

Plantar Fascia Ultrasound Images Characterization and Classification Using Support Vector Machine

Abdelhafid Boussouar¹, Farid Meziane¹ and Lucy Anne Walton²

¹School of Computing, Science and Engineering

²School of Health and Society, Directorate of Radiology

University of Salford, Salford M5 4W T, UK.

A.Boussouar1@edu.salford.ac.uk;

F.Meziane@salford.ac.uk

L.A.Walton@salford.ac.uk

Abstract. The examination of plantar fascia (PF) ultrasound (US) images is subjective and based on the visual perceptions and manual biometric measurements carried out by medical experts. US images feature extraction, characterization and classification have been widely introduced for improving the accuracy of medical assessment, reducing its subjective nature and the time required by medical experts for PF pathology diagnosis. In this paper, we develop an automated supervised classification approach using the Support Vector Machine (Linear and Kernel) to distinguish between symptomatic and asymptomatic PF cases. Such an approach will facilitate the characterization and the classification of the PF area for the identification of patients with inferior heel pain at risk of plantar fasciitis. Six feature sets were extracted from the segmented PF region. Additionally, features normalization, features ranking and selection analysis using an unsupervised infinity selection method were introduced for the characterization and the classification of symptomatic and asymptomatic PF subjects.

The performance of the classifiers was assessed using confusion matrix attributes and some derived performance measures including recall, specificity, balanced accuracy, precision, F-score and Matthew's correlation coefficient. Using the best selected features sets, Linear SVM and Kernel SVM achieved an F-Score of 97.06 and 98.05 respectively.

Keywords: Plantar fascia ultrasound images, Features selection and characterization, SVM, k-folded cross validation, Matthew's correlation coefficient.

1 Introduction

Ultrasound (US) imaging offers a significant potential in the diagnosis of plantar fascia (PF) injuries and monitoring treatments. It offers a real-time effective imaging technique that can reliably confirm structural changes, such as thickening, rupture and identify changes in the internal echo structure associated with diseased or damaged tissues. PF US images are usually examined and analysed by health specialists based on visual perceptions and some manual biometric measurements, such as the thickness

estimation of the PF region, to identify the presence of any kind of lesions and abnormalities such as plantar fasciitis (inflammation of the plantar fascia).

As reported in the literature, thickening, bi-convexity, rough surface, heterogeneous texture, decreased echogenicity, loss edge sharpness and hypoechoic deformities of the PF are considered as part of the diagnostic criteria and characteristic features of symptomatic PF; whereas, surface smoothness, texture homogeneity and uniform hyperechogenicity are characteristics of asymptomatic PF subjects [8,15,19,24] and this is shown in Fig 1.

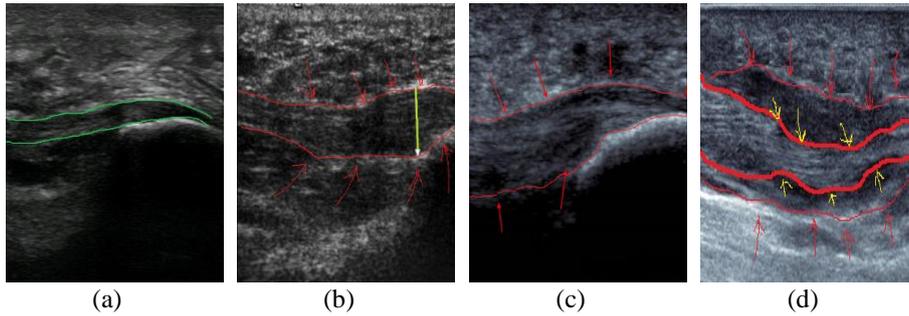


Fig.1. Asymptomatic and Symptomatic PF region comparison: (a) Asymptomatic PF region (green contours), (b-d) Symptomatic PF region: (b) and (c) a thickened PF sections (red arrows) compared to a normal PF in (a) due to planar fasciitis disorder, (d) a huge partial tear of the PF region: the outer red contour clearly shows a surrounding inflammation (plantar fasciitis), while the inner contour (bold red) shows the irregular outline and disrupted PF region fibres.

