

TEXTURE ANALYSIS OF T1-WEIGHTED TURBO SPIN-ECHO MRI FOR THE DIAGNOSIS AND FOLLOW-UP OF COLLAGEN VI-RELATED MYOPATHY

Rafael Rodrigues¹ Marta Gómez-García de la Banda² Mickael Tordjman² David Gómez-Andrés³
Susana Quijano-Roy² Robert-Yves Carlier² Antonio M. G. Pinheiro¹

¹ Instituto de Telecomunicações and Universidade da Beira Interior, Covilhã, Portugal

² APHP - Raymond Poincaré University Hospital, Garches, France

³ Hospital Universitari Vall d'Hebron, Barcelona, Spain

ABSTRACT

Muscle texture analysis in Magnetic Resonance Imaging (MRI) has revealed a good correlation with typical histological changes resulting from neuromuscular disorders. In this research, we assess the effectiveness of several features in describing intramuscular texture alterations in cases of Collagen VI-related myopathy. A T1-weighted Turbo Spin-Echo MRI dataset was used ($N_{\text{subj}} = 26$), consisting of thigh scans from subjects diagnosed with Ullrich Congenital Muscular Dystrophy or Bethlem Myopathy, with different severity levels, as well as healthy subjects. A total of 355 texture features were studied, including attributes derived from the Gray-Level Co-occurrence Matrix, the Run-Length Matrix, Wavelet and Gabor filters. The extracted features were ranked using the Support Vector Machine Recursive Feature Elimination (SVM-RFE) algorithm with Correlation Bias Reduction, prior to cross-validated classification with a Gaussian kernel SVM.

Index Terms— MRI, Texture analysis, Computer-aided diagnosis, COL6-related myopathy, SVM-RFE.

1. INTRODUCTION

Neuromuscular diseases include an extensive range of disorders, which cover a variety of phenotypes [1, 2]. Ullrich Congenital Muscular Dystrophy (UCMD) is a type of muscular dystrophy, caused by mutations in the Collagen VI (COL6) genes. The observed symptoms include generalized muscle weakness, contractures and distal joint hyperlaxity. In the most severe cases, children may present symptoms at birth, never be able to walk, and develop spinal deformities and life-threatening respiratory insufficiency [3, 4].

The pattern of muscle involvement in UCMD is similar to its milder variant, i.e. Bethlem Myopathy [1, 5]. A distinctive manifestation in T1-weighted (T1w) Magnetic Resonance Imaging (MRI) is the appearance of alternating bands of hypo- (preserved muscle) and hyperintensity (fat/connective tissue infiltration), in a striped pattern [3]. The thigh mus-

cles of subjects diagnosed with UCMD or Bethlem Myopathy often show a diffuse involvement of all muscles. T1w characteristic patterns may be seen in the *Vastus lateralis* – hyperintensity rim at the muscle periphery – and the *Rectus femoris* – hyperintense central area [5].

MRI is a standard diagnostic tool for neuromuscular diseases, which has shown a high correlation with their typical manifestations. However, there are still a lot of challenges regarding the development of computer-aided diagnosis (CAD) methods for such disorders, particularly due to their individual rareness [6, 2].

To the best of our knowledge, there is no published research on CAD methods for COL6-related myopathies using MRI, or other non-invasive imaging techniques. Recently, the authors in [4] proposed the first CAD method for these pathologies, which uses a Convolutional Neural Network (CNN) to classify confocal microscopy images of fibroblast cultures derived from forearm tissue samples.

In this paper, a CAD method for COL6-related myopathy in T1w Turbo Spin-Echo (TSE) MRI is proposed. The aim of this research is to identify effective texture-based biomarkers of those pathologies, with a focus on early diagnosis.

2. RELATED WORK

An early effort using MRI texture analysis towards the development of CAD methods for muscle diseases was presented in [7]. The authors compared the visual analysis by a group of radiologists with a computerized texture analysis of T1w MRI, using features from the intensity histogram, gradient, the Run-Length Matrix (RLM), and the Gray-Level Co-occurrence Matrix (GLCM). The proposed method outperformed the radiologists, with a sensitivity of 0.70 and a specificity of 0.86, against 0.56 and 0.71, respectively. Also, the reported results suggested a larger variability in texture observation by humans.

