Detection of Diabetes Mellitus Using Iris Images and a Multi-Feature SVM Approach: A Pilot Study

Basavaraj Hiremath[1], M. Lakshminarayana[2], B. Dhananjay[3], Dr. C K Narayanappa[4]

[1,2,3]Assistant Professor, [4]Professor & HOD, [1,2,3,4]Department of Medical Electronics Engineering,

[1,2,3,4]Ramaiah Institute of Technology, Bangalore, Karnataka, India.

Abstract

Iris based diagnosis is one of the most important approach used in naturopathy to detect various diseases with the aid of different attributes such as those of iris pattern, color and other characteristics. The present work emphasized the identification of diabetic condition based on the attributes of the iris images in an early stage so as to ascertain a subjective risk of an onset of Diabetic condition in the subjects, in the near future.

To assess the preciseness of the software approach developed using image processing and pattern classification, 60 iris images (30 normal and 30 diabetic subjects) were used, which were available in an online database (published by Indian Institute of Technology, Delhi, India). These images were acquired from a clinical grade iris cope. The software paradigm developed included various image processing approaches implemented upon the iris images such as those of preprocessing (noise removal and image enhancement), Image segmentation and normalization for suitable feature extraction (contrast, dissimilarity, difference entropy, homogeneity and difference variance) from the gray level co-occurrence matrix. These features were then fed to the support vector machine (SVM) classifier. 46 images (23 normal and 23 diabetic) were used for training the classifier and 14 images (7 normal and 7 diabetic) images were used to test the same. An overall accuracy of 90% was achieved using the presently developed paradigm. The entire process was developed using MATLAB software tool. Such a non-invasive approach could be extremely useful in the assessment of a plausible early onset of diabetic conditions in the subjects. Further, various other approaches could be developed to achieve a higher accuracy of diabetic classification in the near future.

Keywords*: Iris based diagnosis, Diabetic, Image processing, Classification, Support vector machine.*

Introduction

The iris of human beings provides an insight into the well-being of the body. It is the sole organ originating from the germ layers during embryogenesis, despite being externally visible [1]. The science of the study of iris, termed as iridology relates different changes observed in the iris to be a cue to the variations in the respective internal organs of the human body, on a case-to-case basis [2]. Prominent changes such as those pertaining to discoloration of the lacunae or clouds in the iris are often observed with aging factor. Such changes are analyzed so as to ascertain the physical health of the individual and also to suggest appropriate measures to counter numerous chronic pathological aspects [3].

The present research work proposes a quantitative non-invasive approach with the aid of iris images for the detection of diabetic mellitus at various stages for which the image of the iris is acquired and divided into a standard sixty sectors with every sector being associated with a given organ in the body. Efficient features are extracted from these images for further processing in order to ascertain the onset of diabetes mellitus as well as the pre-diabetic conditions along with the presence of diabetes mellitus in the subjects considered. This is depicted in figure 1, Type 2 Diabetes mellitus, also termed as T2DM is associated with hyperglycemia due to the actions pertaining to insulin in the body. As per the latest health report published by the WHO (2020), over 7% of the population seems to be affected by this disorder worldwide [4]. Chronic versions of T2DM is known to result in various long term disorders with organ dysfunctions with certain life threatening conditions in terms of hyperglycemia. Such subjects are also known to be affected by various cardiovascular disorders as well. It is hence inferred that T2DM needs to be managed with appropriate life style oriented modifications to benefit the affected subjects [5].

Materials and methods

The iris images have been acquired using a standard Iris scope with a 5-megapixel clarity with a 30X magnification for the screening of diabetic subject. Preprocessing as well as enhancement of the images are performed so as to improve the quality of the images. As shown in figure 1, the region of interest in the iris image for diabetic condition is the 36-43 minute of the right iris corresponding to liver. Mapping from a Cartesian to polar coordinate system aided in feature extraction. The spatial features were obtained from the gray level co-occurrence matrix. Among them Contrast, Correlation, Dissimilarity, Homogeneity, Difference variance and Difference Entropy based features were considered due to minimum overlap. The software tool used for this work was MATLAB R2017a.

