World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:16, No:9, 2022

ECG Based Reliable User Identification Using Deep Learning

R. N. Begum, Ambalika Sharma, G. K. Singh

Abstract**—**Identity theft has serious ramifications beyond data and personal information loss. This necessitates the implementation of robust and efficient user identification systems. Therefore, automatic biometric recognition systems are the need of the hour, and electrocardiogram (ECG)-based systems are unquestionably the best choice due to their appealing inherent characteristics. The Convolutional Neural Networks (CNNs) are the recent state-of-the-art techniques for ECG-based user identification systems. However, the results obtained are significantly below standards, and the situation worsens as the number of users and types of heartbeats in the dataset grows. As a result, this study proposes a highly accurate and resilient ECG-based person identification system using CNN's dense learning framework. The proposed research explores explicitly the caliber of dense CNNs in the field of ECG-based human recognition. The study tests four different configurations of dense CNN which are trained on a dataset of recordings collected from eight popular ECG databases. With the highest False Acceptance Rate (FAR) of 0.04% and the highest False Rejection Rate (FRR) of 5%, the best performing network achieved an identification accuracy of 99.94%. The best network is also tested with various train/test split ratios. The findings show that DenseNets are not only extremely reliable, but also highly efficient. Thus, they might also be implemented in real-time ECG-based human recognition systems.

Keywords**—**Biometrics, dense networks, identification rate, train/test split ratio.

I. Introduction

BIOMETRICS has become a fundamental tool for user identification and authentication today, as it fully relies on a specific user. These are attributes that are unique to a particular user (for instance, finger print, hand geometry, face, iris, gait, keystroke, ECG, etc.), and can be employed to identify a user from the rest of the population. While most of the biometrics come with deficiencies like duplicity (finger print, iris, etc.), imitation (voice, gait, etc.), artificial disguise (face with makeup or surgery), ECG comes with several advantages. Its main advantage is being proof of aliveness that assures the physical presence of the user. Another advantage is that it is very difficult to counterfeit. Moreover, it gives information about the physiological and clinical states of the person concerned [1], [2].

The variations in potential generated by the excitable cells of cardiac muscles of the heart over time, manifest a series of waves, which is exhibited by the ECG. The depolarization of the two bottom chambers of the heart is depicted by the QRS complex while their repolarization is reflected by the T wave. A healthy heart beats at a rate of 60-100 beats/min; although conditions like mental, physical, and age can alter it to some extent [3]. The physiological and geometrical attributes of ECG can reveal the identity of an individual [4]. These attributes come entirely from the person's heart signature; such as cardiac muscle, its orientation, activation order, and conductivity [5]. Fig. 1 shows the ECG with its important parameters.

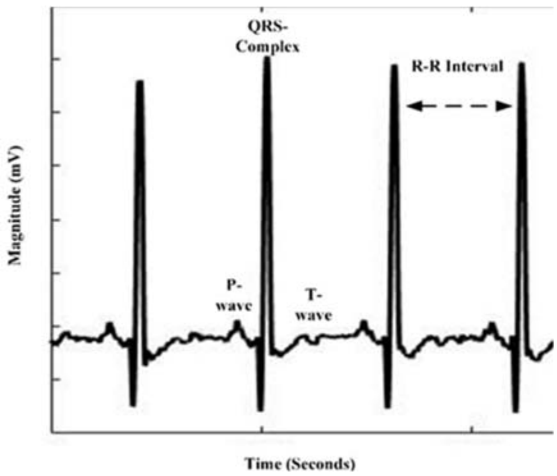

Fig. 1 Important parameters of an ECG waveform

ECG-based automatic user recognition systems have been approached with three techniques, namely, fiducial (dealing with fiducial points of the ECG waveform like onset and offset of a wave, intervals segment, etc.), non-fiducial (attributes based on ECG as a waveform, e.g., frequency characteristics) and hybrid (both fiducial and non-fiducial features). A recent development in this field is the implementation of Deep Learning (DL) approaches, especially CNN to extract features, and classify them to their respective categories. The main advantages of CNN are that it does not need hand-crafted features for classification and identification, and can be generalized for similar tasks, efficient for a large number of classes, and impressive classification accuracy.

Person identification using ECG is an active research area as it offers inherent properties of permanence, universality, and uniqueness. The area is about two decades old when Biel et al. [6] first introduced it to identify 20 individuals using the fiducial approach. Slowly it gained momentum and several

R. N. Begum is a Research Scholar in the Department of Electrical Engineering, Indian Institute of Technology Roorkee, India (corresponding author, phone: 8473890150, 9401044816; e-mail: rbegum@ee.iitr.ac.in, aecritu7@gmail.com).

Dr. Ambalika Sharma and Prof. G. K. Singh are with the Department of Electrical Engineering, Indian Institute of Technology Roorkee, India. (e-mail: ambalika.sharma@ee.iitr.ac.in, gksngfee@gmail.com).

mathematical algorithms over the years have been tried on it, in a quest to achieve outstanding results. Mathematical approaches like Discriminant Analysis [7], [8], Principal Component Analysis (PCA) [9], and Normalized Euclidean Distance [10] have been successfully implemented to identify individuals. Approaches like Wavelet Transform (WT) [11]-[12], Discrete Cosine Transform (DCT) [3], [13], Ziv-Merhav Cross Parsing [14], Dynamic Time Warping (DTW), Fisher's Linear Discriminant Analysis (FLDA) [15], Adaboost, Bagging classifier [16], Optimum patch forest [17], Random forest [18], Support Vector Machines (SVM) [19] and Artificial Neural Networks (ANN) [20], [21] have also shown promising results in human identification using ECG.