To overcome the lack of comprehensive human MRI data, many studies have focused on the Golden Retriever Muscular Dystrophy (GRMD) [8, 9, 10], which is considered to be the

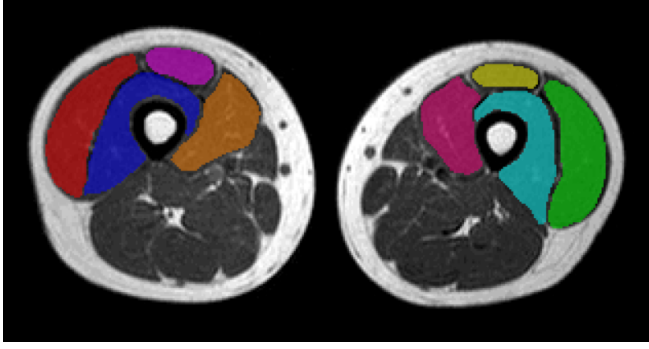


Fig. 1: T1w Turbo Spin-Echo MRI (axial slice) with segmentation of target muscles (*Quadriceps femoris*) in both thighs.

animal model with the highest similarity to Duchenne Muscular Dystrophy (DMD) observed in humans [6]. Duda *et al.* [9] used GLCM-based features to detect GRMD at early stages (2-3 months old), achieving an accuracy above 0.90.

The authors in [11] used a Support Vector Machine (SVM) to classify cases of DMD in T1w and T2w images. Several Wavelet-based features were extracted and Principal Component Analysis was used for feature selection. A mean accuracy of 0.93 was obtained with T1w images and level 2 decomposition features.

In [12], a CNN was tested on a dataset of Chemical Shift-Based Water/Fat Separation MRI, combined with class activation mapping. The authors reported an accuracy of 0.92 for the classification of DMD and Congenital Muscular Dystrophy. The obtained activation maps helped to identify muscles with higher discriminative power in the classification model, depending on the disease.

3. METHODS

3.1. MRI data acquisition and scoring

Whole Body MRI scans of 26 subjects, with ages between 5 and 36 years old, were performed at the Medical Imaging Department of the Raymond-Poincaré Hospital (Garches, France), using a 1.5T scanner (Philips Medical Systems, Eindhoven, The Netherlands). 17 subjects were diagnosed with UCMD or Bethlem Myopathy, while 9 were considered healthy, i.e. showed no perceivable muscle alterations in the MRI.

In this study, we used images from a T1w Turbo Spin-Echo (TSE) sequence of the thighs, in the axial plane, with TR/TE = 631/16ms and slice thickness = 5/6mm. A variable number of slices was retrieved for each subject, ranging from 6 to 12, to maximize the cross-sectional area of target muscles. The *Quadriceps femoris* muscle group – *Vastus lateralis* (VL), *Vastus medialis* (VM), *Vastus intermedius* (VI) and *Rectus femoris* (RF) – was manually segmented in both thighs (Fig. 1), resulting in 392 regions of interest (ROI) per muscle.

The degree of involvement was quantified at the muscle level by expert physicians into a 4-level scale (no symptoms, mild, moderate and severe), based on the recommendations in [5].

3.2. Texture features extraction

A total of 355 features were extracted in order to describe muscle texture in the T1w TSE MR images, which include features derived from the GLCM, the Run-Length Matrix (RLM), the Autoregressive (AR) model, and the Wavelet transform, as well as histogram and gradient-based features. These were extracted using the *MaZda* software¹ [13]. In addition, a set of Gabor filters were used, with ROI-based features computed from their response magnitude. Gabor feature extraction was implemented on MATLAB R2020b.

1) *Gray-Level Co-occurrence Matrix*: GLCM-based features [9, 10] are among the most widely studied features for texture description. The GLCM encodes pairwise spatial relations between gray-level intensities, along four different directions (0°, 45°, 90° and 135°) and considering distances up to 5 pixels. For each direction-distance coordinate, 11 Haralick features were computed - contrast, correlation, entropy, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, inverse difference moment and the second angular moment. The final GLCM descriptor included 220 features.