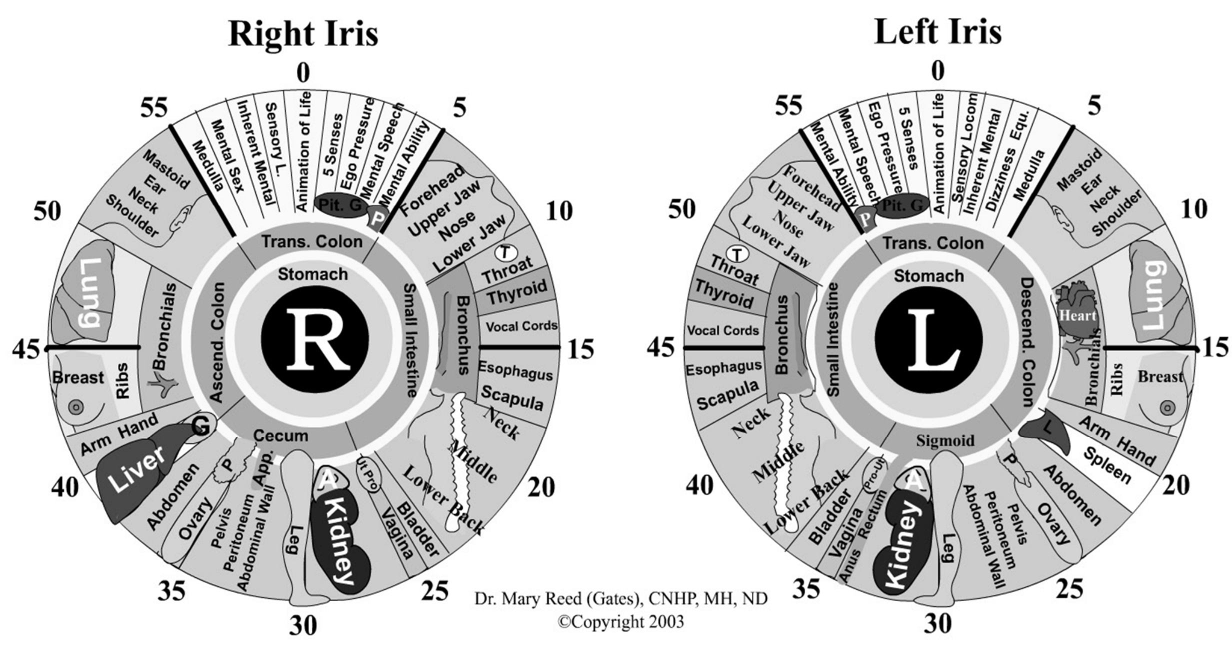

Figure 1: Sector based depiction of iris

Experimental paradigm

The entire process flow is depicted in figure 2.

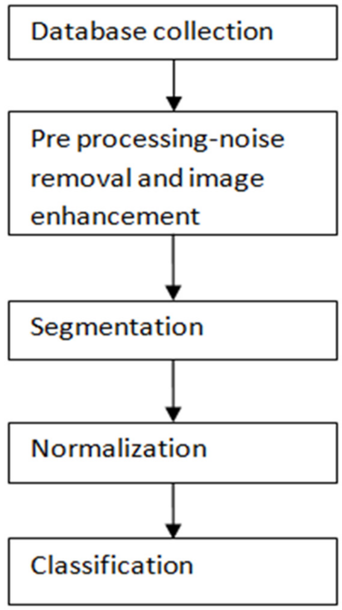

Figure 2: Proposed methodology for the iris based diabetic detection

Data acquisition

An Iris scope was used to acquire the images of iris from T2DM affected and control set of subjects. A magnification factor of 30X was preset incorporating an illumination contour with the aid of an LED light source (having a resolution of 2560 x 1920). 60 images were obtained (30 conditioned and 30 control set of images) for this study. A sample Iris image is as shown in figure 3.

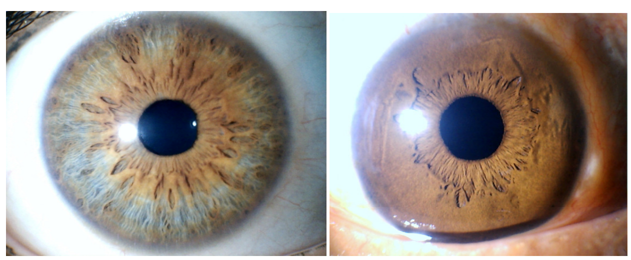

Figure 3: Iris images acquired from an iris scope

Preprocessing

Preprocessing involves performing certain operations on an image in order to enhance or extract required information from the image. It involves steps such as removal of noise and image enhancement to render an image more suitable than the original for processing. Preprocessing of images helps to address the issues of noise and reflections in the image.