With the emergence of the DL era, the last couple of years witnessed several studies of person identification using DL techniques. For instance, Page et al. [22] applied deep neural network (DNN) on ECG–ID data, while Wieclaw et al. [23] applied it on self-collected data, and achieved a classification accuracy of 99.93% and 96% respectively. Studies conducted by Zhang et al. [24], [25] applied CNN to self-collected data and attained an accuracy of 98.4%. Eduardo et al. [26] implemented a deep auto-encoder on a dataset acquired from a local hospital, while a CNN was applied by Labati et al. [27] on E-HOL-03-0202-003 (Intercity Digital Electrocardiogram Alliance - IDEAL) database and PTB Diagnostic ECG Database. In yet another study, 33 women were classified by a deep CNN using ECG obtained from OM signal apparel with an accuracy of 95.95% [28]. Other studies include DNN employed on the ECG-ID [29], deep LSTM used on PTB and ECG-ID [30], CNNs used on MITDB, FANTASIA, NSRDB, and QT-DB [31], Ensemble of Deep CNNs employed on NSRDB [32], [33], CNN applied on spectrograms generated from ECG signals of Fantasia and ECG-ID databases [34], etc. It is evident that most of the above studies explored various DL approaches for possible implementation in ECG-based user identification systems and stressed more on identification accuracy. Some of them were more dedicated to signal processing techniques viz. transforming the signals, conversion from 1D to 2D data, and so on to make them easily classifiable by the DL networks. While others were focused on the easy acquisition of data. But, these studies did not address issues like the efficiency of the networks when a large population of varied types of ECG is considered. Also, none of these studies went deeper into their implemented DL techniques and thus, could not address the reasons for the drop in identification rate properly.

Various DL approaches that could be utilized for ECG-based user identification systems have been studied in the literature. However, none of them dug deeper into the models to solve the underlying technical flaws. Furthermore, the majority of the research relied on self-collected data and had a small number of participants. Others relied on two or three public ECG databases, which were unable to gather the vivid variations in an ECG waveform. A robust, dependable user identification system should be able to recognize all conceivable waveform changes both within and between classes. This necessitates a collection of recordings that include signals of both male and female participants over a wide age range. The dataset should include data from a variety of abnormalities collected over a lengthy period. Another shortcoming of the existing studies is that the precision of the identification was not adequate. Therefore, in the proposed study, the primary contributions are: 1) To use CNN architectures based on a dense learning framework to develop a reliable and efficient user identification system. This study looked at four distinct dense network topologies and chose the highest performing model for investigating the impact of training/testing split ratio on the performance of the model. In addition, each layer of the network is thoroughly analyzed, and the number of learnable parameters is computed. Table VIII includes a full summary of the network's numerous operations at each layer. 2) To use a dataset that is a collection of all possible variations that an ECG waveform might have. This is done to ensure the robustness of the system. To achieve it, ECG data of various heart conditions, wide age range, and long-period records taken over several sessions have been considered from eight different public databases. 3) To achieve an unbeatable identification rate in a broad population of different types of ECG waveforms.

The rest of the paper is organized as follows: Section II discusses the proposed method while Section III discusses the experimentation and implementation. Section IV presents the results with its discussions, while Section V provides a conclusion with future prospects of the study.

II. Methodology

The aim of the proposed study is to identify users using a robust and efficient CNN model employing dense learning framework. The presented approach consists of three stages: 1) dataset acquisition, 2) data pre-processing and 3) classification. Fig. 2 exhibits methodology of the proposed study.

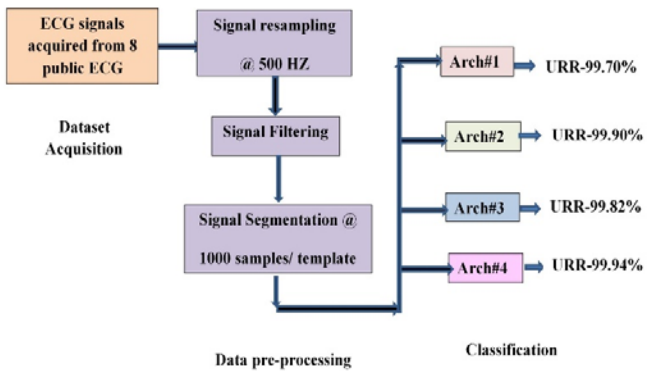

Fig. 2 Proposed Methodology of user recognition (Arch-Architecture model; URR-User Recognition Rate)

A. Dataset Acquisition

A dataset of ECG recordings from eight popular ECG databases, namely MIT-BIH Arrhythmia Database (MITDB) [35], MIT-BIH Normal Sinus Rhythm (NSRDB), Physikalisch-Technische Bundesanstalt Diagnostic ECG Database (PTBDB) [36], QT Database (QTDB) [37], ECG-ID Database (ECG-ID DB) [38], Intracardiac Atrial Fibrillation Database (IAFDB), Creighton University Ventricular Tachyarrhythmia Database

(CUDB) [39], and MIMIC-II/III (Medical Information Mart for Intensive Care) waveform database which are available in Physionet ATM [40] is collected. The challenges observed in the recordings of both genders; wide age range; varied clinical, and physiological health conditions; and recordings of lengthy periods over several days, and weeks are also considered. Table I presents detailed information on the dataset acquired from the above-mentioned databases.

TABLE I
NUMBER OF USERS CONSIDERED FROM EACH DATABASE WITH THEIR DETAILS

S. No.	Database	No. of subjects	Gender: Male, Female	Heart condition: Normal, Diseased	Age range
1	PTB DB	38	29, 9	4, 34	29-82
2	NSRDB	13	3, 10	13, 0	20-45
3	MITDB	10	5, 5	6, 4	24- 87
4	QT-DB	13	9, 4	9, 4	32- 83
5	IAF DB	7	4, 3	0, 7	58-81
6	CUDB	20	Not provided	0, 20	Not Provided
7	ECG-ID DB	25	12, 13	19, 6	21, 49
8	MIMIC-II/III	124	Not provided	0, 124	Not provided

B. Data Pre-processing

In order to obtain uniform and equi-spaced sampled signal, the acquired dataset is resampled at 500 Hz. The resampled data have been filtered to get rid of power line interference (PLI), baseline wander (BLW) and electromyographic (EMG) noise with suitable filtering techniques for each database. The filtered signal then undergoes segmentation at 1000 samples per template (image) to be used for the dense CNNs. Figs. 3 (a)-(c) show specimens of the resampled signal, filtered signal and segmented signal, respectively.

