Deep learning based biometric authentication using electrocardiogram and iris

Ashwini Kailas, Geevagondanahalli Narayanappa Keshava Murthy
Department of Electronics and Instrumentation Engineering, Siddaganga Intitute of Technology, Visveswaraya Technological University, Tumakuru, India

Article Info

ABSTRACT

Authentication systems play an important role in wide range of applications. The traditional token certificate and password-based authentication systems are now replaced by biometric authentication systems. Generally, these authentication systems are based on the data obtained from face, iris, electrocardiogram (ECG), fingerprint and palm print. But these types of models are unimodal authentication, which suffer from accuracy and reliability issues. In this regard, multimodal biometric authentication systems have gained huge attention to develop the robust authentication systems. Moreover, the current development in deep learning schemes have proliferated to develop more robust architecture to overcome the issues of tradition machine learning based authentication systems. In this work, we have adopted ECG and iris data and trained the obtained features with the help of hybrid convolutional neural network- long short-term memory (CNN-LSTM) model. In ECG, R peak detection is considered as an important aspect for feature extraction and morphological features are extracted. Similarly, gabor-wavelet, gray level co-occurrence matrix (GLCM), gray level difference matrix (GLDM) and principal component analysis (PCA) based feature extraction methods are applied on iris data. The final feature vector is obtained from MIT-BIH and IIT Delhi Iris dataset which is trained and tested by using CNN-LSTM. The experimental analysis shows that the proposed approach achieves average accuracy, precision, and F1-core as 0.985, 0.962 and 0.975, respectively.

Corresponding Author:

Ashwini Kailas
Department of Electronics and Instrumentation Engineering, Siddaganga Intitute of Technology,
Visveswaraya Technological University
Tumakuru, India
Email: ashwini0429@gmail.com

1. INTRODUCTION

There is a tremendous growth in the field of security and privacy preserving techniques by means of authentication systems. These security systems are widely adopted in various real-time online and offline applications such as biomedical systems, cloud computing, computer vision. Generally, these authentication systems are classified as password-based authentication, multifactor authentication, certificate-based authentication and token-based authentication. However, these systems have their different applications, advantages and disadvantages. But when considering the human authentication system, then identifying the liveliness of human plays an important role, Therefore, in this work we focus on biometric authentication system. The term "biometrics" refers to our biological, behavioural, or physical traits, and it is seen as a legitimate substitute for passwords, signatures, and other forms of identification. Evidently, biometric

systems are designed to automatically recognise or validate certain persons based on their physical, physiological, or behavioural attributes, such as their iris, face, stride, keystroke dynamics, or other characteristics. In order to increase security, convenience, and social inclusion as well as to offer prospective applications in several scientific and industrial domains, biometric technologies are widely being implemented. The current advancements in machine learning techniques also increased the reliability of these authentication systems.

The traditional authentication systems were based on the unimodal systems which considered single biometric system for authentication however these systems suffer from various issues such as noisy data, inter-class similarity, intra-class variability in different applications. Some of unimodal authentication researches include face authentication [1], [2], Iris authentication [3], [4], ECG authentication [5], [6] and fingerprint authentication [7], [8]. However, these systems suffer from accuracy and afore mentioned related issues. Therefore, need of robust authentication system gained attention due their significant applications. In order to handle these issues, several researches have been presented which are mainly focused on developing the multimodal authentication system, for instance, Zhang *et al*. [9] presented a robust multimodal authentication system which is based on the voice and face data. Tarannum *et al*. [10] used a combination of IRIS, facial and fingerprint data to achieve the increased authentication performance. Currently, optimization-based methods also play important role in classification tasks by reducing the error in attributes and finding the best solution for dimension reduction and feature selection. Based on this concept, Sujatha *et al*. [11] used iris, finger vein, finger print data along with genetic algorithm-based optimization strategy. Similarly, the traditional classification methods suffer from several issues such as these methods rely on data pre-processing, overfitting, under fitting, speed, scalability and accuracy. Therefore, deep learning-based methods have gained huge attention in machine learning based applications. Hammad and Wang [12] presented a combination of ECG and fingerprint biometrics to develop the convolution neural network. Zhao *et al*. [13] used palm print and dorsal hand vein and introduced deep learning model for authentication. Alkeem *et al*. [14] adopted deep learning for ECG based authentication systems.

Electrocardiogram (ECG) verification, in comparison, is difficult to compromise without the user's knowledge. ECG biometric feature collection necessitates the use of specialized apparatus, such as an electrocardiograph, which makes this technique challenging to replicate. The ECG technique usually takes 10 seconds or longer to record ECG signals and reach an adequate degree of accuracy for authentication. ECG signal helps asses the heart's electrical conductivity as well as its cardio vascular alterations. There are two main types of data that an ECG provides. First, a doctor measures the time intervals on the ECG to ascertain how long the electrical wave takes to travel through the heartbeat. The amount of time it takes a wave to move from one area of the heart to another area of the body reveals whether electrical activity is regular, slow, quick, or irregular. Secondly, by monitoring the amount of electrical activity flowing through the heart muscle, a cardiologist may determine whether areas of the heart are too big or overworked. The normal ECG is depicted in Figure 1. According to this figure, we can observe several entities in the given ECG signal such as different waves such as P, Q, R, S and T wave. Based on this, several attributes can be extracted such as PR interval, ST interval, QRS interval and Q-T interval.
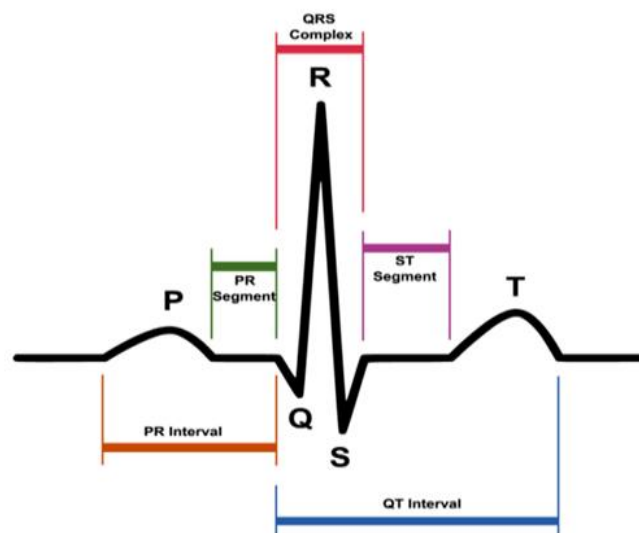

Figure 1. Normal ECG signal [15]

