creating the final vessels centerlines image.

The vessel segmentation method is described in the paper *"A New Vessel Enhancement Transform on Retinal Blood Vessels Segmentation"* [57]. The proposed approach performs a vessel enhancement based on the intensity differences between a pixel and its neighbourhood. False vessels due to bright regions, like exudates, are removed based on new criteria. Furthermore, a new approach to discard false vessel regions that usually appear in the border of the OD due to peripapillary atrophy is also presented. This is achieved using the derivatives of the local maxima over scales of the magnitude of the gradient. Section 6.1 describes the retinal vessel centerline detection method,

and section 6.2 describes the retinal vessel segmentation method. Finally, section 6.3 summarizes the concluding remarks.

6.1 Retinal vessel centerline detection

The retinal vessel centerline detection consists of the seven main processing stages, described in the next subsections:

- Image preprocessing
- Scale-space definition
- Vessel detector
- Vessel enhancement
- Vessels thresholding
- Directional vessel centerline detector
- Final vessel centerline detection

6.1.1 Image preprocessing.

The green channel component of fundus images, designated by $I_G(\mathbf{x})$ is used as input to the proposed algorithm [97]. No contrast normalization or enhancement is performed on the $I_G(\mathbf{x})$ image. If $I_G(\mathbf{x})$ has either a width or height smaller than 1000 pixels, it is proportionally resized using a bi-cubic interpolation, leading to a dimension of 1000 pixels in the smaller direction.

6.1.2 Scale-space definition

As referred in [272], the linearity of the structure depends on the observation scale. This is particularly relevant in the smaller and thinner retinal vessels that are more difficult to detect and extract, thus indicating the need for a scale-space approach. The scale-space defined in appendix B is applied with scales defined by $\sigma = \{2.8,\ 2.9,\ 3\}$. The results, at each scale σ, are then combined to form the final vessel centerline image.

6.1.3 Vessel detector

By considering the image $L(\mathbf{x};\sigma)$ as an intensity surface, Wang and Brady [267] proposed a corner detection algorithm based on the surface curvature measurement. Furthermore, if the corner points are also edge points, the corner detector is defined as,

$$K(\mathbf{x};\sigma) = (\nabla^2 L(\mathbf{x};\sigma))^2 - c|\nabla L(\mathbf{x};\sigma)|^2. \quad (6.1)$$

$K(\mathbf{x};\sigma)$ enhances regions where a rapid change in the edge direction occurs [267]. The parameter c defines how edge-phobic is $K(\mathbf{x};\sigma)$, and is set to $c = 1$. To enable a better response of the vessel structures relatively to its surroundings, the square of the first term in equation (6.1) is removed, resulting in,

$$\tilde{K}(\mathbf{x};\sigma) = \nabla^2 L(\mathbf{x};\sigma) - c|\nabla L(\mathbf{x};\sigma)|^2. \quad (6.2)$$

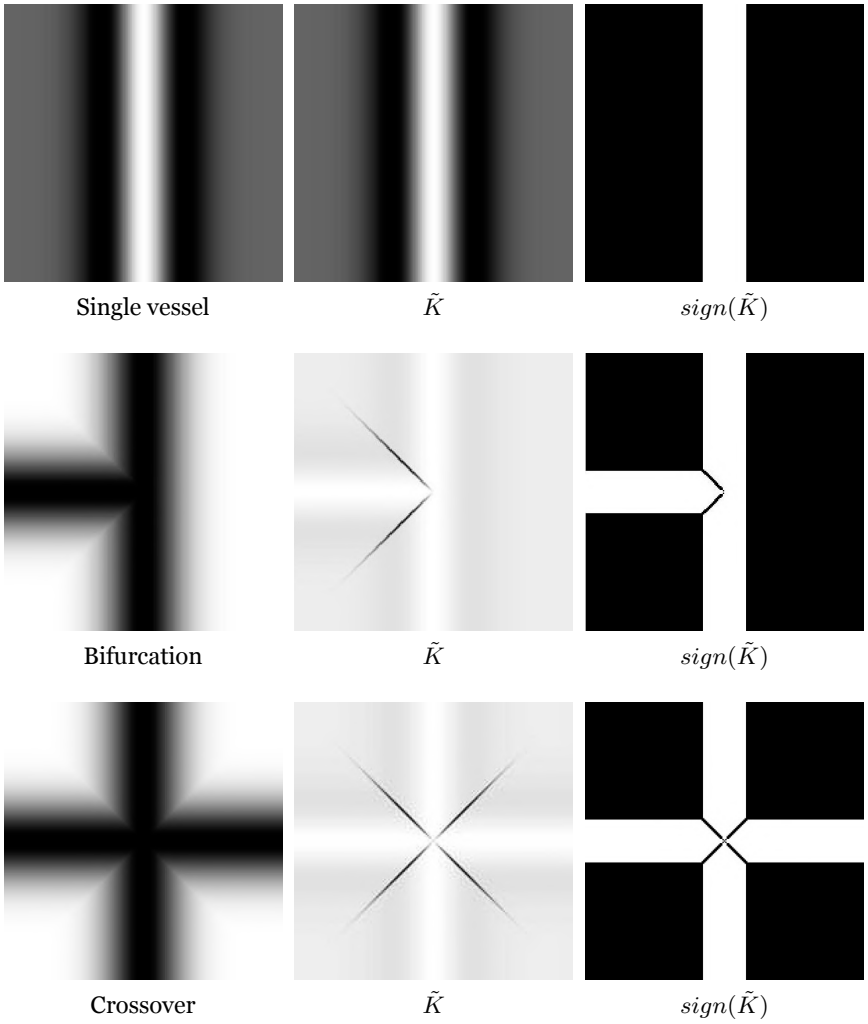


Figure 6.1: $\tilde{K}(\mathbf{x}, \sigma)$ for a synthetic single vessel, vessels bifurcation and vessels crossover, created with a standard deviation of 2. $sign(.)$ stands for the signum function.

Equation 6.2 reveals to be a good vessel detector because it shows good preservation of elongated structures and good discrimination between a vessel and non-vessel structures. Figure 6.1 shows an example of the application of $\tilde{K}(\mathbf{x}; \sigma)$. A strong response in the center of the vessels is observed in a single vessel, a vessels bifurcation and a vessels crossover. Figure 6.3 a) shows an example of $\tilde{K}(\mathbf{x}; \sigma)$ for a fundus image.

6.1.4 Vessel enhancement

The vessels may be considered as piecewise linear segments. Since the cross-section of the vessels has a Gaussian profile, a 2D Gaussian-shaped filter can be used to enhance the retinal vasculature. According to Chaudhuri *et al.* [196] and considering a given location $\mathbf{x} = (x, y)$, the 2D Gaussian matched filter kernel can be defined as,

$$m(x, y) = -e^{\frac{x^2}{2s^2}}, \quad \forall\, |y| \leq W/2. \quad (6.3)$$

W is the length of the segment for which the vessel is assumed to have a fixed orientation, and s defines the spread of the vessel intensity profile [196]. For the vessels detection, the kernel is rotated to match all possible vessel orientations and the maximum response from the filter bank

Figure 6.2: Some binary Φ filters for the $0^o, 45^o, 90^o$ and 135^o directions with $l = 5$.

is registered.
To further enhance the difference between vessels and non-vessels structures, a few modifications are made to the initial matching filter algorithm. Instead of applying the matching filtering directly on the $I_G(\mathbf{x};\sigma)$ image, it is applied to the defined vessel detector $\tilde{K}(\mathbf{x};\sigma)$. Moreover, the used kernels were also modified. Instead of the Gaussian matching filter kernel, a set of 2D binary filters $\Phi(\mathbf{x})$ composed by a white line are used. Since the vessels can have any orientation, $\Phi(\mathbf{x})$ must be rotated at several angles θ. Furthermore, to achieve further enhancement, several line lengths l are also considered. Hence, $\Phi(\mathbf{x}) \equiv \Phi(\mathbf{x}, l, \theta, d)$, where d is the corresponding filter size for each l (see Figure 6.2 as an example). After extensive testing, the following set of values were defined, $l = \{3,\ 5,\ 7,\ 9,\ 11\}$ with the filter size $d = \{5\times5,\ 7\times7,\ 9\times9,\ 13\times13,\ 15\times15\}$ respectively. θ varied from 0^o to 165^o in steps of 15^o. The maxima response between $\tilde{K}(\mathbf{x};\sigma)$ and the filters bank $\Phi(\mathbf{x}, l, \theta, d)$ is defined by,

$$C(\mathbf{x};\sigma) = \max_{\forall\theta,l}\{\tilde{K}(\mathbf{x};\sigma) * \Phi(\mathbf{x}, l, \theta, d)\}. \quad (6.4)$$

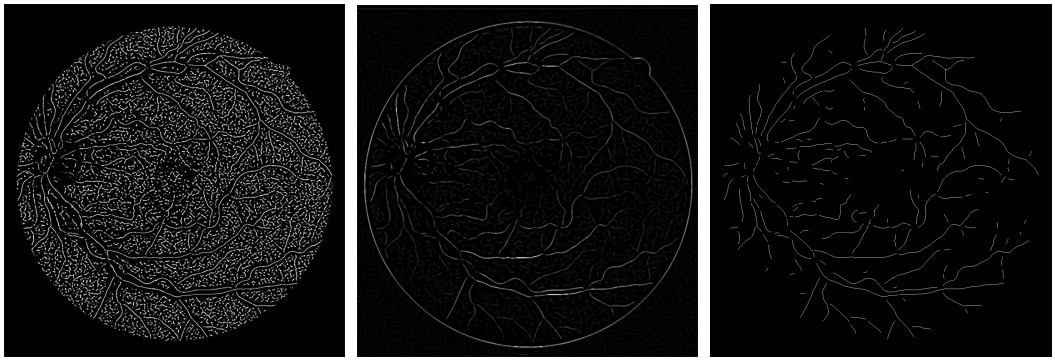
An example of the convolution of equation 6.4 is represented in Figure 6.3 b). Furthermore, as can be seen in Figure 6.3, $C(\mathbf{x};\sigma)$ offers a better distinction between vessels and non-vessels than $\tilde{K}(\mathbf{x};\sigma)$. It is also observed that even the smaller and thinner vessels appear well contrasted in relation to their surroundings. These observations indicate that a simple threshold could select the majority of the vessels structures.

6.1.5 Vessels thresholding

In order to perform the selection of the vessels in, $C(\mathbf{x};\sigma)$ a dynamically computed threshold α must be defined. The threshold α is defined based on the fact that only the highest intensity pixels in $C(\mathbf{x};\sigma)$ correspond to the vessel centerlines. Hence, all $C(\mathbf{x};\sigma)$ values larger than zero are selected, and sorted in ascending order in a vector $\mathbf{v}(x)$. The value of $\mathbf{v}(x)$ that is smaller than the $\beta\%$ maxima values of $\mathbf{v}(x)$ is selected as the threshold α. After extensive testing, β was set to 7%. All the pixels in $C(\mathbf{x};\sigma)$ smaller than α are set to 0. In order to refine the results, a morphological *thinning* operation followed by the removal of all the regions smaller than 12 pixels is applied [273]. The image that results from this thresholding operation followed by the morphological refinement is designated by $C_\alpha(\mathbf{x};\sigma)$. An example is represented in Figure 6.3 c).

6.1.6 Directional vessel centerline detector

From the observation of Figure 6.3 c) some connections between vessels are missing, particularly in bifurcations and crossovers regions. To overcome this problem, it is determined the directional



(a) (b) (c)

Figure 6.3: a) Output of the vessel detector $\tilde{K}(\mathbf{x};\sigma)$. b) Enhanced version of $\tilde{K}(\mathbf{x};\sigma)$, $C(\mathbf{x};\sigma)$. c) Thresholded version of $C(\mathbf{x};\sigma)$, $C_\alpha(\mathbf{x};\sigma)$

vessel centerline, i.e., the vessel centerline for a specific direction. Subsequently, this information can be used to establish the missing crossovers and bifurcations in the $C_\alpha(\mathbf{x};\sigma)$ image. Since the vessels can occur in any direction, it is necessary to select a reduced number of orientations that shall be analysed. As indicated by Mendonça and Campilho [191] the set $\phi = \{0^o,\, 45^o,\, 90^o,\, 135^o\}$ are a good trade-off between accuracy and computation time. Hence, the initial image $L(\mathbf{x};\sigma)$ is processed with a set of directional filters capable of detecting retinal vessels centerlines in the ϕ orientations. The kernels used in this stage are the difference of offset Gaussians filters (DoOG filters), which have revealed good immunity to noise due to their larger size [191, 274]. The kernel used for detecting the vertical vessel centerline candidates are the row gradient filter, given by,

$$\begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -2 & -4 & 0 & 4 & 2 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix}$$

The kernel used for detecting the 45^o vessel centerline candidates are the 135^o gradient filter, given by,

$$\begin{bmatrix} 0 & 0 & 2 & 4 & 1 \\ -2 & -2 & 0 & 2 & 2 \\ -1 & -4 & -2 & 0 & 0 \end{bmatrix}.$$

The other two kernels are just rotated versions of these filters. The Convolution of $L(\mathbf{x};\sigma)$ with the DoOG kernels measures the gradient component in a specific direction, generating the directional images $D_\phi(\mathbf{x};\sigma)$ with $\phi = \{0^o,\, 45^o,\, 90^o, 135^o\}$ (see Figure 5.6 a) and d) as an examples). Each one of these four directional images $D_\phi(\mathbf{x};\sigma)$ are searched on a pixel-by-pixel basis for specific combinations of signs of the surrounding pixels. Hence, and considering the $D_{90^o}(\mathbf{x};\sigma)$ image as an example, the pixel position $\mathbf{x} = (x, y)$ is considered as a vessel candidate if the neighbourhood pixels in the normal direction have the signal combinations that result in $M_{90^o}(\mathbf{x};\sigma)$ as defined in

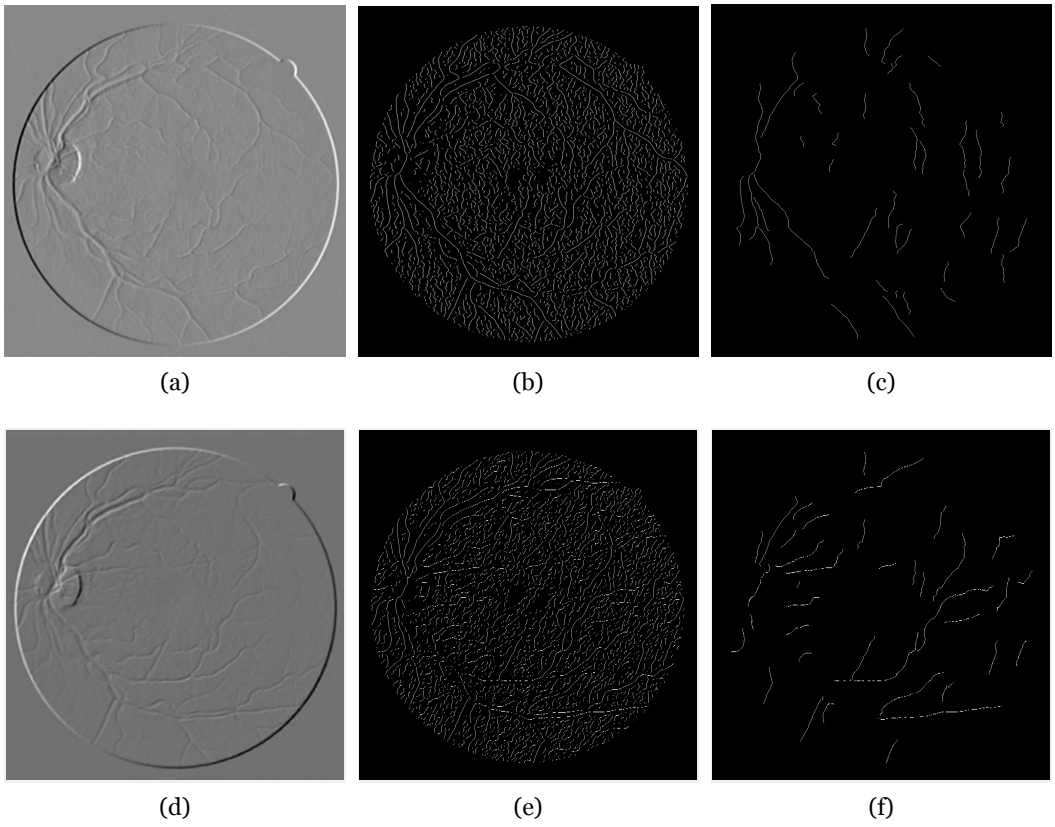


Figure 6.4: a) and d) Result of the DoOG filter in the vertical direction $D_{90^o}(\mathbf{x};\sigma)$, and in the 45^o direction, $D_{45^o}(\mathbf{x};\sigma)$. b) and e) Vessel centerlines at the vertical and 45^o orientations, c) and f) Vertical and 45^o vessel centerlines after the Algorithm 4.

the following,

$$M_{90^o}(\mathbf{x};\sigma) = \begin{cases} 1, & if \begin{cases} D_{90^o}(x,y-2,\sigma) > 0 & \wedge \\ D_{90^o}(x,y-1,\sigma) > 0 & \wedge \\ D_{90^o}(x,y+2,\sigma) < 0 & \wedge \\ D_{90^o}(x,y+1,\sigma) < 0 \end{cases} \\ 0, & otherwise \end{cases} . \quad (6.5)$$

Following the same reasoning, the $M_{90^o}(\mathbf{x};\sigma)$ is defined as,

$$M_{45^o}(\mathbf{x};\sigma) = \begin{cases} 1, & if \begin{cases} D_{45^o}(x-2,y-2,\sigma) > 0 & \wedge \\ D_{45^o}(x-1,y-1,\sigma) > 0 & \wedge \\ D_{45^o}(x+2,y+2,\sigma) < 0 & \wedge \\ D_{45^o}(x+1,y+1,\sigma) < 0 \end{cases} \\ 0, & otherwise \end{cases} . \quad (6.6)$$

$M_{90^o}(\mathbf{x};\sigma)$ and $M_{45^o}(\mathbf{x};\sigma)$ are the vessels centerlines candidates in the vertical and the 45^o orientation, as represented in Figure 6.4 b) and e) respectively. This procedure can be easily extend to the remaining orientations, creating $M_{0^o}(\mathbf{x};\sigma)$ and $M_{135^o}(\mathbf{x};\sigma)$. In order to refine the results, a morphological *thinning* operation is applied to all $M_\phi(\mathbf{x};\sigma)$ [273].

6.1.7 Final vessel centerline calculation

As previously mentioned, some bifurcations and crossovers in $C_\alpha(\mathbf{x};\sigma)$ are missing. To overcome this problem, it was determined the directional vessel centerlines $M_\phi(\mathbf{x};\sigma)$ with $\phi = \{0^o, 45^o, 90^o, 135^o\}$. Since we have four different orientations, each $M_\phi(\mathbf{x};\sigma)$ must be individually combined with $C_\alpha(\mathbf{x};\sigma)$. Furthermore, it is also observed from Figure 6.3 c) and Figure 6.4 b) that the images $C_\alpha(\mathbf{x};\sigma)$ and $M_\phi(\mathbf{x};\sigma)$ are formed by individual regions, where each region represents a possible vessel centerline candidate. Based on these considerations, the Algorithm 4 was implemented (see Figure 6.4 c) and f) for the 90^o and 45^o cases).

Algorithm 4: Bifurcations and crossovers closing algorithm.

Input: $C_\alpha(\mathbf{x};\sigma)$, $M_\phi(\mathbf{x};\sigma)$
Output: Processed $M_\phi(\mathbf{x};\sigma)$

for $\sigma \in \{2.8, 2.9, 3.0\}$ **do**
  for $\phi \in \{0^o, 45^o, 90^o, 135^o\}$ **do**
    Considering that $M_\phi(\mathbf{x};\sigma)$ is formed by N disjoint regions $M_j(\mathbf{x})$,
    $M_\phi(\mathbf{x};\sigma) = \bigcup_{j=1}^{N} M_j(\mathbf{x})$
    for $j = 1$ ***to*** N **do**
      if $M_j(\mathbf{x})$ *intersects less than 2 regions with* $C_\alpha(\mathbf{x};\sigma)$ **then**
        $M_j(\mathbf{x})$ is removed from $M_\phi(\mathbf{x};\sigma)$
      end
    end for
  end for
end for

The final vessel centerline image for the image $I_G(\mathbf{x};\sigma)$ given by the union of all $M_\phi(\mathbf{x};\sigma)$ with $C_\alpha(x,\sigma)$ according to,

$$F(\mathbf{x}) = \Big(\bigcup_{\sigma} C_\alpha(\mathbf{x};\sigma)\Big) \bigcup \Big(\bigcup_{\sigma,\phi} M_\phi(\mathbf{x};\sigma)\Big) \quad (6.7)$$

In Figure 6.5 some qualitative results are shown for the proposed algorithm.

6.1.8 Results

The automatic method for the vessel centerline detection was tested on the images of Drive and Stare datasets, which are described in appendix A. The method performance is described next.

6.1.8.1 Retinal vessel centerline detection evaluation

To evaluate the retinal vessel centerline detection algorithm, two publicly available datasets were used, namely, the Drive dataset [43], composed of 40 images, and the Stare dataset [42], composed of 20 images. In both datasets, two human vessel segmentation ground truths are available. The first human observer groundtruth is used on both datasets as reference.
The evaluation method relies on the fact that an ideal vessel extraction should match a thinned version of the manual segmentation. The two datasets do not give manual annotations on vessel centerlines. Thus, a skeleton extraction method [198] on the annotations of retinal vessels was applied to produce the ground truth maps for the centerline extraction.
The method reliability is based on the selection of the threshold value α, which in turn is defined by the percentage value β. Hence, β was increased in steps of variable width to analyse its influence

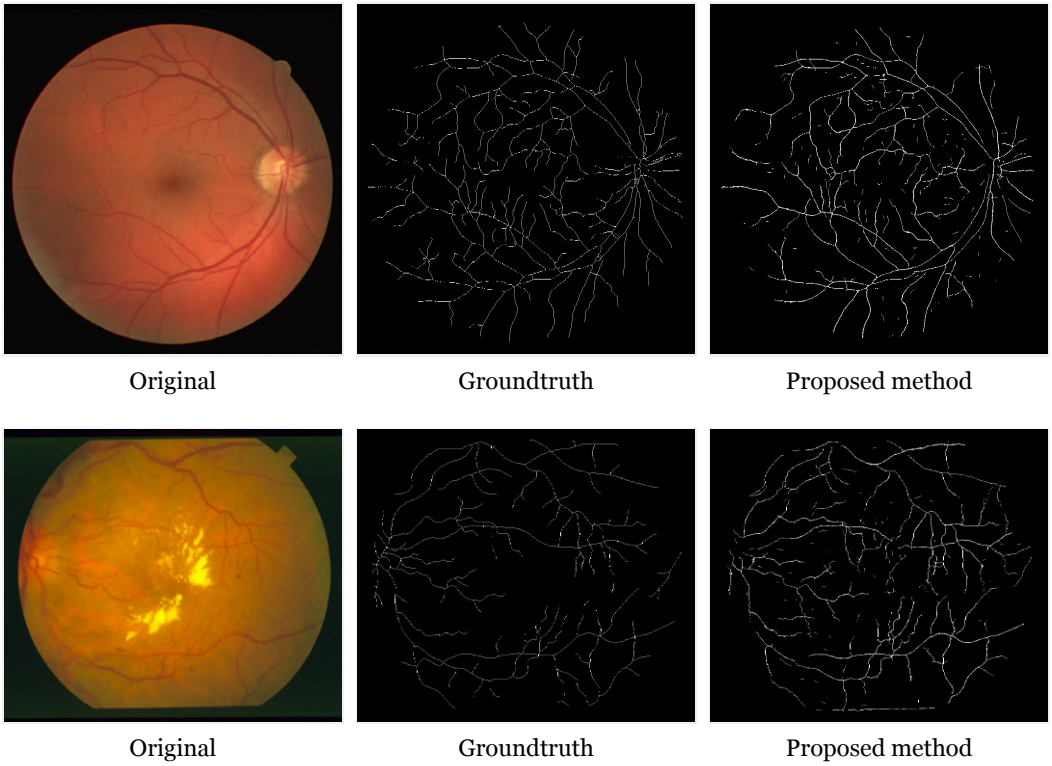


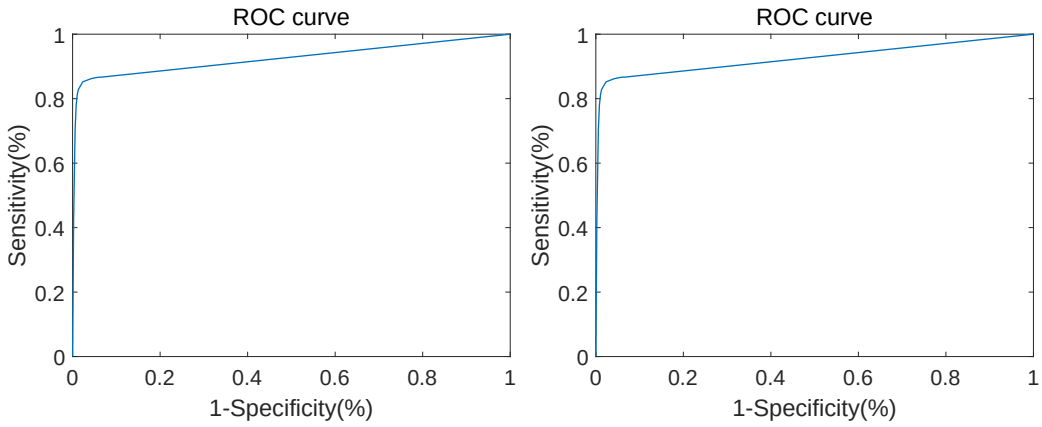
Figure 6.5: Examples of retinal vessels centerline detection result from a sample of the Drive dataset (top row) and of the Stare dataset (bottom row).

on detection performance. The following values $\{35, 25, 15, 9, 7, 5, 3, 1\}$ were tested for β.