Despite the advantages of US imaging, the acquired images interpretation and analysis are time consuming. This is mainly due to the large number of patients, the large medical data history accumulated in the DICOM systems and the large number of physicians required for the analysis and interpretation. The exploration of such massive medical data requires highly efficient and sophisticated techniques capable of finding the class separation between asymptomatic and symptomatic ultrasound images of the plantar fascia. These techniques are highly required to classify different PF US images into normal and abnormal subjects and to prune the huge accumulated data and take into consideration only the symptomatic data with the possibility of plantar fasciitis or other disorders. Therefore, it is a requirement to devise an automated system to characterize and classify PF US images that allows better abnormalities detection and easier interpretation during medical image analysis.

This paper proposes an automated supervised classification approach using linear-SVM and Kernel-SVM to facilitate the detection and the characterization of the plantar fascia region for the classification of PF US images dataset into symptomatic PF subjects and asymptomatic subjects; and the possibility of the identification of patients with normal plantar fascia but at risk of plantar fasciitis disorder.

2 The Methodology and proposed approach

The proposed PF classification model consists of the following modules as shown in Fig 2. (i) preprocessing phase employing speckle noise reduction filtering and image enhancement operations to reduce the effects of undesirable speckle noise phenomenon and improve the contrast of the PF US images using dual tree complex wavelet transform with soft thresholding (DT-CWT_S) and contrast-limited adaptive histogram equalization filter (CLAHE), respectively; (ii) artificial neural networks supervised segmentation phase applying different features measures, a features ranking module and trained radial basic function neural network (RBF-NN) classifier as discussed in [3], to automatically segment the PF region and calculate its thickness; (iii) texture features extraction and analysis introducing 6 sets of feature extraction measures (for extracting a total of 40 features), features ranking and selection operation using an unsupervised infinity feature selection method [17] to select and analyse the most discriminating and suitable features for the classification process; (iv) the classifier module using Linear-SVM and Kernel-SVM to distinguish between asymptomatic and symptomatic plantar fascia subjects; and (v) classification performance analysis using different performance measures such as recall, specificity, balanced accuracy, precision, F-score and MCC.

2.1 Materials and PF data collection

Following ethical approval from the University of Salford Research's Ethics Panel (ST1617-48), written informed consent was collected from all patients' participants. Various PF US images, acquired from a patient's footprint area in the prone position were used in the classification approach; more specifically, a total of 284 (252 normal and 32 abnormal taken from diabetic patients with plantar fasciitis) PF US images, were acquired from a patient's footprint area in the prone position were used. The images were obtained from 45 patients for different PF anatomical structures including rear-foot, mid-foot and fore-foot sections with 256 gray levels, a size dimension of 512x512 pixels and a resolution of 28.35 pixels/centimetre. These images were obtained from the Health Sciences Department, University of Salford, acquired by two expert clinicians according to a precise protocol using a Venue 40 musculoskeletal US system (GE Healthcare, UK) with a 5-13 MHz wideband linear array probe 12.7 mm X 47.1 mm. All the methods used in the proposed approach were implemented using Matlab R2017b (The MathWorks Inc., Natwick, USA).

2.2 Preprocessing

The preprocessing phase aims to prepare the PF US images for further processing including segmentation and classification and improve their accuracy, efficiency, and scalability. This is achieved by (i) minimizing the effects of the multiplicative speckle noise without losing any valuable information (such as tiny lines, edges) using a selected dual tree complex wavelet-based despeckling filter (DT-CWT_S) [13,16]. DT-CWT_S filter integrates homomorphic transformation (using log compression and exponent decompression to transform the multiplicative noise to an additive one) and

multi-scale DT-CWT decomposition and composition employing the BayesShrink sub-band thresholding using soft thresholding to reduce or suppress the speckle noise (noisy coefficients) in PF US images. DT-CWT_S has demonstrated a superior edge preserving behaviour and a good visual appearance [4]; (ii) enhancing the PF region contrast and visually improving the global appearance of the PF US images using CLAHE method after the despeckling operation to adjust the intensity of the PF region and to avoid noise amplification in PF US images using different implemented steps as reported in [26].