Similarly, iris is a small, annular structure in the eye, regulates the pupil's size and diameter, hence regulating how much light reaches the retina and Figure 2 shows a sample iris image which shows that complete eye image is comprised of pupil, iris, eyelid, collarette and sclera. Thus, identification of iris becomes an important aspect to perform various tasks on iris Therefore, in this work; we focus on developing a multimodal authentication system by combining ECG and iris image modalities.
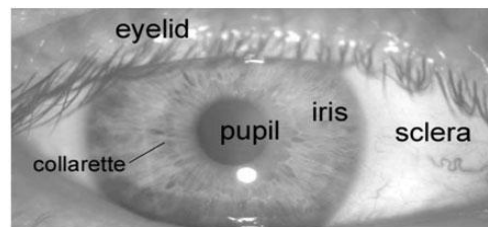

Figure 2. Eyelid image [16]

ECG and Iris-based authentication systems have their own set of challenges that must be overcome for their effective implementation. One of the primary challenges in ECG-based authentication systems is the need for high-quality ECG signals. Poor quality signals can lead to inaccurate detection of R-peaks, which can significantly impact the accuracy of the authentication system. Additionally, ECG signals can be affected by factors such as noise, motion artifacts, and electrode placement, which can further reduce the accuracy of the system. Iris-based authentication systems also face several challenges. One significant challenge is the need for high-quality iris images, as poor-quality images can result in inaccurate feature extraction. Furthermore, factors such as occlusion, dilation, and age-related changes in the iris can affect the accuracy of the authentication system. Another challenge is the need for proper illumination and focus during image capture, as variations in lighting and focus can impact the quality of the iris image.

The main aim of this research is to develop a novel and robust multimodal authentication system by considering ECG and Iris data. In order to achieve this, the proposed model uses feature extraction method where R peak detection and morphological feature extraction methods are employed on ECG signals. For iris, Gabor wavelet, gray level co-occurrence matrix (GLCM) and gray level difference matrix (GLDM) feature extraction methods are employed. Later, the obtained features are combined together to formulate the feature vector. Later, we trained a deep learning classifier by using convolutional neural network long short-term memory CNN-LSTM approach.

As discussed before, the multimodal authentication plays important role in various real-time security applications. Several biometrics combinations have been employed to improve the robustness of authentication systems where ECG and Iris based authentication systems have been adopted widely. In this section, we present a brief literature review about these multimodal authentication systems. Regouid *et al*. [17] focused on ECG, ear and Iris to develop a multimodal biometric authentication system to overcome the challenges of traditional unimodal systems. ECG facilitates the liveness information, ear biometrics helps to obtain the rich and stable information, and iris features helps to ensure the promising reliability and accuracy. This scheme performs normalization and segmentation as pre-processing step. Later, 1D-LBP, Shifted-1D-LBP and 1D-MR-LBP features are extracted from ECG signal. Ear and iris images are transformed into 1D signal. Finally, the K nearest neighbors (KNN) and radial basis function (RBF) classifiers are applied to classify the users as genuine or imposter.

Jiang *et al*. [18] focused on incorporating the authentication for body area sensor networks by combining iris and ECG features. This process follows two-level authentication where first level focus on iris authentication and later ECG authentication is performed to improve the overall security. El-Rahiem *et al*. [19] presented a multimodal authentication system by using ECG and finger vein. The complete process is divided into three stages as: pre-processing where data normalization and filtering techniques are applied, feature extraction phase uses deep CNN model for feature extraction from ECG and finger vein. Finally, different classifiers such as KNN, support vector machine (SVM), artificial neural network (ANN), random forest (RF) and naïve bayes (NB) are used to classify the obtained features. Moreover, this model uses multi-canonical correlation analysis (MCCA) to increase the speed of authentication.

Jadhav *et al*. [20] presented a multimodal biometric authentication system by using the combination of palm-print, iris, and face. The feature extraction process uses deep learning model for robust feature extraction from raw input images. Further, modified group search optimization (MGSO) approach is

employed to obtain the optimized and dimensionality reduced features. The classification phase uses teacher learning based deep neural networks (TL-DNN) model to reduce the classification error.

Hammad *et al*. [5] developed a combination of ECG and fingerprint-based authentication system. CNN model is used for generating the combined feature set and biometric templates are generated from these features. These features are trained and classified by using Q-Gaussian multi support vector machine (QG-MSVM) to increase the classification performance. Singh and Tiwari. [21] combined ECG, sclera, and fingerprint to develop a multimodal biometric system. This combination is carried out in two modules: i) decision-level fusion where combined Whale Optimization (WOA-ANN) is used to generate the sequential model and ii) score-level fusion model which uses salp swarm optimization-deep belief network (SSA-DBN) model.

Zhang *et al*. [9] developed face and voice based multimodal authentication system. This model uses local binary pattern (LBP) method for face feature extraction and voice activity detection (VAD) for voice sample analysis to increase the voice detection accuracy. This model also uses a feature fusion strategy to fuse the face and voice attributes efficiently.

Amritha *et al*. [22] emphasized on score level and feature level fusion to improve the accuracy of multimodal authentication system. This model includes ECG, face, and fingerprint for feature extraction and fusion. The score and feature level fusion are performed separately and later obtained scores are normalized by employing overlap extrema-based min–max (OVEBAMM) method. Huang *et al*. [23] reported that the performance of ECG based authentication system which is affected by the different types of noises and sample variation. To overcome these issues, authors presented local binary pattern-based feature extraction along with the robust multi-feature collective non-negative matrix factorization (RMCNMF) approach to address the noise and sample variation issues. This process helps to learn the latent semantic space with the help of Convolution non-negative matrix factorization (CNMF) method. Kumar *et al*. [24] focused on face and gait biometrics as multimodal authentication system. This approach uses principal component analysis (PCA) and deep neural network approach. The deep learning model replaced Euclidean distance with cross entropy function. The PCA is used for feature extraction and reconstruction of faces and DNN helps to improve the accuracy where final matching score is obtained by applying SoftMax function.

