

CLASSIFICATION OF OCT IMAGES FOR DETECTING DIABETIC RETINOPATHY DISEASE USING MACHINE LEARNING

Marwan Aldahami and Umar Alqasemi

Dept. of Electrical and Computer Engineering, King Abdulaziz University, P.O.Box 80200, Jeddah 21589, Saudi Arabia

ABSTRACT

Optical Coherence Tomography (OCT) imaging aids in retinal abnormality detection by showing the tomographic retinal layers. OCT images are a useful tool for detecting Diabetic Retinopathy (DR) disease because of their capability to capture micrometer-resolution. An automated technique was introduced to differentiate DR images from normal ones. 214 images were subjected to the experiment, of which 160 images were used for classifiers' training, and 54 images were used for testing. Different features were extracted to feed our classifiers, including statistical features and local binary pattern (LBP) features. The experimental results demonstrated that our classifiers were able to discriminate DR retina from the normal retina with Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) of 100%. The retinal OCT images have common texture patterns and using a powerful tool for pattern analysis like LBP features has a significant impact on the achieved results. The result has better performance than previously proposed methods in the literature.

KEYWORDS

Image classification, diabetic retinopathy, support vector machine, optical coherence tomography, retina, machine learning.

1. INTRODUCTION

Diabetic Retinopathy (DR) is a microvascular disease that affects retinal vessels. It is considered one of the most common diseases that causes vision loss for diabetic patients. In the year 2000, about 4 million DR cases were estimated in the United States. In 2010, the number of cases was 7.7 million, which is expected to be 14.6 million by 2050. The National Health and Nutrition Examination Survey (NHANES) conducted a visual analysis in 2005-2008. NHANES estimated that 29% of people aged 40 years or more had simultaneous DR. Additionally, it has been estimated that the prevalence of DR in the sample measured was 4% [1]. NHANES also declared that DR prevalence is approximately 32% in males, while it is approximately 26% in females and 39% in non-Hispanic black people versus 25% in non-Hispanic white people [1].

Optical Coherence Tomography (OCT) is an emerging technology that allows for the study of blood flow within the eye's vascular structure [2]. OCT is a non-invasive technique that uses a low-coherence light to produce high-resolution, cross-sectional, and micrometer-scale images. The principle of OCT is based on optics theory by measuring the transmitted and reflected optical signals that contain time-of-flight information, which yields spatial information about scanned tissue [3].

The early stage of DR is known as non-proliferative diabetic retinopathy (NPDR). During this stage, the retina vasculature begins a process of change wherein vascular permeability and capillary occlusion increase. The advanced stage of DR is called proliferative diabetic retinopathy (PDR), in which vitreous hemorrhage is present. During this stage, the abnormal vessels commence bleeding into the vitreous humor and may result in tractional retinal detachment in the patient's eye, which causes vision impairment. By obtaining OCT angiography images, ophthalmologists are able to detect diabetic eyes that have a potential risk of developing retinopathy [4]. Developing an OCT computer-aided detection (CADE) software will assist ophthalmologists in diagnosing patients in an accurate, fast, and safe manner that will protect diabetic eyes from vision loss at early stages.

In this paper, we have designed an automated system to classify the OCT into normal images and images with DR. The system was trained by inserting 160 images with their class and tested by inserting 54 images. Two types of classifiers were used: Support Vector Machine (SVM) and k-Nearest Neighbor (kNN).

2. LITERATURE REVIEW

Recent studies have addressed the automated classification of OCT images by extracting the images' features and using algorithms for classification or by segmentation. Priya et al. [5] proposed a system to diagnose diabetic retinopathy disease by using 350 fundus images. The used images were collected from "Aravind Eye Hospital and Postgraduate Institute of Ophthalmology". The images were produced from the fundus camera in RGB form. The authors started by preprocessing the images to make them suitable for the machine learning system. The images were converted into gray-scale images. Then, in order to enhance the images' contrast, they were subjected to adaptive histogram equalization. After that, the Matched Filter Response (MFR) and Discrete Wavelet Transform (DWT) were applied to reduce the noise and the images' size. The authors extracted some features from the images subjected to their study, such as the blood vessels, NPDR hemorrhage, and PDR exudates, by using image segmentation. They applied three classifiers: Probabilistic Neural Network (PNN), Bayesian and Support Vector Machine (SVM) classifiers. The best results achieved for the SVM classifier were 96, 98 and 97.6 for specificity, sensitivity and accuracy, respectively. 28.6% of the dataset was used as a training set, while the remaining 72.4% was used for testing. If more data had been used for training, that could improve the performance.