C. Classification

In the proposed study, four network models based on the dense learning framework of CNN are investigated for classifying the ECG waveform images. The CNN are very successful in the task of classification and are one of the greatest innovations in the sphere of computer vision. It is called so because hidden layers of the networks are subjected to the operation of convolution. CNN has precisely earned its fame in the discipline of image processing and speech recognition. The hierarchy of features is instinctually learned by sequentially convolving the input signal with learnable kernels to carry out the task of classification/identification. It was first proposed by LeCun et al. [41] for recognizing ZIP codes, which was based on weight sharing [42] and extended by Wolpert et al. [43]. The greatest boons of CNN are that it does not require training hand-crafted features, can be generalized for an analogous task, and employed for recognizing an enormous number of classes. CNNs comprise two basic sections namely, Feature Extraction and Selection section; and the Classification/Identification section. The operation of feature extraction is carried out in convolutional layers, whereas feature selection is carried out in pooling layers. The task of classification is pursued in fully connected (dense) layers.

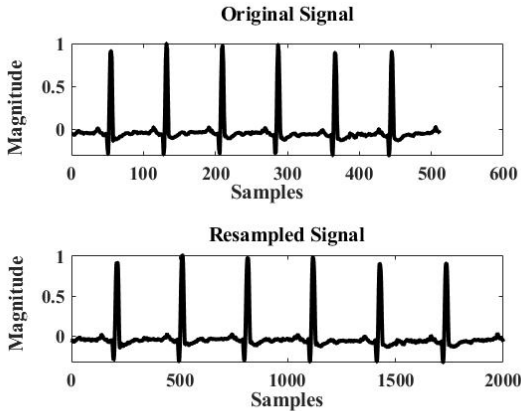

Fig. 3 (a) Specimen showing original and resampled signal from NSRDB

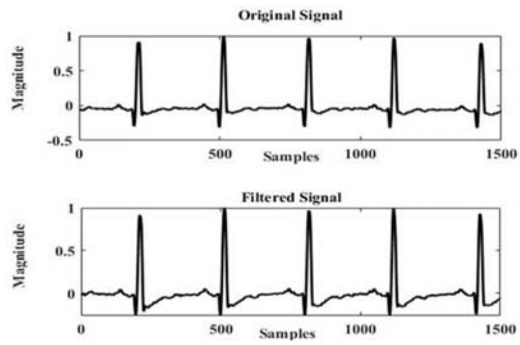

Fig. 3 (b) Specimen showing unfiltered and filtered signal from NSRDB

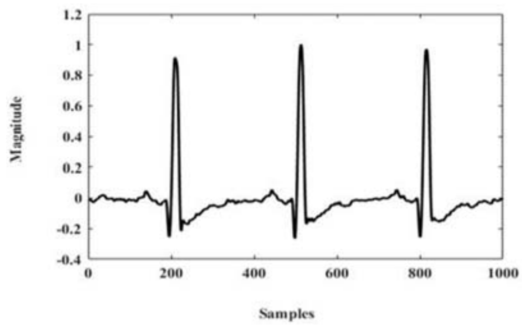

Fig. 3 (c) Specimen showing segmented signal of 1000 samples (1 image or template) from NSRDB

1. Convolutional Layer

The convolutional layers extract the attributes from signals by convolving them with appropriate kernels. The numbers and size of the kernels (filters) to be used are decided by the designer of the network. The number of feature maps derived is equal to the number of kernels used. Starting from the top left corner, the kernel traverses element by element convolving over the signal until it reaches the top right corner. It, then, descends an element and starts over again. The operation of convolution is linear, and can be applied in 1 D, 2D or 3D mode. Equation (1) represents the convolution operation for 2D mode.

$$y[i,j] = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} h[m,n]\,.\,x[i-m,j-n] \qquad (1)$$

where, *x* represents input image matrix and *h* represents filter matrix. The indices *i* and *j* are pixel positions of the image, whereas *m* and *n* are filter value positions, which usually lie in the range of -1 to 1.

The outcome of the convolution operation is subjected to non-linear activation functions like sigmoid, tangent, hyperbolic or ReLU (Rectified Linear Unit Function). The most preferred activation function in CNN is ReLU which adds non-linear behavior to the network [44]. It makes the network train faster, and is defined by,

$$f(x) = max\,(0, x) \tag{2}$$

2. Subsampling/Pooling Layers

The features maps obtained from convolutional layers are dimensionally reduced using subsampling/pooling layers. This is achieved by down-sampling the feature map to minimize learnable parameters and enhance faster training. Two approaches commonly used are Max-Pooling and Average Pooling, where maximum and average value respectively, of a subsection of feature map is selected for the next level.

3. Fully Connected/Dense Layers

The output of feature extraction and selection block is flattened into one dimensional array of numbers and applied to dense layers. Every input node of the dense layer is connected to the output node with a learnable weight. The output of every dense layer is again subjected to non-linear activation function, the ReLU. The last fully connected layer is followed by a soft-max function.

4. Proposed Dense Learning Based CNNs (DenseNets)

In the present study, a couple of DenseNet architectures is explored to discover an efficient model for user recognition in terms of architectural complexity, training time required and prediction accuracy achieved. Efficient parameter usage, deep supervision and feature reuse are the assets of a DenseNet. In DenseNets, a layer *l*, has direct connections from all preceding layers, and thus receives input as feature maps from all preceding layers. Equation (3) explains the working of a DenseNet:

$$X_l = H_l\,([X_0, X_1, X_2 \dots . X_{l-1}]) \tag{3}$$

where, X_l is the output at a layer *l*, H_l is any transformation (for instance, convolution followed by batch normalization and ReLU), [...] is input at layer *l* and $X_0, X_1, X_2 \dots . X_{l-1}$ are outputs at the respective layers [45]-[47].

In the proposed study, DenseNets are configured by varying the number of convolutional layers, concatenation layers and number of filters. Table II presents the details of each architecture explored.

III. Experimentation

A. Dataset

The proposed study attempts to recognize the ECG templates (images) of 250 users collected from eight popular ECG public databases. Table III gives the detailed figures of users considered in various heart conditions (ECG Beat Patterns).