2. METHOD

This section presents the proposed solution for feature extraction and classification for multimodal authentication system. Figure 3 depicts the complete process of proposed approach The proposed ECG feature extraction phase includes deep learning-based R Peak detection and further, several morphological features [25] are obtained based on the R peak. Later, Iris feature extraction stage includes Gabor-wavelet, GLCM, GLDM and PCA based feature extraction. These features are further processed through the CNN-LSTM classifier module to authenticate the users.
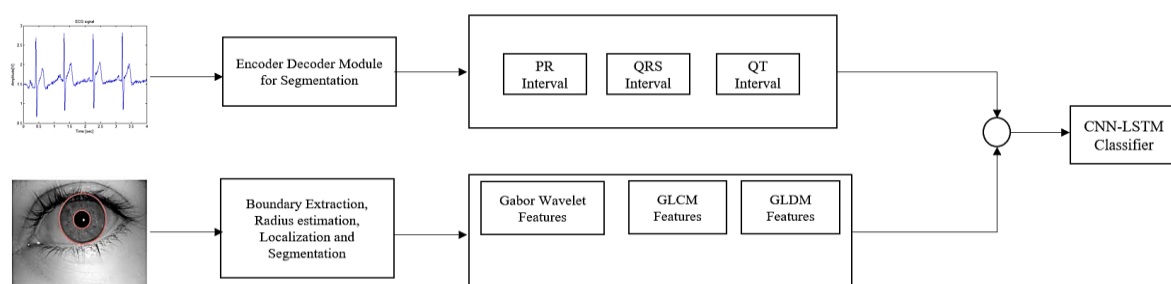

Figure 3. Complete architecture of proposed authentication system

2.1. ECG feature extraction

This section presents the proposed solution for ECG feature extraction. In ECG signal processing, R Peak detection plays important role therefore, we present the R peak detection method. Currently, several methods have been described to detect the R peaks. In this work, we adopted the deep learning method and present an encoder and decoder-based CNN architecture for peak detection by segmenting the input signal. Figure 4 depict the architecture of encoder and decoder module. The encoder path is known as contracting path which contains repeated convolution operations followed by rectified linear unit (ReLU) and max pooling. This structure is used to perform the down-sampling operations. Similarly, the expanding path performs up sampling operations followed by convolutions and ReLU operations.
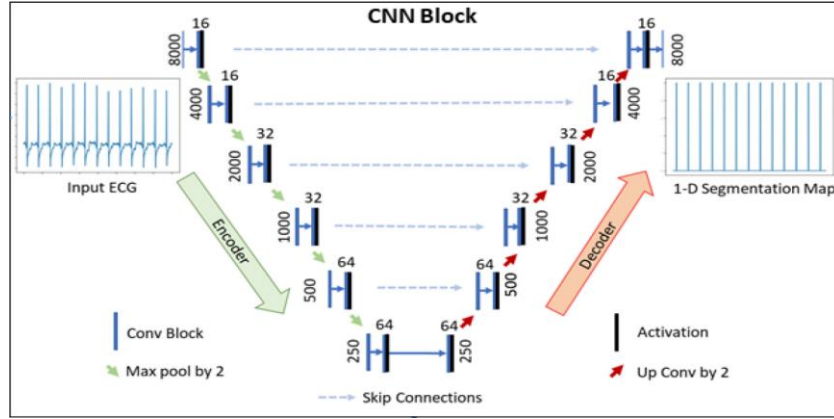
Figure 4. Encoder/decoder modules in CNNs

In this work, we formulated the peak detection problem as 1D segmentation problem for R peak segmentation by using deep learning approach. The input dataset contains i input ECG signals which are normalized between +1 and -1. The output of this module is obtained as 1D segmentation map which contains the R peak locations in the given signals. The segmentation map obtained from processing the signal through encoder and decoder module is expressed as (1):

$$z^{(i)} = E\left(x^{(i)}, \theta_1\right) \hat{y}^{(i)} = D\left(z^{(i)}, \theta_2\right) \tag{1}$$

Where E represents the encoder block, D represents the decoder block, θ_1 & θ_2 are the weight vector which are optimized by incorporating the binary cross entropy loss between actual and predicted R peaks. The loss function is given as (2):

$$Loss = \frac{1}{N}\sum_{i=1}^{N} y_i \times \log \hat{y} + (1 - y_i) \times \log(1 - y_i) \tag{2}$$

−   Based on the R peaks, we obtain RR interval which can be used to approximate the P and T waves. The T wave lie next to 1st R peak whereas P-wave is present in the 2nd R peak in the current RR interval.
−   In order to select the T-wave, 15% of RR interval is added to its 1st R Peak location and continue adding 55% of RR interval to the same location.
−   In order to select the P-wave, 65% portion of RR interval is added to the 1st R peak location and continues adding till 95% to the same location.
−   P and T wave peaks are identified based on the highest value in their corresponding window.
−   The Q peak is estimated by identifying the minimum value in window starting from 20ms before R peak.
−   In order to select the S peak, the lowest value in R Peak is identified from its peak to 20ms after the peak.
        Based on these peaks, we estimate several attributes such as PR interval, QRS detection, QT interval, QT corrected and vent rate (BPM). These parameters can be computed as follows:
−   PR Interval: it can be computed as (3):

$$t_{pr}(i) = \frac{\left(R_{loc(i)} - P_{loc(i)}\right)}{f_s} \tag{3}$$

Where f_s denotes the sampling frequency, R_{loc} represents the locations of R peak, P_{loc} represents the locations of P peaks
−   QRS duration: it is computed as (4):

$$t_{qrs}(i) = \frac{\left(S_{loc(i)} + x\right) - \left(Q_{loc(i)} - x\right)}{f_s} \tag{4}$$

Where x is the immediate 5 ms from signal which are added to the S_{loc} and subtracted from Q_{loc} because QRS duration is identified from start of Q till end of S peak.

−   QT interval: this interval can be computed as (5):

$$t_{qt}(i) = \frac{T_{loc(i)}+(t_{rr(i)}\times 0.13)-(Q_{loc(i)}-x)}{f_s} \tag{5}$$

Where T_{loc} shows the locations of T peaks, a constant 0.13 is multiplied with t_{rr}, and x is subtracted from the Q_{loc}

2.2. Iris feature extraction

This subsection describes he proposed approach for iris segmentation and feature extraction. The iris segmentation plays important role where we consider and inner and outer boundaries of iris. The Figure 5 depicts the inner and outer boundary region. In this, we extend our previous feature extraction technique for iris image authentication. Figure 5(a) depicts the original image whereas Figure 5(b) shows the identification of inner and outer boundaries which can be useful in segmentation process.
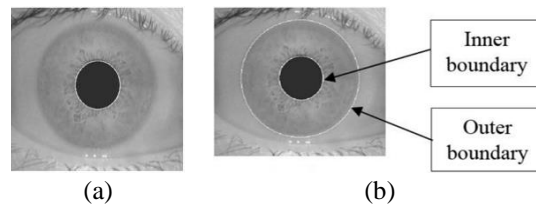

Figure 5. Inner and outer boundary representation, (a) original image and (b) boundary detection

For simplicity, it is assumed that the iris image acquisition devices extract the square region of image where center of iris image is closer to the center of square region. In order to obtain the boundary information, we initiate the iterative process from the center point of image which is denoted by $P(X_I, Y_I)$. Then, the profile operations are performed in vertical and horizontal directions which are denoted as V_P and H_P, respectively. The center point is denoted by $P(X_0, Y_0)$. In order to find the boundaries, we focus on estimating the radius which is computed as (6).