Segmentation of images

Segmentation often relates to a process of ghettoizing a given image with regard to the constituent of the same as required. This is accomplished by isolating the iris from the rest of the eye followed by segmenting out the components related to liver (ie 36 to 43 minute of the iris) as mentioned in the standard Iris chart (Theodore Kriege). In the present context, the area between 36’-43’ of 2^{nd} and 3^{rd}major zone of the right iris were segmented [5], as shown in figure 4.

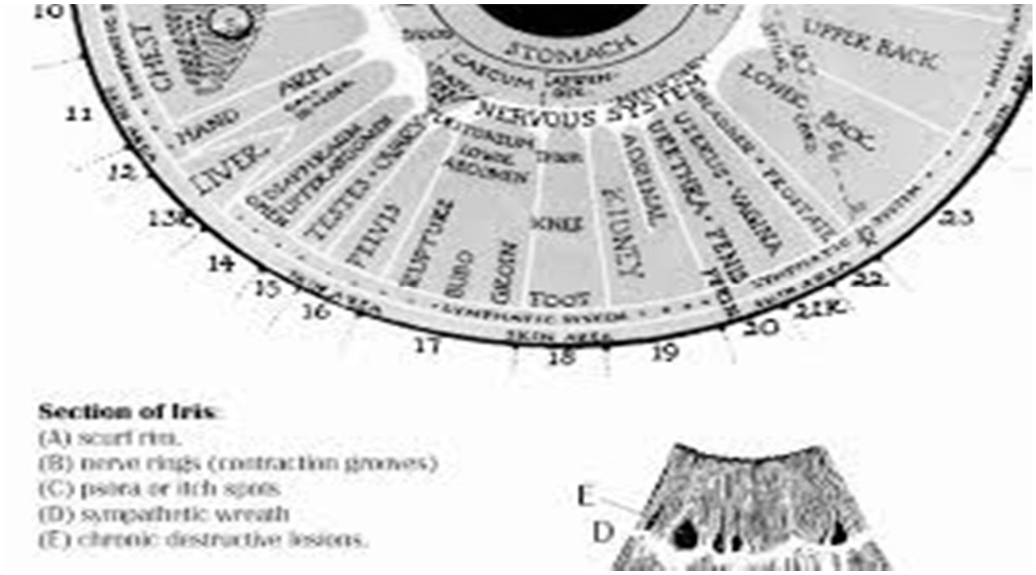

Fig .4. Chart of the iris depicting the patterns and the zones of liver

Steps involved in the process of segmentation

- Binary thresholding: Conversion of a gray scale image into its binary equivalent
- Pupil identification: the maximum area of the binary image is determined as the pupil
- Thresholding: The most frequency seen radius is termed to be the pupil radius, the center of which corresponds to the iris center.
- $Center\ of\ the\ pupil = centroid\ of\ the\ segmented\ image\ x\ radius = mean\frac{(major\ axis+minor\ axis)}{2}$
- Determination of the outer boundary is achieved by scanning from the centre until the point of occurrence of the threshold value
- Masking is done to mask out the area of the iris images at the boundaries as well as that of the pupil.
- Region of interest is ascertained by encompassing the sector of the iris area corresponding to that of the liver with the aid of masking.