For any fundus image, we can determine the number of true positive, false positive, and false negative vessel centerlines extracted for any β value. These values are defined by using a comparison against the ground truth centerline maps, as in [205]. A true positive is defined as any detected point within two pixels of the thinned groundtruth segmentation. False negatives are counted as the number of points in the thinned groundtruth segmentation not within two pixels of a detected centerline point. The tolerance of two pixels was suggested in [205] to account for localization errors in both the *thinned* manual segmentation and in the centerline peak location.

For each, β the Sensitivity and Specificity are determined for each image and then averaged for all the images in each dataset. Figure 6.6 shows the ROC curves for the Drive and the Stare dataset. For both Drive and Stare datasets, the best results were achieved for $\beta = 7\%$. The results of the method for the two datasets are presented in Table 6.1 in terms of Sensitivity, Specificity, Accuracy, and AUC. Moreover, results of other relevant methods are also reported for comparison. The higher accuracy, obtained for both datasets, reveals high stability and independence of the proposed rule-based method from image properties (with no training phase). The developed model reveals reliability for the retinal vessel centerline detection. This is particularly relevant for the Stare dataset, considering it has images with pathological clinical manifestations, making the accurate detection of the vessel centerline challenging.

Figure 6.7 and 6.8 show some vessel centerline examples of the proposed method.



(a) Drive dataset (b) Stare dataset

Figure 6.6: Vessel centerline ROC curves for the Drive and Stare datasets.

Table 6.1: Performance of the proposed and relevant retinal vessel centerline detection methods.

Drive Dataset					
Method	Acc(%)	Sen(%)	Spe(%)	AUC(%)	Evaluation Method
Proposed method (2014) (2)	**98.07**	**79.70**	**99.96**	90.12	All images
Kromm *et al.* (2020) [17]	90.70	75.20	84.20	88.20	All images
Xu *et al.* (2020) [227]	95.37	74.22	96.37	**96.12**	All images

Stare Dataset					
Method	Acc(%)	Sen(%)	Spe(%)	AUC	Evaluation Method
Proposed method (2014)	**98.36**	**82.85**	**99.90**	92.53	All images
Xu *et al.* (2020) [227]	96.89	77.30	97.45	**97.81**	5-by-5 Repeated crx val

Sen - Sensitivity; Spe - Specificity; Acc - Accuracy; crx val - Cross validation; Best values appear in bold.

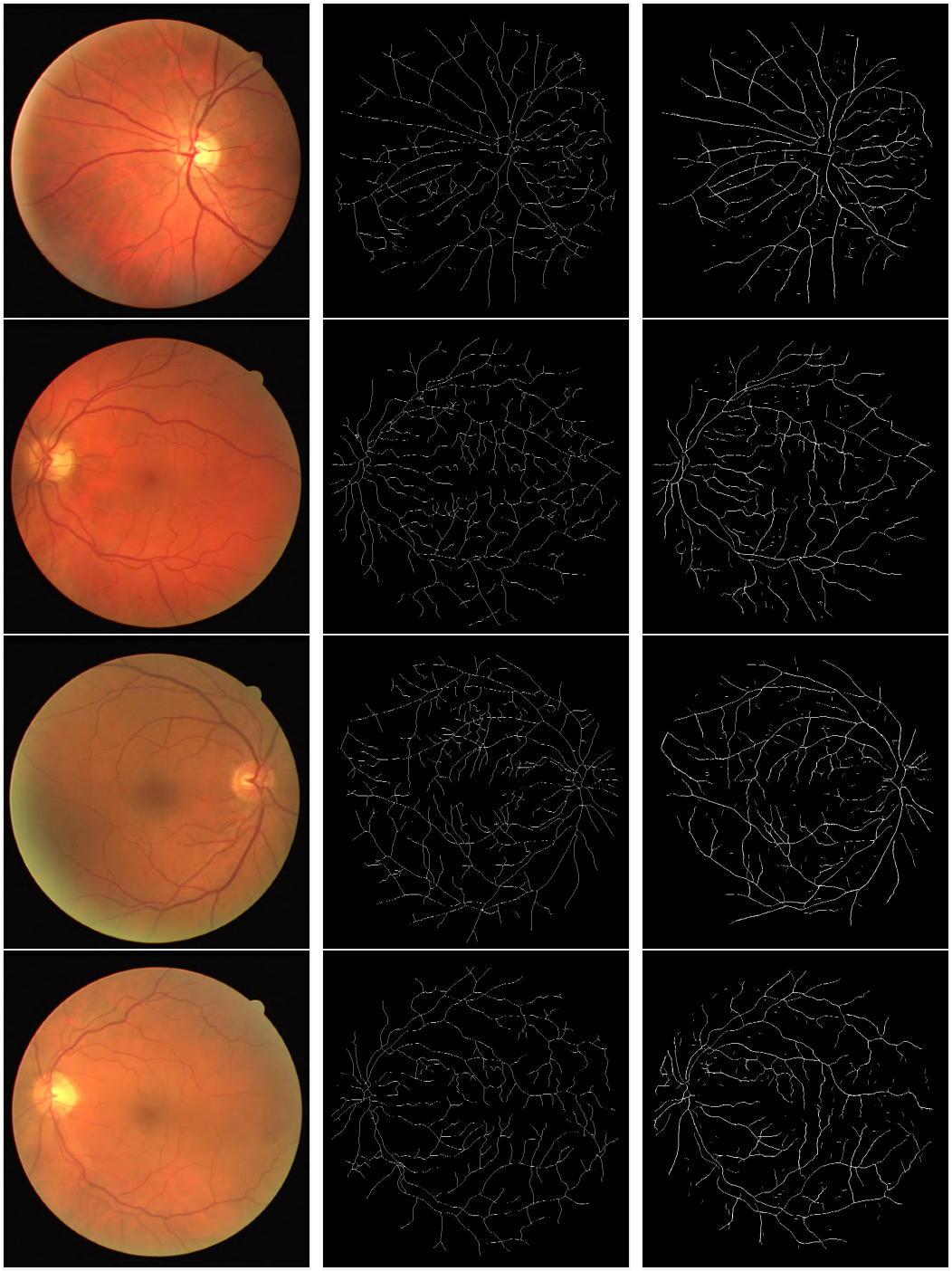
6.2 Retinal vessels segmentation

The retinal vessel segmentation method is divided into three stages;

- Image preprocessing
- Retinal vessels enhancement
- Retinal vessels segmentation

6.2.1 Image preprocessing

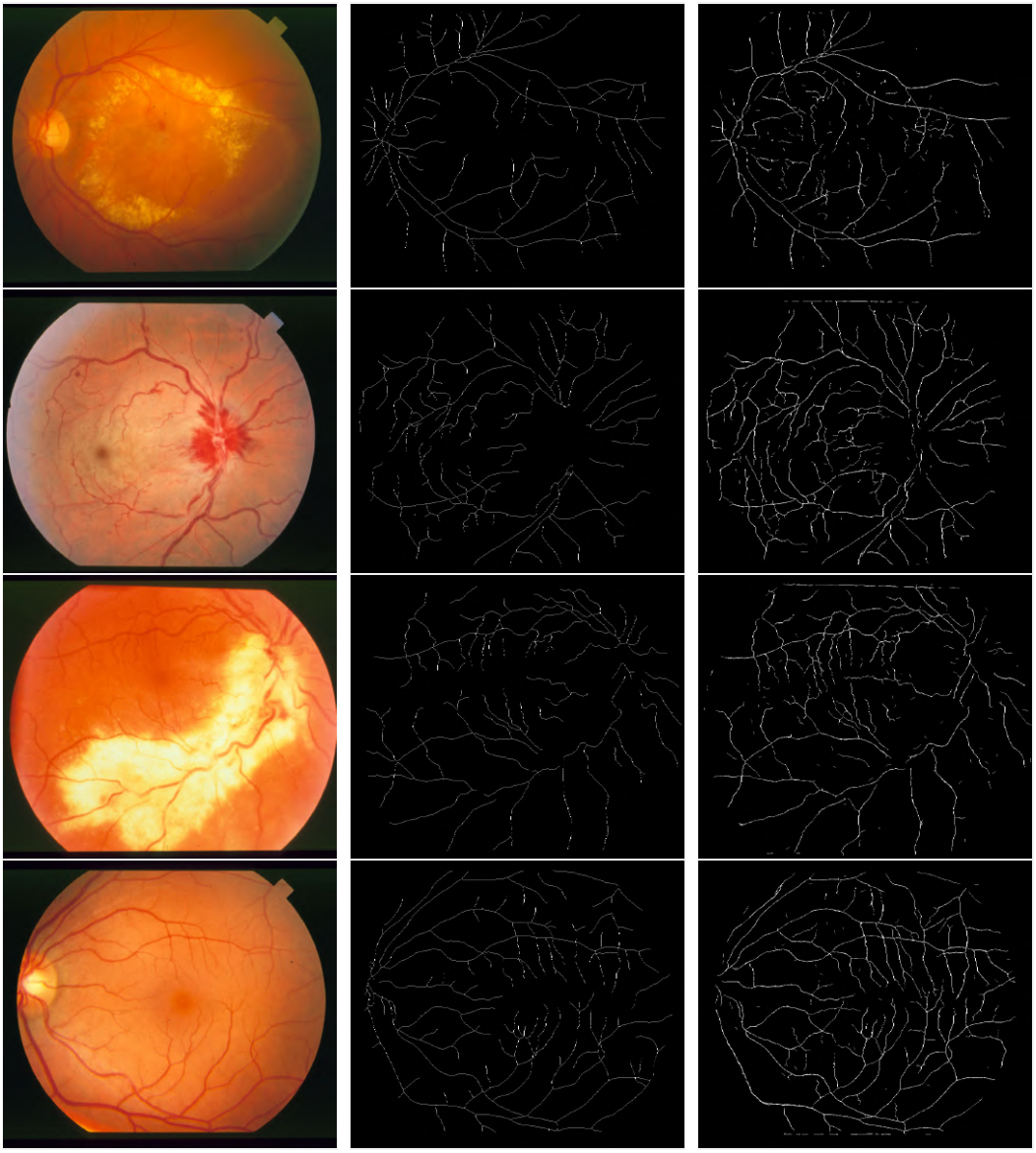
The image preprocessing stage normalizes the images in terms of contrast and non-uniform illumination. The scale-space and differential operators used in this work are also defined.
The green channel component $I_G(\mathbf{x})$ (where $\mathbf{x}$ represents the location (x, y)) of each RGB fundus image is selected [4] (see Figure 6.9 a) as an example). If not provided, the Black-and-White masks $BW(\mathbf{x})$, that define the region of interest in $I_G(\mathbf{x})$, are created by selecting all the pixels in the red channel of the RGB image larger than 38 [65]. Next, if the image dimensions are large, $I_G(\mathbf{x})$ is proportionally resized using bicubic interpolation, such that its smaller dimension is 700 pixels. The size normalization allows the definition of a set of algorithm parameters that are effective in all datasets. To remove background illumination variations, a shade-correction algorithm is performed [193, 210].



Original Groundtruth Proposed method

Figure 6.7: Examples of retinal vessels centerline detection result from the Drive dataset.

First, an inpainting algorithm performs a smooth interpolation of the peripheral pixels of $I_G(\mathbf{x})$. This is done with a discretization of the Laplace's equation $\nabla^2 I(\mathbf{x}) = 0$ [275], creating the image $I_P(\mathbf{x})$ (see an example in Figure 6.9 b)). Next, a median filtering with a mask size 89×89 is applied to $I_P(\mathbf{x})$. This large filter is required to create an image representation of the lightning variation in, $I_P(\mathbf{x})$ and it is designated by $I_M(\mathbf{x})$. The shade-correction algorithm ends by subtracting $I_M(\mathbf{x})$ from $I_P(\mathbf{x})$, creating the image $I_S(\mathbf{x}) = I_P(\mathbf{x}) - I_M(\mathbf{x})$. 4 The next step is the detection and removal of bright regions in $I_S(\mathbf{x})$. These bright regions may represent exudative lesions, central arterial

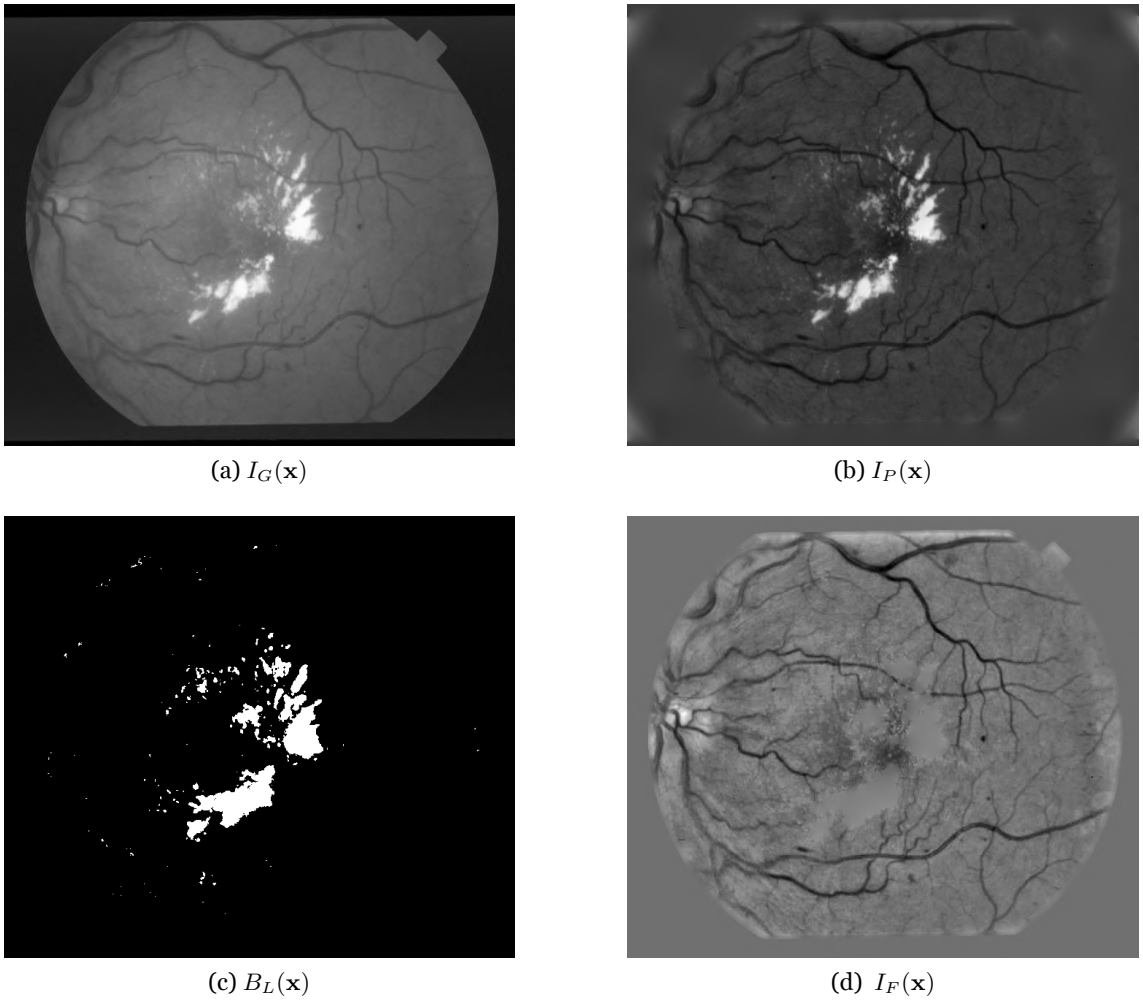


Original Groundtruth Proposed method

Figure 6.8: Examples of retinal vessels centerline detection results from the Stare dataset.

reflex, or belong to the OD. In any case, they can complicate vessel segmentation. Initially, the bright region's image, $B_L(\mathbf{x})$ is created by thresholding $I_S(\mathbf{x})$ and selecting all pixels larger than 12. Next, the OD is detected using the method proposed in [55]. A binary image $OD(\mathbf{x})$ where the pixel of the OD center has a value of 1 and all the remaining pixels have a value of 0 results. This point is then dilated by a 'disk' structuring element of size 60 [169]. This value is suitable to cover both smaller and larger discs. All the regions in $B_L(\mathbf{x})$ that intersect slightly or full the OD region in $OD(\mathbf{x})$ are removed.

Finally, the peripheral regions in $B_L(\mathbf{x})$ are also removed. This is achieved by considering an inverted eroded $BW(\mathbf{x})$ image with a structuring disk element of size 20. All the regions in $B_L(\mathbf{x})$ that partially or fully intersect this inverted image are removed (see an example in Figure 6.9 c)). The image preprocessing stage is finalized with the inpainting of $I_S(\mathbf{x})$ in the $B_L(\mathbf{x})$ regions, using the previously described inpainting method. This creates the final processed image, $I_F(\mathbf{x})$ exemplified in Figure 6.9 d).



(a) $I_G(\mathbf{x})$ (b) $I_P(\mathbf{x})$

(c) $B_L(\mathbf{x})$ (d) $I_F(\mathbf{x})$

Figure 6.9: Image preprocessing steps.

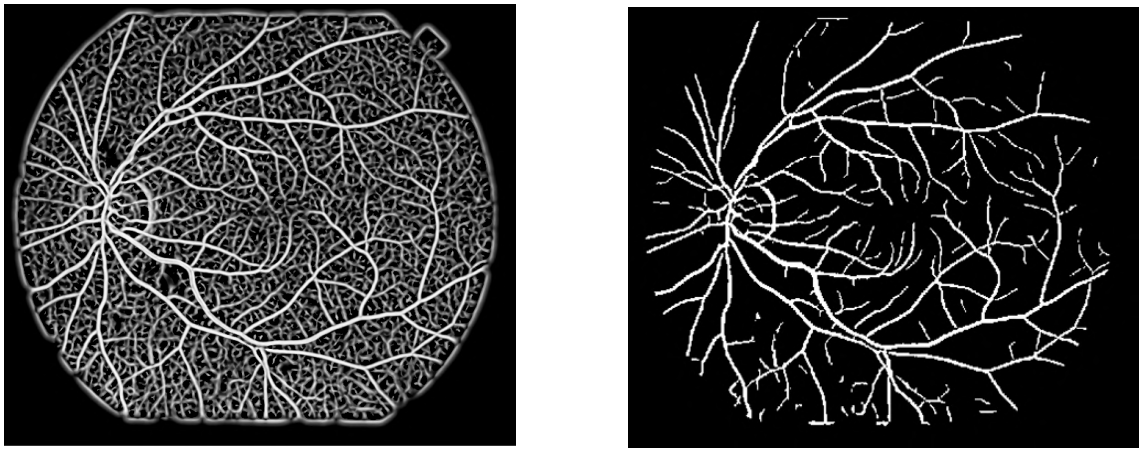
6.2.1.1 Differential operators definition in scale-space

Retinal blood vessels have a wide range of scales, which makes a scale-space analysis ideal for their detection. The scale-space defined in appendix B was used in the proposed method. In this section N_s was set to 20. The first order derivatives are represented by $L_x(\mathbf{x}; t_n)$ and $L_y(\mathbf{x}; t_n)$. To obtain a normalized response at each scale t_n, $L_x(\mathbf{x}; t_n)$ and $L_y(\mathbf{x}; t_n)$ are multiplied by the normalization factor $(t_n\sigma)^\alpha$, where α is the order of the derivative [39].

6.2.2 Retinal vessel enhancement

In this section, a new retinal vessel enhancement transform is defined. The maximum principal curvature $\lambda_2(\mathbf{x}; t_n)$ is commonly used for vessels structures enhancement [276]. To improve the enhancement, a cumulative maximum over scales of the principal curvature $\lambda_2(\mathbf{x}; t_n)$ is proposed. The scales t_u where $u \in \{1, ..., N\}$ are considered. For all scales t_u, such that $u \leq n$, the proposed cumulative maximum over scales of the principal curvature $\lambda_2(\mathbf{x}; t_n)$ is given by,

$$\Lambda_u(\mathbf{x}) = \max_{u \leq n} \left(\frac{\lambda_2(\mathbf{x}; t_u)}{(t_u\sigma)^2} \right), \quad (6.8)$$



(a) $E(\mathbf{x})$ (b) $V(\mathbf{x})$

Figure 6.10: Retinal vessel enhancement.

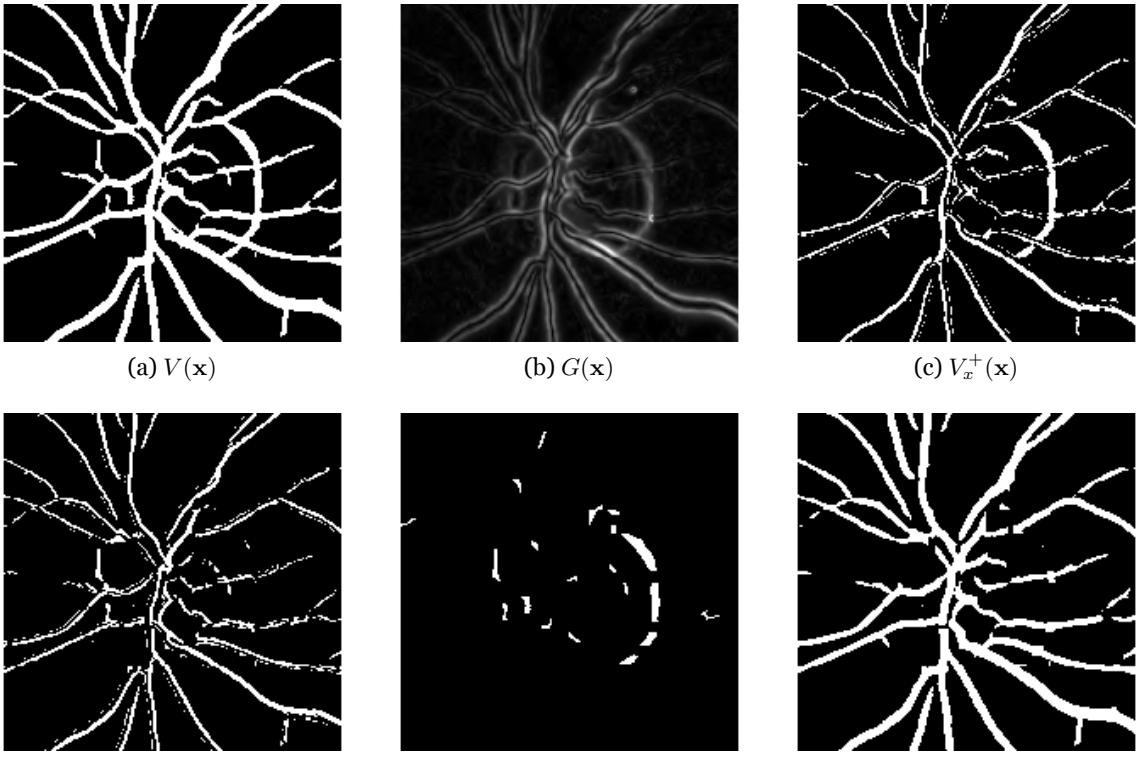
Retinal vessels have a large variation of intensity values in a given direction, while in the perpendicular direction the intensity values variation is small. If a difference is made between $\Lambda_u(\mathbf{x})$ and $\Lambda_u(\mathbf{x'})$, where **x'** is a pixel location in a neighbourhood of $\mathbf{x}$, a response is obtained. To further improve the vessels enhancement, it is necessary to consider all neighbourhood pixels of $\mathbf{x}$. The neighbourhood itself has to be widened due to the various widths that vessels can have. Hence, the best approach to this problem is to use polar coordinates that directly enable a radial analysis. Moreover, a scale-space approach is also taken because it allows reducing the effects of noise. Considering the polar formalism in scale-space $\mathbf{x}_{(d,\theta)}$ defines a neighbourhood pixel for every pixel location $\mathbf{x}$, for a range ρ and angle θ. The final retinal vessel enhancement operator is given by,

$$E(\mathbf{x}) = \sum_{t_u} \left(\sum_{\theta} \sum_{\rho} \left(\Lambda_u(\mathbf{x}_{(\rho,\theta)}) - \Lambda_u(\mathbf{x}) \right) \right). \quad (6.9)$$

The parameters ρ and θ vary according to the set of values $\rho = \{1, 2, ..., 5\}$ and $\theta = \{22.5^o, 45^o, ..., 360^o\}$. Figure 6.10 a) represents an example of $E(\mathbf{x})$.