$$r_s^* = \arg \min_{r \in R} D_s(r) \tag{6}$$

Where D represents the left, right, top and bottom points of the image. Further, this image is processed to obtain the partial derivatives and normalized contour integrals. This includes Gaussian smoothing function, and smoothing scale which examines entire image repeatedly over the given parameter set (x_0, y_o, r). This can be expressed as (7).

$$max_{(r,x_0,y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint \frac{I(x,y)}{2\pi r} ds \right| \tag{7}$$

In order to obtain the normalized output, all points which are inside the iris region are mapped to polar coordinates (r, θ) where r represents the interval $[0,1]$ and θ denotes the angle as $[0, 2\pi]$. Based on this remapping, we obtain the two circular edges which are represented in Figure 6.
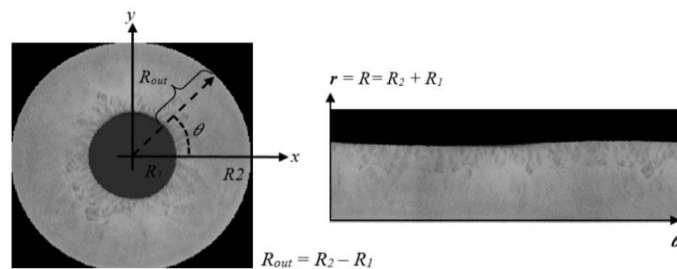

Figure 6. Radius estimation

This helps to obtain the inner and outer radius on the acquired image which is helpful in cropping and segmenting the iris region. Figure 7 shows the outcome of inner and outer boundary localization and segmentation. The localized image Figure 7(a) demonstrates the identified boundaries of iris and removing pupil and other components generates the segmented image as depicted in Figure 7(b).
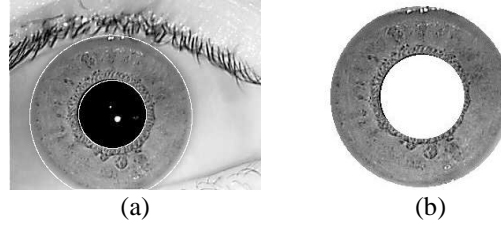

(a)   (b)

Figure 7. Iris localization and segmentation, (a) localization and (b) segmentation

2.2.1. Gabor-wavelet feature extraction

This portion of image is considered as region of interest (ROI) of the image which is used for further feature extraction process. In this work we apply Gabor –wavelet feature, GLCM [26], GLDM [27] and PCA [28]. The 1D-Gabor filter can be expressed as (8).

$$G(\omega) = \exp\left(\frac{-(\log(\omega/\omega_0))^2}{2(\log(\sigma/\omega_0))^2}\right) \tag{8}$$

Where ω_0 denotes the central frequency. In next stage, we convert the polar coordinates to Cartesian coordinates, thus, the resulting frequency domain can be expressed as (9).

$$G(\omega) = \exp\left(\frac{-(\ln(u1/f_0))^2}{2(\log(\sigma_u/f_0))^2}\right) * \exp\left(\frac{-v_1^2}{2\sigma_v^2}\right) \tag{9}$$

Where f_0 denotes the central frequency, σ_u is the bandwidth controller for $u1$, σ_v bandwidth controller for v_1, θ denotes the orientation of filters. Based on this, the output of Gabor filter is correlated with the input image and Gabor kernel $p\_k(x)$ and the Gabor filter can be expressed as (10).

$$a_k(x_0) = \int\int I(x)p_k(x - x_0)dx \tag{10}$$

Where Gabor filter can be expressed as (11).

$$p_k(x) = \frac{k^2}{\sigma}\exp\left(-\frac{k^2}{2\sigma^2}x^2\right)\left(\exp(ikx) - \exp\left(-\frac{\sigma^2}{2}\right)\right) \tag{11}$$

Where k denotes the vector used to describe the wave. In this work we use $\sigma = \pi$ and five spatial frequencies with varied wavenumbers as $k_i = \left(\frac{\pi}{2}, \frac{\pi}{4}, \frac{\pi}{8}, \frac{\pi}{16}, \frac{\pi}{32}\right)$ along with 8 orientations from 0 to π.

2.2.2. GLCM and GLDM feature extraction

Further, we apply gray level co-occurrence matrix (GLCM) based texture feature extraction which is applied on the segmented iris ROI. To extract the GLCM features, we consider different offsets ranging as 0°, 45°, 90°, and 135°. With the help of these offsets, we compute contrast, correlation, energy, homogeneity and entropy features. Table 1 shows the expression to compute features.

Similarly, we consider GLDM feature extraction model which is based on the absolute difference between two pixels given in a gray level and separated by a displacement δ which is denoted as motion vector as (12).

$$\delta = |S(x,y) - S(x + \Delta x, y + \Delta y)| \tag{12}$$

The probability density function can be expressed as (12).

$$D(i|\delta) = Prob[S_\delta(x,y) = i] \tag{13}$$

Where Δx and Δy are integer parameters, $S_\delta(x, y)$ denotes the input image. In order to formulate the feature vector, we concatenate contrast, entropy, angular second moment from PDF. Further, the obtained feature is processed through the PCA approach which has stages:
− In first phase, it computes the mean of each vector as (14):

$$x_m = \frac{1}{N}\sum_{k=1}^{N} xk \tag{14}$$

further, the obtained mean is subtracted from the data vectors to generate the zero mean vectors as (15):

$$x_z = x_i - x_m \tag{15}$$

Where x_z denotes the zero mean vector, x_i represents the each element of the column vector and x_m is the mean of each column vector.
− In next stage, covariance matrix is computed as (16):

$$c = [x_z T * x_z] \tag{16}$$

later, we compute Eigen vectors and Eigen values which are further multiplied with the zero mean vector to produce the feature vector. The Eigen vector is represented as (17):

$$[c - \gamma i]e = 0 \tag{17}$$

and, the obtained feature vector is expressed as (18):

$$f_i = [x_z]e \tag{18}$$

This complete process generates a fused feature vector which is used for training purpose.

Table 1. Feature computation formulas

Feature	Description	Formula
Contrast	It measures the local information variation in the image	$f_1 = \sum_{i,j} \lvert i - j \rvert^2 p(i,j)$
Correlation	It measures the joint probability occurrence of specified pixel	$f_2 = \sum_{i,j} \frac{(i - \mu i)(j - \mu i)p(i,j)}{\sigma_i \sigma_j}$
Energy	It provides the details of squared elements contain in the GLCM matrix. it is also known as angular second moment or uniformity	$f_3 = \sum_{i,j} p(i,j)^2$
Homogeneity	It measures the nearness of GLCM component distribution	$f_4 = \sum_{i,j} \frac{p(i,j)}{1 + \lvert i - j \rvert}$
Entropy	It is used to determine the statistical attributes	$f_5 = -\sum_{k=0}^{M-1} p_k \log_2(p_k)$

2.3. CNN-LSTM classifier construction

In this section we briefly describe the CNN and LSTM model and present a combined architecture for classifier construction. The current advancements in machine learning field have reported the importance of deep learning approaches where CNN has been adopted widely to achieve the efficient classification accuracy [5]. The CNN architectures are based on the feedforward neural network process and it is capable to extract the robust features from convolution structures. In this work, we adopt a single-dimensional convolutional layer with m number of filters and k number of kernels. For each input signal, the corresponding attributes are fed into the corresponding different layer and the obtained output is then processed through the bidirectional long short-term memory (BiLSTM) block to obtain the final classification.