TABLE II
ARCHITECTURAL LAYOUT OF THE DENSENETS EXPLORED IN THE PRESENT STUDY TO BE USED FOR USER IDENTIFICATION

Sl. No.	DenseNet Architectures	Number of convolutional layers	Number concatenation layers	Number of filters in each convolutional layer
1	*Arch#1*	5	3	3
2	*Arch#2*	5	3	16 in first 4 conv. Layers and 8 in 5th conv. layer
3	*Arch#3*	6	4	10
4	*Arch#4*	6	5	10

TABLE III
NUMBER OF USERS FROM VARIOUS CATEGORIES OF ECG BEAT PATTERNS

S. No.	Type of heart beats	Number of users
1	Bundle Branch Block (BBB)	18
2	Cardiomyopathy	10
3	Ventricular Tachyarrhythmia (VA)	23
4	Atrial Fibrillation (AF)	7
5	Miscellaneous (Misc.)	141
6	Normal	51

B. Implementation Details

The present study is implemented on MATLAB 2021a with an execution environment of GPU. User identification is carried out by all the four proposed network models. All the 4 models were tested for 120 users at first. The best performing model is then tested for 250 users as well. The weights of the proposed models are updated by backpropagation algorithm using Stochastic Gradient Descent (SGD) as the optimizing algorithm. Input size of 224 x 224 x 3 is considered, while hyper parameters like Minibatch size, learning rate, filter size and number of filters are set at 14, 0.0001, 3 x 3 and 10 for each convolution layer respectively. The dataset is split at a ratio of 7:3 for training and testing for all the 4 models.

C. Evaluation Metrics

The networks are evaluated based on their performance measured with the following metrics:

i) Identification Accuracy: Accuracy refers to how close a measurement is to the true label. In a set of measurements, it can be stated as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}\, x\ 100\% \tag{4}$$

where TP, TN, FP and FN denote true positive, true negative, false positive and false negative occurrences respectively.

ii) Training time: It is the amount of time it takes to train a network with specific hyper-parameters on a given dataset. The best performing network is then tested for sensitivity, specificity, precision, FAR and FRR.

iii) Sensitivity: The number of correctly categorized positive samples divided by the total number of positive samples is known as sensitivity, and it is represented mathematically as:

$$Sensitivity\ of\ each\ class = \frac{TP}{TP+FN} \quad (5)$$

iv) Specificity: It is the number of correctly classified negative samples to the total number of negative samples and is given by:

$$Specificity\ of\ each\ class = \frac{TN}{TN+FP} \quad (6)$$

v) False Acceptance Rate (FAR): It is the percentage of negative samples that are categorized as positive. It is calculated as:

$$FAR = 1 - Specificity \quad (7)$$

vi) False Rejection Rate (FRR): The rate at which positive samples are categorized as negative samples is known as the FRR. It is expressed as:

$$FRR = 1 - Sensitivity \quad (8)$$

vii) Precision: It represents the ratio of positive samples classified correctly to the total number of samples that are predicted positive. It is given by:

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

IV. Results and Discussions

TABLE IV
PREDICTION ACCURACY ON TRAINING & TESTING DATASETS AND TRAINING TIME REQUIRED BY THE PROPOSED NETWORKS FOR THE TASK OF USER IDENTIFICATION

Sl. No.	Network	Overall training accuracy (%)	Overall testing accuracy (%)	Training time
1	Arch#1	99.70	99.42	136 mins 45s
2	Arch#2	99.88	99.84	114 mins 52s
3	Arch#3	99.82	99.80	119 mins 18s
4	Arch#4	99.96	99.94	136 mins 41s

TABLE V
PREDICTION RATE (PR) OF VARIOUS TYPES OF ECG BEATS AS PREDICTED BY THE PROPOSED NETWORKS FOR THE TASK OF USER IDENTIFICATION

Sl. No.	ECG Beat Pattern	Arch#1 PR (%)	Arch#2 PR (%)	Arch#3 PR (%)	Arch#4 PR (%)
1	Normal	98.97	99.71	99.76	99.90
2	BBB	99.87	100	100	100
3	Cardiomyopathy	100	100	100	100
4	VA	99.59	99.79	99.69	100
5	AF	99.32	100	99.32	99.66
6	Misc.	100	100	100	100

The objective of the proposed study is to develop a resilient and efficient user-identification system that can identify various types of ECG beats of a large population accurately with a minimum training time. To implement the objective, four distinct dense networks were investigated whose architectural details are given in Table II. The results of the study are given in Table IV exhibiting the performance in terms of user identification rate on training and, testing datasets as well as training time required. The results of the prediction accuracy achieved by them in different ECG beat patterns are given in Table V. A bar graph showing the prediction accuracy obtained by all four proposed models is presented in Fig. 4, whereas Fig. 5 presents a pie chart showing the training time required by each proposed network.