6.2.3 Retinal vessels segmentation

In this section, the segmentation of retinal vessels is described. First, $E(\mathbf{x})$ is thresholded with two different values, T_1 and T_2, defining a hysteresis selection where only a percentage of its maxima are selected. The threshold value T_1 creates the binary image $V_{T_1}(\mathbf{x})$ and its regions are considered vessel seeds. The value T_2 creates the binary image $V_{T_2}(\mathbf{x})$ and its regions represent a more complete definition of the vessel's branches in terms of shape, length, and width.
Regions in $V_{T_1}(\mathbf{x})$ smaller than 30 pixels are removed to avoid overseeding and a final over-segmentation. Furthermore, vessel seeds that are too close to bright regions in $B_L(\mathbf{x})$ most likely represent artefacts and are removed. For any vessel seed region, the mean distance of its perimeter pixels to the bright regions in $B_L(\mathbf{x})$ is computed. If it is closer than eight pixels, the vessel seed is removed from $V_{T_1}(\mathbf{x})$. The next step is a normalization between 0 and 1 of $\Lambda_{u=n}(\mathbf{x})$, followed by a histogram equalization with 64 uniform intervals. A binarization is then performed using a threshold of 0.85 (selects 54 intervals), resulting in $\Lambda_B(\mathbf{x})$ [78]. An initial vessel segmentation is obtained by selecting all the regions in $V_{T_1}(\mathbf{x})$ that intersect the regions in $V_{T_2}(\mathbf{x})$ and also the regions in $\Lambda_B(\mathbf{x})$. An example of $V(\mathbf{x})$ is represented in Figure 6.10) b).



(a) $V(\mathbf{x})$ (b) $G(\mathbf{x})$ (c) $V_x^+(\mathbf{x})$

(d) $V_x^-(\mathbf{x})$ (e) $R(\mathbf{x})$ (f) $V_F(\mathbf{x})$

Figure 6.11: Peripapillary atrophy regions removal.

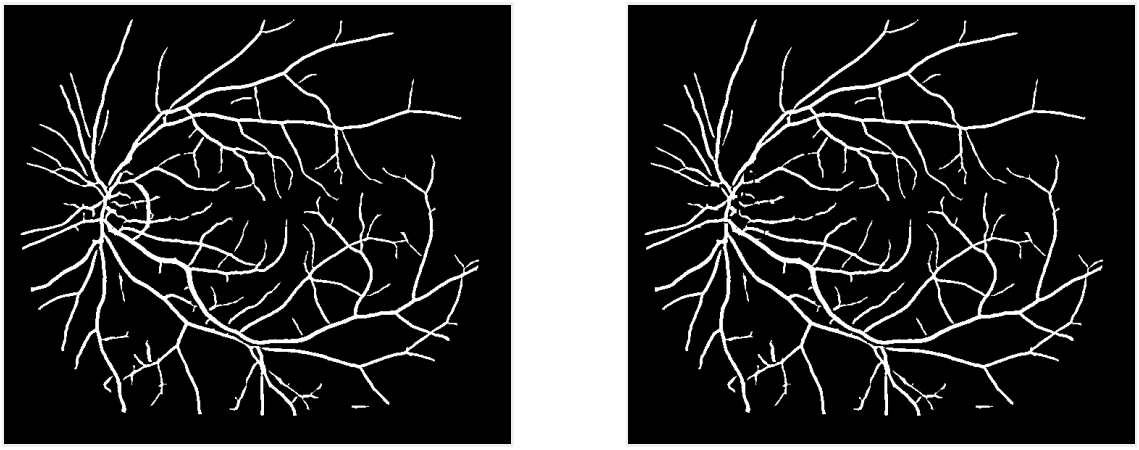
Usually, along the OD borders appear darker regions of peripapillary atrophy. Due to their low pixel values, they are misidentified as a vessel (see example in Figure 6.11 a)). Even if the OD was previously removed before the retinal vessel enhancement, these regions will still be present because of the required feature preservation. These false regions are removed based on the normalized maximum over different scales of the gradient magnitude, given by,

$$G(\mathbf{x}) = \max_{t_n} \left(\frac{|\nabla L(\mathbf{x}; t_n)|}{(t_n \sigma)} \right), \quad (6.10)$$

and the sign of its derivatives $G_x(\mathbf{x})$ and $G_y(\mathbf{x})$. The process is first explained to the x direction. $G_x(\mathbf{x})$ is calculated by the convolution of $G(\mathbf{x})$ with the mask $[-1\ 0\ 1]$. The vessels in $V(\mathbf{x})$ have a Gaussian profile. This means that, if a vessel is divided along its centerline, $G_x(\mathbf{x})$ has positive or negative values on each side of the centerline [191]. Darker regions near the OD border do not exhibit a Gaussian profile, meaning that its regions have either negative or positive values in $G_x(\mathbf{x})$. These regions in $V(\mathbf{x})$ are detected by the $sgn(G_x(\mathbf{x}))$, where $sgn(.)$ stands for the sign function, according to,

$$\begin{aligned} V_x^+(\mathbf{x}) &= V(\mathbf{x}) sgn(G_x(\mathbf{x})) : G_x(\mathbf{x}) > 0, \\ V_x^-(\mathbf{x}) &= V(\mathbf{x}) sgn(G_x(\mathbf{x})) : G_x(\mathbf{x}) < 0. \end{aligned} \quad (6.11)$$

$V_x^+(\mathbf{x})$ and $V_x^-(\mathbf{x})$ represents the positive and negative values of the darker regions near the OD border (see examples in Figure 6.11 c) and d)). When $V_x^+(\mathbf{x})$ and $V_x^-(\mathbf{x})$ are compared to $V(\mathbf{x})$, only half of the vessels are present, but darker regions near the OD border either suffers no change or totally disappear. A line scan search is performed in $V(\mathbf{x})$ to determine the regions that either suffer no change or completely disappears in $V_x^+(\mathbf{x})$ and $V_x^-(\mathbf{x})$. These regions are saved in $R_x(\mathbf{x})$.



(a) (b)

Figure 6.12: Full retinal vasculature segmentation example with (a) and without (b) peripapillary region.

The above process is repeated along the y direction creating the image $R_y(\mathbf{x})$. The darker regions near the OD border are given by $R(\mathbf{x}) = R_x(\mathbf{x}) + R_y(\mathbf{x})$. Spurious regions in $R(\mathbf{x})$ smaller than 10 pixels are discarded. Furthermore, the regions in $R(\mathbf{x})$ that do not intersect the regions in $OD(\mathbf{x})$ are discarded (see example in Figure 6.11 e)). The final segmented vessels $V_F(\mathbf{x})$, is given by,

$$V_F(\mathbf{x}) = V(\mathbf{x}) - R(\mathbf{x}). \quad (6.12)$$

In Figure 6.11 f) it is possible to observe an example of the OD region with the final peripapillary atrophy region were removed. Figure 6.12 represents an example of the full retinal vasculature segmentation before and after the peripapillary removal.

6.2.4 Results

The automatic method for the vessel segmentation was tested on the images of Drive, Stare, and HRF datasets which are described in appendix A. The performance of the vessel segmentation approach is described next.

6.2.4.1 Retinal vessel segmentation evaluation

To evaluate the retinal vessel segmentation algorithm, three publicly available datasets were used, namely, the Drive dataset [43], composed of 40 images, the Stare dataset [42], composed of 20 images, and the HRF dataset [277], composed of 45 images. For the Drive and the Stare dataset, the first human observer ground truth is used as the evaluation reference [39].
To evaluate the performance, the vessel segmentation resulting in Equation 6.12 is compared to its groundtruth. The evaluation of the proposed method is performed in terms of Sensitivity, Specificity, Accuracy, and Area under the curve (AUC), as described in section 3.3.3. The method reliability is based on the selection of the threshold value T_1 and T_2. Hence, both T_1 and T_2 were evaluated in steps of variable width to analyse their influence on the detection performance. The percentage set of values $\{10, ..., 40\}\%$ were tested for T_1 and T_2.
After, extensive testing T_2 value was set to 20%. Other larger values were evaluated with no significant difference in the final performance.
T_1 is used as a free-running parameter. For each T_1 the Sensitivity and Specificity are determined

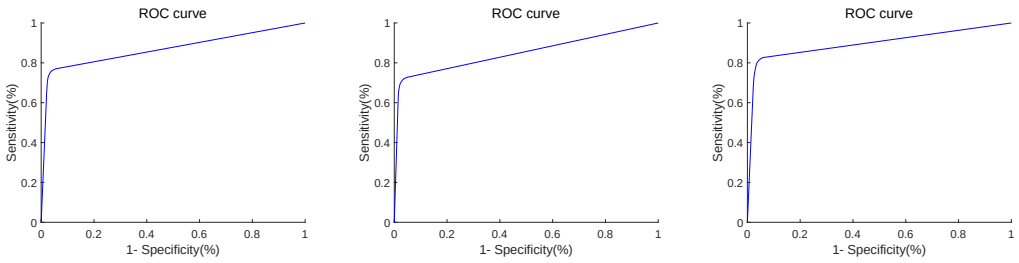
Table 6.2: Performance of the proposed and relevant retinal vessel segmentation methods.

Drive Dataset					
Method	Acc(%)	Sen(%)	Spe(%)	AUC(%)	Evaluation Method
Proposed method (2018)	94.16	70.44	**99.97**	83.98	All images
Khawaja et al. (2019) [211]	95.53	80.43	97.30	-	All images
Shukla et al. (2020) [14]	94.76	70.15	98.36	-	All images
Orlando et al. (2016) [27]	-	78.97	96.84	-	All images
Li et al. (2020) [37]	**97.69**	81.45	98.83	**98.95**	All images
Xu et al. (2020) [232]	95.71	**83.39**	97.50	98.21	All images

Stare Dataset					
Method	Acc(%)	Sen(%)	Spe(%)	AUC	Evaluation Method
Proposed method (2018)	94.23	75.08	**99.96**	86.45	All images
Singh et al. (2016) [200]	92.70	79.39	-	91.40	All images
Rad et al. (2017) [278]	92.89	73.42	96.08	-	All images
Khawaja et al. (2019) [211]	95.45	80.11	96.94	-	All images
Shukla et al. (2020) [14]	95.73	70.17	98.63	-	All images
Orlando *et al.* (2016) [27]	-	76.80	97.38	-	crx val(Leave-one-out)
Li *et al.* (2020) [37]	**97.97**	**85.05**	98.89	**99.24**	All images(50%Test)
Xu *et al.* (2020) [232]	96.64	84.63	98.02	98.81	5-by-5 Repeated crx val

HRF Dataset					
Method	Acc(%)	Sen(%)	Spe(%)	AUC(%)	Evaluation Method
Proposed method (2018)	94.81	80.30	**99.96**	88.85	All images
Panda *et al.* (2016) [221]	94.20	**81.59**	95.25	-	All images
Annunziata *et al.* (2016) [193]	95.81	71.28	98.36	-	All images
Yu *et al.* (2012) [206]	95.66	79.38	-	-	All images
Orlando *et al.* (2016) [27]	-	78.74	95.84	-	crx val(Leave-one-out)
Park *et al.*(2020) [230]	**97.00**	69.48	99.31	-	All images(30%Test)

Sen - Sensitivity; Spe - Specificity; Acc - Accuracy; crx val - Cross validation; best values appear in bold.



(a) Stare dataset (b) Drive dataset (c) HFR dataset

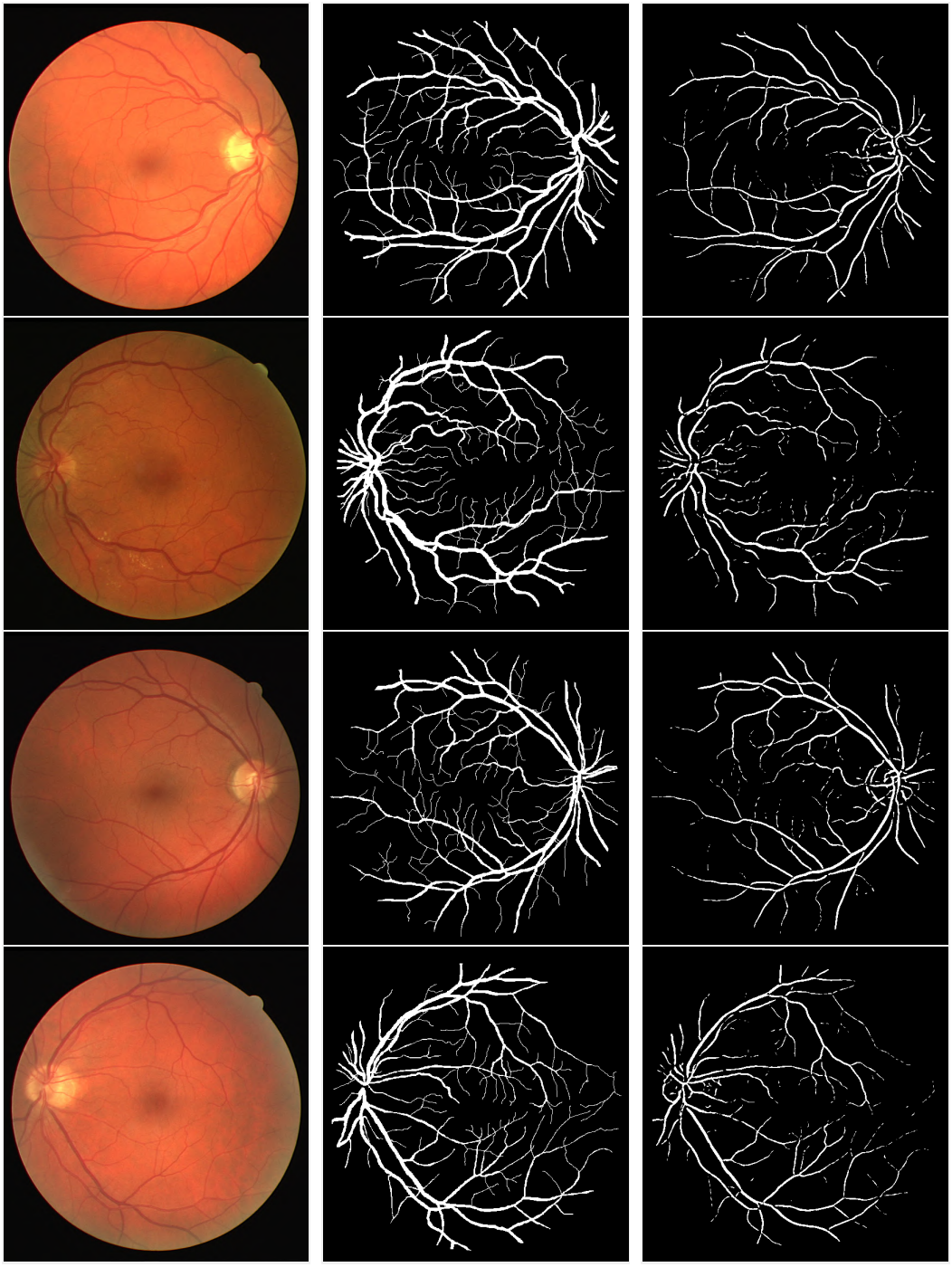
Figure 6.13: Vessel segmentation ROC curves for the Stare, Drive and HFR datasets.

for each image and then averaged for all the images in each dataset. Figure 6.13 shows the ROC curves for the Drive, Stare, and HRF dataset.

The results for the three datasets are represented in, Table 6.2 along with other relevant methods in terms of Sensitivity, Specificity, Accuracy, and AUC.

The proposed approach achieves similar accuracy values with other methods. Furthermore, the proposed rule-based method shows stability and independence from image properties, revealing its reliability.

Overall, on all datasets, a slight decrease in performance is observed for the rule-based methods when compared to machine learning methods. However, the machine learning methods were only tested in a subset, while most of the images are used for training. Moreover, for each dataset, a new training set needs to be defined. The developed method uses the same parametrization for all

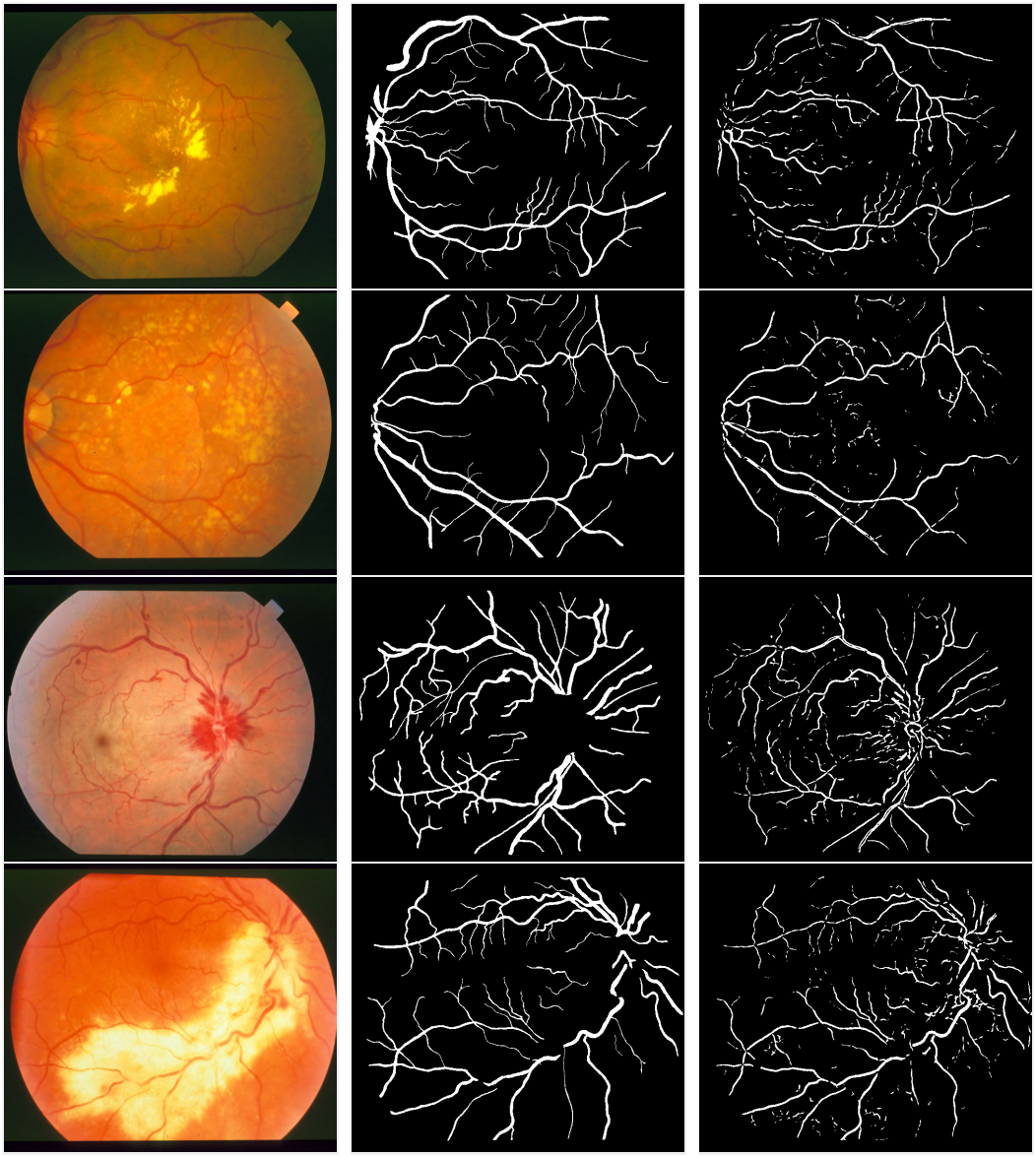


Original Groundtruth Proposed method

Figure 6.14: Examples of retinal vessels segmentation detection result from the Drive dataset.

datasets.

Figure 6.14, 6.15 and 6.16 show some retinal vessel segmentation examples of the proposed method for each studied dataset.



Original Groundtruth Proposed method

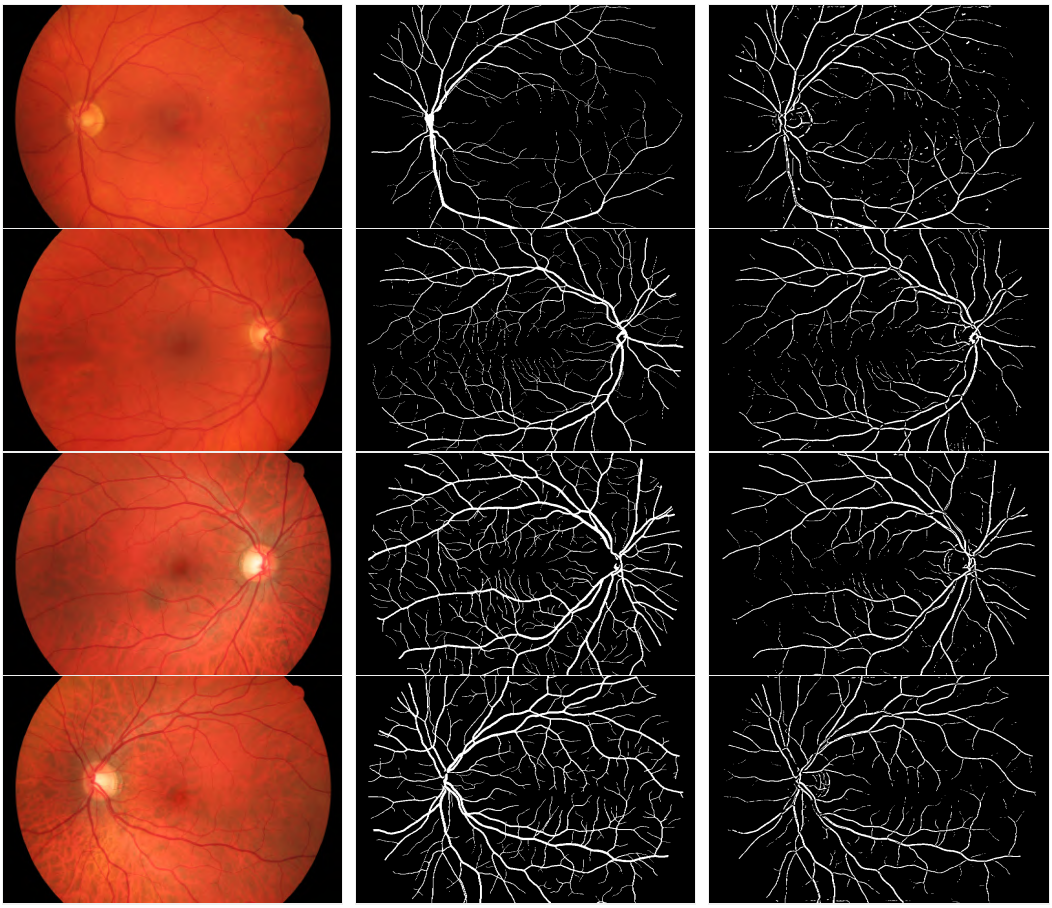
Figure 6.15: Examples of retinal vessels segmentation result from the Stare dataset.

6.3 Concluding remarks

The retinal vessels centerline detection is essential for the automated analysis and characterization of the vascular network. In the previous sections, new and automatic methodologies are described for retinal vessels centerline detection.

Very few works were proposed that specifically deal with the vessels centerline detection. Despite some evaluation limitations given by the manual segmentations, the achieved results suggest that the proposed method is very effective in vessel centerline extraction, providing a good basis for vessels segmentation and vascular tortuosity evaluation. This is confirmed by visual inspection as in Figures 6.14, 6.15 and 6.16 where some examples for the retinal vessel centerline detection method are shown.

In this work, an effective methodology for the automatic segmentation of the retinal vessels is presented with several important contributions: 1) a new retinal vessel enhancement method and



Original Groundtruth Proposed method

Figure 6.16: Examples of retinal vessels segmentation result from the HFR dataset.

2) a new approach to remove darker regions near the OD borders based on differential operators. The method depends on several fixed values. Furthermore, the final segmentation result exhibits some disconnected vessels as well as small vessels that are not detected.

Chapter 7

Retinal microaneurysms detection

Diabetic retinopathy is a major cause of visual loss in the developed world. However, when detected and treated in early stages, blindness can be prevented [6, 247]. The earliest visible ophthalmological sign of DR is the appearance of small and round red dots in fundus images, designated by microaneurysms (MAs) [68]. MAs result from the swelling of small retinal blood vessels formed due to the weakness of the vessel walls [6]. Furthermore, the appearance and disappearance of the MAs, designated by MA turnover, allows controlling the evolution of this disease [279].

The detection of MAs in fundus images might be fundamental, as these images result from a non-invasive procedure widely used in screening procedures [280]. This makes the automatic detection of MAs of utmost importance for early diagnosis and progression control of DR [31].

Although the recently developed research, automatic detection of MAs is far from being trivial. MAs appear in fundus images as small, round dark dots not larger than the thinnest visible vessels, which makes them difficult to distinguish from retinal vessels. Variations of size, shape, and illumination between fundus images, make the automatic detection of MAs a very challenging task [48].

MAs detection is typically divided into three stages [31]. First, image preprocessing improves image quality by removing non-uniform illumination and noise reduction [26]. The second stage aims at the extraction of as many MA candidates as possible [36]. The last stage labels the MA candidates as true or false [31, 26].

In this chapter, a novel multiscale method for the MAs automatic detection is proposed, which can be classified as a rule-based approach. The material herein described is based in the papers *"Microaneurysms detection using a novel neighbourhood analysis"* and *"Microaneurysms Detection in Retinal Images using a multiscale Approach"*. The main motivation was the development of an automatic method for the detection of MAs in fundus images. Furthermore, the proposed method is also evaluated in the detection of DR.

As it is common, also the developed method is divided into three stages, namely, image preprocessing, microaneurysm candidates detection and retinal vessels detection, and labelling of the microaneurysm candidates. Apart from the initial image preprocessing, all stages of the algorithm are designed using a multiscale framework, providing the mechanism for effective microaneurysms detection.

Four publicly available datasets are used for performance evaluation. The proposed method proves to be competitive against other state-of-the-art approaches in both microaneurysms detection and DR detection.

This chapter starts by describing the microaneurysms detection method in section 7.1. Afterward, section7.2 presents the results in images of three different datasets, where a comparison with other results is also included. A discussion of the proposed method is made in section 7.3. Finally, section7.4 summarizes concluding remarks.