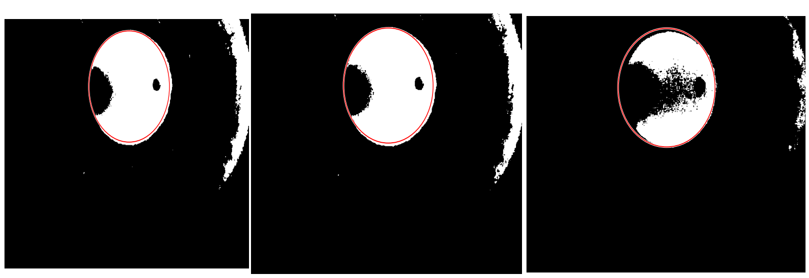

Fig.5 Iterations of Thresholding

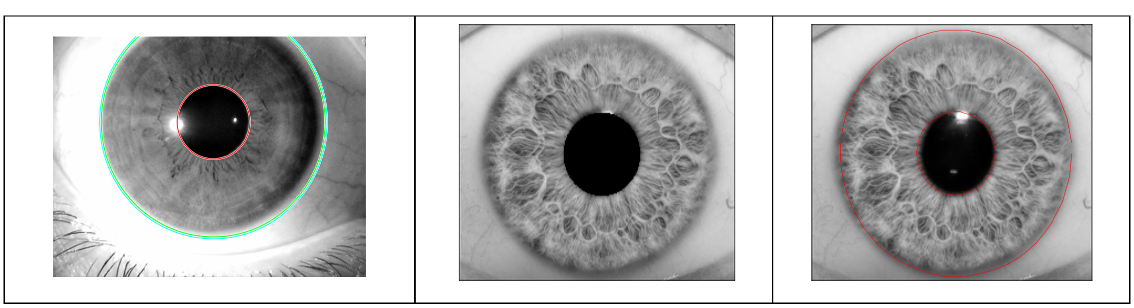

Fig.6: Iris detection

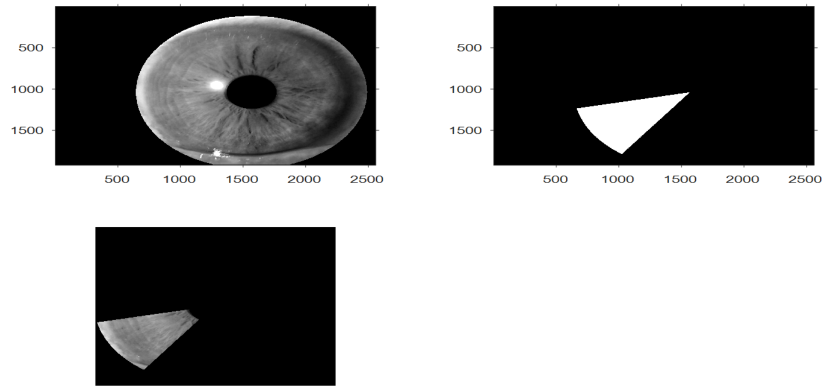

Fig.7 Region of interest of the iris image corresponding to the liver

Normalization

Normalization converts the circular coordinate system into rectangular axis using polar mapping approach, as shown in figure 8.

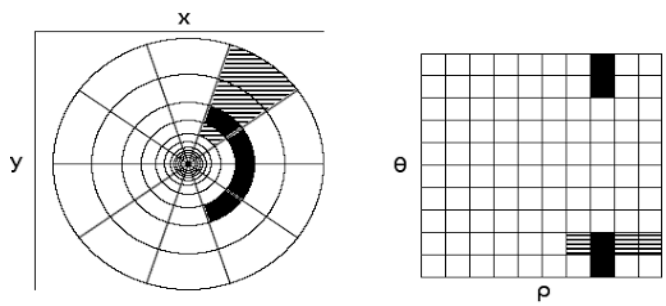

Figure 8: Illustration of polar mapping

Figure 9: Normalization of Segmented ROI Sector

Extraction of the desired features

Feature extraction depicts a process of comprehending the visual aspects of a given image so as to quantify as well as retrieve the same, as required. This aids in the representation of the underlying characteristics of the texture of the given images in a much simplified approach. Such a representation will help to increase the accuracy of classification of the given image datasets. The present work confined to the extraction of image features such as those of difference entropy, contrast, homogeneity, correlation, difference variance and dissimilarity. This was accomplished using the Gray Level Co-occurrence Matrix (GLCM) technique. These features were further used to classify the given subjects into Diabetic (Conditioned) and non-diabetic (Control) datasets.

Contrast

Contrast quantifies the variations in the local gray levels of a given image. A higher contrast value is an indicator of the presence of noise/edges in the image as is given as follows

$$\sum_{n=0}^{Ng-1}(i-j)^2\{\sum_{i=1}^{Ng}\sum_{j=1}^{Ng}p(i,j)\}$$

Where,

p(i,j) = (i,j)th entry of normalized GLCM

Ng = total number of gray levels in the image

Correlation

Correlation denotes the dark levels in relation with the pixels at predefined positions in the image and is given as follows.

$$\sum_{i=1}^{Ng}\sum_{j=1}^{Ng}\frac{ijp(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y}$$