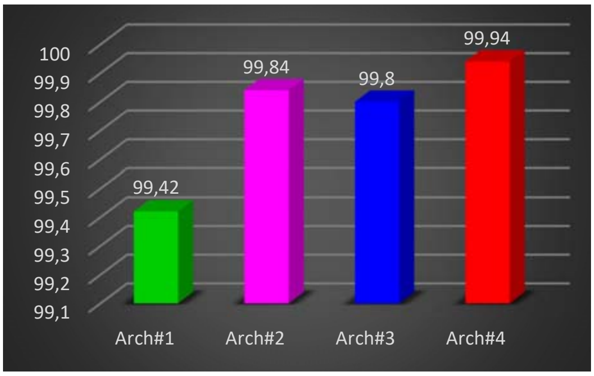

Fig. 4 Prediction accuracy of testing dataset by the proposed networks

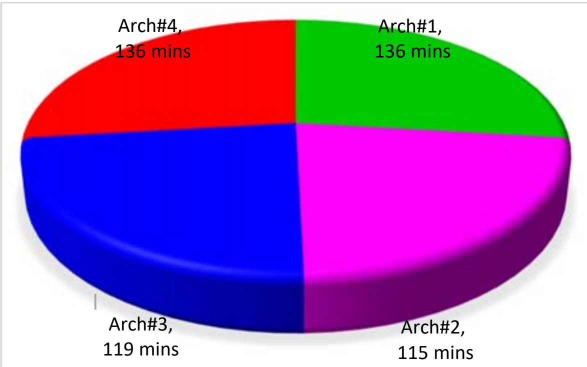

Fig. 5 Training time required by the proposed networks

A. Discussions on the Performance of Proposed DenseNet Models for User Identification

As reported in Table II, the task of user recognition is carried out with the proposed DenseNets comprising of just five or six convolutional layers. It is seen from Tables II and IV that given the present specifications, neither the change in the number of convolutional and concatenation layers nor the change in the number of filters in the convolution layers has a significant impact on the performance of the networks. With the given specifications, the training time of all the networks is approximately the same. But, it is also seen that among all the models, Arch#4 predicts the users of both diseased and normal heartbeats with highest accuracy. Thus, Arch#4 is also analyzed for sensitivity, specificity, precision, FAR, and FRR. Fig. 6 shows the line graph of the Sensitivity, Specificity, and Precision of Arch#4 while Table VI gives the FAR and FRR. As reported in the table, Arch#4 has an FRR of approximately 2% and 5% for User No. 24 and User No. 93, respectively, which can be explained by the fact that one template from User No. 24 and two templates from User No. 93 are misclassified in the User No. 12 and the User No. 30, respectively. This also explains the reason for FAR of 0.02% and 0.04% for User No.

12 and User No. 30, respectively. It is worth mentioning here that User No. 12 and User No. 24 belong to the diseased category whereas, User No. 30 and User No. 93 belong to the normal category of ECG.

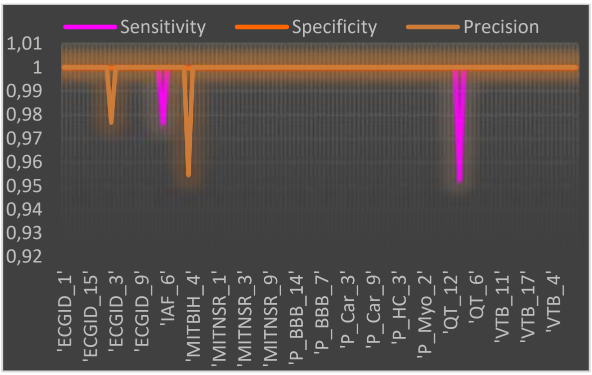

Fig. 6 Sensitivity, specificity and precision of Arch#4

Fig. 7 shows the architectural layout of Arch#4. The effect of train/test split ratios is also studied for Arch#4 to rule out the possibility of overfitting and the results are presented in Table VII.

The details of each layer of Arch#4 and their computations when trained for 250 users are given in Table VIII. A sample input image to the network is displayed in Fig. 8, along with its convolution with the kernels at the first convolution layer in Fig. 9.

TABLE VI
SENSITIVITY, SPECIFICITY, PRECISION, FRR AND FAR OF ARCH#4

User No.	Sensitivity	Specificity	Precision	FRR	FAR
1-11	1	1	1	0	0
12	1	0.9998	0.9767	0	0.0002
13-23	1	1	1	0	0
24	0.9762	1	1	0.023	0
25-29	1	1	1	0	0
30	1	0.9996	0.9545	0	0.0004
31-92	1	1	1	0	0
93	0.9524	1	1	0.0476	0
94-120	1	1	1	0	0

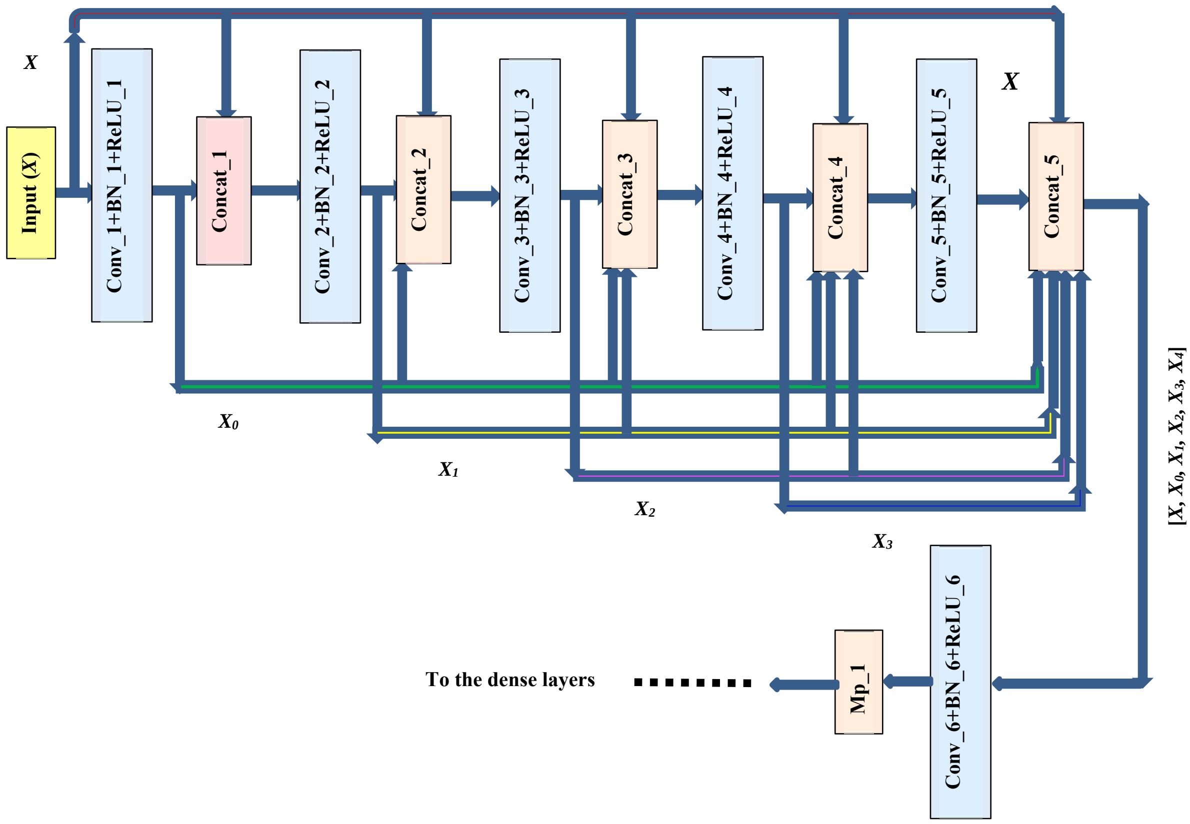

Fig. 7 Pictorial layout of Arch#4 CNN based on dense learning framework (X-input; X_0, X_1, X_2, X_3, X_4 – feature maps at respective layers; Conv- Convolution operation, BN- Batch normalization, Mp- Max pooling; […]- Concatenation of feature maps)