Mahendran Gandhi et al. [6] used a gray-level co-occurrence matrix (GLCM) to extract the input features that feed the SVM classifier. They tried to build an automated method by using morphological operators and SVM classifiers on non-dilated color fundus retinal images to detect the exudates. The used images were five fundus images in JPEG format with size 2196 x 1958 pixels. The SVM classifier was used to examine the disease's severity, whether the effect on the patient's eye was moderate or severe. The used classifiers' results were that all five images were diagnosed with an abnormality, with three severely affected by exudates and two moderately affected.

Sohini Roychowdhury et al. [7] introduced a computer-aided screening system named DREAM. DREAM uses fundus images collected from two databases: the DIARETDB1 dataset and the MESSIDOR dataset in order to differentiate the DR images from the normal ones and generate a severity grade. Some classifiers were used, such as AdaBoost, Support Vector Machine (SVM), the Gaussian Mixture Model (GMM), and k-Nearest Neighbor (kNN). AdaBoost helped in the reduction of extracted features to 30 selected features out of 78. The feature reduction decreased the average computing time from 59.54s to 3.46s. The DREAM system achieved a sensitivity, specificity and AUC of 100%, 53.1% and 0.904, respectively.

Ahmed El Tanboly et al. [8] developed a DR detection system by using OCT images in three stages by using different segmentation and classification techniques. They extracted three main features to quantify the following from the segmented OCT images: "reflectivity", "curvature", and "thickness" of retinal layers. The segmented layers are characterized by a function used to describe the random distribution, called cumulative-probability distribution function (CDF), of its extracted features. The used classifier has been trained in order to select the distinctive features of retinal layers and to detect the DR by using their CDFs. The system results were 83%, 92% and 100% for sensitivity, accuracy and specificity, respectively.

Mohammed Ghazal et al. [9] proposed a CADe system for detecting NPDR in the early stages by using OCT images. The built system consists of four primary stages: preprocessing, feature extraction, system training, then diagnosis and testing. The preprocessing stage contains the segmentation of retinal OCT images into 12 layers. The extracted layers are aligned by using layer number 6 outer nuclear layer ONL as reference. The output of the preprocessing stage is fed into the convolutional neural networks (CNNs) for training and evaluation. The best results have been acquired by using the proposed CNN, where the sensitivity, specificity and accuracy were 100%, 88% and 94%, respectively. It has not been revealed how accurate the alignment with the y-axis is and how this affects the final result.

Peyman Gholami et al. [10] proposed an automated classification method to identify eyes with an ocular disease like DR, Age-related Macular Degeneration (AMD) or Macular Hole (MH), or normal eyes, from processing OCT images. The images were collected at Sankara Nethralaya (SN) eye hospital, Chennai, India. The images were preprocessed by removing the noise by using a wavelet-based denoising technique. Additionally, the images were down sampled from 512 x 1024 pixels into 500 x 750 pixels. The authors relied on extracting the Local Binary Pattern (LBP) features to feed the used classifiers. The used feature selection reduced the used features from 375000 to 16383 features. They used SVM, random forest and multiphase method classifiers. The achieved result from classifying the normal and abnormal images was AUC 98.6%, where the used multiphase classifier achieved AMD, DR, and MH AUC as 100%, 95% and 83%, respectively. The proposed system has perfect AMD detection, but it needs more improvement for detecting MH.

Muhammad Awais et al. [11] worked on a system for separating Diabetic Macular Edema (DME) OCT images from normal ones. The used images were collected from the Singapore Eye Research Institute (SERI). They used a pre-trained CNN, and the features were extracted at different layers by using a model involving visual graphic geometry with 16 weight layers (VGG16). They carried out four experiments with noise removal, image cropping, both and neither. The best results were obtained by using images with no noise removal or cropping and applying the kNN classifier. The results were 93% accuracy, 87% sensitivity and 100% specificity. The authors do not reveal the number of images used in the experiments.

Xuechen Li et al. [12] developed an automated system called "OCTD_Net" for separating DR images from normal ones by using OCT images. The proposed system classifies whether the image is normal or with DR and assigns value 1 for the DR patients with changes in the thickness and reflection of retinal layers and 0 for DR images that do not display any significant changes. The used features in their system were the optical reflection of retinal layers (gray-level intensity of OCT images) and the retinal layers' thickness (pixels). Softmax was the used classifier in their experiment. The system was found to have a sensitivity of 0.90, an accuracy of 0.92 and a specificity of 0.95. The advantage of this system compared with others is the ability to classify the severity of DR where present.