Where

μ_x = Mean of the row sum of the GLCM

μ_y= Mean of the column sum of the GLCM

σ_x = Standard deviation of the row sum of the GLCM

σ_y= Standard deviation of the column sum of the GLCM

Difference entropy

This entity provides an insight into the disorders pertaining to the gray level difference distribution of a given image and is given as follows

$$-\sum_{i=0}^{Ng-1} p_{x-y}(i)\log\{p_{x-y}(i)\}$$

Difference variance

This provides the dispersion of gray level difference distribution of a given data and is given as follows

$$\text{Difference variance}=\sum_{i=2}^{2Ng}(i-[\sum_{i=2}^{2Ng} ip_{x-y}(i))$$

Where

$$p_{x-y}(k)=\sum_{i=1}^{Ng}\sum_{j=1}^{Ng}p(i,j) \text{ where k=0,1,2,3… Ng-1}$$

Homogeneity

Homogeneity is the measure of smoothness of the gray level distribution of the image

$$\sum_{i=1}^{Ng}\sum_{j=1}^{Ng}p(i,j)/(1+(i-j)^2)$$

Classification

Based on the features selected classification was performed to differentiate Diabetic and Non Diabetic images. Support vector machine (SVM) was used which is a supervised learning algorithm for classification and regression problems. It operates by plotting each feature on n dimensional space with the features being the value of the coordinates and then approximates a hyper plane which best segregates these features into classes. In cases where linear hyper plane cannot be used to separate features into classes kernels are employed. Kernels are function which transforms a low dimensional input space into a higher dimension. They convert a non-separable problem into a separable one.

Support Vector Machine (SVM) is employed to classify a given set of data based on the hyperplane developed with the of a pre-defined training dataset. SVM, being a linear classifier, can be transformed so as to cater to various non-linear set of data with the aid of different types of kernels. The present work encompasses the deployment of linear, quadratic and Random Basis Function (RBF) kernels for classification, as shown below.

$$K(X_i , Y_i) = X_i^T Y_i + C$$

$$K(X_i , Y_i) = 1 - \frac{||X_i - Y_i||^2}{||X_i - Y_i||^2 + C}$$

$$d(X) = \sum_{N=1}^{svnum} \alpha\, y\, K(X_i , Y_i) + b$$

$$\text{Where } K(X_i , Y_i) = e^{-\frac{||X_i - Y_i||^2}{2\sigma^2}}$$

Where

X_i = i^{th} vector in the input (X is the input data vector)

Y_i = the class to which the element belongs to

C = Capacity constant

α = hyperparameter

y = entire class

b = bias

In the present approach, the Kernel employed was Gaussian kernel (which is a type of RBF kernel). Training set included 46 images with 13 diabetic images and 13 normal images and testing was performed with 14 images. An overall accuracy of 90% was obtained.

Results and Discussions

The features obtained from the iris images are tabulated in table 1, which depicts the difference between the features with respect to the conditioned and the control set of datasets.

Table.1. Dominant GLCM features selected for classification

Feature	Diabetic		Non-Diabetic	
	max	min	max	min
Contrast	0.037888	0.00983	0.834701	0.158757
Correlation	0.992481	0.960522	0.927481	0.556674
Dissimilarity	0.035163	0.009833	0.484043	0.196118
Homogeneity	0.995083	0.964799	0.932123	0.778952
Difference variance	0.037888	0.009833	0.834701	0.158757
Difference entropy	0.160931	0.055234	0.956693	0.426146

Table 2 provides the confusion matrix obtained for the testing and training data using SVM classifier with Gaussian kernel.

Table 2 Confusion matrix

	Diabetic	Normal
Diabetic	29	1
Normal	4	26

Major findings

The present work highlighted the importance of segmentation technique as applied to iris image based assessment of T2DM disorder wherein the thresholding was continuously iterated within the range of 0.08 to 0.3 with a step of 0.01 so as to encompass the entire dataset after which the most frequently occurring radius was considered as the pupil radius. This was followed by masking of pupil and the sclera to obtain the Iris region.

Conclusions and Future work

The present work highlights the importance of dominant features like contrast, correlation, dissimilarity, homogeneity, variance and entropy which have been used for the classification. SVM classification was used for the classification. Through the available dataset and using SVM classifier, the developed algorithm was able to classify diabetic or not with an accuracy of 90%.