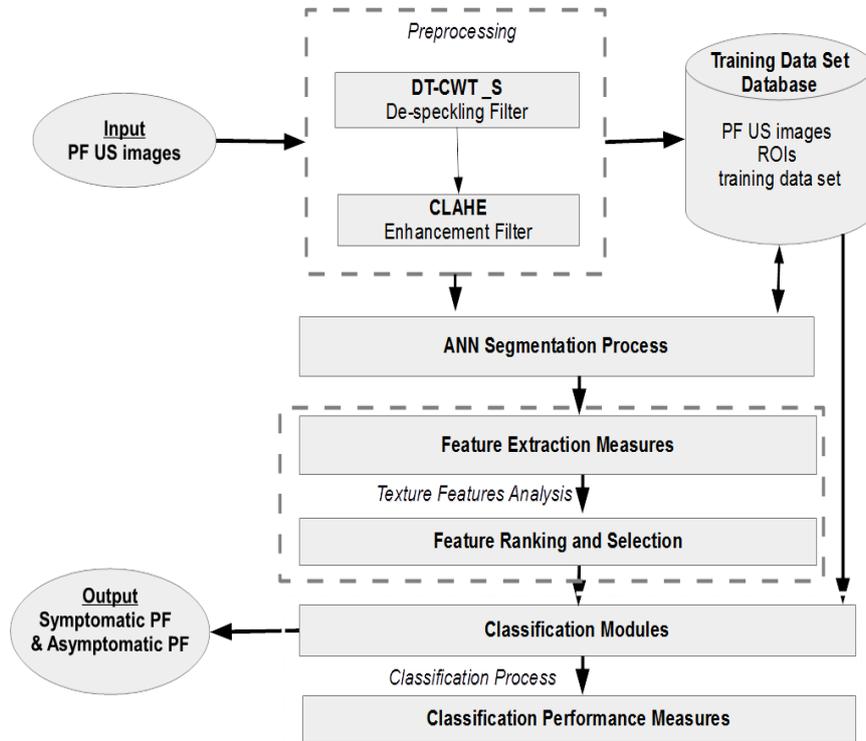


Fig 2. Flowchart illustrating the plantar fascia classification system based on a texture features analysis and different classifiers modules

2.3 Segmentation

Automated segmentation is one of the most important tasks in medical image processing and analysis, including, pattern recognition, supervised or unsupervised subjects classification and novelty detection; it is mainly used to locate the desired region of interest objects in the input images dataset. As reported in [3], an automated ANNs supervised segmentation approach was introduced in this study to segment different PF regions. The proposed segmentation approach uses the radial basic function neural network (RBF-NN) classifier [9] to automatically segment the PF region and estimate its thickness. Two different quantitative evaluation metrics namely: the region-based

metrics and distance based metrics were used to evaluate the segmentation method. For the region-based metrics, the highest precision obtained was 97.70 ± 1.20 , whereas for the distance-based metrics it was 0.10 ± 0.07 . The full description of the segmentation process and steps and analysis are reported in detail in [3].

2.4 Feature extraction

In most classification tasks, feature extraction is an important step to extract the relevant information (reduced input dataset representation) from the input dataset in order to perform the remaining tasks. Thus, the main goal of feature extraction in this classification study is to extract a set of textual features from the PF segments (using different measures) that discriminate between one input pattern and another, and then feed this into different classifiers for a classification task.