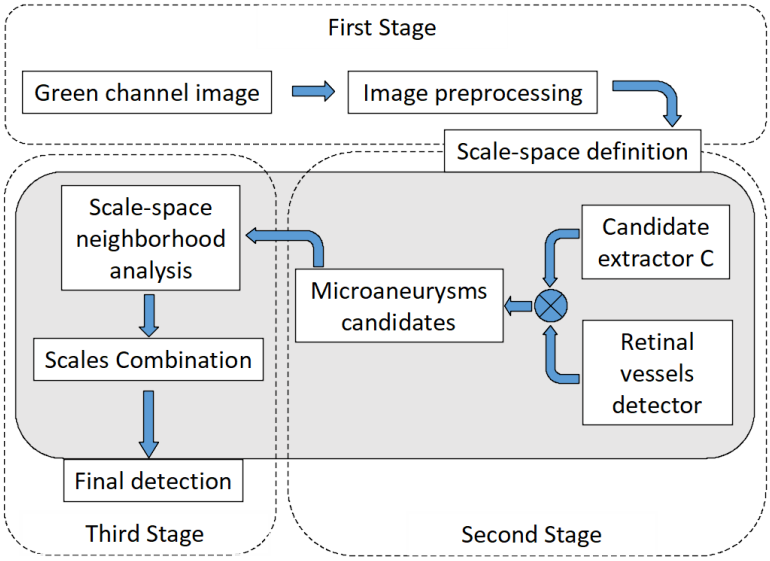


Figure 7.1: Workflow for the proposed MAs detection method.

7.1 Proposed method

The proposed method is divided into three stages (see Figure 7.1): 1) Preprocessing stage that normalizes images size and contrast; 2) multiscale MA candidate extraction and retinal vessel removal; 3) MA candidates labelling as true or false through a multiscale neighbourhood analysis [78] and MA detection using a multiscale approach. The algorithm parameters are defined using the ROC training dataset [48]. This dataset consists of a set of fundus images of various sizes, with variable contrast and quality, similar to those found in clinical screenings [281]. The validation of the method is pursued using E-Ophtha-MA [49] and Latim [47] datasets.

7.1.1 Image preprocessing

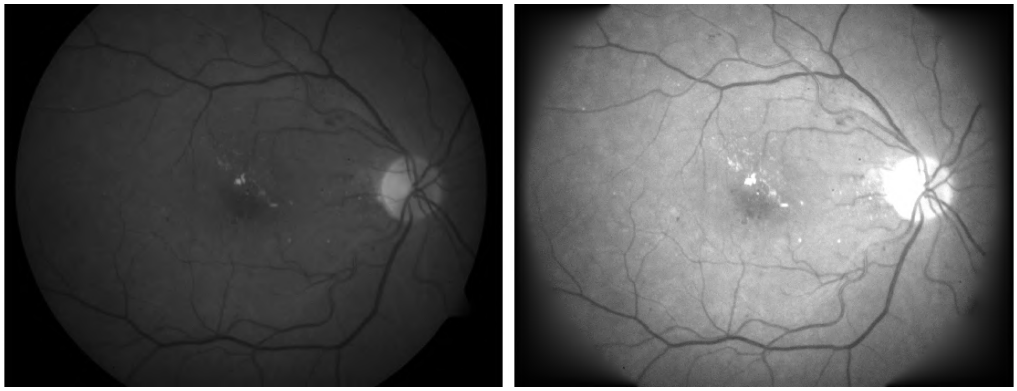
The green channel component $I_G(\mathbf{x})$ (where $\mathbf{x}$ represents the location (x, y)) of each RGB fundus image is selected [4, 36].

Fundus images have different resolutions, so $I_G(\mathbf{x})$ is proportionally scaled using a bi-cubic interpolation such that its smaller and larger dimensions stand between 1000 and 1500 pixels. If both conditions cannot be met, the image reduction is such that the smaller dimension is set to 1000 pixels. This size normalization ensures that the algorithm parameters are the same between all datasets [247].

Next, the scaled $I_G(\mathbf{x})$ values are mapped to cover the full range available (8-bit unsigned integers, with values ranging from 0 to 255). Moreover, 1% of the pixels are saturated at both low and high-intensity values, increasing the contrast and creating the image $I_C(\mathbf{x})$.

There is also the need to define the region of interest (ROI), from the dark background of $I_C(\mathbf{x})$. A binary mask of the ROI designated by $I_{BW}(\mathbf{x})$ is created by selecting all the pixels of the fundus image which have an intensity larger than 20 in the red channel [65].

The transition of intensity values in the peripheral region between the ROI and the dark background of $I_C(\mathbf{x})$ must be smoothed [193, 57]. To do this, an inpainting algorithm, described in [193], performs a smooth interpolation in this transition region at $I_C(\mathbf{x})$. An example of a resulting preprocessed image is represented in Figure 7.2 b).



(a) Original image $I_G(\mathbf{x})$ (b) Pre-processed image

Figure 7.2: Image preprocessing example.

The developed method requires OD removal. For that, the OD center was manually marked at each image and dilated using a *"disk"* structuring element of size 100 pixels, resulting in the binary image $I_{OD}(\mathbf{x})$. This manual process can be replaced by an automatic and reliable algorithm [55, 167]. The preprocessed image multiplied by this mask results in the image $I(\mathbf{x})$, used for the MA detection algorithm presented next.

7.1.2 Multiscale construction

MAs detection requires a selective noise reduction process that retains their contrast [282, 22] and variable circular shape with a Gaussian profile [241, 68]. These geometric characteristics are particularly suited for a multiscale approach [5].
In this chapter, it is used the scale-space defined in appendix B. The parameter t is defined at N_s=7. To preserve MAs shape and size, a slow evolution is required for the scale-space procedure. Due to the existing noise in the lower scales t_0 and t_1 are not considered and only scales t_2 to t_7 are used. Furthermore, the minimum and maximum principal curvatures of the Hessian $H(\mathbf{x};t_n)$ at $(\mathbf{x};t_n)$ defined by $\lambda_1(\mathbf{x};t_n)$ and $\lambda_2(\mathbf{x};t_n)$ [283] are also computed in appendix B.

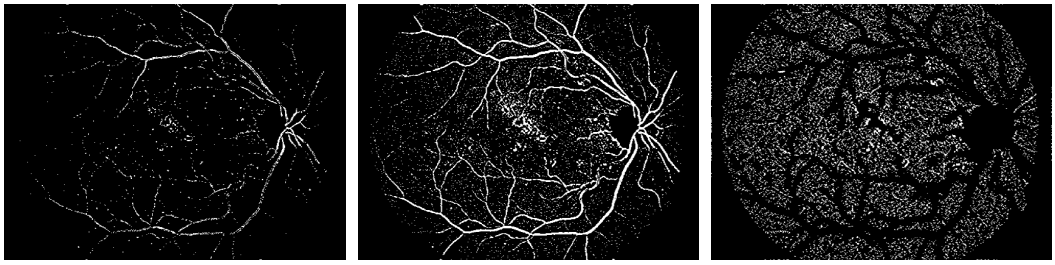
7.1.3 Variable size derivative masks

Derivative masks are usually defined as $[-1\ 0\ 1]$, as in previous chapters. To consider intensity variations from the farthest pixels, variable size derivative masks are considered. Five derivatives masks are defined according to $[-1\ \mathbf{D}(2i-1)\ 1]$, where $\mathbf{D}(2i-1)$ represents a vector with $2i-1$ zeros, with $i = \{1, 2, ...5\}$. Hence, as an example, i =3 will result in $[-1\ 0\ 0\ 0\ 0\ 0\ 1]$. $F_{x|i}(\mathbf{x})$ represents the derivative of a given function $F(\mathbf{x})$ in the x direction. The subscript $y|i$ represents the transpose derivative in the y direction.

7.1.4 Multiscale enhancement and MAs candidate extraction

Lindeberg *et al.* [284] showed that the Laplacian $\nabla L(\mathbf{x};t_n)$ exhibits a good blob response, thus capable of MAs enhancement. To further enhance the MAs response $\nabla L(\mathbf{x};t_n)$ is raised to an exponent of three followed by the transformation,

$$T(\mathbf{x};t_n) = (\nabla^2 L(\mathbf{x};t_n))^3 - \min_{\forall \mathbf{x} \in \{I_{BW}=1\}} (\lambda_1(\mathbf{x};t_n), (\nabla^2 L(\mathbf{x};t_n))^3). \quad (7.1)$$



(a) MA extraction $S(\mathbf{x})$ (b) Vessel detection $V_1(\mathbf{x})$ (c) Vessels removal $V(\mathbf{x})$

Figure 7.3: Example of the MA candidates selection steps.

The transformation in equation 7.1 was first proposed in our earlier study [55] and it is capable of dark structures enhancement. Furthermore, Bennink *et al.* [272] presented a 3D lineness filter for vessel detection. Since MAs are also dark structures, the lineness filter can be modified to provide MAs enhancement. The combination of equation 7.1 and the lineness filter is given by,

$$A(\mathbf{x};t_n) = \left(\frac{\lambda_1^2(\mathbf{x};t_n)}{\lambda_2^2(\mathbf{x};t_n)}\right)\frac{T(\mathbf{x};t_n)}{L(\mathbf{x};t_n)}. \quad (7.2)$$

Furthermore, equation 7.2 is computed with previously described variable size derivative masks, according to,

$$F(\mathbf{x};t_n) = \prod_{i=1}^{5} A_{|i}(\mathbf{x};t_n). \quad (7.3)$$

Equation 7.3 gives a strong enhancement to possible MAs. To select the most relevant candidates, at each scale t_n, only a percentage μ of the higher values of $F(\mathbf{x};t_n)$ are selected, resulting in the binary image $\hat{F}(\mathbf{x};t_n)$. Expanding this to all scales and adding the resulting images, results in the multilevel image $J(\mathbf{x})$,

$$J(\mathbf{x}) = \sum_{n=2}^{N_s} \hat{F}(\mathbf{x};t_n). \quad (7.4)$$

The values of $J(\mathbf{x}) \in \{0,...,6\}$ since only 6 scales t_n are used. It is expected that MAs will be present on the largest number of scales. Hence, only regions in equation 7.4 with $J(\mathbf{x}) > \beta$, were $\beta \in \{1,...,6\}$ were considered, resulting in the binary MAs candidate extractor $S(\mathbf{x})$ (an example is represented in Figure 7.3 a)).

7.1.5 Retinal vessels removal

Retinal vessel crossings are locally circular or have a blob-like shape, which can be erroneously identified as MAs [285]. The vessel detector used in this work uses a multiscale approach based on the mean curvature described in section B. At each scale t_n the multiscale mean curvature $K(\mathbf{x};t_n)$ [261, 78] is defined as,

$$K(\mathbf{x};t_n) = \frac{\lambda_1(\mathbf{x};t_n) + \lambda_2(\mathbf{x};t_n)}{2}. \quad (7.5)$$

To improve the detection of MAs $K(\mathbf{x};t_n)$ is computed with the previously described variable size derivative masks and summed together, resulting in,

$$K_d(\mathbf{x};t_n) = \sum_{i=1}^{5} \frac{\lambda_{1|i}(\mathbf{x};t_n) + \lambda_{2|i}(\mathbf{x};t_n)}{2}. \quad (7.6)$$

Furthermore, a normalization between 0 and 1 followed by a histogram equalization (64 levels) [78] is performed at $K_d(\mathbf{x};t_n)$ resulting in $K_{he}(\mathbf{x};t_n)$. This procedure further enhances dark regions like vessels and MAs. The multiscale vessel detector $V_1(\mathbf{x})$ is given by the summation across all scales of $K_{he}(\mathbf{x};t_n)$,

$$V_1(\mathbf{x}) = \sum_{n=2}^{N_s} K_{he}(\mathbf{x};t_n). \quad (7.7)$$

A binary image of $V_1(\mathbf{x})$, designated by $V(\mathbf{x})$ is obtained by thresholding $V_1(\mathbf{x}) \geq \alpha$ (an example is represented in Figure 7.3 b)). Larger and elongated regions in $V(\mathbf{x})$ are likely to be vessel structures [241] and can be removed according to size and eccentricity criteria. The eccentricity varies between 0 (circles) and 1 (line segment). Regions in $V(\mathbf{x})$ larger than δ pixels or with an eccentricity greater than or equal to ε are represented with 0 (see example in Figure 7.3 d)). Furthermore, regions in $V(\mathbf{x})$ that intersect $I_{OD}(\mathbf{x})$ are also removed (see example in Figure 7.3 c)).

7.1.6 Final extraction of MA candidates

The final extractor of MAs candidates $C(\mathbf{x})$ is defined by a logic AND operation between the MA candidate extractor $S(\mathbf{x})$ and the vessel detector $V(\mathbf{x})$, given by,

$$C_I(\mathbf{x}) = S(\mathbf{x}) \cdot V(\mathbf{x}). \quad (7.8)$$

The shape and dimension of the MAs resulting from equation 7.8 is imprecise due to the logic AND operation. In previous work [78] it was verified that regions in $S(\mathbf{x})$ provide a good representation of MAs shape and dimension. Hence, all regions in $C_I(\mathbf{x})$ are reshaped according to the regions in $S(\mathbf{x})$, resulting in $C(\mathbf{x})$.

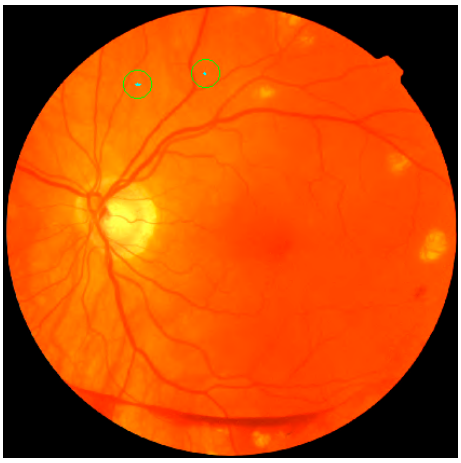
7.1.7 MA labelling with neighbourhood analysis

This stage labels the MAs candidate $C(\mathbf{x})$ as true or false. One of the main clues for the selection of MAs, is the difference in intensity values between the MA and its surroundings [282, 256]. A two-stage neighbourhood analysis, initially reported in [78], is used to perform MAs labelling. For simplicity, the process is explained for a single MA candidate. Moreover, the labelling is performed at each scale t_n.

The first stage analyses the candidate shape, and discards false positives that do not agree with a predefined template (see Algorithm 1). It takes as input the edge descriptor $E(\mathbf{x};t_n)$, the watershed of $E(\mathbf{x})$, $W(\mathbf{x})$, and the MAs candidate extractor $C(\mathbf{x})$. $E(\mathbf{x};t_n)$ is given by [29],

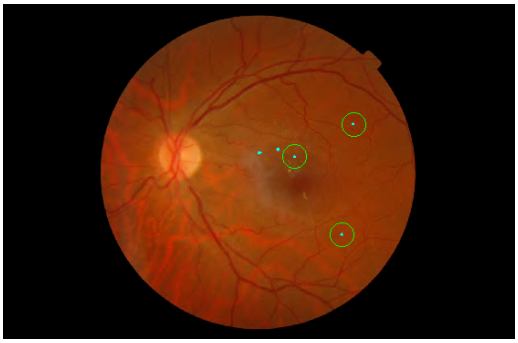
$$\begin{aligned} E(\mathbf{x},t_n) &= L_x^2(\mathbf{x},t_n)L_{xx}(\mathbf{x},t_n) + 2L_x(\mathbf{x},t_n)L_y(\mathbf{x},t_n)L_{xy}(\mathbf{x},t_n) \\ &\quad + L_y^2(\mathbf{x},t_n)L_{yy}(\mathbf{x},t_n). \end{aligned} \quad (7.9)$$

The regions in $W(\mathbf{x};t_n)$ define the neighbourhood of the MA candidate. Following Algorithm 5 some neighbourhood regions might be merged with the initial MA candidate region. This merging



(a)

(b)



(c)

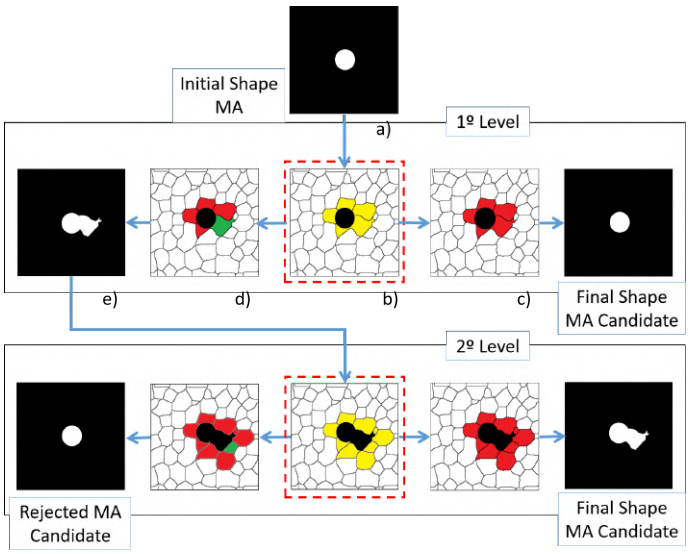
Figure 7.4: MAs detection examples for a) ROC, b) e-Ophtha-MA and c) LAtim datasets. Cyan regions represent detected MAs. The green circles represent the groundtruth.

is made two times at most (controlled by the accumulator j). This MA candidate reshape leads to an improved MA labelling. However, additional reshaping steps elongates the MA candidate, and it was concluded that they reduce the labelling efficiency. The result of Algorithm 5 represented by $C_R(\mathbf{x}; t_n)$ provide the reshaped MA candidates.

The second stage labels the regions in $C_R(\mathbf{x}; t_n)$ as true or false. The process is represented in Algorithm 6 with a near vessel MA candidate as an example. Algorithm 6 takes as input the normalized edge descriptor $E(\mathbf{x}; t_n) \in (0, 1)$ defined by $E_N(\mathbf{x}; t_n)$ and the reshaped $C_R(\mathbf{x}; t_n)$. The perimeter pixels of a candidate are successively dilated, and its median value is calculated. If in any of these dilations, the median of the $E(\mathbf{x}; t_n)$ of the neighbourhood regions is larger than the median of the $E(\mathbf{x}; t_n)$ of the dilated perimeter, the candidate is labelled as a true MA (see Algorithm 6 for details).

If in any dilation the median of the $E(\mathbf{x}; t_n)$ of the neighbourhood regions is larger than the median of the $E(\mathbf{x}; t_n)$ of the dilated perimeter, the candidate is labelled as a true MA (see Algorithm 6 for details). This labelling is made at each scale, t_n creating the binary matrix $D(\mathbf{x}, t_n)$ that represents the final MAs detected at each scale t_n. Figure. 7.4 show some detection examples.

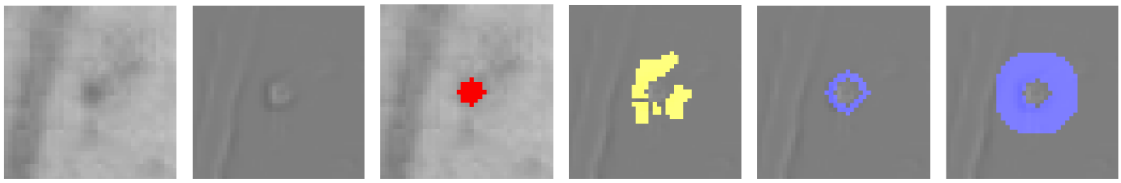
Algorithm 5: MA candidate shape definition.



Input: Edge descriptor $E(\mathbf{x};t_n)$, $W(\mathbf{x};t_n) = watershed(E(\mathbf{x};t_n))$ and $C(\mathbf{x})$
Output: Shape modified MAs $C_R(\mathbf{x};t_n)$

for *each scale* t_n **do**
  for *each region* $P \in C(\mathbf{x})$ **do** `// See (a)`
    $j = 0$
    Remove P from $W(\mathbf{x};t_n)$ `// See (b)`
    Find the adjacent regions of P designated by P_a `// yellow patches at (b)`
    for *each* P_a *calculate the* $E(\mathbf{x};t_n)$ *median value* $\rightarrow \tilde{p_a}$ **do**
      if *all* $\tilde{p_a} <= 0$ **then** `// red patches at (c)`
        Go to 2^{nd} line `// evaluate another` P
      end
      if *any* $\tilde{p_a} > 0$ **then** `// green patch at (d)`
        Merge P with $P_a \rightarrow P$ `// See (e)`
        if $j < 1$ **then**
          $j = j + 1$ and go to 4^{th} line
        else
          Remove P from $C(\mathbf{x})$ and go to 2^{nd} line `// Evaluate another` P
        end
      end
    end for
  end for
  $C_R(\mathbf{x};t_n) \leftarrow C(\mathbf{x})$
end for

Algorithm 6: MAs labelling.



a) $E(\mathbf{x};t_n)$ b) $E_N(\mathbf{x};t_n)$ c) P d) P_a e) $D_P, \pi = 2$ f) $D_P, \pi = 7$

Input: Normalized edge descriptor $E_N(\mathbf{x};t_n)$, $C_R(\mathbf{x};t_n)$ and $W(\mathbf{x};t_n)$ `// See (a,b)`
Output: Final MAs detected matrix $D(\mathbf{x};t_n)$

for *each* t_n **do**
  for *each region* $P \in C(\mathbf{x})$ **do** `// See (c)`
    Find the adjacent regions of $P \rightarrow P_a$ defined in $W(\mathbf{x};t_n)$ `// See (d)`
     Calculate the $E_N(\mathbf{x};t_n)$ median value of $P_a \rightarrow \tilde{p_a}$
    for $\pi = 2$ *to* 7 **do**
      Evaluate $(P \oplus b_\pi) - P \rightarrow D_P$ `// See (e,f)`
       Calculate the $E_N(\mathbf{x};t_n)$ median value $D_P \rightarrow \tilde{d_p}$
      if $\tilde{p_a} > \tilde{d_p}$ **then**
        Label P as MA and go to 2^{nd} line
      else if $\tilde{p_a} <= \tilde{d_p} \wedge \varepsilon = 7$ **then**
        Remove P from $C_R(\mathbf{x};t_n)$ and go to 2^{nd} line
      end
    end for
  end for
  $D(\mathbf{x};t_n) \leftarrow C_R(\mathbf{x};t_n)$
end for

$\oplus$ - morphological dilatation

7.2 Validation and experimental results

7.2.1 Datasets

For the evaluation of the developed method for MAs detection, three publicly available datasets are considered, notably, E-Ophtha-MA [49], Latim [47] and ROC Training [48]. The publicly available dataset Messidor [46] is used for the DR detection performance analyses. These are described in appendix A In the following a short description of these datasets is given.

1) *E-Ophtha-MA* [49] - Results from a teleophthalmology project in DR. It consists of 381 images for MAs detection. 148 images contain 1306 MAs and 233 have no lesions. The MAs were outlined and confirmed by a group of ophthalmologists. All images are in JPEG format with different resolutions (1440×960, 1504× 1000, 2504×1696, 2048×1360).

2) *Latim* [47] - Contains 36 fundus images compiled by the Brest University Hospital. Images have a resolution of (2240×1488) stored in TIFF format. The MAs were annotated by a single DR expert.

3) *ROC* [48] - Consists of 504 Training images and 50 Test images. The images were taken from a DR screening program and acquired with different cameras, fields of view, and resolutions, and stored in JPEG format. Image resolutions vary into $(768 \times 576$, 1058×1061 and $1389 \times 1383)$. Only the Training set has the correspondent ground truth images annotated by four different retinal experts. A region is a true MA when it is marked by at least two experts.

Table 7.1: Sensitivities of the proposed method for the ROC Training dataset in different scales and scales combinations.

Scales	Sensitivity against FP/I							F_{Score}
	1/8	1/4	1/2	1	2	4	8	
t_2	0.051	0.083	0.127	0.198	0.227	0.319	0.355	0.194
t_3	0.045	0.078	0.120	0.207	0.259	0.350	0.395	0.208
t_4	0.045	0.080	0.127	0.210	0.262	0.360	0.405	0.213
t_5	0.048	0.083	0.130	0.208	0.264	0.364	0.408	0.215
t_6	0.0391	0.0722	0.1201	0.2043	0.2635	0.3594	0.4158	0.211
t_7	0.041	0.076	0.123	0.202	0.262	0.349	0.4038	0.208
t_5, t_6	0.052	0.086	0.140	0.214	0.273	0.365	0.417	0.221
t_5, t_4	0.042	0.078	0.135	0.211	0.266	0.365	0.414	0.216
t_4, t_5, t_6	0.058	0.091	0.145	0.215	0.277	0.374	0.427	0.227

4) *Messidor* [46] - Consists of 1200 images with resolutions $(440\times960$, 2240×1488 and $2304\times1536)$ stored in TIFF format. Each image has a DR grade given by R0, R1, R2 and R3. R0 corresponds to the absence of DR. R1 and R2 correspond to mild and severe cases of non-proliferative DR. R3 stands for proliferative DR. The grading depends on the number of MAs, haemorrhages, and neovascularization [75].

7.2.2 Influence of the parameters

The developed rule-based method has several thresholds, notably μ, β, α, ε and δ, which have influence in its performance. The parameters β and μ were defined for the candidate extractor $S(\mathbf{x})$ in section 7.1.4. The parameters α, ε and δ were defined for the vessel extractor $V(\mathbf{x})$ in section 7.1.5. The reliability of the developed algorithm requires a very stable behaviour for this parameterization. The influence of these parameters is studied in the following, using μ as the free-running parameter. In this sense, μ is the parameter that establishes the sensitivity of the method.