Khaled Alsaih et al. [13] proposed a system for separating DME OCT images from normal ones. They used 32 OCT volumes containing more than 3800 images. The used technique was based on the evaluation of intraretinal cystoid space formation, hard exudates, retinal thickening, and subretinal fluid. The used features were local binary pattern (LBP) features and extraction of a histogram of oriented gradients within a multiresolution approach, bag of words (BoW) representations and principal component analysis (PCA). The used classifiers were random forest and SVM. The achieved sensitivity and specificity were both 87.50%. Because of the missing detection of positive cases, the system is not reliable enough to be used in clinical applications.

Yo-Ping Huang et al. [14] proposed a method for detecting DR by using fundus images. They ranked the DR attributes by applying a fuzzy analytical network process from most to least DR-relevant factors. The transformed fuzzy neural network classifier was used to improve the classification process. The associated rules among the selected attributes of DR were extracted to reveal their degree of severity and importance. The used novel system with the newly proposed models B and C has improved the classification quality for both training and testing sets where the achieved AUC is 1.0.

Table 1. A comparison of the proposed methods by comparing the key feature(s) used and the used classifiers.

Table 1. Comparison between the previous works

#	Author s	Year	Key Features	Classifiers
				PNN
1	[5]	2013	Thresholding Morphological processing algorithms	Bayesian Classification
				SVM
2	[6]	2013	entropy, contrast, correlation, energy, homogeneity and dissimilarity	SVM
				GMM kNN SVM AdaBoost
3	[7]	2014	30 features used	
4	[8]	2017	Quantifying "reflectivity", "curvature", and "thickness"	Deep Fusion Classification Network (DFCN)
5	[9]	2019	Patches extraction	CNN
6	[10]	2018	LBP	random forest with SVM
7	[11]	2017	-	kNN
8	[12]	2019	Thickness and re- flection of retinal layers	Softmax
9	[13]	2017	LBP	PCA SVM
10	[14]	2019	-	Transformed Fuzzy Neural Network

3. METHODOLOGY

In our project, we used 214 OCT retinal images: 107 images were normal retina images, and the other 107 images were DR images. The images were selected to be 50% normal images and 50% DR images. The whole DR images were selected and the normal images were selected randomly from the dataset, whereas the first 107 images were selected. All used images were taken from

Scholars Portal Dataverse- University of Waterloo Dataverse- [15] - an open-source database that contains different retinal OCT images-. The OCT images database was published by a group of scholars from the University of Waterloo. The selected database contains OCT retinal images for different diseases. The images are in a high resolution (jpeg format) that can be downloaded and used with no need for preprocessing. The database contains more than 500 spectral-domain OCT volumetric scans, divided into five datasets: Normal 206 images, Age-related Macular Degeneration 55 images, Macular Hole 102 images, Diabetic Retinopathy 107 images and Central Serous Retinopathy 102 images [16]. The images were collected from a raster scan protocol with size of 512x1024 pixels. They were obtained from Sankara Nethralaya eye hospital, India [17].

The normal images and DR images were downloaded from the database [18] and 214 images were selected for the experiment. The dataset was divided into two groups: a group representing about 75% of the images was used for classifiers' training, while about 25% of images were used for testing. Five statistical features were used to feed the classifiers with the required information that helped distinguish between the DR and normal images. The used statistical features are the matrix's mean, standard deviation, mode, variance and median. Additionally, the mean of derivatives, the standard deviation of derivatives, mode of derivatives and median of derivatives were used. All these features were computed using MATLAB built-in functions. P-value for the selected features is less than 5% (P-value <0.05).

The results were not satisfactory, and the system needed further improvement. Therefore, local binary pattern (LBP) features were used, and the results displayed a significant improvement. The LBP features detect the uniform local binary pattern of textures in gray-scale images [19]. Rotationally invariant feature information was used. 8 neighborhoods for each pixel were used in the computation. There are many parameters for LBP features found in MATLAB, like the number of neighbors, radius of circular pattern, rotation invariance flag, type of normalization and interpolation method.

After adjusting the used features, the used classifiers SVM and kNN were also adjusted in order to improve the classifiers' performance by changing the SVM kernel function, polynomial-SVM kernel order, SVM optimization routine, kNN neighbors and kNN distance metric. Fig. 1 presents the flowchart for the used system. The used classifiers' functions are built-in functions in MATLAB.