In this stage six different sets of features (40 features in total) were extracted from the segmented PF region including: (i) Haralick spatial gray level dependence matrices (SGLDM) [10] where a total of 12 SGLDM features were computed and averaged for a selected distance $d=1$ (3×3 matrices) and four different orientation angles 0° , 45° , 90° , and 135° ; (ii) Region based features; (iii) Neighbourhood gray tone difference matrix (NGTDM) for a kernel window of 3×3 [2]; (iv) Histogram based features or first-order statistics (FOS) [5,21]; (v) Statistical feature matrix (SFM) [5,25]; (vi) Laws' texture energy measures [5,11,25].

All the features extracted may have some redundancy, thus we introduced a feature selection and analysis stage to reduce this redundancy and to select the most discriminating feature sets.

2.5 Feature normalization

All features in this study were normalized using the mean variance normalization (MVN) approach which helps in reducing any non-linear distortion and scaling all features so they fall within a specified range (e.g. $[0 \ 1]$ or $[-1 \ 1]$) to prevent high measurement values (especially region based measurements) from outweighing other feature values with smaller values (e.g. SGLDM features) [7]. The normalized features $NX_{j,n}$ are computed by calculating the difference between the features and their mean values, and then dividing them by their standard deviation values as given by equation 1 [7].

$$NX_{i,n} = \frac{x_{j,n} - \mu_{j,n}}{\sigma_{j,n}} \quad (1)$$

where $\mu_{j,n}$ is the mean value of the feature vector x_i and $\sigma_{j,n}$ its standard deviation.

2.6 Feature ranking and selection

A common deficiency in most pattern recognition and classification tasks is the high dimension of the extracted feature space compared to the number of the input samples (40 features \times 284 observations). This will lead to some common problems such as: over-fitting, poor generalization and high computation cost. In order to minimize the aforementioned problems, a combination feature ranking and feature selection unsupervised infinity techniques [17,18] were introduced to reduce the correlated

measurements and to select the most discriminating features. Different selected feature sets were analysed to choose the best discriminating features for SVM classification modules based on high F-score values.

2.7 Classification

Following feature ranking and selection analysis, the feature classification approach was implemented using Linear-SVM and Kernel-SVM. To obtain a good classification result, three main conditions were taken into consideration during the classification process: (1) careful selection of features; (2) a good classifier; and (3) suitable training samples [22]. All the analysed feature sets described in the previous sections, were treated as input vectors to the selected classifier modules and their results were evaluated using different classification measures. To overcome the over-fitting problem and to validate the robustness of the classifiers, a cross-validation task was also introduced using the k-folded (k=10 folds) approach to randomly select the training and testing instance classes.

3 Support vector machines

Support vector machines (SVM) [23] are widely used in bioinformatics and medical studies for pattern recognition related problems [12]. The main concept of SVM is that, firstly, it differentiates between two class samples according to the optimal maximum margin (distance between each set) hyperplane (or decision boundary) search result [21]; secondly, if the hyperplane fails to split the previous linear class samples, the SVM makes use of different kernel functions such as polynomial kernel, Gaussian-RBF and sigmoids-NN instead of linear SVM [6,14,23]. This aims to achieve high dimensional feature space when translating original data samples [20]. In this study, both Linear-SVM and Kernel-SVM classifiers were tested and the Gaussian-RBF kernel function is used in the Kernel-SVM main function. For the PF US 2D training dataset T_S with N_L labelled instances (X_j, Y_j) , where X_j denotes the feature instances and Y_j is the class label with 1 for normal and -1 for abnormal PF class, and N is the total number of samples (252 normal and 32 abnormal samples with 40 extracted features).

4 Experimental results and discussion

For the classification experimental results, a total of 284 (252 symptomatic and 32 asymptomatic) US images of the PF regions (rear-foot, mid-foot and fore-foot sites) were analysed. Six different sets of features representing a total of 40 features were computed both from symptomatic and asymptomatic US images of the PF segments. For all extracted features, feature selection approach was introduced, and their means, weights and ranking orders were computed and analysed for normal and abnormal PF US images.