2) *Run-Length Matrix*: The RLM [8, 10] counts pixel runs along four different orientations - 0°, 45°, 90° and 135°. $P(g, l)$ refers to the number of runs with length l and gray level g . From the resulting P matrix of each orientation, the gray level non-uniformity, the run-length non-uniformity, the short run emphasis moment, the long run emphasis moment, and the fraction of image in runs were computed, yielding a total of 20 features.

3) *Autoregressive model*: The AR model describes pixel intensities as the weighted sum of neighboring pixels, with reference to the image mean gray-level intensity. [13]. The 5 computed features were the weights of neighboring pixels $\Theta_{1,2,3,4}$, and the minimized error variance σ_{AR} .

4) *Histogram and Gradient-based features*: The gray-level intensity histogram of the images was computed, from which the following 9 features were derived: mean, variance, skewness, kurtosis, and percentiles 1%, 10%, 50%, 90% and 99%. The mean, variance, skewness and kurtosis of the absolute gradient were also computed, as well as the percentage of non-zero gradient pixels. [13].

5) *Wavelet-based features*: A 4-level Wavelet decomposition was performed using the Haar wavelet [11]. Then, a ROI-based energy feature was computed for every sub-band (LL , LH , HL and HH) and scale, as follows:

$$E_w = \frac{\sum_{i=1}^I \sum_{j=1}^J (D_{i,j})^2}{n} \quad (1)$$

¹version 4.6, Institute of Electronics, Technical University of Lodz

where $D_{i,j}$ denotes the Wavelet coefficients for a given band and scale, and n is the number of pixels within a given ROI, at the considered scale. For each ROI, 16 energy features were computed (4 bands \times 4 scales).

6) *Gabor-based features*: The Gabor transform detects local texture information at specific orientations and spatial frequencies. The spatial kernel of a 2D Gabor filter can be defined by the following equation [14, 15]:

$$g(x, y) = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} \cdot e^{i\left(2\pi \frac{x'}{\lambda} + \varphi\right)} \quad (2)$$

where $x' = x \cos \theta - y \sin \theta$ and $y' = x \sin \theta + y \cos \theta$. The θ parameter is the orientation of the normal to the sinusoidal carrier wave and λ represents its wavelength. We used a spatial aspect ratio $\gamma = 0.5$ and a phase offset $\varphi = 0$. The standard deviation of the Gaussian envelope is given by $\sigma = 0.56 \lambda$, to keep a constant spatial frequency bandwidth of 1 octave [15].

The MR images were filtered using 2D Gabor filters at four different orientations ($\theta \in \{\pi/4, \pi/2, 3\pi/4, \pi\}$) and wavelengths ($\lambda \in \{2, 4, 6, 8\}$ pixels/cycle). The ROI-based mean, variance, skewness and kurtosis were computed from the Gabor response magnitude, along with an energy feature, as described in Eq. (1), for a total of 80 features.

3.3. Classification of texture features

3.3.1. SVM Recursive Feature Elimination

Feature selection is commonly used in pattern recognition for dimensionality reduction. Besides improving the classification performance, by avoiding data overfitting, it also ranks features according to some criteria, providing insight on the most relevant features for a specific problem.

The SVM Recursive Feature Elimination (SVM-RFE) algorithm was initially described in [16]. It is an iterative embedded backward elimination method, which uses the weights of the SVM separating hyperplane as ranking criteria. In this work, we followed the non-linear approach proposed in [17], using an SVM with a Gaussian kernel (LIBSVM implementation [18], $C = 1$, $\gamma = 2^{-6}$):

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2} \quad (3)$$

As it is not possible to directly determine the weight vector \mathbf{w} for non-linear cases, given the high-dimensionality, a different criterion was proposed in [17]. The algorithm computes the Lagrangian formulation of the separating hyperplane (Eq. (3)) with and without the k^{th} feature. The resulting difference is used as ranking criterion, with smaller differences leading to a lower ranking.

$$L_D = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (4)$$

Table 1: ROI distribution according to severity scoring [5].