In parallel to SVM and kNN classifiers, the Artificial Neural Network (ANN) has been used in order to validate the result by using different approach. The used ANN has 21 input layers, 2 hidden layers and each one has 21 layers, and 1 out layer.

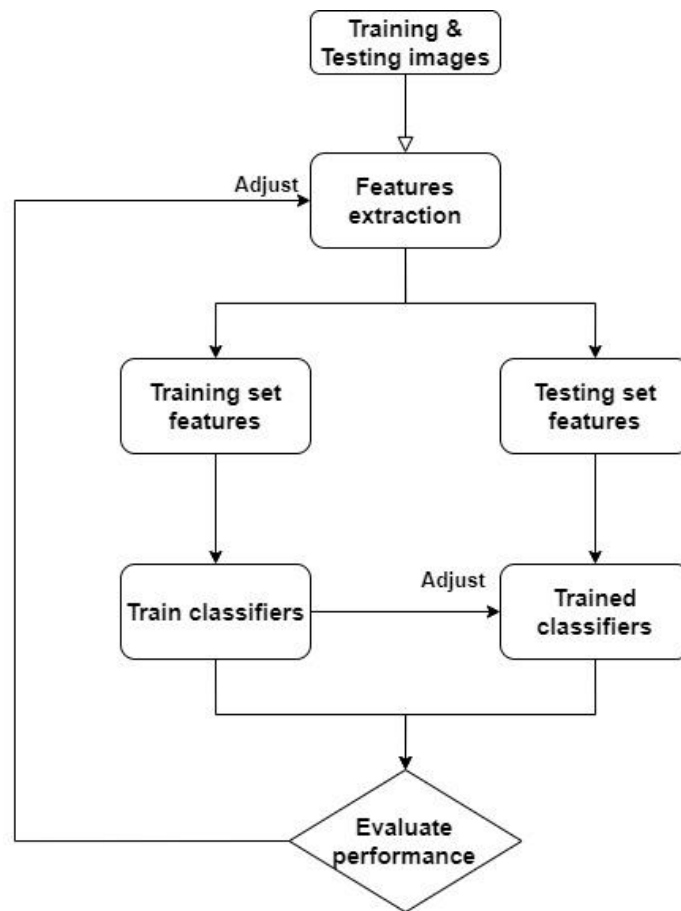


Figure 1. Flowchart of the used system

4. CLASSIFIERS

SVM is a widely-used supervised classifier developed by V. Vapnik and his team (AT&T Bell Labs.). SVM can be used to train the linear, neural network, polynomial, or radial basis functions classifiers [20]. SVM represents the data in space and gives binary classification to the classified data. In this study, our system assigns the value 0 to normal images and 1 for DR images. Different kernels can be used to enhance the classifier's performance.

kNN is one of the common pattern recognitions methods used to classify data. It depends mainly on nearby sample observation. The unclassified sample will be relying on the class of the class represented by the majority of its k nearest neighbors in the training set [21].

ANN classifier is a model built to simulate the biological nervous system in humans. In the biological neural network, the dendrites in neurons receive the signal and the axon transmit it to other neurons [22]. In the ANN model, the input data is transmitted to hidden layers that contain many neural networks. After the processing stage in hidden layers, the output is collected from the output layer that has a prediction to the received information in the first stage. The hidden layers are sets of weighted inputs and produce an output through an activation function. This recurrent process extracts the features and generates the final output, which is the image classification in this paper.

A summary of the used classifiers is found in Table 2.

Table 2. Summary of the used classifiers

Classifier	Parameters
Support Vector Machine	Linear, Radial Basis Function (RBF) and Polynomial
k-Nearest Neighbour	The number of nearest neighbours in the predictors are 1, 3 and 5.
NN	Artificial Neural Network

5. RESULTS

The following performance parameters have been used in a confusion matrix in order to evaluate the classifiers' performance:

- True Positive (TP): event values with a correct prediction.
- False positive (FP): event values with incorrect predictions.
- True negative (TN): no-event values with a correct prediction.
- False negative (FN): no-event values with incorrect prediction.