4.1 Feature extraction and selection analysis

The main reason for feature selection analysis is to eliminate similar or highly co-dependent features and to find the best discriminatory features that predict the best classification results. Feature selection analysis results of the top ranked features calculated from the 284 US images of the segmented PF region are represented in Fig. 3. For each feature, the weight predictor was calculated and its rank order was assigned accordingly. The best features (with the highest weight and ranked predictor) were found to be LS, Contrast, Variance, LE, Energy, SumSquare, AngSecMoment, LL, EE, DiffVariance, Strength, ES, Complexity, Correlation, DiffEntropy, SS, SumAverage, MajAxLength, Periodicity, Business, Mean, Skewness, Kurtosis, Orientation, Roughness, ConvexArea, Extent, EquivDiameter and Area.

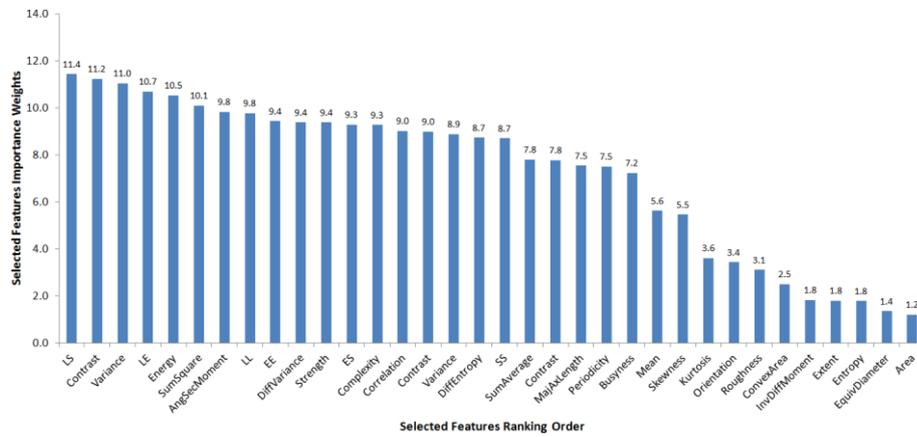


Fig. 2. Graph representation of 34 ranked predictors (features importance) based on their importance weights

Furthermore, to determine the best features for each classifier, the F-score measures were computed for the Linear-SVM and Kernel-SVM classifiers using different selected feature sets (from 1 to 40, starting with the highest ranked features). Six selected feature sets were defined using the highest F-score measure. The F-score result obtained by Kernel-SVM and Linear SVM were 98.05% and 97.06% respectively when using 34-features. It is also clearly evident from this analysis, that there are differences in feature weights values between asymptomatic and symptomatic PF subjects. From this interpretation, symptomatic PF texture tends to be darker with high contrast, high variance, high shape measures (high thickness) (due to the accumulation of the inflammation fluid), more extent, high convex area (due to irregularity of the PF surface and outline disruption), high complexity (more heterogeneous), low strength, less periodicity, more roughness and low grey intensity. While on the other side, asymptomatic PF texture are brighter with low contrast, low variance, less shape measures, less extent, less convex area, low complexity (more homogeneous), high strength, more periodicity, more smoothness and high grey intensity.

4.2 Classification analysis

For the classification task, Linear-SVM and Kernel-SVM classifiers were trained and tested using the same training and testing datasets. To overcome the over-fitting problem during the training stage and to assess the performance of various classification modules, 10-fold cross-validation was introduced. The main concept of the cross-validation approach is that each sample is added in both training and testing samples. For 10-fold cross validation approach, datasets (252 asymptomatic PF subjects and 32 symptomatic subjects) were randomly partitioned into 10 different equal splits (folds) (i.e. 10-1=9 folds were used for training task and the remaining fold is used for testing, with an iteration of 10 times dropping one-fold out for testing each time). Six different classification performance measures were computed and analysed for the two classifiers. For each classification the mean value of the 10-cross validations was computed. The results of the classifications using the best selected features are summarized in Table 1.