	<i>Vastus lateralis</i>	<i>Vastus intermedius</i>	<i>Vastus medialis</i>	<i>Rectus femoris</i>
Healthy	132	132	132	132
Mild	44	56	80	92
Moder./Sev.	216	204	180	168

To reduce the time-consumption of the SVM-RFE algorithm, features were initially ranked in batches with 10% of the current remaining features. However, to increase the ranking accuracy for the best features, batch processing was disabled for less than 150 remaining features.

A Correlation Bias Reduction algorithm (CBR) was also introduced in [17]. The hyperplane weights for groups of highly correlated features may be incorrectly estimated, as they will likely have a similar relevance in the SVM model. As a result, those features may be removed, despite being potentially useful. At each iteration, the CBR algorithm analyses the batch of features to be removed and identifies the one with the highest ranking criterion (F_t). Then, it searches for features that correlate with F_t , with a coefficient above a given threshold T_c (set as 0.9 in this work), both in the removed and the remaining sets. If none of the features with a correlation coefficient above T_c is present in the remaining set, F_t is ranked again in the next iteration.

SVM-RFE has been primarily designed for binary classification scenarios. However, the algorithm used in this work extends SVM-RFE for multi-class tasks, by adding the weight vectors of the binary sub-classifiers.

3.3.2. Feature ranking and classification

The moderate and severe samples were grouped into the same severity class, as preliminary K-Means clustering showed that a clear borderline between such cases was hard to define. Table 1 shows the distribution of severity scores for the segmented muscles, according to this 3-level scale.

To address the class imbalance in the MRI dataset, data was randomly sampled with a constraint so that the number of observations of each class was identical. As described below, the experiment was divided in two stages and both were ran 10 times with resampled datasets, to assess the stability of the proposed method.

Initially, the SVM-RFE + CBR algorithm was ran, yielding 10 feature rankings for each target muscle. On a second stage, different subsets, including the first common N_f features across those 10 rankings ($N_f \in \{10, 20, \dots, 90, 100\}$), were classified with the Gaussian kernel SVM [18].

A leave-one-out cross-validation was set at subject level for SVM performance evaluation. In each fold, the samples from one subject were classified, leaving samples from the

Table 2: SVM performance measures (recall and precision) for individual classes, using the smaller feature subsets to achieve the highest overall accuracy (see Fig. 2). The bottom row shows the size of the feature subset in each case.

	<i>Vastus lateralis</i>		<i>Vastus intermedius</i>		<i>Vastus medialis</i>		<i>Rectus femoris</i>	
	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
Healthy	0.96 ± 0.02	1.00 ± 0.00	1.00 ± 0.00	0.98 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	0.98 ± 0.01	0.98 ± 0.01
Mild	0.99 ± 0.01	0.82 ± 0.03	0.93 ± 0.02	0.92 ± 0.03	0.77 ± 0.03	0.72 ± 0.02	0.91 ± 0.01	0.88 ± 0.03
Moder./Sev.	0.80 ± 0.05	0.99 ± 0.01	0.92 ± 0.04	0.94 ± 0.02	0.70 ± 0.03	0.75 ± 0.03	0.90 ± 0.04	0.92 ± 0.01
No. Feats	80		50		70		50	

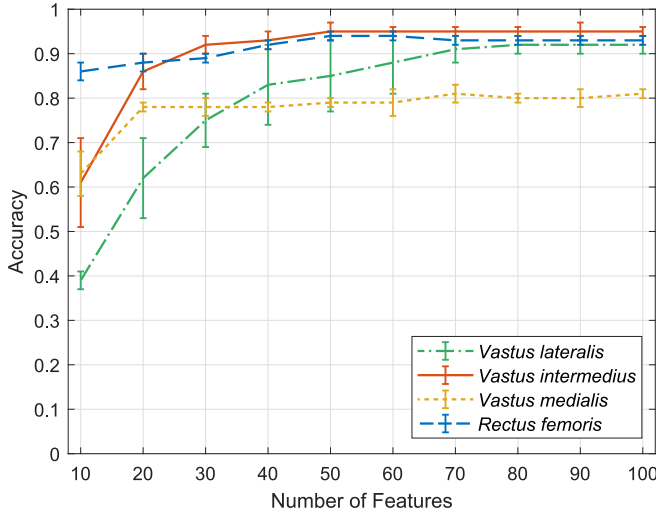


Fig. 2: SVM overall accuracy vs. number of features selected from the SVM-RFE + CBR ranking.

remaining subjects for SVM training. For each value of N_f , the mean and standard deviation of the performance measures across all repetitions were used for the model evaluation.