The Accuracy (ACC) is the ratio of the total number of correct predictions, and it can be calculated by the equation:

$$ACC = (TP+TN)/(TP+TN+FP+FN) \quad (1)$$

The Sensitivity (Sens) is the ratio of positive cases that were correctly identified, as determined using the equation:

$$Sens = TP/(TP+FN) \quad (2)$$

Specificity (Spec) is the ratio of actual negative cases that were correctly predicted, as determined using the equation:

$$Spec = TN/(TN+FP) \quad (3)$$

Receiver Operator Characteristic (ROC) curve is a fundamental graphical tool used to show the diagnostic performance of an operator (physician or machine) in two or more class problems. The Area Under the ROC Curve (AUC) is commonly computed to give a sense of how good the overall performance of the operator is over all the cases diagnosed. Table 3 presents the achieved classifiers' performance.

Table 3. The classifiers' performance of training and testing datasets

Classifier	Image group	Sens%	Spec%	Acc%	AUC%
Linear-SVM	Training	100	100	100	100
	Test	100	100	100	100
RBF-SVM	Training	100	100	100	100
	Test	100	100	100	100
Polynomial-SVM	Training	100	100	100	100

(Kernel order 2)	Test	100	96.43	98.15	98.15
	Training	100	100	100	100
Polynomial-SVM (Kernel order 3)	Test	100	96.43	98.15	98.15
	Training	100	100	100	100
kNN (k=1)	Test	100	100	100	100
	Training	97.53	98.73	98.13	98.13
kNN (k=3)	Test	100	100	100	100
	Training	100	100	100	100
kNN (k=5)	Test	100	100	100	100
	Training	100	100	100	100
ANN	Test	100	100	100	100
	Training	100	100	100	100

6. DISCUSSION

This paper discussed an automated system for detecting DR disease by using retinal OCT images. Using the technology in the medical field can shorten the time, improve the quality and decrease errors. Using the CADe system in diagnosing DR will assist in detecting the disease in the early stages. It gives ophthalmologists the opportunity to treat or control the disease and save the patients' vision loss. The proposed system is capable of separating the OCT normal images from the DR images with 100% accuracy.

The proposed system achieved 100% accuracy, which means all images classified as DR images are valid abnormal cases. The system has 100% sensitivity, whereby it can detect all positive cases without neglecting any positive cases or classifying the negative case as positive. In addition, the system has the same performance in regard to detecting negative cases, where it achieved 100% specificity. The proposed algorithm was executed in approximately 24 seconds by a personal computer. The computer has 16GB RAM, Intel i7 with x64-bit processor with a speed of 1.8GHz.

7. CONCLUSION

Diabetic retinopathy disease is one of the leading causes of vision loss. Early DR detection prevents disease progression and helps ophthalmologists to treat or control the disease. In this study, we have proposed an automated system that helps ophthalmologists to better diagnose DR from OCT images. The proposed system is implemented in MATLAB R2018a. The system displayed better performance than previous studies: the accuracy, sensitivity, specificity and AUC were 100% by using SVM, kNN and ANN classifiers, as shown in Table 3. We found that using features for detecting the images' texture, such as the used features in our study (LBP features) significantly improved the system performance because the retina OCT images have common texture patterns.

We recommend further work to be applied on a vast number of images with more classification levels by separating the normal OCT images from the abnormal ones and classifying the found abnormality from ocular diseases. This work needs to use more sophisticated techniques, like

deep learning, in order to deal with a large amount of data. Our study is limited to the collected images from the University of Waterloo. Therefore, we recommend applying the used approach to different images group.

LIST OF ABBREVIATIONS

OCT: Optical Coherence Tomography; DR: Diabetic Retinopathy; LBP: local binary pattern; ROC: receiver operating characteristic; AUC: Area Under Curve; CADE: computer-aided detection; SVM: support vector machine; kNN k-Nearest Neighbor.