The LSTM architecture consist of three different gates as input gate i_g, forget gate f_g and output gate o_g. Moreover, it contains a dedicated memory cell c_t which helps to learn the long-term dependencies in the given sequence, and hidden state h_t. These components of LSTM are represented as (19):

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i)$$

$$f_t = \sigma\left(w_f[h_{t-1}, x_t] + b_f\right)$$

$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o)$$

$$c_t = f_t \odot c_{t-1} + t_t \odot \tanh(w_c[h_{t-1}, x_t] + b_c)$$

$$h_t = o_t \odot \tanh(c_t) \tag{19}$$

Where w_i, w_o and w_f represents the weights of neurons. Similarly, b_i, b_o and b_f represents the biases, σ denotes the sigmoid function and $\odot$ represents the elementwise multiplication and $\tanh()$ is the hyperbolic tangent function. The traditional LSTM model is extended to formulate the bidirectional LSTM where two LSTMs operate simultaneously in forward and backward direction. The backward LSTM module is used to capture the past contextual information whereas forward LSTM captures the future information. The final hidden state can be expressed as (20):

$$h_t = \mu\left(\overrightarrow{h_t}, \overleftarrow{h_t}\right) \tag{20}$$

3. RESULTS AND DISCUSSION

In previous section, we have described the complete proposed approach to process the ECG and Iris data to formulate a robust multimodal authentication system. In this section, we present the experimental analysis of proposed approach and compare its obtained performance with existing schemes. In this experiment, we have considered two different datasets as IIT Delhi Database and MIT-BIH dataset from Physionet.

3.1. Dataset details

IIT Delhi Database: this dataset is considered for Iris image analysis. It has been acquired in Biometrics Research Laboratory by using JIRIS, JPC1000, digital CMOS camera. All of the images in the currently accessible database are in bitmap (*.bmp) format and obtained from 224 users. The database has 176 men and 48 women, all of whom fall into the age range of 14 to 55. A total of 224 separate folders, each linked to an integer identity or number, make up the database's 1120 photos. All of the photographs in this collection have a resolution of 320 by 240 pixels and were taken inside. Figure 8 depicts some sample images from this dataset.

MIT BIH dataset: The MIT-BIH dataset encompasses 48 hours recording of two-channel ambulatory ECG signals obtained from 47 subjects. Out of these 47 subjects. Out of these 48 hours, 23 readings are chosen form 24 hours' ambulatory signals which includes mixed population of inpatient and outpatient as 60% and 40% subjects respectively. Remaining 25 readings are obtained from the same set of the signal which includes clinically vital arrhythmias. For simplification, these signals were processed through digitization phase which generates the digitized signal at 360 samples per second per channel. This digitization is done with 11-bit resolution over 10mV range. These signals are annotated by expert cardiologists and entire dataset contains 110,000 annotations. Figure 8 from Figures 9(a)-9(f) depicts some sample of original signal, different noise types and combined signal with different noises.
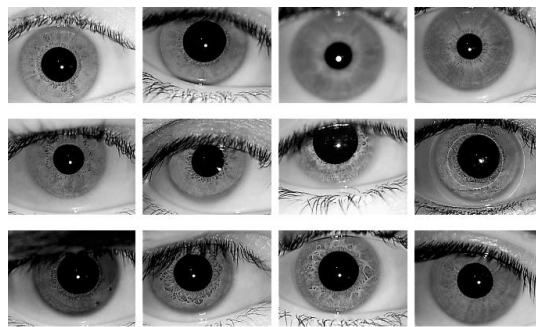

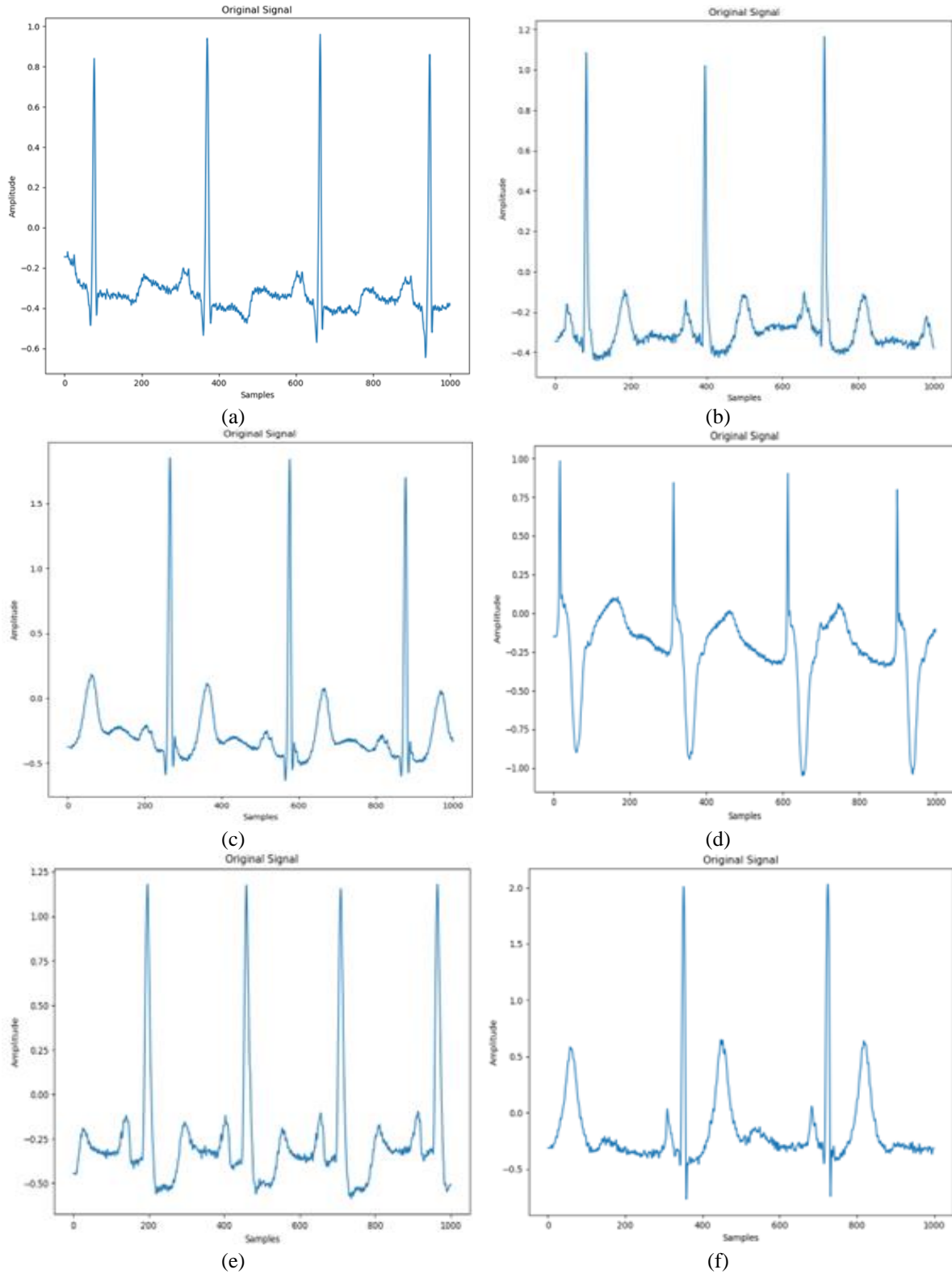Figure 8. Sample images from IITD Iris dataset