Table 1. Linear-SVM and Kernel-SVM classification performance measures using the best selected feature sets

| Classifier | Recall | Specificity | B-Accuracy | Precision | F-Score | MCC¹ |
|-------------------|---------------|--------------------|-------------------|------------------|----------------|------------------------|
| Linear-SVM | 95.75 | 84 | 89.88 | 98.41 | 97.06 | 71.46 |
| Kernel-SVM | 96.18 | 100 | 98.09 | 100 | 98.05 | 81.32 |

In all the six classification performances, Kernel-SVM performed better than Linear SVM. For specificity and Precision, Kernel-SVM achieved 100%.

5 Conclusions

In this study we developed a new automatic supervised classification system for discriminating different ultrasound plantar fascia images using Linear-SVM and Kernel-SVM classifiers. This will help medical experts to improve the efficiency of the PF pathology diagnosis and minimize the time required for the diagnosis. Six different feature set measures were used to extract and analyse the texture features. Additionally, the infinity selection method was successfully adopted to rank and characterize asymptomatic and symptomatic features, based on their weights importance. The results of the feature selection stage revealed that the top selected features can represent the characteristics of asymptomatic and symptomatic PF subjects ultrasound images well. The Inf selection method to select the best features is quite effective. In order to define and compare the best features, the F-score measure was independently computed for both classifiers (Linear-SVM and Kernel-SVM) using different selected feature sets (1-40). The best selected feature set for every classifier were fed to the related classifier as the input vector for the classification task. In the experiments, different performance evaluation measures were used to assess the classification capability of the two classifiers using their best selected features. The results have shown that Kernel-SVM

¹ The Matthew's correlation coefficient

outperformed Linear-SVM in all the performance evaluation measures. For specificity and precision, Kernel-SVM achieved 100%.

The classification results are very high for this first experiment. However, we experimented with a further two classifiers to have an idea on the performance of the Linear and Kernel SVM algorithms. The radial basic function neural network (RBF-NN) classifier and the linear discriminant analysis (LDA) were used with the best selected features set. The results are summarised in Table 2.

Table 2. RBF-NN and LDA Classification performance measures using the best selected feature sets

| Classifier | Recall | Specificity | B-Accuracy | Precision | F-Score | MCC² |
|-------------------|---------------|--------------------|-------------------|------------------|----------------|------------------------|
| RBF-NN | 98.82 | 96.67 | 97.74 | 99.60 | 99.21 | 92.82 |
| LDA | 97.62 | 81.25 | 89.43 | 97.62 | 97.63 | 78.87 |

The results show that for precision, specificity and B-Accuracy, Kernel-SVM still outperforms all other classifiers. RBF-NN obtained better results in Recall, F-Score and MCC. LDA in general performed worst then Kernel-SVM and RBF-NN. We are aiming in the next phase of the development of the system to use more classifiers.

References

1. Abe, S. Support Vector Machines for Pattern Classification. *Advances in Computer Vision and Pattern Recognition*. Springer London, 2010.
2. Amadasun, M. and King, R. Textural features corresponding to textural properties. *IEEE Transactions on systems, man, and Cybernetics* 19(5), pages 1264-1274, 1989.
3. Boussouar, A., Meziane, F. and Crofts, G. Plantar fascia segmentation and thickness estimation in ultrasound images. *Computerized Medical Imaging and Graphics* 56, pages 60-73, 2017.
4. Boussouar A. Thickness estimation, automated classification and novelty detection in ultrasound images of the plantar fascia tissues, *PhD thesis*, School of Computing, Science and Engineering, University of Salford, UK, 2019.
5. Christodoulou, C.I., Pattichis, C.S., Pantziaris, M. and Nicolaides, A. Texture-based classification of atherosclerotic carotid plaques. *IEEE transactions on medical imaging*. 22(7), pages 902-912, 2003.
6. Cortes, C. and Vapnik, V. Support-vector networks. *Machine learning* 20(3), pages 273-297, 1995
7. Dougherty, G. *Pattern Recognition and Classification: An Introduction*. SpringerLink: Bücher. Springer New York, 2012.
8. Fabrikant, J.M. and Park, T.S. Plantar fasciitis (fasciosis) treatment outcome study: Plantar fascia thickness measured by ultrasound and correlated with patient self-reported improvement. *The Foot* 21(2), pages 79-83, 2011.
9. Ham, F.M. and Kostanic, I. *Principles of Neurocomputing for Science and Engineering*, McGraw-Hill Higher Education, 2000.
10. Haralick, R.M., Shanmugam, K. and Dinstein, I.H. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6), pages 610-621, 1973.