REFERENCES

- [1] R. P. Singh, M. J. Elman, S. K. Singh, A. E. Fung, and I. Stoilov, "Advances in the treatment of diabetic retinopathy," *Journal of Diabetes and its Complications*, p. 107417, 2019.
- [2] A. C. S. Tan et al., "An overview of the clinical applications of optical coherence tomography angiography," *Eye*, vol. 32, no. 2, pp. 262-286, 2018/02/01 2018.
- [3] D. Huang et al., "Optical coherence tomography," *science*, vol. 254, no. 5035, pp. 1178-1181, 1991.
- [4] T. E. de Carlo et al., "Detection of microvascular changes in eyes of patients with diabetes but not clinical diabetic retinopathy using optical coherence tomography angiography," *RETINA*, vol. 35, no. 11, pp. 2364-2370, 2015.
- [5] R. Priya and P. Aruna, "Diagnosis of diabetic retinopathy using machine learning techniques," *ICTACT Journal on soft computing*, vol. 3, no. 4, pp. 563-575, 2013.
- [6] M. Gandhi and R. Dhanasekaran, "Diagnosis of diabetic retinopathy using morphological process and SVM classifier," in *2013 International Conference on Communication and Signal Processing*, 2013, pp. 873-877: IEEE.
- [7] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "DREAM: diabetic retinopathy analysis using machine learning," *IEEE journal of biomedical and health informatics*, vol. 18, no. 5, pp. 1717-1728, 2013.
- [8] A. ElTanboly et al., "A computer-aided diagnostic system for detecting diabetic retinopathy in optical coherence tomography images," *Medical physics*, vol. 44, no. 3, pp. 914-923, 2017.
- [9] M. Ghazal, S. Ali, A. Mahmoud, A. Shalaby, and A. S. El-Baz, "Accurate detection of non-proliferative diabetic retinopathy in optical coherence tomography images using convolutional neural networks," *BioRxiv*, p. 667865, 2019.
- [10] P. Gholami, M. S. Hassani, M. K. Parthasarathy, J. S. Zelek, and V. Lakshminarayanan, "Classification of optical coherence tomography images for diagnosing different ocular diseases," in *Multimodal Biomedical Imaging XIII*, 2018, vol. 10487, p. 1048705: International Society for Optics and Photonics.
- [11] M. Awais, H. Müller, T. B. Tang, and F. Meriaudeau, "Classification of sd-oct images using a deep learning approach," in *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, 2017, pp. 489-492: IEEE.
- [12] X. Li, L. Shen, M. Shen, F. Tan, and C. S. Qiu, "Deep learning based early stage diabetic retinopathy detection using optical coherence tomography," *Neurocomputing*, vol. 369, pp. 134-144, 2019.
- [13] K. Alsaih, G. Lemaitre, M. Rastgoo, J. Massich, D. Sidibé, and F. Meriaudeau, "Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images," *Biomedical engineering online*, vol. 16, no. 1, p. 68, 2017.
- [14] Y.-P. Huang, H. Basanta, T.-H. Wang, H.-C. Kuo, and W.-C. Wu, "A Fuzzy Approach to Determining Critical Factors of Diabetic Retinopathy and Enhancing Data Classification Accuracy," *International Journal of Fuzzy Systems*, vol. 21, no. 6, pp. 1844-1857, 2019.
- [15] Optical Coherence Tomography Image Retinal Database. Available: <https://dataverse.scholarsportal.info/dataverse/OCTID>
- [16] P. Gholami, P. Roy, M. K. Parthasarathy, and V. Lakshminarayanan, "OCTID: Optical coherence tomography image database," *Computers & Electrical Engineering*, vol. 81, p. 106532, 2020.
- [17] P. Gholami, M. Kuppuswamy Parthasarathy, P. Roy, and V. Lakshminarayanan, "OCTID citation," V1 ed: Scholars Portal Dataverse, 2018.
- [18] J. J. Salazar et al., "Anatomy of the Human Optic Nerve: Structure and Function," 2018.

- [19] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on pattern analysis and machine intelligence, vol. 24, no. 7, pp. 971-987, 2002.
- [20] E. Osuna, R. Freund, and F. Girosit, "Training support vector machines: an application to face detection," in Proceedings of IEEE computer society conference on computer vision and pattern recognition, 1997, pp. 130-136: IEEE.
- [21] T. Denoeux, "A k-nearest neighbor classification rule based on Dempster-Shafer theory," IEEE transactions on systems, man, and cybernetics, vol. 25, no. 5, pp. 804-813, 1995.
- [22] Y. Narayan, "Hb vsEMG signal classification with time domain and Frequency domain features using LDA and ANN classifier," Materials Today: Proceedings, vol. 37, pp. 3226-3230, 2021.

AUTHORS

Eng. Marwan Aldahami M.Sc. Graduate Student of Biomedical Engineering at the Dept. of Electrical and Computer Engineering, King Abdulaziz University, Jeddah, Saudi Arabia. Bachelor's degree in Biomedical Technology from King Saud University, Riyadh, Saudi Arabia.



Dr. Umar S. Alqasemi, Associate Professor of Biomedical Engineering at the Dept. of Electrical and Computer Engineering, King, Abdulaziz University, Jeddah 21589, Saudi Arabia. PhD and MSc degree in Biomedical Engineering from UConn, Storrs, USA. Research work in ultrasound, optical, and photoacoustic imaging, medical imaging recognition, bioelectronics, and digital and analog signal and image processing.