Figure 9. Sample ECG signals obtained from MIT-BIH dataset, (a) signal 100, (b) signal 101, (c) signal 102, (d) signal 103, (e) signal 104, and (f) signal 105

This dataset is further divided in a ratio where 70% data is used for training and 30% data samples are used for testing purpose. We perform random shuffling the given dataset to ensure that the data points are in a random order. This helps in preventing any ordering biases that might exist in the data. The proposed

classification model uses certain hyperparameters to improve the learning process. The complete process is performed for k fold cross-validation with k=5. These hyperparameters are demonstrated in Table 2.

Table 2. Hyperparameters

Simulation Parameter	Considered Value
Loss Function	Cross Entropy
Optimizer	Adam
Batch Size	32
Number of Epochs	200
Learning Rate	0.001
Dropout	0.5

3.2. Performance measurement parameters

The performance of proposed approach is measured based on confusion matrix calculation. The confusion matrix generated with the help of true positive, false positive, false negative and true negative. Table 3 shows a sample representation of confusion matrix.

Table 3. Confusion matrix

Actual class	Predicted class	
	Genuine User	Imposter User
Genuine User	True Positive	False Negative
Imposter User	False Positive	True Negative

Based on this confusion matrix, we measure several statistical performance parameters such as accuracy, precision, F1-score by using proposed approach. Accuracy is the measurement of correct instance classification out of total number instances. The accuracy is measured:

$$Acc = \frac{TP+TN}{TP+TN+FP+FN}$$

Then, we compute precision of the proposed approach. It is computed by taking the ratio of true positive and (true and false) positives.

$$P = \frac{TP}{TP+FP}$$

Finally, we compute the F-measure based on the sensitivity and precision values, which is expressed:

$$F = \frac{2*P*Sensitivity}{P+Sensitivity}$$

3.3. Comparative analysis with combined feature extraction process

With the help of aforementioned performance measurement parameters, we measure the performance of proposed approach. In order to classify the ECG and iris dataset, we consider 50 samples where these data samples are stored in each folder which is allocated their corresponding class. We measure the performance in terms of precision, sensitivity, specificity, F1-score and accuracy for different classifiers. Moreover, we evaluate the performance for ECG and Iris separately and later measured the performance for combined feature extraction. Table 4 shows the obtained classification performance.

Table 4. Classification performance for different classifiers

Classifier	Precision			F1-score			Sensitivity			Specificity			Accuracy		
	Iris	ECG	combine	Iris	ECG	combine	Iris	ECG	combine	Iris	ECG	combine	Iris	ECG	combine
SVM	0.925	0.910	0.928	0.942	0.951	0.961	0.941	0.951	0.941	0.941	0.935	0.925	0.931	0.955	0.940
KNN	0.930	0.923	0.936	0.951	0.952	0.961	0.945	0.955	0.950	0.942	0.930	0.928	0.941	0.9450	0.951
Neural Network	0.935	0.941	0.940	0.952	0.952	0.962	0.961	0.955	0.950	0.945	0.938	0.940	0.945	0.936	0.955
Naïve Bayes	0.938	0.945	0.951	0.955	0.951	0.965	0.962	0.961	0.962	0.941	0.941	0.965	0.940	0.941	0.951
Random Forest	0.941	0.952	0.951	0.958	0.963	0.965	0.961	0.951	0.952	0.942	0.945	0.961	0.945	0.955	0.960
Decision Tree	0.945	0.948	0.952	0.96	0.962	0.965	0.966	0.9612	0.955	0.945	0.955	0.962	0.955	0.951	0.971
Bagged ensemble	0.946	0.951	0.958	0.960	0.968	0.969	0.968	0.9712	0.951	0.955	0.956	0.965	0.950	0.958	0.965
Proposed	0.952	0.955	0.962	0.965	0.975	0.975	0.979	0.9750	0.978	0.9650	0.965	0.9710	0.961	0.9680	0.985

3.4. Comparative analysis with different attacks on the input data

In order to show the robustness of proposed approach, we incorporated several attacks to the original data and measured the performance of proposed model. For ECG signal, we consider three different types of noises as baseline wander, muscle artifacts and power line interference. On the other hand, we consider different attacks on iris data such as cropping, rotation, adding noise, image blurring. Figure 9 depicts different attacks on image data where Figure 10(a) shows the original image, Figure 10(b) depicts the image with cropping attack, Figure 10(c) depicts the image after rotation attack, Figure 10(d) is the noisy image and Figure 10(e) shows the image after motion blur attack.

For this experiment, we measure the outcome of proposed approach and compared with existing classification schemes. Table 5 shows the comparative performance in terms of precision and F1-score for different attacks such as Cropping, Rotation, Noise, and Blurring.