4. RESULTS AND DISCUSSION

Fig. 2 shows the cross-validated SVM accuracy as a function of the number of selected features, N_f . In all cases, the accuracy initially increased, reaching more stable values for higher N_f . For each muscle, the lowest N_f to yield maximum accuracy, within the studied range, was 50 for *VI*, 50 for *RF*, 70 for *VM*, and 80 for *VL*. The recall and precision measures of the cross-validated SVM classification with those selected subsets are shown in Table 2.

The selected subsets achieved a very good performance in separating healthy from disease cases in general, with near-perfect recall and, most importantly, precision rates (first row of Table 2). Overall accuracy in the best scenario was above 0.90 for three of the target muscles (*VL*, *VI* and *RF*). Nonetheless, the separation between normal and affected samples was perfect (recall and precision = 1) for *VM*.

GLCM-based features were the most frequent across the top performing subsets of the four muscles (*VL*: 61/80, *VI*: 41/50, *VM*: 43/70, *RF*: 37/50), with a clear predominance of the Inverse Difference Moment and the Sum Average features. Histogram features were also present in the four subsets - mean and several percentile features - and were generally high ranked, with at least 3 features in the top 20.

Gabor features were very relevant in the best subset for *VM*, with 11 attributes - mean, skewness, kurtosis and energy - at every tested wavelength ($\lambda \in \{2, 4, 6, 8\}$). They were also selected for *VL* (2 skewness features, $\lambda = 8$), *VI* (2 mean features, $\lambda = 8$) and *RF* (5 skewness features, $\lambda \in \{2, 4\}$).

The energy features from Wavelet low-pass sub-bands (*LL*), at different decomposition levels, were selected for all muscles. In the case of *VM*, features from the *HH* sub-band at level 3 and the *LH* sub-band at level 1 were also selected.

Some RLM-based features were selected, namely the gray level non-uniformity in the 4 directions (*VL*), the short run emphasis moment at 90° , gray level non-uniformity at 135° and the fraction at 45 and 90° (*VM*), and also the horizontal short run emphasis moment and fraction (*RF*).

Finally, σ_{AR} was selected for *VI* and *VM*, along with Θ_1 and Θ_2 (*VM*), and Θ_3 (*RF*). None of the gradient-based features were part of the top-performing feature subsets.

5. CONCLUSION AND FUTURE WORK

In this paper, we presented the first computer-assisted approach to the diagnosis of COL6-related myopathy using MRI. The obtained results were encouraging and suggest that texture analysis might be suitable for clinical application on the diagnosis and follow-up of these disorders. The Gaussian kernel SVM performed very well in distinguishing healthy from mild cases, with relatively few selected features from the *Quadriceps femoris* target muscles.

In future research, we will study a wider range of texture features, as well as the influence of muscle combinations and different feature selection methods. Moreover, we plan to test deep learning methods with this dataset, which is expected to be expanded with new data.

6. COMPLIANCE WITH ETHICAL STANDARDS

This study was performed in line with the principles of the Declaration of Helsinki by the World Medical Association, as revised in 2000.

7. ACKNOWLEDGMENTS

The authors have no conflicts of interest to declare. This research was co-funded by the COST Action BM1304 and FCT/MCTES under the project UIDB/EEA/50008/2020. Author R. Rodrigues was funded by "Fundação para a Ciência e a Tecnologia" (FCT), under the individual doctoral grant SFRH/BD/130858/2017.