REFERENCES

[1] Kumar, A., Gangwar, R., Ahmad Zargar, A., Kumar, R. and Sharma, A., 2024. Prevalence of diabetes in India: A review of IDF diabetes atlas 10th edition. Current diabetes reviews, 20(1), pp.105-114.

[2] Önal, M.N., Güraksin, G.E. and Duman, R., 2023. Convolutional neural network-based diabetes diagnostic system via iridology technique. Multimedia tools and Applications, 82(1), pp.173-194.

[3] Esteves, R.B., Morero, J.A.P., de Souza Pereira, S., Mendes, K.D.S., Hegadoren, K.M. and Cardoso, L., 2021. Parameters to increase the quality of iridology studies: A scoping review. European Journal of Integrative Medicine, 43, p.101311.

[4] Ogurtsova, K., Guariguata, L., Barengo, N.C., Ruiz, P.L.D., Sacre, J.W., Karuranga, S., Sun, H., Boyko, E.J. and Magliano, D.J., 2022. IDF diabetes Atlas: Global estimates of undiagnosed diabetes in adults for 2021. Diabetes research and clinical practice, 183, p.109118.

[5] Chaskar, U.M. and Sutaone, M.S., 2012, January. On a methodology for detecting diabetic presence from iris image analysis. In 2012 International Conference on Power, Signals, Controls and Computation (pp. 1-6). IEEE.

[6] Yang, S. and Deng, A., 2023. Fast and accurate computation of polar harmonic Fourier moments for image description. JOSA A, 40(9), pp.1714-1723.

[7] Das, D. and Naskar, R., 2024. Image splicing detection using low-dimensional feature vector of texture features and Haralick features based on Gray Level Co-occurrence Matrix. Signal Processing: Image Communication, 125, p.117134.

[8] Jan, F., Alrashed, S. and Min-Allah, N., 2024. Iris segmentation for non-ideal Iris biometric systems. Multimedia Tools and Applications, 83(5), pp.15223-15251.

[9] Daugman, J., 2003. The importance of being random: statistical principles of iris recognition. Pattern recognition, 36(2), pp.279-291.

[10] Prasad, G., Gaddale, V.S., Kamath, R.C., Shekaranaik, V.J. and Pai, S.P., 2024. A study of dimensionality reduction in GLCM feature-based classification of machined surface images. Arabian Journal for Science and Engineering, 49(2), pp.1531-1553.

[11] Malgheet, J.R., Manshor, N.B. and Affendey, L.S., 2021. Iris recognition development techniques: a comprehensive review. Complexity, 2021(1), p.6641247.

[12] Jan, F. and Min-Allah, N., 2020. An effective iris segmentation scheme for noisy images. Biocybernetics and Biomedical Engineering, 40(3), pp.1064-1080.

[13] Abdul Basith Khan, M., Hashim, M.J., King, J.K., Govender, R.D., Mustafa, H. and Al Kaabi, J., 2020. Epidemiology of type 2 diabetes—global burden of disease and forecasted trends. Journal of epidemiology and global health, 10(1), pp.107-111.

[14] American Diabetes Association Professional Practice Committee and American Diabetes Association Professional Practice Committee:, 2022. 2. Classification and diagnosis of diabetes: Standards of Medical Care in Diabetes—2022. Diabetes care, 45(Supplement_1), pp.S17-S38.

[15] G. T. Orpilla, J. R. C. M. Salazar and N. B. Linsangan, "Raspberry Pi-Based Dual Iris Recognition System Using Coiflet Wavelet," 2024 4th International Conference on Neural Networks, Information and Communication Engineering (NNICE), Guangzhou, China, 2024, pp. 634-638.

[16] S. N. Pasha, A. D, R. Anbunathan, G. C and M. G K, "An Efficient Novel Approach for Iris Recognition and Segmentation Based on the Utilization of Deep Learning," 2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT), Bengaluru, India, 2023, pp. 1-5.

[17] M. J. Ferdous, N. A. Turzo and P. Sarker, "Iris-Pupil Thickness based Biometric Approach for Determining Age Group using Linear SVM," 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), Dhaka, Bangladesh, 2020, pp. 275-279.

[18] M. A. E. Abdalla, A. A. Abdo and A. O. Lawgali, "Utilizing Discrete Wavelet Transform and Discrete Cosine Transform for Iris Recognition," 2020 20th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), Monastir, Tunisia, 2020, pp. 283-286.