Similarly, we have measured the performance in terms of Sensitivity and Specificity by considering aforementioned attacks. The obtained performance is presented in Table 6. Finally, we present the outcome of proposed approach in terms of classification accuracy and compare it with traditional classifiers. Table 7 shows the obtained performance. Table 6 shows that the proposed approach is able to achieve desired accuracy for four different types of attacks and outperforms traditional classification method.
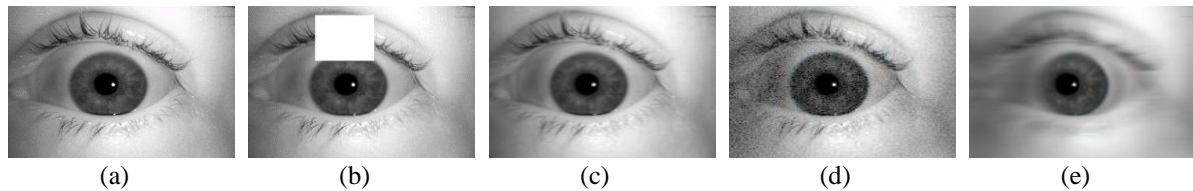

| (a) | (b) | (c) | (d) | (e) |

Figure 10. Sample attack images, (a) original image, (b) image cropping, (c) image rotation attack, (d) noisy image, and (e) motion blur attack

Table 5. Precision and F1-score performance for different attacks

Classifier	Precision				F1 score			
	Cropping	Rotation	Noise	Blurring	Cropping	Rotation	Noise	Blurring
SVM	0.825	0.830	0.851	0.861	0.821	0.830	0.815	0.821
KNN	0.803	0.841	0.833	0.835	0.803	0.8125	0.8210	0.841
Neural Network	0.821	0.844	0.850	0.840	0.821	0.8210	0.8350	0.851
Naïve Bayes	0.833	0.850	0.855	0.855	0.833	0.8315	0.8410	0.856
Random Forest	0.850	0.8350	0.825	0.861	0.856	0.8410	0.8331	0.862
Decision Tree	0.855	0.841	0.803	0.850	0.881	0.8511	0.8516	0.852
Bagged ensemble	0.870	0.856	0.825	0.861	0.901	0.8920	0.8620	0.860
Proposed Model	0.9210	0.892	0.912	0.910	0.921	0.9210	0.918	0.901

Table 6. Sensitivity and Specificity performance for different attacks

Classifier	Sensitivity				Specificity			
	Cropping	Rotation	Noise	Blurring	Cropping	Rotation	Noise	Blurring
SVM	0.7915	0.8011	0.8215	0.810	0.8345	0.8511	0.8410	0.810
KNN	0.812	0.8050	0.835	0.805	0.855	0.820	0.842	0.845
Neural Network	0.805	0.810	0.840	0.815	0.82	0.795	0.844	0.813
Naïve Bayes	0.815	0.829	0.841	0.8351	0.83	0.83	0.851	0.818
Random Forest	0.8351	0.815	0.844	0.8401	0.866	0.843	0.855	0.836
Decision Tree	0.8401	0.808	0.805	0.830	0.850	0.844	0.842	0.849
Bagged ensemble	0.854	0.851	0.815	0.841	0.830	0.858	0.845	0.865
Proposed Model	0.9012	0.8915	0.923	0.915	0.905	0.911	0.846	0.895

Table 7. Classification accuracy performance for different attacks

Classifier	Accuracy			
	Cropping	Rotation	Noise	Blurring
SVM	0.845	0.813	0.821	0.84
KNN	0.861	0.835	0.835	0.835
Neural Network	0.868	0.840	0.840	0.840
Naïve Bayes	0.895	0.855	0.835	0.855
Random Forest	0.890	0.861	0.840	0.861
Decision Tree	0.885	0.840	0.855	0.850
Bagged ensemble	0.895	0.855	0.861	0.8910
Proposed Model	0.934	0.8950	0.921	0.9325

4. CONCLUSION

In this work, we have mainly focused on development of multimodal biometric authentication system with improved accuracy and robustness. The proposed approach considers ECG and iris data due to their significant advantages in biometric authentication for physiological and physiological behavior. The proposed model performs deep learning based segmentation method for R peak segmentation from ECG signals. Further, different morphological features are extracted with the help of this peak data. Similarly, wavelet, GLCM and GLDM feature extraction processes are applied to extract the robust features from iris images. Finally, CNN-LSTM based hybrid classifier is used to learn these patterns and classify the users as genuine or imposter. The comparative study shows the proposed approach achieves average performance as 0.962, 0.975, 0.978, 0.9710 and 0.985 in terms of Precision, F1-score, Sensitivity, Specificity, and Accuracy. However, this process is tested on limited number of attacks such as Cropping, Rotation, Noise, and blurring with less intensity of attack whereas real-time attacks could be more intense. Moreover, this work can be extended by incoproating multimodal authentication system along with liveliness detection of subjects.

REFERENCES

[1] W. Xu *et al.*, "RFace: anti-spoofing facial authentication using COTS RFID," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, May 2021, pp. 1–10. doi: 10.1109/INFOCOM42981.2021.9488737.

[2] J.-H. Im, S.-Y. Jeon, and M.-K. Lee, "Practical Privacy-preserving face authentication for smartphones secure against malicious clients," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2386–2401, 2020, doi: 10.1109/TIFS.2020.2969513.

[3] M. K. Morampudi, M. V. N. K. Prasad, and U. S. N. Raju, "Privacy-preserving iris authentication using fully homomorphic encryption," *Multimedia Tools and Applications*, vol. 79, no. 27–28, pp. 19215–19237, Jul. 2020, doi: 10.1007/s11042-020-08680-5.

[4] A. R. Khan *et al.*, "Authentication through gender classification from iris images using support vector machine," *Microscopy Research and Technique*, vol. 84, no. 11, pp. 2666–2676, Nov. 2021, doi: 10.1002/jemt.23816.

[5] M. Hammad, P. Pławiak, K. Wang, and U. R. Acharya, "ResNet-attention model for human authentication using ECG signals," *Expert Systems*, vol. 38, no. 6, Sep. 2021, doi: 10.1111/exsy.12547.

[6] M. Ingale, R. Cordeiro, S. Thentu, Y. Park, and N. Karimian, "ECG biometric authentication: A comparative analysis," *IEEE Access*, vol. 8, pp. 117853–117866, 2020, doi: 10.1109/ACCESS.2020.3004464.

[7] X. Yin, S. Wang, M. Shahzad, and J. Hu, "An IoT-oriented privacy-preserving fingerprint authentication system," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 11760–11771, Jul. 2022, doi: 10.1109/JIOT.2021.3131956.

[8] W. Yang, S. Wang, K. Yu, J. J. Kang, and M. N. Johnstone, "Secure fingerprint authentication with homomorphic encryption," in *2020 Digital Image Computing: Techniques and Applications (DICTA)*, Nov. 2020, pp. 1–6. doi: 10.1109/DICTA51227.2020.9363426.

[9] X. Zhang, D. Cheng, P. Jia, Y. Dai, and X. Xu, "An efficient android-based multimodal biometric authentication system with face and voice," *IEEE Access*, vol. 8, pp. 102757–102772, 2020, doi: 10.1109/ACCESS.2020.2999115.

[10] A. Tarannum, Z. U. Rahman, L. K. Rao, T. Srinivasulu, and A. Lay-Ekuakille, "An efficient multi-modal biometric sensing and authentication framework for distributed applications," *IEEE Sensors Journal*, vol. 20, no. 24, pp. 15014–15025, Dec. 2020, doi: 10.1109/JSEN.2020.3012536.