TABLE VII
PREDICTION RATE (PR) OF VARIOUS TYPES OF ECG BEATS AS PREDICTED BY THE PROPOSED NETWORK FOR THE TASK OF USER IDENTIFICATION

Sl. No.	Training dataset (%)	Testing dataset (%)	Training dataset accuracy (%)	Testing dataset accuracy (%)	Training time required
1	90 (15120 frames)	10 (1680 frames)	99.82	100	84 mins 11s
2	80 (13440 frames)	20 (3360 frames)	99.85	99.91	102 mins 10s
3	70 (11760 frames)	30 (5040 frames)	99.96	99.94	136 mins 41s
4	60 (10080 frames)	40 (6720 frames)	99.66	99.85	160 mins 55s
5	50 (8400 frames)	50 (8400 frames)	99.95	99.89	198 mins 46s
6	40 (6720 frames)	60 (10080 frames)	99.71	99.77	221 mins 11s
7	30 (5040 frames)	70 (11760 frames, 98)	99.91	99.63	252 mins 38s
8	20 (3360 frames)	80 (13440 frames)	99.95	99.21	287 mins 25s
9	10 (1680 frames)	90 (15120 frames)	99.93	97.41	321 mins 32s

TABLE VIII
DETAILS OF VARIOUS LAYERS OF *ARCH#4*

Layer No.	Name of Layer	Type	Activations	Learnables
1	Input (224x224x3 images with ‘zerocenter’ normalization)	Image input	224x224x3	-
2	Conv_1 (10 3x3x3 convolutions with stride [1 1] and padding ‘same’)	Convolution	224x224x10	Weights 3x3x3x10 Bias 1x1x10
3	BN_1 (Batch Normalization with 10 channels)	Batch Normalization	224x224x10	Offset 1x1x10 Scale 1x1x10
4	relu_1 (ReLU)	ReLU	224x224x10	-
5	Concat_1 (Concatenation of 2 inputs along dimension 3)	Concatenation	224x224x13	-
6	Conv_2 (10 3x3x13 convolutions with stride [1 1] and padding ‘same’)	Convolution	224x224x10	Weights 3x3x13x10 Bias 1x1x10
7	BN_2 (Batch Normalization with 10 channels)	Batch Normalization	224x224x10	Offset 1x1x10 Scale 1x1x10
8	relu_2 (ReLU)	ReLU	224x224x10	-
9	Concat_2 (Concatenation of 3 inputs along dimension 3)	Concatenation	224x224x23	-
10	Conv_3 (10 3x3x23 convolutions with stride [1 1] and padding ‘same’)	Convolution	224x224x10	Weights 3x3x23x10 Bias 1x1x10
11	BN_3 (Batch Normalization with 10 channels)	Batch Normalization	224x224x10	Offset 1x1x10 Scale 1x1x10
12	relu_3 (ReLU)	ReLU	224x224x10	-
13	Concat_3 (Concatenation of 4 inputs along dimension 3)	Concatenation	224x224x33	-
14	Conv_4 (10 3x3x33 convolutions with stride [1 1] and padding ‘same’)	Convolution	224x224x10	Weights 3x3x33x10 Bias 1x1x10
15	BN_4 (Batch Normalization with 10 channels)	Batch Normalization	224x224x10	Offset 1x1x10 Scale 1x1x10
16	relu_4 (ReLU)	ReLU	224x224x10	-
17	Concat_4 (Concatenation of 5 inputs along dimension 3)	Concatenation	224x224x43	-
18	Conv_5 (10 3x3x43 convolutions with stride [1 1] and padding ‘same’)	Convolution	224x224x10	Weights 3x3x43x10 Bias 1x1x10
19	BN_5 (Batch Normalization with 10 channels)	Batch Normalization	224x224x10	Offset 1x1x10 Scale 1x1x10
20	relu_5 (ReLU)	ReLU	224x224x10	-
21	Concat_5 (Concatenation of 6 inputs along dimension 3)	Concatenation	224x224x53	-
22	Conv_6 (10 3x3x53 convolutions with stride [1 1] and padding ‘same’)	Convolution	224x224x10	Weights 3x3x53x10 Bias 1x1x10
23	BN_6 (Batch Normalization with 10 channels)	Batch Normalization	224x224x10	Offset 1x1x10 Scale 1x1x10
24	relu_6 (ReLU)	ReLU	224x224x10	-
25	Mp_1 (1x2 max pooling with stride [2 2] and padding [0 0 0 0])	Max Pooling	112x112x10	-
26	Fc_1 (500 fully connected layer)	Fully Connected	1x1x500	Weights 1000x125440 Bias 1x1x10
27	Relu_7 (ReLU)	ReLU	1x1x500	-
28	Fc_2 (300 fully connected layer)	Fully Connected	1x1x300	Weights 750x1000 Bias 750x1
29	Relu_8 (ReLU)	ReLU	1x1x300	-
30	Fc_3 (200 fully connected layer)	Fully Connected	1x1x200	Weights 500x750 Bias 500x1
31	Relu_9 (ReLU)	ReLU	1x1x200	-
32	Fc_4 (120 fully connected layer)	Fully Connected	1x1x120	Weights 250x500 Bias 250x1
33	Softmax (softmax)	Softmax	1x1x120	-
34	Classoutput (crossentropyex with 120 class)	Classification Output	-	-