8. REFERENCES

- [1] L. Ten Dam, A. J. van der Kooi, C. Verhamme, M. Wattjes, and M. de Visser, "Muscle imaging in inherited and acquired muscle diseases," *Eur. J. Neurol.*, vol. 23, no. 4, pp. 688–703, 2016.
- [2] S. Quijano-Roy and R. Y. Carlier, "Neuroimaging in non-dystrophic myopathies," in *Clinical Neuroradiology: The ESNR Textbook*, F. Barkhof, H. R. Jäger, M. M. Thurnher, and À. Rovira, Eds., pp. 2145–2183. Springer International Publishing, 2019.
- [3] S. Quijano-Roy, D. Avila-Smirnow, R. Y. Carlier, and WB-MRI Muscle Study Group, "Whole body muscle MRI protocol: pattern recognition in early onset NM disorders," *Neuromuscul. Disord.*, vol. 22, pp. S68–S84, 2012.
- [4] A. Bazaga, M. Roldán, C. Badosa, C. Jiménez-Mallebrera, and J. M. Porta, "A Convolutional Neural Network for the automatic diagnosis of collagen VI-related muscular dystrophies," *Appl. Soft Comput.*, vol. 85, pp. 105772, 2019.
- [5] E. Mercuri *et al.*, "Muscle MRI in Ullrich congenital muscular dystrophy and Bethlem myopathy," *Neuromuscul. Disord.*, vol. 15, no. 4, pp. 303–310, 2005.
- [6] J. D. De Certaines *et al.*, "Application of texture analysis to muscle MRI: 1-What kind of information should be expected from texture analysis?," *EPJ Nonlinear Biomed. Phys.*, vol. 3, no. 1, pp. 1–14, 2015.
- [7] S. Herlidou, Y. Rolland, J. Bansard, E. Le Rumeur, and J. De Certaines, "Comparison of automated and visual texture analysis in MRI: characterization of normal and diseased skeletal muscle," *Magn Reson Imaging*, vol. 17, no. 9, pp. 1393–1397, 1999.
- [8] Z. Fan *et al.*, "Characteristics of magnetic resonance imaging biomarkers in a natural history study of Golden Retriever Muscular Dystrophy," *Neuromuscul. Disord.*, vol. 24, no. 2, pp. 178–191, 2014.
- [9] D. Duda, M. Kretowski, N. Azzabou, and D. Jacques, "MRI texture analysis for differentiation between healthy and golden retriever muscular dystrophy dogs at different phases of disease evolution," in *IFIP Int. Conf. Comput. Inf. Syst. Ind. Manage.* Springer, 2015, pp. 255–266.
- [10] A. Eresen, L. Alic, S. M. Birch, W. Friedeck, J. F. Griffin IV, J. N. Kornegay, and J. X. Ji, "Texture as an imaging biomarker for disease severity in Golden Retriever Muscular Dystrophy," *Muscle & Nerve*, vol. 59, no. 3, pp. 380–386, 2019.
- [11] M.-H. Zhang, J.-S. Ma, Y. Shen, and Y. Chen, "Optimal classification for the diagnosis of Duchenne Muscular Dystrophy images using Support Vector Machines," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 11, no. 9, pp. 1755–1763, 2016.
- [12] J. Cai, F. Xing, A. Batra, F. Liu, G. A. Walter, K. Vandenborne, and L. Yang, "Texture analysis for muscular dystrophy classification in MRI with improved class activation mapping," *Pattern Recognit.*, vol. 86, pp. 368–375, 2019.
- [13] P. M. Szczypiński, M. Strzelecki, A. Materka, and A. Klepaczko, "MaZda — a software package for image texture analysis," *Comput. Methods Programs Biomed.*, vol. 94, no. 1, pp. 66–76, 2009.
- [14] Y. Zhang, W. Li, L. Zhang, X. Ning, L. Sun, and Y. Lu, "Adaptive learning Gabor filter for finger-vein recognition," *IEEE Access*, vol. 7, pp. 159821–159830, 2019.
- [15] P. Kruizinga and N. Petkov, "Nonlinear operator for oriented texture," *IEEE Trans Image Process*, vol. 8, no. 10, pp. 1395–1407, 1999.
- [16] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using Support Vector Machines," *Machine Learning*, vol. 46, pp. 389–422, 2002.
- [17] K. Yan and D. Zhang, "Feature selection and analysis on correlated gas sensor data with recursive feature elimination," *Sens. Actuators B Chem.*, vol. 212, pp. 353–363, 2015.
- [18] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans Intell Syst Technol*, vol. 2, no. 3, pp. 27:1–27:27, 2011.