² The Matthew's correlation coefficient

11. Laws, K.I. Rapid texture identification. *Proc. SPIE 0238, Image Processing for Missile Guidance*, pp. 238-244, 1980.
12. Martínez-Trinidad, J., Ochoa, J. and Kittler, J. Progress in Pattern Recognition, Image Analysis and Applications. *11th Iberoamerican Congress on Pattern Recognition, CIARP 2006*, Cancún, Mexico, November 14-17, Proceedings, Springer, 2006.
13. Michailovich, O. and Tannenbaum, A. Despeckling of medical ultrasound images. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 53(1), pages 64-78, 2006.
14. Osuna, E., Freund, R. and Girosi, F. *Support vector machines: Training and applications*, 1997.
15. Park, J.W., Yoon, K., Chun, K.S., Lee, J.Y., Park, H.J., Lee, S.Y. and Lee, Y.T. Long-term outcome of low-energy extracorporeal shock wave therapy for plantar fasciitis: Comparative analysis according to ultrasonographic findings. *Annals of Rehabilitation Medicine*, 38(4), pages 534-540, 2014.
16. Rabbani, H., Vafadust, M., Abolmaesumi, P. and Gazor, S. Speckle noise reduction of medical ultrasound images in complex wavelet domain using mixture priors. *IEEE Transactions on Biomedical Engineering*, 55(9), pages 2152-2160, 2008.
17. Roffo, G., Melzi, S. and Cristani, M. Infinite feature selection. *IEEE International Conference on Computer Vision (ICCV)*, pages 4202-4210, 2015.
18. Roffo, G., Melzi, S., Castellani, U. and Vinciarelli, A. Infinite latent feature selection: A probabilistic latent graph-based ranking approach. *IEEE International Conference on Computer Vision (ICCV)*, 2017.
19. Saber, N., Diab, H., Nassar, W. and Razaak, H.A. Ultrasound guided local steroid injection versus extracorporeal shockwave therapy in the treatment of plantar fasciitis. *Alexandria Journal of Medicine*, 48(1), pages 35-42, 2012.
20. Shi, X., Cheng, H., Hu, L., Ju, W. and Tian, J. Detection and classification of masses in breast ultrasound images. *Digital Signal Processing*, 20(3), pages 824-836, 2010.
21. Umbaugh, S. *Computer Imaging: Digital Image Analysis and Processing*. A CRC Press book. Taylor & Francis, 2005.
22. Unger, H., Meesad, P. and Boonkrong, S. Recent Advances in Information and Communication Technology, *Advances in Intelligent Systems and Computing*. Springer International Publishing, 2015.
23. Vapnik, V. *The nature of statistical learning theory*. Springer science & business media, 2013.
24. Wearing, S.C., Smeathers, J.E., Sullivan, P.M., Yates, B., Urry, S.R. and Dubois, P. Plantar fasciitis: are pain and fascial thickness associated with arch shape and loading? *Physical therapy*, 87(8), pages 1002-1008, 2007.
25. Wu, C.M., Chen, Y.C. and Hsieh, K.S. Texture features for classification of ultrasonic liver images. *IEEE Transactions on medical imaging*, 11(2), pages 141-152, 1992.
26. Zuiderveld, K. Contrast Limited Adaptive Histogram Equalization, Chapter VIII.5, *Graphics Gems IV*. P.S. Heckbert (Eds.), Cambridge, MA, Academic Press, pages 474-485, 1994.