The evaluation is made in terms of a free-response operating characteristic curve (F_{ROC}), where lesion sensitivity (True positive (TP) per number of Positives) values are plotted against the FP/I (False positives per image) [48]. The several sensitivities and FP/I values are calculated by varying the threshold $\mu \in \{0.1, 0.2, ..., 0.9, 1, 2, ..., 3\}$ of the MAs candidate extractor $S(\mathbf{x})$ defined in section 7.1.4. For the F_{ROC} curve comparison [48], the F_{Score} score defined as the average sensitivity at seven low predefined FP/I $\{1/8, 1/4, 1/2, 1, 2, 4, 8\}$ values is used. The area under the F_{ROC} curve is defined as F_{AUC}. The curve fitting is obtained using the trapezoidal integration and normalization method proposed in [31].

The ROC training dataset is considered the most challenging for MAs detection due to the heterogeneity of its images in terms of resolution, contrast, and illumination. Hence, it was used to study and define parameterization.

Figure 7.5 represent the F_{ROC} curves for the evaluation of each threshold parameter. The best parameter values and resulting F_{Score} appear in bold. Before any specific analysis is observed that β, α, ε, and δ result in very similar performances for the studied values, revealing the robustness of the method. The parameter β depends on the number of scales. As only six scales are used, it will be inside the range $\{1, ..., 6\}$. Figure 7.5 a) shows that the best result is achieved when β =5, with a F_{Score} =0.227 provides better performance for lower FP/I values. $\beta = 6$ leads to very similar performance. The parameter α also assumes values in the range, $\{1, ..., 6\}$ as also depends on the number of scales. The F_{Score} curves with α =3,...,6 reveal very similar performance. The best

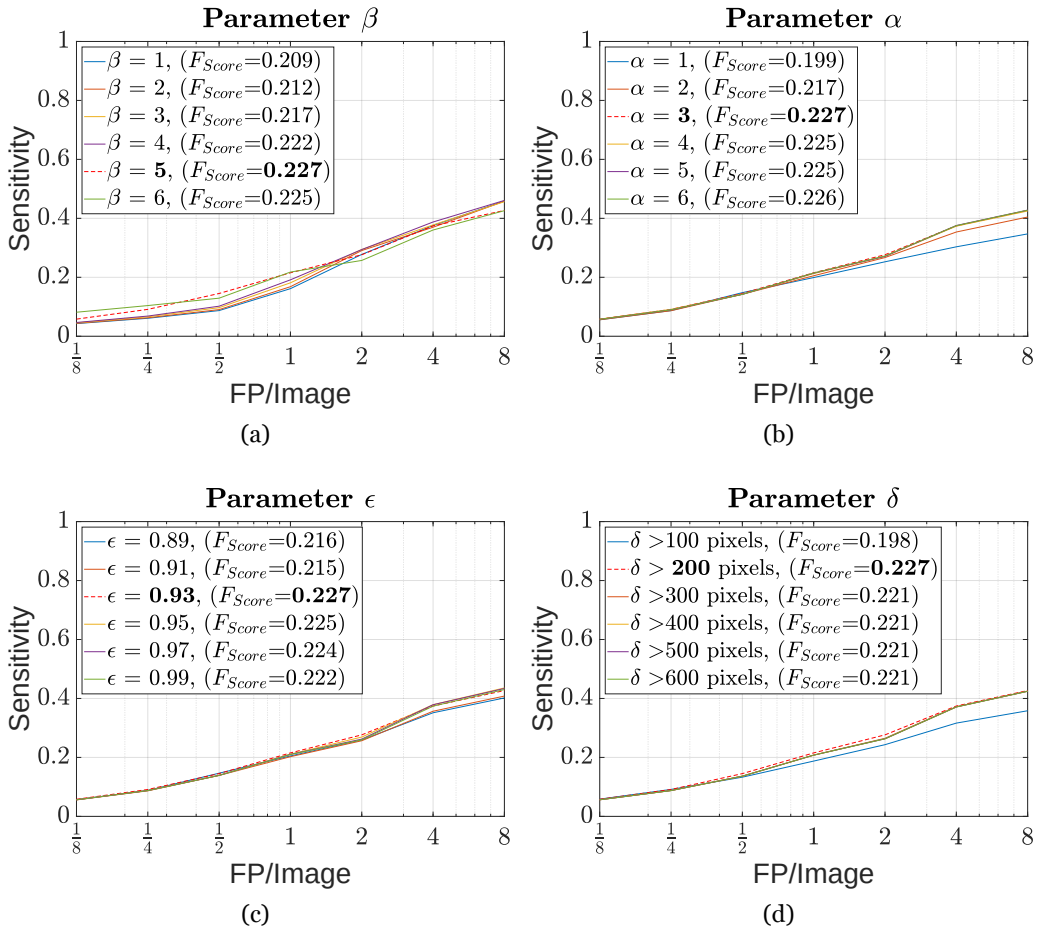


Figure 7.5: Parameter variation analysis. a) β, b) α, c) ε and d) δ parameters. Best parameter and F_{Score} values appear in bold.

value is α =3 (see Figure 7.5 b)). The parameter ε is a threshold for the vessels eccentricity. As vessels do not have a large eccentricity, being almost straight lines, ε was evaluated in the range {0.89,...,0.99}. The performance peak results for ε =0.93 (F_{Score} =0.227), although the defined model reveals very similar performance with this parameter variations. The threshold δ represents a dimensional criterion for the region to be considered as non-MA. The tested values ranged from {100,...,600} pixels. The best performance was obtained for $\delta \geq$ 200 pixels (see Figure. 7.5 d)). However, for higher values of δ, the performance values tend to decrease, although F_{Score} =0.221 does not change.

7.2.3 Lesion-level analysis

The lesion-level evaluation is made in terms of the same free-response operating characteristic curve (F_{ROC}) and the final score F_{Score} previously defined. The algorithm performance is evaluated at each individual scale and with different scales combinations. A comparison with other methods is also made using three publicly available datasets.

Figure 7.6 a) represents the F_{ROC} curve for single scale evaluation. It can be seen that the detection performance is very similar across all scales. This is also seen in Table 7.1 where the F_{Score} between scales also shows similar values. The best performing scale is, t_5 with a F_{Score} =0.215. The scales t_4 and t_6 with an F_{Score} of 0.213 and 0.211 respectively also perform very similarly.

Table 7.2: MA detection comparison on several datasets.

Methods	Sensitivity against FP/I							F_{Score}	F_{AUC}	Evaluation Method
	1/8	1/4	1/2	1	2	4	8			
E-Ophtha-MA Dataset										
Proposed method (2021)	**0.173**	**0.249**	**0.353**	**0.446**	**0.545**	**0.637**	**0.699**	**0.443**	**0.602**	All images
Du *et al.* (2020) [20]	0.227	0.309	0.407	0.488	0.622	0.739	0.820	0.516	-	10-Fold crx val
Melo *et al.* I (2020) [63]	0.209	0.284	0.365	0.449	0.551	0.551	0.551	0.423±0.003	0.515	10-by-10 Repeated crx val(10%)
Melo *et al.* II (2020) [63]	0.145	0.197	0.303	0.404	0.499	0.598	0.598	0.392	0.525	10-Fold crx val(10% Test)
Eftekhari *et al.* (2019) [26]	0.091	0.258	0.401	0.534	0.579	0.667	0.771	0.471	0.637	10-Fold crx val(10% Test)
Chudzik *et al.* (2018) [253]	0.185	0.313	0.465	**0.604**	**0.716**	**0.801**	**0.849**	**0.562**±0.233	**0.640**	2-Fold crx val(50% Test)
Dashtbozorg *et al.* (2017) [31]	**0.358**	**0.417**	**0.471**	0.522	0.558	0.605	0.638	0.510	0.575	10-Fold crx val(10% Test)
Wu *et al.* (2017) [22]	0.063	0.117	0.172	0.245	0.323	0.417	0.573	0.273	0.386	All images(50% Test)
ROC Training Dataset										
Proposed method (2021)	0.058	0.091	0.145	0.215	0.277	0.374	0.427	0.227	0.308	All images
IRIA Group* (2011) [6]	0.041	**0.160**	**0.192**	**0.242**	**0.321**	**0.397**	**0.493**	**0.264**	**0.368**	8-Fold crx val(10% Test)
OKMedical I (2010) [32]	**0.060**	0.096	0.138	0.189	0.267	0.311	0.345	0.201	-	All images
Du *et al.* (2020) [20]	0.155	0.166	0.197	0.260	0.331	0.436	0.504	0.293	-	10-Fold crx val
Melo *et al.* I (2020) [63]	0.073	0.090	0.115	0.149	0.204	0.262	0.327	0.174±0.005	0.240	10-by-10 Repeated crx val(10%)
Melo *et al.* II (2020) [63]	0.077	0.092	0.113	0.149	0.205	0.283	0.348	0.181	0.254	10-Fold crx val(10% Test)
Eftekhari *et al.* (2019) [26]	0.047	0.173	0.351	**0.552**	**0.613**	**0.722**	**0.769**	**0.461**	**0.660**	10-Fold crx val(10% Test)
Chudzik *et al.* (2018) [253]	0.039	0.141	0.174	0.243	0.306	0.385	0.193	0.298±0.116	-	2-Fold crx val(50% Test)
Wang *et al.* (2017) [250]	**0.273**	**0.379**	**0.398**	**0.481**	**0.545**	0.576	0.598	**0.464**	**0.548**	All images(50% Test)
Dashtbozorg *et al.* (2017) [31]	**0.435**	**0.443**	**0.454**	0.476	0.481	0.495	0.506	**0.471**	0.484	10-Fold crx val(10% Test)
Wu *et al.* (2017) [22]	0.037	0.056	0.103	0.206	0.295	0.339	0.376	0.202	0.302	10-Fold crx val(10% Test)
Dai *et al.* (2016) [35]	0.219	0.257	0.338	0.429	0.528	0.598	**0.662**	0.433	0.553	2-fold crx val (50% test)
Lazar *et al.* (2013) [68]	0.037	0.055	0.103	0.162	0.196	0.223	0.285	0.152	-	All images(MA/Non-MA patches)
Fujita Lab (2009) [256]	0.050	0.068	0.106	0.146	0.195	0.252	0.323	0.163	-	All images(25% Test)
ROC Test Dataset										
DRSCREEN (2012) [5]	0.173	**0.275**	**0.380**	**0.444**	**0.526**	***0.599***	**0.643**	**0.434**	**0.551**	All images
OkMedical I (2010) [32]	**0.198**	0.265	0.315	0.356	0.394	0.466	0.501	0.357	0.430	All images
Latim (2008) [47]	0.166	0.230	0.318	0.385	0.434	0.534	0.598	0.381	0.489	All images
Lazar *et al.* (2013) [68]	0.251	0.312	0.350	0.417	0.472	0.542	0.615	0.423	0.510	All images
Adal *et al.* (2014) [36]	0.204	0.255	0.297	0.364	0.417	0.478	0.532	0.364	0.446	All images
Pereira *et al.* (2014) [247]	0.053	0.083	0.135	0.187	0.276	0.407	0.540	0.240	0.366	All images
GIB Valladolid (2009)[237]	0.190	0.216	0.254	0.300	0.364	0.411	0.519	0.322	-	All images
Niemeijer *et al.* (2005) [34]	0.243	0.297	0.336	0.397	0.454	0.498	0.542	0.395	0.469	All images
Latim Dataset										
Proposed Method (2021)	0.049	**0.058**	**0.090**	**0.142**	**0.203**	**0.371**	**0.535**	**0.207**	**0.322**	All images
Pereira *et al.* (2014) [247]	-	0.050	0.077	0.130	0.215	0.308	0.449	0.205	-	All images

† - Values are from [239]; crx val - Cross validation. Best values appear in bold.

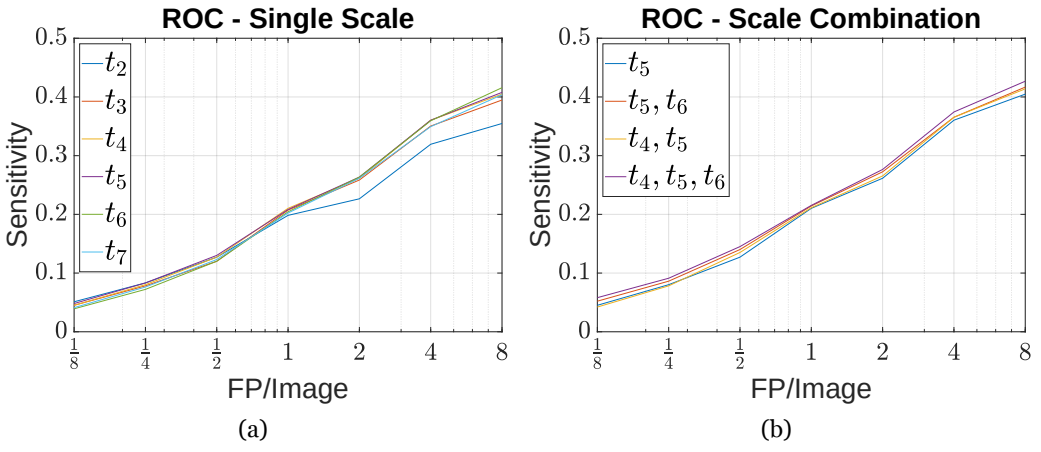
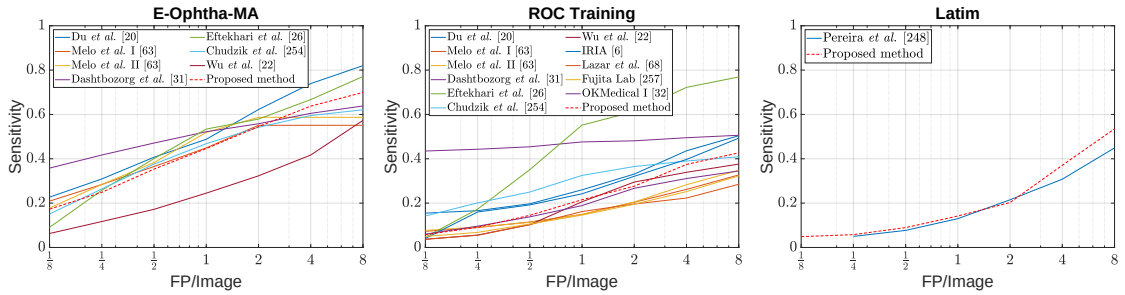


Figure 7.6: Performance comparison of the proposed algorithm in ROC dataset. In a) the labelling is performed at the specified scales, and in b) the labelling is performed using a scale combination.

It was also evaluated whether a scale combination improves MAs detection. In this scenario, a region is considered an MA if it is present at all scales of a given scale combination. Several scales combination were tested, but only the ones that provide the best performance in terms of F_{Score} are shown in Figure 7.6 b) and in Table 7.1. The selection criteria were that a scale combination must



(a) (b) (c)

Figure 7.7: F_{ROC} plots comparing several methods in a) E-Ophtha dataset, b) ROC dataset and c) Latim dataset.

achieve a F_{Score} value larger than the obtained with the individual scales. Relevant scale combination are (t_4, t_5), (t_5, t_6) and (t_4, t_5, t_6). Scale combination (t_4, t_5) and (t_5, t_6) achieved an F_{Score} of 0.216 and 0.221 respectively. The combination (t_4, t_5, t_6) achieved an F_{Score} of, 0.227 improving the best single scale result (t_5 that result in an F_{Score} of 0.215). Due to its better performance this scale combination is used in the remaining evaluation.

The ROC plots of the E-ophtha-MA, ROC training, and Latim datasets are shown in Figure 7.7. Quantitative results are displayed in Table. 7.2. Other relevant methods are also reported for comparison. The proposed method results in F_{Score} =0.443, 0.227 and 0.207, and F_{AUC} =0.602, 0.308 and 0.322, respectively for the three datasets.

7.2.4 Image-level analysis

The Messidor eye fundus dataset is classified in terms of DR or non-DR images, i.e. (R0 vs. (R1,R2,R3)) [75, 6]. This dataset allows the evaluation of the performance of the proposed method for a screening procedure, considering the detection of MAs. To compare the results with other methods, it is used the area under the curve AUC and the Sensitivity at an operating point of 50% of Specificity, as in [24, 75]. Sensitivity and Specificity are defined as, Sensitivity=TP/(TP+FN) and Specificity=TN/(TN+FP) respectively, where TP is the number of true positive images, TN of true negative, FP of false positive, and FN of false negative. An image classification only depends on the detection or not of MAs, and it is independent of the performance of the MAs detection in each image.

The ROC plots are shown in Figure 7.8 and the results are summarized in Table 7.3. The decision of the detection of DR was tested considering at least one MA is detected or when at least two MAs are detected. The proposed method achieves AUC=0.874 and a Sensitivity=0.936 at a Specificity of 50% when at least one MA is detected. An AUC=0.879 and a Sensitivity=0.931 at a Specificity of 50% result when a DR classification is considered if at least 2 MAs are detected. It is important to analyse the MA detection model when at least two MAs are detected because many MA detection models tend to detect only one false positive in fundus images without any MA. According to the British Diabetic Association, a screening system must have a Sensitivity of 95% and specificity of 80% [286]. The proposed method achieves a sensitivity of 98.73% and 98.18% for the specificity of 80% respectively for at least one or two MAs detection.

Table 7.3: DR classification AUC and Sensitivity (defined at a Specificity of 50%) values for the Messidor dataset.

Messidor		
Method	AUC	Sensitivity
Proposed Method (2021)	0.874	0.936
Proposed Method (2021) ($\#MA \geq 2$)	0.879	0.931
Giancardo *et al.* (2013) [74]	0.854	0.910
Sanchéz *et al.* (2011) [75]	0.876	0.922
Orlando *et al.* (2018) [24]	0.893	0.911
Seoud *et al.* [7]	0.899	0.939
Expert A [75]	0.922	0.945
Expert B [75]	0.865	0.912

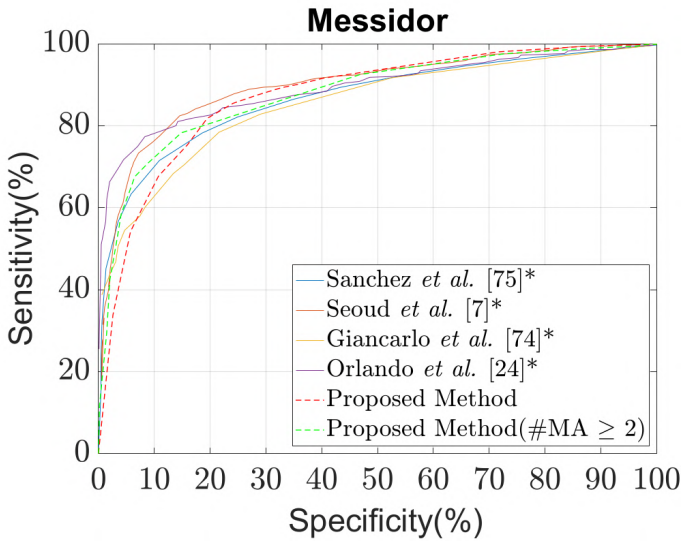


Figure 7.8: DR classification ROC plots for several methods using the Messidor dataset. * plots obtained from [74, 7, 75, 24, 5].

7.3 Discussion

The proposed method was evaluated in terms of lesion-level detection and image-level classification.

7.3.1 Lesion-level performance

A very effective multiscale method for MAs detection was developed. This method slightly improves the performance while significantly improving the FP/I ratios when multiple scales are used, instead of a single scale. In practice, some wrongly detected regions in individual scales are suppressed using multiscale analysis.

The best performance was found for the scales t_4, t_5 and t_6. Niemeijer *et al.* [34] used a σ =2 to extract features, which is very similar to scale t_5 (2.06). Zhang *et al.* [32] defined a multiscale correlation with σ values ranging from 1.1 to 1.5. Ram *et al.* [6] defined four scales with a σ ={1,2,4,8}. The developed method used the complete dataset for evaluation. This makes a comparison with most of the previously developed methods difficult, as they only use a part of the dataset. The state-of-the-art methods report average values of partial evaluations, typically using a 10-fold cross-validation, which benefits the final results. Those methods are also based on machine learning,

and the reported performances used the larger part of the tested dataset for training. The proposed MAs detection method was designed using the ROC training dataset and was used without any change in e-Ophtha-MA and Latim datasets. It provided the best results in experiments that used the complete datasets. The evaluation was not possible using the ROC testing dataset since it is no longer possible to submit new results and the ground truth was not made available.

7.3.2 Image-level performance

The proposed method was also tested to classify a DR using the fundus images based on the presence or absence of MAs. The proposed method exhibit a similar performance to Sanchéz *et al.*, Orlando *et al.* and Seoud *et al.* [7], as well as a performance comparable to the Human expert B. When the decision of a DR case is defined when at least two MAs are detected reveals more reliability than basing the decision on the detection of a single MA. This is caused because the method tends to define an FP MA in images without any MA. The achieved results demonstrate that the proposed method can effectively distinguish DR from non-DR cases. This is important since the grading performance of automated systems is highly dependent on MAs detection [5].

7.4 Conclusion

In this chapter, a new multiscale approach to detect MAs was described and evaluated in several publicly available datasets. Furthermore, the classification of fundus images revealing a DR based on the MAs detection was also considered. This chapter also presents an assessment of the most appropriate scales for the MAs detection, complementing previous work that defines these scales empirically. It was also revealed that a multiscale analysis can improve the MAs detection. Although recent research has been focusing on deep learning methods, rule-based methods can also be important for the definition of features, that can significantly improve the outcome of these methods [24]. Moreover, machine learning based methods tend to require new training for different datasets, while the developed method reveals to be reliable independently of the dataset. The achieved results show that the proposed approach is very competitive when compared with the current state of the art MAs detection methods, particularly in high-resolution images. Similarly, the use of the proposed method for DR screening revealed similar performances with the current state-of-the-art methods as well as with human examiners. The DR classification was only based on the MAs detection and its performance could be improved if other pathological clinical manifestations like exudates and haemorrhages are considered.

Chapter 8

Conclusions and future work

This chapter presents the conclusions resulting from the research developed in the context of this thesis. Furthermore, some research topics that can be addressed in the future, are also described.

8.1 Conclusions

This thesis describes the studies developed on retinal fundus image analysis. Two different retinal anatomical structures were considered, namely the detection of the vascular network and the OD. The retinal pathological clinical manifestations addressed in this thesis were the segmentation of exudates and the detection of MAs. For the exudates segmentation, two different methods were developed. The first method segmented the exudates using a scale-space curvature framework. The second method is based on the noise map distribution. Both methods exhibit good performance, achieving results similar to other state-of-the-art techniques.
For the OD detection, a method based on a new technique named cumulative sum fields was developed. This method was tested with eight different datasets, and results are comparable to other state-of-the-art methods. However, the best-performing methods in the literature were evaluated on a smaller number of datasets [58, 59, 60, 61, 62].
The detection of the blood vessel's centerline was obtained by a modified corner detector combined with binary filters and dynamic thresholding. This method was evaluated in two publicly datasets, with better results than other state-of-the-art methods.
The segmentation of the vascular network was achieved with a new scale-space blood vessel enhancement method. The performance evaluation is made on three publicly available datasets with competitive results.
Finally, the method for MAs detection relies on a multiscale framework for both MA detection and labelling. The labelling uses an approach based on the MAs neighbourhood. The achieved results show that this method is very competitive when compared with the current state-of-the-art methods, particularly in high-resolution images.
Based on the developed methods, several individual improvements or contributions were proposed. Among them, the most prominent are: A new enhancement method of vessel structures, a new cumulative sum fields technique for the OD detection, a scale-space modified corner detector for the vessels detection, exudates extraction using a noise map distribution, a new procedure to detect and remove peripapillary atrophy regions around the OD border, and a neighbourhood analysis for MAs labelling only dependent on pixel intensity.
One of the main objectives was the creation of a somehow uniform and common approach for the detection and segmentation of the retinal structures and pathological clinical manifestations. In several described cases a scale-space framework approach was used, notably for exudates segmentation, retinal vasculature segmentation, and MAs detection.
The state-of-the-art methods, using scale information, either extract region candidates [31, 32, 33] or obtain multiscale features [34, 35, 37, 6], with scale values usually empirically defined.
Overall, the developed research led to a more solid understanding of which are the most appropriate scales (individual or combination) to carry out the detection and segmentation of retinal

structures and pathological clinical manifestations. This knowledge can have an important impact in the development of new features and descriptors that can be used in the improvement of the generalization of machine learning methods [63, 24].
For the exudates segmentation and the OD detection method, no preprocessing stage was performed for non-illumination equalization and contrast normalization. This is important for pixel information preservation and is a major difference from other works [64, 65]. These findings contradict some works in the literature [66, 67], where a preprocessing stage is considered fundamental to perform retinal structures and pathological clinical manifestations detection and segmentation. The developed method's robustness was verified by its evaluation on several datasets. Particularly, the OD detection method proved to be effective in pathological fundus images.
Since MAs are the first clinical manifestation in fundus images [68], its detection enabled the evaluation of DR detection. The evaluation for DR detection revealed a similar performance with the current state-of-the-art methods as well as with human examiners. This also suggests the potential for the MAs detection method to be used as a DR screening tool.
In summary, although recent research for automated analysis of fundus images has been focused on machine learning approaches, rule-based methods can provide important contributions for the development of automated computational analysis methods. Particularly, the research work developed in this thesis gives important insights about new techniques, a uniform scale-space framework, and scale information that can be used in the development of new features or descriptors with a higher discriminative power [63, 24].