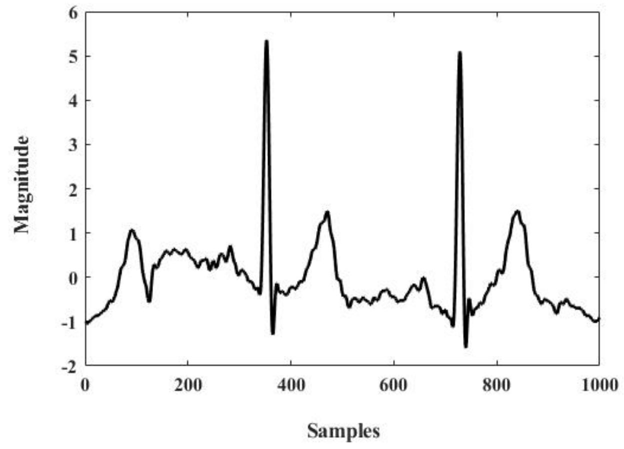

Fig. 8 An ECG-ID image used as input to Arch#4

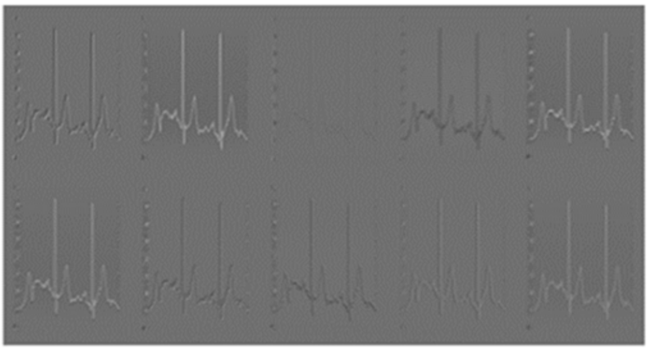

Fig. 9 Output of the sample ECG-ID image at conv_1 (first convolution layer) showing 10 feature maps from 10 filter

TABLE IX
PERFORMANCE OF ARCH#4

Sl. No.	Number of users	Overall training accuracy (%)	Overall testing accuracy (%)	Training time
1	120	99.96	99.94	136 mins 41 secs
2	250	100	99.92	336 mins 39 secs

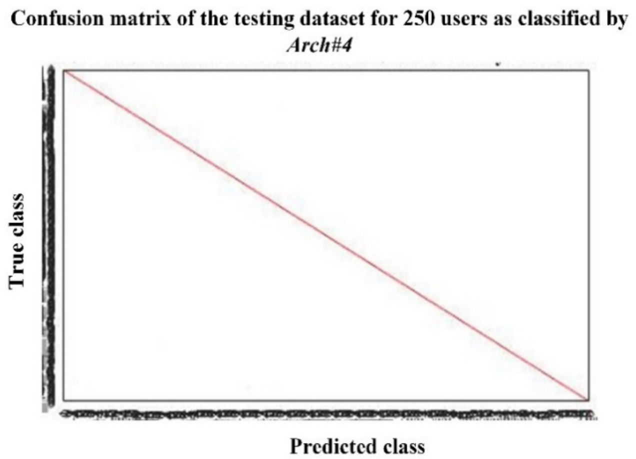

Fig. 10 Confusion matrix for 250 users by Arch#4

Table IX shows the impact on the performance of Arch#4 when the number users is increased from 120 to 250. Fig. 10 depicts the confusion matrix of the testing dataset that classified 250 users.

It is noticed from Table VII that the proposed Arch#4 is consistent in its performance with almost all the split ratios, ruling out the possibility of overfitting. It is also noticed that from Table IX that the proposed model is unabated even by the substantial increase in the number of users with various types of ECG beats. The overall testing accuracy obtained for 250 users is 99.92% which is only 0.02% lesser than what it obtained with half the number of users.

B. General Discussions

TABLE X
COMPARATIVE ANALYSIS OF RECENT ECG BASED USER RECOGNITION STUDIES USING DL WITH THE PROPOSED STUDY

Sl.No.	Reference	Method	Number of Subjects	Accuracy (%)
1	Page et al. (2015) [22]	DNN	90	99.85
2	Wieclaw et al. (2017) [23]	DNN	18	88.97
3	Zhang et al. (2018) [24]	CNN	10 (Single arm)	98.4
4	Donida et al. (2019) [27]	Deep CNN	52	100
5	Pourbabaee et al. (2017) [28]	Deep CNN	33	95.95
6	Lynn et al. (2018) [29]	DNN	20	94
7	Bogdanov et al. (2018) [30]	Deep LSTM	290	99
8	Deshmane et al. (2018) [31]	CNN	40 (Fantasia)	96.95
9	Kim et al. (2019) [32]	Ensemble of deep CNN	18 (MITNSR)	98.9
10	Abdeldayem et al. (2019) [1]	Deep CNN	488	94.9
11	Hammad et al. (2020) [48]	ResNet Attention network	PTBDB (not specified)	98.85
			CYBHi DB (not specified)	99.27
12	**Proposed study**	**Dense Framework (Arch#4)**	**120**	**99.94**
			250	**99.92**

The main contribution of this study is the development of a reliable and efficient user identification system that obtains very high identification accuracy with varied possible types of ECG waveforms in a big population. To address these issues, the proposed research investigates four potential dense learning framework architectures. Our previous study [46] highlighted the capacity of a dense network for ECG-based user identification, in which the network was tested for 60 users and its performance was compared to three other networks to ensure its trustworthiness. In the proposed study, a dataset of 35000 templates (images) of 250 people of both genders with diverse cardiac conditions, a wide age range, and extended ECG recordings recorded over several days was gathered to add variation to the sample. Only if a network can reliably identify all of the differences in a lot can it be said to be efficient. Arch#4 was judged to be the most successful of all the models. The proposed model was compared to other recently published models in the literature. A brief comparison of existing ECG-based user recognition experiments employing DL with the proposed study is presented in Table X. Most of the previous investigations [22]-[24], [27]-[29], [31], [32] used a relatively small dataset that was either self-collected or derived from only two or three public databases. Conclusions drawn on such studies are not reliable. Moreover, the identification rate dropped significantly when the dataset was extended [1]. In addition to that, the identification accuracy achieved was also not satisfactory.

The proposed approach, in contrast to earlier research, focuses on making the user identification system robust and efficient. As a result, the proposed study creates a robust DL model that comprehensively learns all of the ECG waveform patterns in the lot. Furthermore, with the proposed basic architectural arrangement, the network achieves an excellent identification rate of 99.92% for 250 users, making it incredibly efficient. The suggested Arch#4 has a training duration of roughly 5 hours due to the decreased amount of computable parameters.