[11] E. Sujatha, J. Sathiya Jeba Sundar, P. Deivendran, and G. Indumathi, "Multimodal biometric algorithm using iris, finger vein, finger print with hybrid GA, PSO for authentication," 2021, pp. 267–283. doi: 10.1007/978-981-15-8335-3_22.

[12] M. Hammad and K. Wang, "Parallel score fusion of ECG and fingerprint for human authentication based on convolution neural network," *Computers & Security*, vol. 81, pp. 107–122, Mar. 2019, doi: 10.1016/j.cose.2018.11.003.

[13] S. Zhao, W. Nie, and B. Zhang, "Deep learning-based hyperspectral multimodal biometric authentication system using palmprint and dorsal hand vein," in *AI and Deep Learning in Biometric Security*, First edition. | Boca Raton, FL: CRC Press, 2021. |: CRC Press, 2021, pp. 1–22. doi: 10.1201/9781003003489-1.

[14] E. Al Alkeem *et al.*, "Robust deep identification using ECG and multimodal biometrics for industrial internet of things," *Ad Hoc Networks*, vol. 121, p. 102581, Oct. 2021, doi: 10.1016/j.adhoc.2021.102581.

[15] R. Avanzato and F. Beritelli, "Automatic ECG diagnosis using convolutional neural network," *Electronics*, vol. 9, no. 6, p. 951, Jun. 2020, doi: 10.3390/electronics9060951.

[16] G. Kaur, A. Girdhar, and M. Kaur, "Enhanced iris recognition system – an integrated approach to person identification," *International Journal of Computer Applications*, vol. 8, no. 1, pp. 1–5, Oct. 2010, doi: 10.5120/1182-1630.

[17] M. Regouid, M. Touahria, M. Benouis, and N. Costen, "Multimodal biometric system for ECG, ear and iris recognition based on local descriptors," *Multimedia Tools and Applications*, vol. 78, no. 16, pp. 22509–22535, Aug. 2019, doi: 10.1007/s11042-019-7467-x.

[18] D. Jiang, G. Zhang, O. W. Samuel, F. Liu, and H. Xiao, "Dual-Factor WBAN Enhanced Authentication System Based on Iris and ECG Descriptors," *IEEE Sensors Journal*, vol. 22, no. 19, pp. 19000–19009, Oct. 2022, doi: 10.1109/JSEN.2022.3198645.

[19] B. A. El-Rahiem, F. E. A. El-Samie, and M. Amin, "Multimodal biometric authentication based on deep fusion of electrocardiogram (ECG) and finger vein," *Multimedia Systems*, vol. 28, no. 4, pp. 1325–1337, Aug. 2022, doi: 10.1007/s00530-021-00810-9.

[20] S. B. Jadhav, N. K. Deshmukh, and V. T. Humbe, "HDL-PI: hybrid deeplearning technique for person identification using multimodal finger print, iris and face biometric features," *Multimedia Tools and Applications*, vol. 82, no. 19, pp. 30039–30064, Aug. 2023, doi: 10.1007/s11042-022-14241-9.

[21] S. P. Singh and S. Tiwari, "A dual multimodal biometric Authentication system based on WOA-ANN and SSA-DBN Techniques," *Sci*, vol. 5, no. 1, p. 10, Mar. 2023, doi: 10.3390/sci5010010.

[22] S. Amritha Varshini and J. Aravinth, "Hybrid level fusion schemes for multimodal biometric authentication system based on matcher performance," 2021, pp. 431–447. doi: 10.1007/978-981-33-6862-0_35.

[23] P. Huang, L. Guo, M. Li, and Y. Fang, "Practical privacy-preserving ECG-based authentication for IoT-based healthcare," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 9200–9210, Oct. 2019, doi: 10.1109/JIOT.2019.2929087.

[24] A. Kumar, S. Jain, and M. Kumar, "Face and gait biometrics authentication system based on simplified deep neural networks," *International Journal of Information Technology*, vol. 15, no. 2, pp. 1005–1014, Feb. 2023, doi: 10.1007/s41870-022-01087-5.

[25] W. Xu, L. Wang, B. Wang, and W. Cheng, "Intelligent recognition algorithm of multiple myocardial infarction based on morphological feature extraction," *Processes*, vol. 10, no. 11, p. 2348, Nov. 2022, doi: 10.3390/pr10112348.

[26] F. Özbilgin, Ç. Kurnaz, and E. Aydın, "Prediction of coronary artery disease using machine learning techniques with iris analysis," *Diagnostics*, vol. 13, no. 6, p. 1081, Mar. 2023, doi: 10.3390/diagnostics13061081.

[27] R. Gutsche *et al.*, "Radiomics outperforms semantic features for prediction of response to stereotactic radiosurgery in brain metastases," *Radiotherapy and Oncology*, vol. 166, pp. 37–43, Jan. 2022, doi: 10.1016/j.radonc.2021.11.010.

[28] H. Hafeez, M. N. Zafar, C. A. Abbas, H. Elahi, and M. O. Ali, "Real-time human authentication system based on iris recognition," *Eng*, vol. 3, no. 4, pp. 693–708, Dec. 2022, doi: 10.3390/eng3040047.

BIOGRAPHIES OF AUTHORS

Ashwini Kailas 🔵 ⓖ SC ◗ is a Research Scholar currently pursuing her Ph.D. in Siddaganga Institute of Technology, Tumkur affiliated to Visvesvaraya Technological University Belagavi. Prior she has obtained her Masters of Technology in Digital Electronics from Sri Siddhartha Institute of Technology, Tumkur. Her research focus is in the field of image processing, signal processing, machine learning and deep learning. She can be contacted at the mail id: ashwini0429@gmail.com.

Geevagondanahalli Narayanappa Keshava Murthy 🔵 ⓖ SC ◗ is currently working as an Assistant Professor, in the Department of Electronics and Instrumentation Engineering, with 23 years of teaching experience and 2 years of industry experience. He received his B.E degree from the esteemed Siddaganga Institute of Technolgy, Tumakuru in the year 1996, M.E from UVCE, Bangalore in 2001 and was awarded Ph.D. degree in 2016 from VIT University, one among top 12 universities of India. His research interests include Smart sensors, Biomedical Engineering, Biosciences and Cognitive neuroscience. He has more than 13 papers published in peer-reviewed international journals and more than 10 conferences in his credit and has received research award twice for his publications at VIT University. He has served as session chair for few conferences and is BOE member for few colleges. As an extension of his work and knowledge currently he is guiding 2 Ph. D students. He also has delivered lectures in various engineering colleges and Universities. He can be contacted at email: gnk@sit.ac.in.