8.2 Future directions

The research developed during this thesis revealed the potential applicability of scale-space approaches in the automated detection of retinal structures and pathological clinical manifestations. Future directions for improvement should consider clinical applicability.
Even though the MAs detection method exhibits the potential to be used as a DR screening method, other issues must be addressed. First, there is a need to include other pathological clinical manifestations, like exudates, haemorrhages, and neovascularization in a possible DR screening tool. Secondly, the validation of a screening tool requires larger datasets where other factors like different age groups, ethnicity and interaction with other retinal diseases must be included and considered. One of the main handicaps in publicly available datasets is the lack of ethnic heterogeneity. The majority of the publicly available datasets are composed of images representing Caucasian and Asian populations. An exception is the Hamilton Eye Institute Macular Edema Dataset (HEI-MED) [69] that is formed by a reasonable mixture of ethnicities. This lack of heterogeneity of ethnicities in fundus image datasets results in limited applicability generalization of the automated algorithms to a broader population. This is a well-known fact that different ethnicities show different morphological and pathological alterations to the same retinal diseases [2].
One of the major problems for the improvement and applicability of automated methods results from the testing databases annotation reliability and lesion ambiguity present in several datasets. This compromises the development of reliable and effective algorithms that can be effectively used in real-life clinical scenarios. Future research should include peer-reviewing of current datasets and the creation of new annotated datasets.

Bibliography

[1] U. Schmidt-Erfurth, A. Sadeghipour, B. S. Gerendas, S. M. Waldstein, and H. Bogunović, "Artificial intelligence in retina," *Progress in Retinal and Eye Research*, vol. 67, pp. 1–29, Nov. 2018. xi, 1

[2] L. A. Remington, "Chapter 1 - visual system," in *Clinical Anatomy of the Visual System (Second Edition)*, second edition ed., L. A. Remington, Ed. Saint Louis: Butterworth-Heinemann, 2005, pp. 1–8. xi, xvii, xxvii, 1, 7, 9, 116

[3] K. Tobin, E. Chaum, V. Govindasamy, and T. Karnowski, "Detection of anatomic structures in human retinal imagery," *IEEE transactions on medical imaging*, vol. 26, pp. 1729–39, 01 2008. xi

[4] R. J. Winder, P. J. Morrow, and e. a. McRitchie, "Algorithms for digital image processing in diabetic retinopathy." *Comp. Med. Imag. and Graph.*, vol. 33, no. 8, pp. 608–622, 2009. xi, 1, 43, 50, 63, 89, 102, 143

[5] B. Antal and A. Hajdu, "An ensemble-based system for microaneurysm detection and diabetic retinopathy grading," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 6, pp. 1720–1726, June 2012. xi, xxix, 2, 37, 41, 103, 111, 113, 114, 143

[6] K. Ram, G. Joshi, and J. Sivaswamy, "A successive clutter-rejection-based approach for early detection of diabetic retinopathy," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 664–673, 2011. xi, xiii, 2, 3, 37, 41, 101, 111, 112, 113, 115

[7] L. Seoud, T. Hurtut, J. Chelbi, F. Cheriet, and J. M. Langlois, "Red Lesion Detection Using Dynamic Shape Features for Diabetic Retinopathy Screening," *IEEE Transactions on Medical Imaging*, 2016. xi, xxix, 2, 39, 42, 113, 114

[8] T. V. Phan, L. Seoud, H. Chakor, and F. Cheriet, "Automatic screening and grading of age-related macular degeneration from texture analysis of fundus images," *Journal of Ophthalmology*, vol. 2016, pp. 1–11, 01 2016. xi, 2

[9] M. Zapata, D. Royo-Fibla, O. Font, J. Vela, I. Marcantonio, E. U. Moya-Sanchez, A. Sanchez-Perez, D. Garcia-Gasulla, U. Cortés, E. Ayguade, and J. Labarta, "Artificial intelligence to identify retinal fundus images, quality validation, laterality evaluation, macular degeneration, and suspected glaucoma," *Clinical Ophthalmology*, vol. Volume 14, pp. 419–429, 02 2020. xi, 2

[10] A. Mvoulana, R. Kachouri, and M. Akil, "Fully automated method for glaucoma screening using robust optic nerve head detection and unsupervised segmentation based cup-to-disc ratio computation in retinal fundus images," *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, vol. 77, p. 101643, 2019. xi, 2

[11] J. Pruthi, K. Khanna, and S. Arora, "Optic cup segmentation from retinal fundus images using glowworm swarm optimization for glaucoma detection," *Biomed. Signal Process. Control.*, vol. 60, p. 102004, 2020. xi, 2

[12] Q. Abbas, I. Qureshi, and M. Ibrahim, "An automatic detection and classification system of five stages for hypertensive retinopathy using semantic and instance segmentation in densenet architecture," *Sensors*, vol. 21, 10 2021. xi, 2

[13] U. Schmidt-Erfurth, A. Sadeghipour, B. Gerendas, S. Waldstein, and H. Bogunović, "Artificial intelligence in retina," *Progress in retinal and eye research*, vol. 67, pp. 1–29, 2018. xii, 2

[14] A. K. Shukla, R. K. Pandey, and R. B. Pachori, "A fractional filter based efficient algorithm for retinal blood vessel segmentation," *Biomedical Signal Processing and Control*, vol. 59, 5 2020. xii, 2, 31, 36, 96

[15] S. Karkuzhali and D. Manimegalai, "Robust intensity variation and inverse surface adaptive thresholding techniques for detection of optic disc and exudates in retinal fundus images," *Biocybernetics and Biomedical Engineering*, vol. 39, pp. 753–764, 2019. xii, 2, 16, 20, 58, 59

[16] Y. LeCun, P. Haffner, L. Bottou, and Y. Bengio, *Object Recognition with Gradient-Based Learning*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 319–345. xii, 2, 18

[17] C. Kromm and K. Rohr, "Inception capsule network for retinal blood vessel segmentation and centerline extraction," 04 2020, pp. 1223–1226. xii, 2, 34, 36, 89

[18] Y. Peng, S. Dharssi, Q. Chen, T. Keenan, E. Agrón, W. Wong, E. Chew, and Z. Lu, "Deepseenet: A deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs," *Ophthalmology*, vol. 126 4, pp. 565–575, 2019. xii, 2, 11, 12

[19] G. Litjens, T. Kooi, B. Ehteshami Bejnordi, A. Setio, F. Ciompi, M. Ghafoorian, J. van der Laak, B. Ginneken, and C. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, 02 2017. xii, 2, 15

[20] J. Du, B. Zou, C. Chen, Z. Xu, and Q. Liu, "Automatic microaneurysm detection in fundus image based on local cross-section transformation and multi-feature fusion," *Computer Methods and Programs in Biomedicine*, vol. 196, p. 105687, 2020. xii, 2, 41, 111

[21] R. Zheng, L. Liu, S. Zhang, C. Zheng, F. Bunyak, R. Xu, B. Li, and M. Sun, "Detection of exudates in fundus photographs with imbalanced learning using conditional generative adversarial network." *Biomedical optics express*, vol. 9 10, pp. 4863–4878, 2018. xii, 2, 143

[22] B. Wu, W. Zhu, F. Shi, S. Zhu, and X. Chen, "Automatic detection of microaneurysms in retinal fundus images," *Computerized Medical Imaging and Graphics*, 2017. xii, 2, 39, 40, 41, 103, 111, 143, 144

[23] W. Zhou, C. Wu, D. Chen, Y. Yi, and W. Du, "Automatic Microaneurysm Detection Using the Sparse Principal Component Analysis-Based Unsupervised Classification Method," *IEEE Access*, 2017. xii, 2, 38

[24] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images," *Computer Methods and Programs in Biomedicine*, 2018. xii, xvii, xxix, 2, 40, 41, 42, 112, 113, 114, 116

[25] E. Pead, R. Megaw, J. Cameron, A. Fleming, B. Dhillon, E. Trucco, and T. MacGillivray, "Automated detection of age-related macular degeneration in color fundus photography: a systematic review," *Survey of Ophthalmology*, vol. 64, pp. 498 – 511, 2019. xii, 2, 10, 11, 13

[26] N. Eftekhari, H. R. Pourreza, M. Masoudi, K. Ghiasi-Shirazi, and E. Saeedi, "Microaneurysm detection in fundus images using a two-step convolutional neural network," *BioMedical Engineering Online*, 2019. xii, 2, 37, 38, 40, 41, 101, 111

[27] J. Orlando, E. Prokofyeva, and M. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," *IEEE Transactions on Biomedical Engineering*, vol. 64, 03 2016. xii, xiii, 32, 36, 96

[28] A. Witkin, "Scale-space filtering," *IJCAI'83*, vol. 2, pp. 1019–1022, 1983. xiii, 3, 143

[29] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales," *Journal of Applied Statistics*, pp. 224–270, 1994. xiii, 3, 105

[30] Q. Li, J. You, and D. Zhang, "Vessel segmentation and width estimation in retinal images using multiscale production of matched filter responses," *Expert Systems with Applications*, vol. 39, pp. 7600–7610, 7 2012. xiii, 3

[31] B. Dashtbozorg, J. Zhang, and B. ter Haar Romeny, "Retinal microaneurysms detection using local convergence index features," *IEEE Transactions on Image Processing*, vol. PP, 07 2017. xiii, 3, 37, 41, 101, 109, 111, 115, 140

[32] B. Zhang, X. Wu, J. You, Q. Li, and F. Karray, "Detection of microaneurysms using multi-scale correlation coefficients," *Pattern Recognition*, vol. 43, no. 6, pp. 2237 – 2248, 2010. xiii, 3, 29, 37, 38, 41, 111, 113, 115

[33] M. U. Akram, S. Khalid, and S. A. Khan, "Identification and classification of microaneurysms for early detection of diabetic retinopathy," *Pattern Recognition*, 2013. xiii, 3, 115

[34] M. Niemeijer, J. Staal, M. Abramoff, M. Suttorp-Schulten, and B. van Ginneken, "Automatic detection of red lesions in digital color fundus photographs," *IEEE Trans. Med. Imaging*, vol. 24, p. 584–592, 2005. xiii, 3, 37, 41, 111, 113, 115

[35] B. Dai, X. Wu, and W. Bu, "Retinal microaneurysms detection using gradient vector analysis and class imbalance classification," *PLoS ONE*, 2016. xiii, 3, 39, 41, 111, 115

[36] K. M. Adal and *et al.*, "Automated detection of microaneurysms using scale-adapted blob analysis and semi-supervised learning," *Computer Methods and Programs in Biomedicine*, 2014. xiii, 3, 37, 38, 41, 101, 102, 111, 143

[37] K. Li, X. Qi, Y. Luo, Z. Yao, X. Zhou, and M. Sun, "Accurate retinal vessel segmentation in color fundus images via fully attention-based networks," *IEEE journal of biomedical and health informatics*, vol. PP, 10 2020. xiii, 3, 33, 35, 36, 96, 115

[38] C. Duanggate, K. Inthajak, B. Uyyanonvara, S. S. Makhanov, S. A. Barman, and T. H. Williamson, "Automatic optic disc detection for rop images using scale-space theory," *ICE-SIT 2010*, p. 64, 2010. xiii, 3

[39] M. Martinez-Perez, A. Hughes, S. Thom, A. Bharath, and K. Parker, "Segmentation of blood vessels from red-free and fluorescein retinal images." *Med. Image Anal.*, vol. 11, no. 1, pp. 47–61, 2007. xiii, 3, 30, 36, 92, 95

[40] M. Niemeijer, M. D. Abràmoff, and B. van Ginneken, "Fast detection of the optic disc and fovea in color fundus photographs," *Med. Image Anal.*, vol. 13, no. 6, pp. 859 – 870, 2009, includes Special Section on Computational Biomechanics for Medicine. xiii, 3, 23, 78

[41] X. Zhang and G. Fan, "Retinal spot lesion detection using adaptive multiscale morphological processing," vol. 4292, pp. 490–501, 2006. xiii, 3

[42] "Stare project website clemsoa university, clemson,sc[online]," available from: <http://cecas.clemson.edu/∼ahoover/stare/>, Last accessed on 2018/08/20. xiii, xxvii, 10, 13, 21, 28, 87, 95, 139

[43] J. Staal, M. Abramoff, M. Niemeijer, M. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med.*, vol. 23, no. 4, pp. 501–509, 2004. xiii, 10, 21, 28, 74, 87, 95, 139

[44] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "Diaretdb0: Evaluation database and methodology for diabetic retinopathy algorithms," *Technical report*, 2006. xiii, xxvii, 10, 11, 21, 74, 139

[45] ——, "Diaretdb1 diabetic retinopathy database and evaluation protocol," vol. 2007, 01 2007. xiii, 10, 15, 21, 56, 74, 139, 140

[46] E. Decencière and *et al.*, "Feedback on a publicly distributed database: the messidor database," *Image Analysis & Stereology*, vol. 33, no. 3, pp. 231–234, Aug. 2014. xiii, 10, 21, 41, 108, 109, 140

[47] G. Quellec, M. Lamard, P. M. Josselin, G. Cazuguel, B. Cochener, and C. Roux, "Optimal wavelet transform for the detection of microaneurysms in retina photographs," *IEEE Transactions on Medical Imaging*, vol. 27, pp. 1230–1241, 2008. xiii, 10, 37, 41, 102, 108, 111, 140

[48] M. Niemeijer and *et al.*, "Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs," *IEEE Trans. Med.*, vol. 29, no. 1, pp. 185–195, Jan 2010. xiii, xiv, 10, 21, 37, 40, 74, 101, 102, 108, 109, 140

[49] E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Laÿ, and A. Chabouis, "Teleophta: Machine learning and image processing methods for teleophthalmology," *Irbm*, vol. 34, pp. 196–203, 2013. xiii, 10, 21, 37, 74, 102, 108, 141

[50] J. Odstrcilík, R. Kolár, A. Budai, J. Hornegger, J. Jan, J. Gazárek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database," *IET Image Process.*, vol. 7, pp. 373–383, 2013. xiii, xxvii, 10, 21, 28, 29, 35, 36, 74, 141

[51] C. Huang, Y. Zong, Y. Ding, X. Luo, K. Clawson, and Y. Peng, "A new deep learning approach for the retinal hard exudates detection based on superpixel multi-feature extraction and patch-based cnn," *Neurocomputing*, 2020. xiii, 18, 19, 20, 59

[52] K. Wisaeng and W. Sa-ngiamvibool, "Exudates detection using morphology mean shift algorithm in retinal images," *IEEE Access*, vol. 7, pp. 11 946–11 958, 2019. xiii, 16, 19, 20, 59, 143

[53] I. Soares, M. Castelo-Branco, and A. M. G. Pinheiro, "Scale-space curvature detection of retinal exudates with a dynamic threshold," *ISPA'11*, pp. 523–528, 2011. xiv, 3, 15, 43

[54] I. Soares, M. Castelo-Branco, and A. Pinheiro, "Curvature detection and segmentation of retinal exudates," in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, May 2012, pp. 1719–1722. xiv, 3, 43

[55] I. Soares, M. Castelo-Branco, and A. M. G. Pinheiro, "Optic disc localization in retinal images based on cumulative sum fields," *IEEE J. Biomed. Health. Inf.*, vol. 20, no. 2, pp. 574–585, March 2016. xv, 4, 12, 20, 63, 77, 91, 103, 104, 143

[56] I. Soares, M. Castelo-Branco, and A. Pinheiro, "Vessel centerline detection in retinal images based on a corner detector and dynamic thresholding," in *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, Sept 2014, pp. 2020–2024. xv, 4, 81

[57] I. Soares, M. Castelo-Branco, and A. M. G. Pinheiro, "A New Vessel Enhancement Transform on Retinal Blood Vessels Segmentation," in *Eurographics Workshop on Visual Computing for Biology and Medicine*, A. Puig Puig, T. Schultz, A. Vilanova, I. Hotz, B. Kozlikova, and P.-P. Vázquez, Eds. The Eurographics Association, 2018. xv, 4, 13, 34, 81, 102

[58] A. Mahfouz and A. Fahmy, "Fast localization of the optic disc using projection of image features," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3285–3289, 2010. xvi, 21, 22, 24, 25, 26, 67, 71, 76, 77, 79, 115

[59] S. Lu and J.-H. Lim, "Automatic optic disc detection from retinal images by a line operator," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 88–94, 2011. xvi, 21, 25, 26, 77, 78, 79, 115

[60] S. A. Ramakanth and R. V. Babu, "Approximate nearest neighbour field based optic disk detection," *Comput Med Imaging Graph*, vol. 38, no. 1, pp. 49 – 56, 2014. xvi, 23, 25, 26, 76, 77, 79, 115

[61] J. Dietter, W. Haq, I. Ivanov, L. A. Norrenberg, M. Völker, M. Dynowski, D. Röck, F. Ziemssen, M. Leitritz, and M. Ueffing, "Optic disc detection in the presence of strong technical artifacts," *Biomed. Signal Process. Control.*, vol. 53, 2019. xvi, 22, 25, 26, 76, 77, 79, 115

[62] R. Panda, N. Puhan, and G. Panda, "Robust and accurate optic disk localization using vessel symmetry line measure in fundus images," *Biocybernetics and Biomedical Engineering*, vol. 37, pp. 466–476, 2017. xvi, 20, 22, 25, 26, 27, 77, 79, 115

[63] T. Melo, A. M. Mendonça, and A. Campilho, "Microaneurysm detection in color eye fundus images for diabetic retinopathy screening," *Computers in Biology and Medicine*, vol. 126, p. 103995, 2020. xvii, 2, 39, 40, 41, 111, 116

[64] A.-H. Abdel-Razik Youssif, A. Ghalwash, and A. Abdel-Rahman Ghoneim, "Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter," *IEEE Trans. Med.*, vol. 27, no. 1, pp. 11–18, 2008. xvii, 22, 25, 26, 63, 67, 74, 76, 77, 79, 116

[65] F. ter Haar, "Automatic localization of the optic disc in digital colour images of the human retina," *M.S. thesis, Utrecht University, Utrecht, The Netherlands*, 2005. xvii, 64, 79, 89, 102, 116

[66] M. Fraz, M. Badar, A. Malik, and S. Barman, "Computational methods for exudates detection and macular edema estimation in retinal images: A survey," *Archives of Computational Methods in Engineering*, vol. 26, 09 2018. xvii, 15, 19, 116

[67] M. Esmaeili, H. Rabbani, A. Dehnavi, and A. Dehghani, "Automatic detection of exudates and optic disk in retinal images using curvelet transform," *Iet Image Processing*, vol. 6, pp. 1005–1013, 2012. xvii, 116

[68] I. Lazar and A. Hajdu, "Retinal microaneurysm detection through local rotating cross-section profile analysis," *IEEE Trans. Med.*, vol. 32, no. 2, pp. 400–407, Feb 2013. xvii, 38, 39, 40, 41, 101, 103, 111, 116, 143

[69] L. Giancardo, F. Meriaudeau, T. Karnowski, Y. Li, S. Garg, K. Jr, and E. Chaum, "Automatic diabetic macular edema detection in fundus images using publicly available datasets," *Medical Image Analysis*, 01 2011. xvii, 116

[70] S. Hauser, *Harrison's Neurology in Clinical Medicine*, 4th ed. Mcgraw-Hill Education, 2016. xxvii, 8

[71] M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Reviews in Biomedical Engineering*, vol. 3, 2010. xxvii, 7, 8, 9, 10, 11, 12, 13

[72] M. H. Sarhan, S. Albarqouni, M. Yigitsoy, N. Navab, and A. Eslami, "Multi-scale microaneurysms segmentation using embedding triplet loss," pp. 174–182, 2019. xxvii, 11, 12

[73] R. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering." Springer-Verlag, 1998, pp. 130–137. xxviii, 30, 31, 64, 65, 66

[74] L. Giancardo, F. Meriaudeau, T. P. Karnowski, K. W. Tobin, and E. Chaum, "Validation of microaneurysm-based diabetic retinopathy screening across retina fundus datasets," in *Proceedings of CBMS 2013 - 26th IEEE International Symposium on Computer-Based Medical Systems*, 2013. xxix, 42, 113

[75] C. I. Sánchez and *et al.*, "Evaluation of a Computer-Aided Diagnosis System for Diabetic Retinopathy Screening on Public Data," *Investigative Ophthalmology & Visual Science*, vol. 52, no. 7, pp. 4866–4871, 06 2011. xxix, xxxi, 19, 40, 41, 42, 109, 112, 113, 140

[76] X. Ren, Y. Zheng, Y. Zhao, C. Luo, H. Wang, J. Lian, and Y. He, "Drusen segmentation from retinal images via supervised feature learning," *IEEE Access*, vol. PP, pp. 1–1, 12 2017. 2

[77] I. Soares, M. Castelo-Branco, and A. M. G. Pinheiro, "Exudates dynamic detection in retinal fundus images based on the noise map distribution," in *2011 19th European Signal Processing Conference*, 2011, pp. 46–50. 4, 43

[78] I. Soares, M. Castelo-Branco, and A. Pinheiro, "Microaneurysms detection using a novel neighborhood analysis," in *Proceedings of the Ophthalmic Med. Image Anal. First International Workshop, OMIA 2014, Held in Conjunction with MICCAI 2014*, Sept 2014, pp. 65–72. 4, 93, 102, 104, 105

[79] R. Besenczi, J. Tóth, and A. Hajdu, "A review on automatic analysis techniques for color fundus photographs," *Computational and Structural Biotechnology Journal*, vol. 14, pp. 371–384, 2016. 10

[80] S. P. Yoon, D. Grewal, A. C. Thompson, B. W. Polascik, C. Dunn, J. Burke, and S. Fekrat, "Retinal microvascular and neurodegenerative changes in alzheimer's disease and mild cognitive impairment compared with control participants." *Ophthalmology. Retina*, vol. 3 6, pp. 489–499, 2019. 10

[81] C. Czakó, T. Kovács, Z. Ungvari, A. Csiszár, A. Yabluchanskiy, S. M. Conley, T. Csipo, A. Lipecz, H. Horváth, G. Sándor, L. István, T. Logan, Z. Nagy, and I. Kovács, "Retinal biomarkers for alzheimer's disease and vascular cognitive impairment and dementia (vcid): implication for early diagnosis and prognosis," *GeroScience*, vol. 42, pp. 1499 – 1525, 2020. 10

[82] C. Cheung, G. Thomas, W. Tay, M. Ikram, W. Hsu, M. Lee, Q. Lau, and T. Wong, "Retinal vascular fractal dimension and its relationship with cardiovascular and ocular risk factors." *American journal of ophthalmology*, vol. 154 4, pp. 663–674.e1, 2012. 11, 14

[83] N. Salamat, M. S. Missen, and A. Rashid, "Diabetic retinopathy techniques in retinal images: A review," *Artificial intelligence in medicine*, vol. 97, pp. 168–188, 2019. 11

[84] V. Gulshan and *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA - Journal of the American Medical Association*, 2016. 11

[85] N. Thakur and M. Juneja, "Survey on segmentation and classification approaches of optic cup and optic disc for diagnosis of glaucoma," *Biomed. Signal Process. Control.*, vol. 42, pp. 162–189, 2018. 11

[86] E. Duh, J. Sun, and A. W. Stitt, "Diabetic retinopathy: current understanding, mechanisms, and treatment strategies." *JCI insight*, vol. 2 14, 2017. 11

[87] E. Imani, H. Pourreza, and T. Banaee, "Fully automated diabetic retinopathy screening using morphological component analysis," *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, vol. 43, pp. 78–88, 2015. 11, 30

[88] J. Mo, L. Zhang, and Y. Feng, "Exudate-based diabetic macular edema recognition in retinal images using cascaded deep residual networks," *Neurocomputing*, vol. 290, pp. 161–171, 2018. 11

[89] L. Giancardo, F. Mériaudeau, T. Karnowski, Y. Li, S. Garg, K. Tobin, and E. Chaum, "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets," *Medical image analysis*, vol. 16 1, pp. 216–26, 2012. 11, 18

[90] B. Fenner, R. L. M. Wong, W. Lam, G. Tan, and G. Cheung, "Advances in retinal imaging and applications in diabetic retinopathy screening: A review," *Ophthalmology and Therapy*, vol. 7, pp. 333 – 346, 2018. 12

[91] H. Quigley and A. Broman, "The number of people with glaucoma worldwide in 2010 and 2020," *British Journal of Ophthalmology*, vol. 90, pp. 262 – 267, 2006. 12

[92] W. Wong, X. Su, X. Li, C. Cheung, R. Klein, C. Cheng, and T. Wong, "Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis." *The Lancet. Global health*, vol. 2 2, pp. e106–16, 2014. 12

[93] G. Dai, W. He, L. Xu, E. E. Pazo, T. Lin, S. sha Liu, and C. Zhang, "Exploring the effect of hypertension on retinal microvasculature using deep learning on east asian population," *PLoS ONE*, vol. 15, 2020. 13

[94] S. Lemmens, A. Devulder, K. V. Keer, J. Bierkens, P. D. Boever, and I. Stalmans, "Systematic review on fractal dimension of the retinal vasculature in neurodegeneration and stroke: Assessment of a potential biomarker," *Frontiers in Neuroscience*, vol. 14, 2020. 14

[95] K. Kipli, M. E. Hoque, L. Lim, M. Mahmood, S. K. Sahari, R. Sapawi, N. Rajaee, and A. Joseph, "A review on the extraction of quantitative retinal microvascular image feature," *Computational and Mathematical Methods in Medicine*, vol. 2018, 2018. 14

[96] K. B. Khan, A. A. Khaliq, A. Jalil, M. A. Iftikhar, N. Ullah, M. W. Aziz, K. Ullah, and M. Shahid, "A review of retinal blood vessels extraction techniques: challenges, taxonomy, and future trends," *Pattern Analysis and Applications*, vol. 22, pp. 767–802, 8 2019. 14, 29, 31

[97] R. Winder, P. Morrow, I. McRitchie, J. Bailie, and P. Hart, "Algorithms for digital image processing in diabetic retinopathy," *Comput Med Imaging Graph*, vol. 33, no. 8, pp. 608 – 622, 2009. 15, 82

[98] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics Gems IV*, pp. 474–485, 12 1994. 15

[99] I. Figueiredo, S. Kumar, C. M. Oliveira, J. D. Ramos, and B. Engquist, "Automated lesion detectors in retinal fundus images," *Computers in biology and medicine*, vol. 66, pp. 47–65, 2015. 15

[100] S. Banerjee and D. Kayal, "Detection of hard exudates using mean shift and normalized cut method," *Biocybernetics and Biomedical Engineering*, vol. 36, pp. 679–685, 2016. 15, 16

[101] T. Walter, J.-C. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina," *IEEE Trans. Med.*, vol. 21, no. 10, pp. 1236–1243, 2002. 15, 21

[102] G. Matheron, *Random Sets and Integral Geometry*. New York: Wiley, 1975. 15, 21

[103] S. Beucher and F. Meyer, *The Morphological Approach to Segmentation: The Watershed Transformation*, 01 1993, vol. Vol. 34, p. 433–481. 15

[104] P. Soille, *Morphological Image Analysis-Principles and Applications*, 01 2003, vol. 49. 15, 16, 17, 22, 24, 29, 34, 37

[105] D. Welfer, J. Scharcanski, C. M. Kitamura, M. M. D. Pizzol, L. W. B. Ludwig, and D. R. Marinho, "Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach." *Comp. in Bio. and Med.*, vol. 40, no. 2, pp. 124–137, 2010. 16, 20, 22, 25, 77

[106] B. Jähne, H. Haußecker, and P. Geißler, *Handbook of Computer Vision and Applications. Volume 2: Signal Processing and Pattern Recognition*, 01 1999. 16

[107] E. Imani and H. Pourreza, "A novel method for retinal exudate segmentation using signal separation algorithm," *Computer methods and programs in biomedicine*, vol. 133, pp. 195–205, 2016. 16, 18

[108] J.-L. Starck, Y. A, J. Bobin, M. Elad, and D. Donoho, "Morphological component analysis," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 5914, 08 2005. 16

[109] J. Kaur and D. Mittal, "A generalized method for the segmentation of exudates from pathological retinal fundus images," *Biocybernetics and Biomedical Engineering*, vol. 38, pp. 27–53, 2018. 16, 19, 20, 59

[110] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. USA: Kluwer Academic Publishers, 1981. 16

[111] R. Gonzalez and R. Woods, *Digital Image Processing*, 01 2006. 16

[112] N. Otsu, "A threshold selection method from gray-level histograms," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 9, pp. 62–66, 01 1979. 16

[113] H. Yazid, H. Arof, and H. Isa, "Exudates segmentation using inverse surface adaptive thresholding," *Measurement*, vol. 45, pp. 1599–1608, 2012. 16

[114] J. Solomon, K. Crane, A. Butscher, and C. Wojtan, "A general framework for bilateral and mean shift filtering," 04 2014. 16

[115] S. Ali, D. Sidibé, K. Adal, L. Giancardo, E. Chaum, T. Karnowski, and F. Mériaudeau, "Statistical atlas based exudate segmentation," *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, vol. 37 5-6, pp. 358–68, 2013. 16

[116] R. Kirsch, "Computer determination of the constituent structure of biologic images," *Computers and Biomedical Research*, vol. 4, pp. 315–328, 06 1971. 16, 18, 38

[117] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, 05 2002. 16

[118] G. Joshi, J. Sivaswamy, and S. Krishnadas, "Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment," *IEEE transactions on medical imaging*, vol. 30, pp. 1192–205, 06 2011. 16

[119] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-8, pp. 679 – 698, 12 1986. 16

[120] S. S. Kar and S. P. Maity, "Automatic detection of retinal lesions for screening of diabetic retinopathy," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 3, pp. 608–618, 2018. 16

[121] Z. Wang, W. Pan, N. Cuppens-Boulahia, F. Cuppens, and C. Roux, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Inf. Forensic Secur.*, vol. 13, pp. 600–612, 01 2013. 16

[122] S. Sil Kar and S. Maity, "Blood vessel extraction and optic disc removal using curvelet transform and kernel fuzzy c-means," *Computers in Biology and Medicine*, vol. 70, 01 2016. 16

[123] C. Cortes and V. Vapnik, "Support vector network," *Machine Learning*, vol. 20, pp. 273–297, 09 1995. 16

[124] I. Rish, "An empirical study of the naïve bayes classifier," *IJCAI 2001 Work Empir Methods Artif Intell*, vol. 3, 01 2001. 16

[125] R. Quinlan, "Simplifying decision trees," *International Journal of Man-Machine Studies*, vol. 27, pp. 221–234, 09 1987. 16

[126] L. Breiman, "Random Forests." *Machine Learning*, vol. 45, pp. 5–32, Jan. 2001. 16

[127] D. Optiz and R. MacLin, "Popular ensemble methods: An empirical study," *Journal of Artificial Intelligence Research*, vol. 11, pp. 169–198, 01 1999. 16

[128] N. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *American Statistician - AMER STATIST*, vol. 46, pp. 175–185, 08 1992. 16, 23

[129] A. Jain, J. Mao, and K. Mohiuddin, "Artificial neural networks: a tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, March 1996. 16

[130] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943. 16

[131] A. Ramakrishnan, S. Raja, and H. Ram, "Neural network-based segmentation of textures using gabor features," 02 2002, pp. 365 – 374. 16

[132] M. Akram, A. Tariq, M. Anjum, and M. Javed, "Automated detection of exudates in colored retinal images for diagnosis of diabetic retinopathy." *Applied optics*, vol. 51 20, pp. 4858–66, 2012. 16, 20

[133] M. Akram, S. Khalid, A. Tariq, S. Khan, and F. Azam, "Detection and classification of retinal lesions for grading of diabetic retinopathy," *Computers in biology and medicine*, vol. 45, pp. 161–71, 2014. 17

[134] T. Lee, "Image representation using 2d gabor wavelet," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 18, pp. 959 – 971, 11 1996. 17, 30, 32

[135] L. Breiman, "Bagging predictors" machine learning," *Machine learning*, vol. 24, 01 1996. 17

[136] M. M. Fraz, W. Jahangir, S. Zahid, M. M. Hamayun, and S. A. Barman, "Multiscale segmentation of exudates in retinal images using contextual cues and ensemble classification," *Biomedical Signal Processing and Control*, vol. 35, pp. 50–62, 2017. 17, 20

[137] J. Serra, *Image Analysis and Mathematical Morphology*. Academic Press, 1982. 17, 34

[138] X. Zhang, G. Thibault, E. Decencière, B. Marcotegui, B. Laÿ, R. Danno, G. Cazuguel, G. Quellec, M. Lamard, P. Massin, A. Chabouis, Z. Victor, and A. Erginay, "Exudate detection in color retinal images for mass screening of diabetic retinopathy," *Medical image analysis*, vol. 18 7, pp. 1026–43, 2014. 17, 19

[139] R. Kindermann and J. Snell, *Markov Random Fields and Their Applications*. Rhode Island, USA: American Mathematical Society, 01 1980. 17

[140] B. Harangi and A. Hajdu, "Automatic exudate detection by fusing multiple active contours and regionwise classification," *Computers in biology and medicine*, vol. 54, pp. 156–71, 2014. 17, 20

[141] Q. Liu, B. Zou, J. Chen, W. Ke, K. Yue, Z. Chen, and G. Zhao, "A location-to-segmentation strategy for automatic exudate segmentation in colour retinal fundus images," *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, vol. 55, pp. 78–86, 2017. 17, 20

[142] W. Kusakunniran, Q. Wu, P. Ritthipravat, and J. Zhang, "Hard exudates segmentation based on learned initial seeds and iterative graph cut," *Computer methods and programs in biomedicine*, vol. 158, pp. 173–183, 2018. 17, 20

[143] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, 2010. 17

[144] D. Greig, B. Porteous, and A. Seheult, "Exact maximum a posteriori estimation for binary images," *Journal of the Royal Statistical Society, Series B*, vol. 51, p. 271–279, 01 1989. 17

[145] S. Long, X. Huang, Z. Chen, S. Pardhan, and D. Zheng, "Automatic detection of hard exudates in color retinal images using dynamic threshold and svm classification: Algorithm development and evaluation," *BioMed Research International*, 2019. 17, 20, 59

[146] M. Javidi, A. Harati, and H. Pourreza, "Retinal image assessment using bi-level adaptive morphological component analysis," *Artificial intelligence in medicine*, vol. 99, p. 101702, 2019. 18, 20, 59

[147] S. Chen, S. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *International Journal of Control*, vol. 50, 05 2007. 18

[148] P. Prentasic and S. Lonarić, "Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion," *Computer methods and programs in biomedicine*, vol. 137, pp. 281–292, 2016. 18

[149] J. Bridle, *Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition*, 01 1990, pp. 227–236. 18

[150] P. Khojasteh, L. A. P. Junior, T. J. Carvalho, E. R. S. Rezende, B. Aliahmad, J. Papa, and D. Kumar, "Exudate detection in fundus images using deeply-learnable features," *Computers in biology and medicine*, vol. 104, pp. 62–69, 2019. 18, 20, 59

[151] A. Levinshtein, A. Stere, K. Kutulakos, D. Fleet, S. Dickinson, and K. Siddiqi, "Turbopixels: Fast superpixels using geometric flows," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, pp. 2290–7, 12 2009. 18

[152] A. Osareh, B. Shadgar, and R. Markham, "A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, pp. 535–545, 2009. 18

[153] C. Bishop, *Neural Networks For Pattern Recognition*, 01 2005, vol. 227. 18

[154] Y. Meyer, *Wavelets and Operators*, ser. Cambridge Studies in Advanced Mathematics, D. H. Salinger, Ed. Cambridge University Press, 1993, vol. 1. 18, 23

[155] C. Sánchez, R. Hornero, M. López, M. Aboy, J. Poza, and D. Abásolo, "A novel automatic image processing algorithm for detection of hard exudates based on retinal image analysis." *Medical Engineering & Physics*, vol. 30, no. 3, pp. 350–357, 2008. 18

[156] J. Cohen, P. Cohen, S. West, and L. Aiken, *Applied Multiple Regression/Correlation Analysis For The Behavioral Sciences*, 01 2003, vol. 52. 19

[157] C. Agurto, V. Murray, H. Yu, J. Wigdahl, M. Pattichis, S. Nemeth, E. S. Barriga, and P. Soliz, "A multiscale optimization approach to detect exudates in the macula," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, pp. 1328–1336, 2014. 19, 20

[158] H. Wold, "Estimation of principal components and related models by iterative least squares," *Journal of Multivariate Analysis - MA*, vol. 1, 01 1966. 19

[159] C. Pereira, L. Gonçalves, and M. Ferreira, "Exudate segmentation in fundus images using an ant colony optimization approach," *Inf. Sci.*, vol. 296, pp. 14–24, 2015. 19

[160] M. Dorigo, "Optimization, learning and natural algorithms, phd thesis," Ph.D. dissertation, Politecnico di Milano, Italy, 1992. 19

[161] L. Giancardo, F. Meriaudeau, T. Karnowski, Y. Li, K. Tobin, and E. Chaum, "Automatic retina exudates segmentation without a manually labelled training set." *ISBI'11*, pp. 1396–1400, 2011. 19

[162] M. Foracchia, E. Grisan, and A. Ruggeri, "Detection of optic disc in retinal images by means of a geometrical model of vessel structure," *IEEE Trans. Med.*, vol. 23, no. 10, pp. 1189–1195, 2004. 20, 22, 24, 26, 50, 63, 74

[163] A. Giachetti, L. Ballerini, and E. Trucco, "Accurate and reliable segmentation of the optic disc in digital fundus images," *Journal of Medical Imaging*, vol. 1, no. 2, p. 024001, Jul. 2014. 20, 23, 26, 77

[164] L. J. Uribe-Valencia and J. Martínez-Carballido, "Automated optic disc region location from fundus images: Using local multi-level thresholding, best channel selection, and an intensity profile model," *Biomed. Signal Process. Control.*, vol. 51, pp. 148–161, 2019. 21, 23, 25, 26, 27

[165] C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson, "Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images," *Br J Ophthalmol*, vol. 83, pp. 902–910, 1999. 21

[166] A. Oppenheim and R. Schafer, "From frequency to quefrency: A history of the cepstrum," *Signal Processing Magazine, IEEE*, vol. 21, pp. 95 – 106, 10 2004. 21

[167] S. Lu, "Accurate and efficient optic disc detection and segmentation by a circular transformation," *IEEE Trans. Med.*, vol. 30, no. 12, pp. 2126–2133, 2011. 21, 25, 26, 74, 76, 77, 103

[168] M. Reza, "Automatic detection of optic disc in color fundus retinal images using circle operator," *Biomed. Signal Process. Control.*, vol. 45, pp. 274–283, 2018. 21, 25

[169] A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *IEEE Trans. Med.*, vol. 22, no. 8, pp. 951–958, 2003. 21, 24, 25, 26, 74, 91

[170] "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification." vol. 25, no. 9, pp. 1214–1222, 2006. 22

[171] D. Zhang and Y. Zhao, "Novel accurate and fast optic disc detection in retinal images with vessel distribution and directional characteristics," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 1, pp. 333–342, 2016. 22, 25

[172] R. Duda and P. Hart, "Use of the hough transform to detect lines and curves in pictures," *Communications of The ACM - CACM*, vol. 15, 01 1975. 22

[173] M. Lalonde, M. Beaulieu, and L. Gagnon, "Fast and robust optic disc detection using pyramidal decomposition and hausdorff-based template matching," *IEEE Trans. Med.*, vol. 20, no. 11, pp. 1193–1200, 2001. 22, 78

[174] E. Adelson, C. Anderson, J. Bergen, P. Burt, and J. Ogden, "Pyramid methods in image processing," *RCA Eng.*, vol. 29, 11 1983. 22

[175] D. Huttenlocher, G. Klanderman, and W. Rucklidge, "Comparing images using the hausdorff distance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, pp. 850–863, 10 1993. 22

[176] H. Voorhees and T. Poggio, "Detecting textons and texture boundaries in natural images," pp. 250–258, 01 1987. 22

[177] D. Wu, M. Zhang, J.-C. Liu, and W. Bauman, "On the adaptive detection of blood vessels in retinal images," *IEEE transactions on bio-medical engineering*, vol. 53, pp. 341–3, 03 2006. 22

[178] H. Yu, E. Barriga, C. Agurto, S. Echegaray, M. Pattichis, W. Bauman, and P. Soliz, "Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets," *IEEE Trans. Inf. Technol.*, vol. 16, no. 4, pp. 644–657, 2012. 23, 26, 63

[179] S. A. Ramakanth and R. V. Babu, "Featurematch: A general ANNF estimation technique and its applications," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 2193–2205, May 2014. 23

[180] A. Criminisi, P. Pérez, and K. Toyama, "Object removal by exemplar-based inpainting." vol. 2, 01 2003, pp. 721–728. 23

[181] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 959–973, 2003. 23

[182] R. Kamble, M. Kokare, G. Deshmukh, F. A. Hussin, and F. Mériaudeau, "Localization of optic disc and fovea in retinal images using intensity based line scanning analysis," *Computers in biology and medicine*, vol. 87, pp. 382–396, 2017. 23, 25, 26

[183] B. Al-Bander, W. Al-Nuaimy, B. M. Williams, and Y. Zheng, "Multiscale sequential convolutional neural networks for simultaneous detection of fovea and optic disc," *Biomed. Signal Process. Control.*, vol. 40, pp. 91–101, 2018. 23, 24, 26

[184] S. Roychowdhury, D. D. Koozekanani, S. N. Kuchinka, and K. K. Parhi, "Optic disc boundary and vessel origin segmentation of fundus images," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, pp. 1562–1574, 11 2016. 24, 25, 26, 27, 76, 77, 78

[185] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2538–2548, Sept 2012. 24, 33, 36

[186] M. I. Meyer, A. Galdran, A. M. Mendonça, and A. Campilho, "A pixel-wise distance regression approach for joint retinal optical disc and fovea detection," in *MICCAI*, 2018. 24, 26

[187] M. N. Bajwa, M. I. Malik, S. A. Siddiqui, A. Dengel, F. Shafait, W. Neumeier, and S. Ahmed, "Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning," *BMC Medical Informatics and Decision Making*, vol. 19, 7 2019. 24, 25, 26, 27, 77

[188] "Idrid: Diabetic retinopathy – segmentation and grading challenge," *Medical Image Analysis*, vol. 59, p. 101561, 2020. 24

[189] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Med. Image Anal.*, vol. 9, pp. 179–190, 2005. 25, 49

[190] X. Xu, M. Niemeijer, Q. Song, M. Sonka, M. Garvin, J. Reinhardt, and M. Abramoff, "Vessel boundary delineation on fundus images using graph-based approach," *IEEE Trans. Med. Imag*, pp. 1184–1191, 2011. 28, 81

[191] A. Mendonça and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," *IEEE Trans. Med.*, vol. 25, no. 9, pp. 1200–1213, 2006. 28, 29, 36, 67, 81, 85, 94

[192] R. Panda, N. Puhan, and G. Panda, "New binary hausdorff symmetry measure based seeded region growing for retinal vessel segmentation," *Biocybern Biomed Eng*, vol. 36, no. 1, pp. 119 – 129, 2016. 28

[193] R. Annunziata, A. Garzelli, L. Ballerini, A. Mecocci, and E. Trucco, "Leveraging multiscale hessian-based enhancement with a novel exudate inpainting technique for retinal vessel segmentation," *IEEE J. Biomed. Health. Inf.*, vol. 20, no. 4, pp. 1129–1138, July 2016. 28, 30, 36, 81, 89, 96, 102

[194] J. Almotiri, K. Elleithy, and A. Elleithy, "Retinal vessels segmentation techniques and algorithms: A survey," *Applied Sciences*, vol. 8, 01 2018. 28, 31

[195] M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. Rudnicka, C. Owen, and S. Barman, "Blood vessel segmentation methodologies in retinal images - a survey," *Computer methods and programs in biomedicine*, vol. 108, pp. 407–33, 04 2012. 28, 29

[196] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med.*, vol. 8, no. 3, pp. 263–269, 1989. 28, 29, 83

[197] B. Zhang, X. Wu, J. You, Q. Li, and F. Karray, "Hierarchical detection of red lesions in retinal images by multiscale correlation," *SPIE Medical Imaging*, vol. 7260, p. 72601L, 2009. 29

[198] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *Med Imaging, IEEE Trans*, vol. 19, no. 3, pp. 203–210, 2000. 29, 87, 143

[199] E. Gumbel, "Les valeurs extrêmes des distributions statistiques," *Ann. Inst. Henri Poincaré*, vol. 5, pp. 115–158, 01 1935. 29

[200] N. P. Singh and R. Srivastava, "Retinal blood vessels segmentation by using gumbel probability distribution function based matched filter," *Computer Methods and Programs in Biomedicine*, vol. 129, pp. 40–50, 6 2016. 29, 36, 96

[201] O. Chutatape, L. Zheng, and S. Krishnan, "Retinal blood vessel detection and tracking by matched gaussian and kalman filters," in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286)*, vol. 6, 1998, pp. 3144–3149 vol.6. 29

[202] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960. 29

[203] F. Zana and J.-C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1010–1019, Jul 2001. 29

[204] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, C. G. Owen, A. R. Rudnicka, and S. A. Barman, "Retinal vessel extraction using first-order derivative of gaussian and morphological processing," in *Advances in Visual Computing*, G. Bebis, R. Boyle, B. Parvin, D. Koracin, S. Wang, K. Kyungnam, B. Benes, K. Moreland, C. Borst, S. DiVerdi, C. Yi-Jen, and J. Ming, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 410–420. 30, 36

[205] M. Sofka and C. Stewart, "Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures," *IEEE Trans. Med.*, vol. 25, no. 12, pp. 1531–1546, 2006. 30, 34, 64, 88

[206] H. Yu, S. Barriga, C. Agurto, G. Zamora, W. Bauman, and P. Soliz, "Fast vessel segmentation in retinal images using multiscale enhancement and second-order local entropy," *Proceedings of SPIE - The International Society for Optical Engineering*, pp. 45–, 02 2012. 30, 96

[207] C.-I. Chang, K. Chen, J. Wang, and M. Althouse, "A relative entropy-based approach to image thresholding," *Pattern Recognition*, vol. 27, pp. 1275–1289, 09 1994. 30

[208] U. Nguyen, A. Bhuiyan, L. Park, and K. Ramamohanarao, "An effective retinal blood vessel segmentation method using multi-scale line detection," *Pattern Recognit.*, vol. 46, pp. 703–715, 2013. 30, 32

[209] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, "Trainable cosfire filters for vessel delineation with application to retinal images," *Med. Image Anal.*, vol. 19, no. 1, pp. 46 – 57, 2015. 30, 36

[210] D. Marin, A. Aquino, M. Gegundez-Arias, and J. Brav, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Trans. Med.*, vol. 30, pp. 146–158, 2011. 31, 81, 89

[211] A. Khawaja, T. M. Khan, M. A. Khan, and S. J. Nawaz, "A multi-scale directional line detector for retinal vessel segmentation," *Sensors (Switzerland)*, vol. 19, 11 2019. 31, 36, 96

[212] J. Yang, M. Huang, J. Fu, C. Lou, and C. Feng, "Frangi based multi-scale level sets for retinal vascular segmentation," *Computer Methods and Programs in Biomedicine*, vol. 197, 12 2020. 31

[213] B. Lam and H. Yan, "A novel vessel segmentation algorithm for pathological retina images based on the divergence of vector fields," *IEEE transactions on medical imaging*, vol. 27, pp. 237–46, 03 2008. 31

[214] B. Lam, Y. Gao, and A. W.-C. Liew, "General retinal vessel segmentation using regularization-based multiconcavity modeling," *Medical Imaging, IEEE Transactions on*, vol. 29, pp. 1369 – 1381, 08 2010. 31

[215] G. Kovács and A. Hajdu, "A self-calibrating approach for the segmentation of retinal vessels by template matching and contour reconstruction," *Med. Image Anal.*, vol. 29, pp. 24 – 46, 2016. 31

[216] L. Wang, H. Zhang, K. He, Y. Chang, and X. Yang, "Active contours driven by multi-feature gaussian distribution fitting energy with application to vessel segmentation," *PloS one*, vol. 10, p. e0143105, 11 2015. 31

[217] K. Pearson, "On lines and planes of closest fit to points in space," *Philosophical Magazine*, vol. 2, pp. 559–572, 11 1900. 31

[218] M. A. Palomera-Pérez, M. E. Martínez-Pérez, H. Benítez-Pérez, and J. L. Ortega-Arjona, "Parallel multiscale feature extraction and region growing: application in retinal blood vessel detection." *IEEE Trans. Information Technology in Biomedicine*, vol. 14, no. 2, pp. 500–506, 2010. 31

[219] H. Jiang, B. He, D. Fang, Z. Ma, B. Yang, and L. Zhang, "A region growing vessel segmentation algorithm based on spectrum information," *Computational and mathematical methods in medicine*, vol. 2013, p. 743870, 11 2013. 32

[220] Y. Zhao, X. Wang, and F. Shih, "Retinal vessels segmentation based on level set and region growing," *Pattern Recognit.*, vol. 47, pp. 2437–2446, 2014. 32, 36

[221] R. Panda, N. Puhan, and G. Panda, "New binary hausdorff symmetry measure based seeded region growing for retinal vessel segmentation," *Biocybernetics and Biomedical Engineering*, vol. 36, no. 1, pp. 119 – 129, 2016. 32, 36, 96

[222] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE Trans Med Imag*, vol. 26, no. 10, pp. 1357–1365, 2007. 32, 36

[223] S. Roychowdhury, D. Koozekanani, and K. Parhi, "Blood vessel segmentation of fundus images by major vessel extraction and sub-image classification," *IEEE journal of biomedical and health informatics*, vol. 19, 07 2014. 32, 36

[224] H. Permuter, J. Francos, and I. Jermyn, "Gaussian mixture models of texture and colour for image database retrieval," vol. 3, 05 2003, pp. III – 569. 32

[225] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Transactions on Medical Imaging*, vol. 35, pp. 1–1, 03 2016. 33

[226] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang, "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning," *Neurocomputing*, vol. 149, pp. 708–717, 02 2015. 33, 36

[227] E. Cheng, L. Du, Y. Wu, Y. Zhu, V. Megalooikonomou, and H. Ling, "Discriminative vessel segmentation in retinal images by fusing context-aware hybrid features," *Machine Vision and Applications*, vol. 25, pp. 1779–1792, 10 2014. 33, 35, 36, 89

[228] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," 07 2010, pp. 2963 – 2970. 33

[229] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, and W. Gao, "Wld: A robust local image descriptor," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, pp. 1705–20, 09 2010. 33

[230] K.-B. Park, S. Choi, and J. Y. Lee, "M-gan: Retinal blood vessel segmentation by balancing losses through stacked deep fully convolutional networks," *IEEE Access*, vol. PP, pp. 1–1, 08 2020. 34, 36, 96

[231] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv:1411.1784*, 11 2014. 34

[232] R. Xu, T. Liu, X. Ye, F. Liu, L. Lin, L. Li, S. Tanaka, and Y. W. Chen, "Joint extraction of retinal vessels and centerlines based on deep semantics and multi-scaled cross-task aggregation," *IEEE Journal of Biomedical and Health Informatics*, 2020. 34, 35, 96

[233] S. Morales, V. Naranjo, J. Angulo, Á. Legaz-Aparicio, and R. Verdú, "Retinal network characterization through fundus image processing: Significant point identification on vessel centerline," *Signal Process. Image Commun.*, vol. 59, pp. 50–64, 2017. 34

[234] C. Szegedy, A. Toshev, and D. Erhan, "Advances in neural information processing systems," *Deep Neural Networks for Object Detection*, pp. 2553–2561, 01 2013. 34

[235] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," 06 2015, pp. 1–9. 34

[236] M. Javidi, H. Pourreza, and A. Harati, "Vessel segmentation and microaneurysm detection using discriminative dictionary learning and sparse representation," *Computer Methods and Programs in Biomedicine*, vol. 139, 10 2016. 37

[237] C. I. Sánchez, R. Hornero, A. Mayo, and M. García, "Mixture model-based clustering and logistic regression for automatic detection of microaneurysms in retinal images," vol. 7260, 2009, pp. 72 601M–72 601M–8. 37, 38, 41, 111

[238] I. Lazar and A. Hajdu, "Microaneurysm detection in retinal images using a rotating cross-section based model," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, 2011, pp. 1405–1409. 37

[239] L. Giancardo, F. Meriaudeau, T. Karnowski, K. Tobin, Y. Li, and E. Chaum, "Microaneurysms detection with the radon cliff operator in retinal fundus images," *SPIE Medical Imaging*, vol. 7623, p. 29, 2010. 37, 38, 41, 111

[240] G. E. Øien and P. Osnes, "Diabetic retinopathy: Automatic detection of early symptoms from retinal images," in *PROC. NORWEGIAN SIGNAL PROCESSING SYM*, 1995. 37

[241] B. Zhang, F. Karray, Q. Li, and L. Zhang, "Sparse representation classifier for microaneurysm detection and retinal blood vessel extraction," *Information Sciences*, vol. 200, no. 0, pp. 78 – 90, 2012. 37, 38, 41, 103, 105, 143

[242] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Applied and Computational Harmonic Analysis*, vol. 3, pp. 186–200, 04 1996. 37

[243] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical recipes in C. The art of scientific computing*, 01 1992, vol. 542. 37

[244] K. Binmore and J. Davies, "Calculus: Concepts and methods," *Cambridge University Press, Cambridge, U.K.*, 2002. 38

[245] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," vol. 3951, 07 2006, pp. 404–417. 38

[246] X. Zhu, "Semi-supervised learning literature survey," *Comput. Sci.*, 01 2005. 38

[247] C. Pereira and *et al.*, "Using a multi-agent system approach for microaneurysm detection in fundus images," *Artificial Intelligence in Medicine*, vol. 60, no. 3, pp. 179–188, 2014. 38, 41, 101, 102, 111

[248] Y. Shoham, R. Powers, and T. Grenager, "If multi-agent learning is the answer, what is the question?" *Artificial Intelligence*, vol. 171, pp. 365–377, 05 2007. 38

[249] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," vol. 1, 07 2005, pp. 886–893. 39

[250] S. Wang, H. Tang, L. Al turk, Y. Hu, S. Sanei, G. Saleh, and T. Petö, "Localising microaneurysms in fundus images through singular spectrum analysis," *IEEE Transactions on Biomedical Engineering*, vol. 64, pp. 1–1, 06 2016. 39, 41, 111

[251] C. Seiffert, T. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "Rusboost: A hybrid approach to alleviating class imbalance," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 40, pp. 185 – 197, 02 2010. 39

[252] U. Budak, A. Sengur, Y. Guo, and Y. Akbulut, "A novel microaneurysms detection approach based on convolutional neural networks with reinforcement sample learning algorithm," *Health Information Science and Systems*, vol. 5, p. 14, 11 2017. 40, 41

[253] P. Chudzik, S. Majumdar, F. Caliva, B. Al-Diri, and A. Hunter, "Microaneurysm detection using fully convolutional neural networks," *Computer Methods and Programs in Biomedicine*, vol. 158, 02 2018. 40, 41, 111

[254] A. Bria, C. Marrocco, and F. Tortorella, "Addressing class imbalance in deep learning for small lesion detection on medical images," *Computers in biology and medicine*, vol. 120, p. 103735, 2020. 40

[255] T. Inoue, Y. Hatanaka, S. Okumura, C. Muramatsu, and H. Fujita, "Automated microaneurysm detection method based on eigenvalue analysis using hessian matrix in retinal fundus images," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, July 2013, pp. 5873–5876. 41

[256] A. Mizutani and *et al.*, "Automated microaneurysm detection method based on double-ring filter in retinal fundus images," in *SPIE Medical Imaging 2009: 105–112, 2002. Computer-Aided Diagnosis*, vol. 7260, no. 1. N. Karssemeijer and M. L. Giger, Eds, 2009. 41, 105, 111

[257] G. B. Kande, P. V. Subbaiah, and T. S. Savithri, "Feature extraction in digital fundus images," *Journal of Medical and Biological Engineering*, vol. 9, no. 3, pp. 122–130, 2009. 43

[258] C. Sanchez, M. Niemeijer, M. Suttorp Schulten, M. Abramoff, and B. Van Ginneken, "Improving hard exudate detection in retinal images through a combination of local and contextual information," *ISBI'10*, pp. 5–8, 2010. 43, 49

[259] J. Sivaswamy, G. Joshi, and S. Chandra, "An alternative curvature measure for topographic feature detection," *ICVGIP'06*, vol. 4338, pp. 228–239, 2006. 44

[260] S. Garg and J. Sivaswamy, "Automatic Drusen Detection from Colour Retinal Images," in *Proc ICMIT'06*, 2006, pp. 84–88. 46

[261] T. Deserno, Ed., *Biological and Medical Physics, Biomedical Engineering*. Springer, 2011. 47, 48, 104, 144, 145

[262] Z. Sbeh and L. Cohen, "A new approach of geodesic reconstruction for drusen segmentation in eye fundus images," *IEEE Trans. on medical imaging*, vol. 20, no. 12, 2001. 51

[263] H. F. Jaafar, A. K. Nandi, and W. Al-Nuaimy, "Automated detection of exudates in retinal images using a split-and-merge algorithm," in *EUSIPCO2010*, 2010, pp. 1622–1626. 52

[264] K. Ram and J. Sivaswamy, "Multi-space clustering for segmentation of exudates in retinal color photographs." *Proc EMBS'09*, vol. 2009, pp. 1437–1440, 2009. 59

[265] H. Li and O. Chutatape, "A model-based approach for automated feature extraction in fundus images," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, 2003, pp. 394–399 vol.1. 63

[266] C. Kondermann, D. Kondermann, and M. Yan, "Blood vessel classification into arteries and veins in retinal images," pp. 651 247–651 247–9, 2007. 64

[267] H. Wang and M. Brady, "Real-time corner detection algorithm for motion estimation," *Image and Vision Computing*, vol. 13, no. 9, pp. 695 – 703, 1995. 64, 78, 82

[268] L. van Vliet and F. G. A. Faas, "Multi-orientation analysis by decomposing the structure tensor and clustering," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 3, 2006, pp. 856–860. 73

[269] M. Breuß, B. Burgeth, and J. Weickert, "Anisotropic continuous-scale morphology," in *Pattern Recognition and Image Analysis*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2007, vol. 4478, pp. 515–522. 73

[270] P. D. Kovesi, "MATLAB and Octave functions for computer vision and image processing," Centre for Exploration Targeting, School of Earth and Environment, The University of Western Australia, 2000, available from: <http://www.csse.uwa.edu.au/∼pk/research/matlabfns/>. 73

[271] H. Farid and E. Simoncelli, "Differentiation of discrete multidimensional signals," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 496–508, 2004. 73

[272] H. E. Bennink and *et al.*, "A novel 3d multi-scale lineness filter for vessel detection," ser. MICCAI'07, 2007, pp. 436–443. 82, 104, 143

[273] L. Lam, S. Lee, and C. Suen, "Thinning methodologies-a comprehensive survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 14, pp. 869–885, 10 1992. 84, 86

[274] T. S. Yoo, G. D. Stetten, and B. Lorensen, *Basic image processing and linear operators, in Insight into Images*. Ed. Wesley, MA. K. Peters, 2004. 85

[275] C. Leahy, A. O'Brien, and C. Dainty, "Illumination correction of retinal images using laplace interpolation," *Appl. Opt.*, vol. 51, no. 35, pp. 8383–8389, Dec 2012. 90

[276] N. Merten, K. Lawonn, P. Gensecke, O. Gro, and B. Preim, "Lung Vessel Enhancement in Low-Dose CT Scans - The LANCELOT Method," in *Bildverarbeitung für die Medizin (BVM)*. Erlangen: Springer Verlag, 2018, pp. 347–352. 92

[277] J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database," *IET Image Processing*, vol. 7, no. 4, pp. 373–383, 2013, available from: <https://www5.cs.fau.de/research/data/fundus-images/>. 95

[278] A. Ehsani Rad, M. Rahim, H. Kolivand, and I. Amin, "Morphological region-based initial contour algorithm for level set methods in image segmentation," *Multimedia Tools and Applications*, vol. 76, 01 2017. 96

[279] S. Nunes, I. Pires, A. Rosa, L. Duarte, R. Bernardes, and J. Cunha-Vaz, "Microaneurysm turnover is a biomarker for diabetic retinopathy progression to clinically significant macular edema: findings for type 2 diabetics with nonproliferative retinopathy," *Ophthalmologica*, vol. 223, no. 5, pp. 292–7, 2009. 101

[280] M. Habib, R. Welikala, A. Hoppe, C. Owen, A. Rudnicka, and S. Barman, "Detection of microaneurysms in retinal images using an ensemble classifier," *Informatics in Medicine Unlocked*, vol. 9, 05 2017. 101

[281] S. A. Ali Shah, A. Laude, I. Faye, and T. B. Tang, "Automated microaneurysm detection in diabetic retinopathy using curvelet transform," *Journal of Biomedical Optics*, vol. 21, no. 10, p. 101404, 2016. 102

[282] Y. Hatanaka, T. Inoue, S. Okumura, C. Muramatsu, and H. Fujita, "Automated microaneurysm detection method based on double-ring filter and feature analysis in retinal fundus images," in *Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on*, June 2012, pp. 1–4. 103, 105, 143

[283] H. Deng, W. Zhang, E. Mortensen, T. Dietterich, and L. Shapiro, "Principal curvature-based region detector for object recognition," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, June 2007, pp. 1–8. 103, 145

[284] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, pp. 465–470, 1996. 103

[285] R. Srivastava, L. Duan, D. W. Wong, J. Liu, and T. Y. Wong, "Detecting retinal microaneurysms and hemorrhages with robustness to the presence of blood vessels," *Computer Methods and Programs in Biomedicine*, 2017. 104

[286] H. Herbert, K. Jordan, and D. Flanagan, "Is screening with digital imaging using one retinal view adequate?" *Eye (London, England)*, vol. 17, pp. 497–500, 05 2003. 112

Appendix A

Datasets

This appendix summarizes the features of the fundus images datasets which are used for the evaluation of the different algorithms of this thesis. All of these datasets were created by several research groups and are publicly available.

A.1 Stare dataset

The Stare (STructured Analysis of the REtina) [42] dataset consists of 31 images of healthy retinas and 50 images of pathological retinas, which is widely used for benchmarking of the optic disc (OD) localization. Images were captured using a TopCon TRV-50 fundus camera at 35^o field of view. The resolution of the images is 700×605 pixels with 8 bits per RGB channel. Furthermore, 20 images were manually annotated for retinal vessels by two observers.

A.2 Drive dataset

The Drive (Digital Retinal Images for Vessel Extraction) [43] dataset is a publicly available dataset for the comparison of vessel segmentation methods in retinal images. It contains 40 images. The images resolution are 565×584 pixels, with 8 bits per colour plane. The image acquisition was made with a Cannon CR5 non-mydriatic CCD camera with a 45^o field of view. The images forming the Drive dataset were randomly selected from 400 diabetic subjects that participated in a diabetic retinopathy screening program. From the 40 selected image, 33 images do not show diabetic signs, whereas 7 images have signs of mild to early diabetic retinopathy. The images have been grouped into a training subset and a test subset with 20 images each. Each image has the manual segmentation of the vasculature structure. The images are manually annotated by two human observers for the retinal vessels.

A.3 Diaretdbv0 dataset

The Diaretdbv0 (Standard Diabetic Retinopathy dataset Calibration level 0) [44] dataset is a public dataset for benchmarking diabetic retinopathy detection from colour fundus images. This dataset consists of 130 colour fundus images of which 20 are normal and 110 contain signs of the diabetic retinopathy (hard exudates, soft exudates, micronaneurysms, haemorrhages and neovascularization). Images were captured with a 50^o field-of-view digital fundus camera with a resolution of 1500×1152 pixels and unknown camera settings. The data correspond to practical situations, and can be used to evaluate the general performance of diagnosis methods. This data set is referred to as ”calibration level 0 fundus images”. Groundtruth images are provided for all images in the dataset, both for hard and soft exudates [45]. These ground truth images are the result of several markings made by experts for each retinal fundus image. The areas are marked with a confidence level that depends on the number of experts that marked it as exudates or not. The ground truth

images of hard and soft exudates are merged, for each confidence level, since the method does not perform any separation between the two exudate types [31].

A.4 Diaretdbv1 dataset

The Diaretdbv1 (Standard Diabetic Retinopathy dataset Calibration Level 1) [45] dataset is a public dataset for benchmarking diabetic retinopathy detection from colour fundus images. It consists of 89 colour fundus images of which 84 contain at least mild nonproliferative signs (Microaneurysms) of DR, and 5 are considered as Normal which do not contain any signs of the DR. Images were captured using the same 50^o field-of-view digital fundus camera at 1500×1152 with varying imaging settings. The data correspond to a good (not necessarily typical) practical situation, where the images are comparable, and can be used to evaluate the general performance of diagnostic methods. This data set is referred to as "calibration level 1 fundus images". Independent markings from 4 medical experts labelled the micronaneurysms, haemorrhages, and hard and soft exudates. The experts were further instructed to report their confidence and especially annotate the single most representative point for each finding. The ground truth confidence levels, <50% ,>50%,100%, represent the certainty of the decision that a labelled finding is correct.

A.5 Messidor dataset

The Messidor (Méthodes d'Evaluation de Systèmes de Segmentation et d'Indexation Dédiées à l'Ophtalmologie Rétinienne) [46] dataset is a large publicly available dataset for evaluating the computer-assisted diagnoses of DR. It contains 1200 eye fundus color images of the posterior pole acquired by the Hôpital Lariboisière Paris, the Faculté de Médecine St. Etienne and the LaTIM-CHU de Brest (France). The images were acquired using a Topcon TRC NW6 non-mydriatic retinograph with a 45^o FOV, with a resolution of 1440×960, 2240×1488, or 2304×1536 pixels stored in TIFF format. A total of 660 images represent subjects with diabetic retinopathy (DR) and macular edema (ME) Each image has a DR grade given by R0, R1, R2 and R3. R0 correspond to the absence of DR. R1 and R2 correspond to mild and severe cases of non-proliferative DR. R3 stands for proliferative DR. The grading depends on number of MAs, haemorrhages and neovascularization [75].

A.6 Latim dataset

The Latim (Laboratoire de Traitement de l'Information Medicale) [47] dataset is composed of 36 colour fundus images compiled by the Brest University Hospital. The retinal images have a resolution of (2240×1488) pixels, stored in a TIFF format. The MAs were annotated by a single DR expert.

A.7 ROC dataset

The ROC (Retinopathy Online Challenge) [48] dataset is formed by 100 colour fundus photographs selected from a larger set of 150000 photographs acquired in DR screening program The inclusion criteria were that the screening program ophthalmologists had marked the image as containing microaneurysms and had not marked it as ungradable. The images were acquired at different

screening sites, with different camera types, resulting in three different types of images with different resolutions (768×576, 1058×1051 and 1389×1383 pixels). The images were captured using either a Topcon NW 100, a Topcon NW 200, or a Canon CR5-45NM. All images are in JPEG format and compression was set in the camera. The 50 images were randomly split into a training and a test subset each containing 50 images. Only the training subset have the correspondent ground truth images annotated by four retinal experts. A region is a true microaneurysm if it is marked by at least two experts.

A.8 E-ophtha-MA dataset

The E-Ophtha-MA [49] is a dataset of colour fundus images especially designed for scientific research in Diabetic Retinopathy (DR). It was generated from a Teleophthalmology network for DR screening, in a project funded by the French Research Agency (ANR). It consists of 381 images for microaneurysms (MAs) detection. Of these, 148 images contain, 1306 MAs and 233 images have no lesions. The MAs were outlined and confirmed by a group of ophthalmologists. All images are in JPEG format with different resolutions (1440×960, 1504×1000, 2504×1696, 2048×1360 pixels).

A.9 E-ophtha-EX dataset

The E-Ophtha-EX dataset [49] was particularly designed for scientific research in DR. It was generated from a Teleophthalmology network for DR screening, in a project funded by the French Research Agency (ANR). It is composed of 47 images with exudates and 35 images with no lesion. The MAs were outlined and confirmed by a group of ophthalmologists. All images are in JPEG format with different resolutions (1440×960, 1504×1000, 2504×1696, 2048×1360 pixels).

A.10 HRF dataset

The HRF (High-Resolution Fundus) [50] dataset contains 15 images of healthy patients, 15 images of patients with diabetic retinopathy (DR) and 15 images of glaucomatous patients. Gold standard vessel segmentation images are available for each image. The field of view (FOV) masks are also provided. The gold standard data was generated by a group of experts working in the field of retinal image analysis and clinicians from the cooperated ophthalmology clinics The image sizes are 3.304×2.336 pixels, with a training/testing image split of 22/23.

Appendix B

Scale-space framework and differential operators definition

The retina is made up of several anatomical structures of different shapes with a wide range of scales. The OD is a bright, large and round structure [55], while the vascular network is made up of linear elements of various sizes with a Gaussian profile [198]. The linearity of the vessel elements depend on the observation scale. This is particularly relevant in the smaller and thinner retinal vessels that are more difficult to detect and extract, thus indicating the need for a variable scale approach [272].
Furthermore, retinal pathological signs like exudates have irregular shapes with rounded edges. In addition, their size ranges from small regions with a Gaussian profile to regions larger than the OD [21, 52]. Finally, MAs are small, round and dark objects with a Gaussian profile [68].
Hence, an effective detection of these anatomical structures and pathological signs require a selective noise reduction process that retains their contrast [282, 22] and variable shapes [241, 68], making them particularly suited for a scale-space approach [5].
In this appendix, the variables and mathematical concepts related to the scale-space definition and other differential operators are explained. Other necessary variables are defined in each chapter as required.

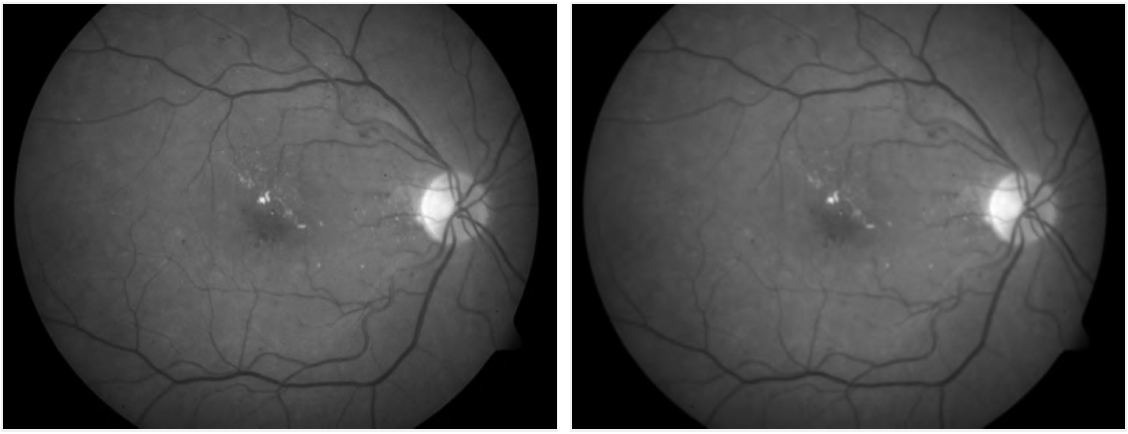
B.1 Scale-space definition

The green channel component $I_G(\mathbf{x})$ (where $\mathbf{x}$ represents the location (x, y)) of each RGB retinal image is selected, since it offers the best contrast and provides the most relevant clinical visual information [4, 36].
Using the definitions of the linear scale-space image representation [28], a family of derived images is defined by the convolution of $I_G(\mathbf{x})$ with the Gaussian filter $g(\mathbf{x}; t)$, given by,

$$L(\mathbf{x}; t) = g(\mathbf{x}) * L(\mathbf{x}; t-1), \forall t \geq 1, \quad \textbf{(B.1)}$$

with $L(\mathbf{x}, 0) = I(\mathbf{x})$, where t represents the scale. $g(\mathbf{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|\mathbf{x}|^2}{\sigma^2}}$. $\sigma^2 = 2t$ is the variance of the Gaussian filter, which defines the scale level. Larger t values, result in smaller bandwidth, leading to stronger smoothing of $I_G(\mathbf{x})$, removing the details that are significantly smaller than t [28]. The retinal features can range from small structures (MAs, exudates in initial stages and thinner vessels) to large structures (broader vessels and larger exudates). Also, a special care is needed for the detection of close features because of the possible feature merging that can occur in larger scales. To preserve the shape and size of these retinal features, a slow evolution is required for the scale-space procedure. The Gaussian mask must be effective in noise reduction and in feature preservation. In this thesis, the following mask was applied,

$$W_G = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$



(a) $L(\mathbf{x};t_1)$, $\sigma_{n=1}=0.84$ (b) $L(\mathbf{x};t_{10})$, $\sigma_{n=10}=2.66$

(c) $L(\mathbf{x};t_{20})$, $\sigma_{n=20}=3.76$ (d) $L(\mathbf{x};t_{30})$, $\sigma_{n=30}=4.61$

Figure B.1: Some examples of the derived fundus images $L(\mathbf{x};t_n)$, at several scales.

$W_G(\mathbf{x})$ corresponds to a Gaussian kernel with a variance $\sigma^2=\sqrt{2}/2$ and size 3×3. Since $\sigma^2=2t$, the parameter t can be defined at N_s different scales according to $t_n=n\frac{\sqrt{2}}{4}, n\in\{0,1,...,N_s-1\}$. The Gaussian scale-space representation is defined as $L(\mathbf{x};t_n)$, $n\in\{0,1,...,N_s-1\}$. Figure B.1 shows some examples of filtered images. As previously defined, $t_0=0$ (for $n=0$) represents the initial image $I_G(\mathbf{x})$. Considering $L(\mathbf{x};t_n)$ as an intensity surface, the local shape characteristics can be described by the Hessian matrix $H(\mathbf{x};t_n)$. The Hessian is a matrix of the second order partial derivatives that describe the local curvature of a function [261]. For any given scale t_n, the Hessian matrix can be written as,

$$H(\mathbf{x};t_n)=\begin{bmatrix} L_{xx}(\mathbf{x};t_n) & L_{xy}(\mathbf{x};t_n) \\ L_{yx}(\mathbf{x};t_n) & L_{yy}(\mathbf{x};t_n) \end{bmatrix}. \quad \text{(B.2)}$$

$L_{xx}(\mathbf{x};t_n)$, $L_{yy}(\mathbf{x};t_n)$ and $L_{xy}(\mathbf{x};t_n)$ are the second-order partial derivatives of the image evaluated at the location $\mathbf{x}$ using the $[-1\ 0\ 1]$ derivative mask and the corresponding transpose [22]. Solving the characteristic equation of the Hessian,

$$|H(\mathbf{x};t_n)-\lambda I(\mathbf{x})|=0, \quad \text{(B.3)}$$

relatively to λ, where $I(\mathbf{x})$ is the identity matrix, the correspondent eigenvalues, can be computed

according to [283],

$$\lambda_1(\mathbf{x};t_n) = \frac{1}{2}\Big(- \sqrt{2L_{xx}(\mathbf{x};t_n)L_{yy}(\mathbf{x};t_n) + L_{xx}(\mathbf{x};t_n)^2 + 4L_{xy}(\mathbf{x};t_n)^2 + L_{xx}(\mathbf{x};t_n)^2} + L_{xx}(\mathbf{x};t_n) + L_{yy}(\mathbf{x};t_n)\Big), \tag{B.4}$$

$$\lambda_2(\mathbf{x};t_n) = \frac{1}{2}\Big(\sqrt{2L_{xx}(\mathbf{x};t_n)L_{yy}(\mathbf{x};t_n) + L_{xx}(\mathbf{x};t_n)^2 + 4L_{xy}(\mathbf{x};t_n)^2 + L_{xx}(\mathbf{x};t_n)^2} + L_{xx}(\mathbf{x};t_n) + L_{yy}(\mathbf{x};t_n)\Big). \tag{B.5}$$

The eigenvalues $\lambda_1(\mathbf{x};t_n)$ and $\lambda_2(\mathbf{x};t_n)$, represent the minimal and maximal principal curvatures at each location $(\mathbf{x}) \in L(\mathbf{x};t_n)$ [261].