V. Conclusion

Imposter activities are on the rise as the number of computer and internet apps grows. As a result, practically in every industry today, an automatic and efficient user identification system is a must. The goal of this research is to use DenseNets to create an efficient ECG-based user recognition system. In terms of training time and prediction accuracy, these networks are quite effective. When the suggested Dense CNN model was trained for 250 users, it achieved a 99.92% identification rate. Dense networks appear to be particularly promising for ECG-based user recognition, according to this research.

The research can be carried out further using various architectural layouts of dense networks and a variety of hyper-parameter adjustments to develop a model with the smallest computable parameters and the highest accuracy.

References

[1] Abdeldayem SS, Bourlai T. A Novel Approach for ECG-Based Human Identification Using Spectral Correlation and Deep Learning. IEEE Trans Biometrics, Behav Identity Sci. 2019;2(1):1–14.
[2] Belgacem N, Amine Naït A, Fethi R. Person Identification System Based on Electrocardiogram Signal Using Lab VIEW. Int J Comput Sci Eng. 2012;4(06):974–81.
[3] Agrafioti F, Hatzinakos D. ECG based recognition using second order statistics. Proc 6th Annu Commun Networks Serv Res Conf CNSR 2008. 2008;82–7.
[4] Hoekema R, Uijen GJH, Van Oosterom A. Geometrical aspects of the interindividual variability of multilead ECG recordings. IEEE Trans Biomed Eng. 2001;48(5):551–9.
[5] Agrafioti F, Hatzinakos D. ECG biometric analysis in cardiac irregularity conditions. Signal, Image Video Process. 2009;3(4):329–43.
[6] Biel L, Pettersson O, Philipson L, Wide P. ECG analysis: A new approach in human identification. IEEE Trans Instrum Meas. 2001;50(3):808–12.
[7] Kyoso M, Uchiyama A. Development of an ECG identification system. In: Annual Reports of the Research Reactor Institute, Kyoto University. 2001. p. 3721–3.
[8] Israel SA, Irvine JM, Cheng A, Wiederhold MD, Wiederhold BK. ECG to identify individuals. Pattern Recognit. 2005;38(1):133–42.
[9] Wang Y, Plataniotis KN, Hatzinakos D. Integrating analytic and appearance attributes for human identification from ECG signals. Biometrics Symp BCC 2006. 2006;
[10] Plataniotis KN, Hatzinakos D, Lee JKM. ECG biometric recognition without fiducial detection. Biometrics Symp BCC 2006. 2006;(May 2014).
[11] Chan ADC, Hamdy MM, Badre A, Badee V. Wavelet distance measure for person identification using electrocardiograms. IEEE Trans Instrum Meas. 2008;57(2):248–53.
[12] Fatemian SZ, Agrafioti F, Hatzinakos D. HeartID: Cardiac biometric recognition. IEEE 4th Int Conf Biometrics Theory, Appl Syst BTAS 2010. 2010;1–5.
[13] Wang Y, Agrafioti F, Hatzinakos D, Plataniotis KN. Analysis of Human Electrocardiogram for Biometric Recognition. EURASIP J Adv Signal Process. 2007;2008(1).
[14] Coutinho DP, Fred ALN, Figueiredo MAT. One-lead ECG-based personal identification using Ziv-Merhav cross parsing. Proc - Int Conf Pattern Recognit. 2010;3858–61.
[15] Venkatesh N, Jayaraman S. Human electrocardiogram for biometrics using DTW and FLDA. Proc - Int Conf Pattern Recognit. 2010;3838–41.
[16] Choi HS, Lee B, Yoon S. Biometric Authentication Using Noisy Electrocardiograms Acquired by Mobile Sensors. IEEE Access. 2016;4:1266–73.
[17] Belgacem N, Fournier R, Nait-Ali A, Bereksi-Reguig F. A novel biometric authentication approach using ECG and EMG signals. J Med Eng Technol. 2015;39(4):226–38.
[18] Tan R, Perkowski M. Toward improving electrocardiogram (ECG) biometric verification using mobile sensors: A two-stage classifier approach. Sensors (Switzerland). 2017;17(2).
[19] Li M, Narayanan S. Robust ECG biometrics by fusing temporal and cepstral information. Proc - Int Conf Pattern Recognit. 2010;(January 2019):1326–9.
[20] Yadav N, Duhan M, Rose A. Biometric Human Recognition using ECG Signals. Iarjset. 2017;4(6):168–71.
[21] Patro K, extraction PK-. Optimized Feature Selection with Mutual Information for ECG based Bio-Metric Recognition system using Genetic Algorithm. ResearchgateNet (Internet). 2018;(May). Available from: https://www.researchgate.net/profile/Kiran_Patro/publication/325153313_Optimized_Feature_Selection_with_Mutual_Information_for_ECG_based_Bio-Metric_Recognition_system_using_Genetic_Algorithm/links/5afae6d2a6fdccacab1770a2/Optimized-Feature-Selection-with-
[22] Page A, Kulkarni A, Mohsenin T. Utilizing deep neural nets for an embedded ECG-based biometric authentication system. IEEE Biomed Circuits Syst Conf Eng Heal Minds Able Bodies, BioCAS 2015 - Proc. 2015;0–3.
[23] Wieclaw L, Khoma Y, Falat P, Sabodashko D, Herasymenko V. Biometrie identification from raw ECG signal using deep learning techniques. Proc 2017 IEEE 9th Int Conf Intell Data Acquis Adv Comput Syst Technol Appl IDAACS 2017. 2017;1:129–33.
[24] Zhang Q, Zhou D. Deep Arm/Ear-ECG Image Learning for Highly Wearable Biometric Human Identification. Ann Biomed Eng. 2018;46(1):122–34.
[25] Zhang Q, Zhou D, Zeng X. PulsePrint: Single-arm-ECG biometric human identification using deep learning. 2017 IEEE 8th Annu Ubiquitous Comput Electron Mob Commun Conf UEMCON 2017. 2017;2018-Janua:452–6.
[26] Eduardo A, Aidos H, Fred A. ECG-based biometrics using a deep auto encoder for feature learning an empirical study on transferability. ICPRAM 2017 - Proc 6th Int Conf Pattern Recognit Appl Methods. 2017;2017-Janua(Icpram):463–70.
[27] Donida Labati R, Muñoz E, Piuri V, Sassi R, Scotti F. Deep-ECG: Convolutional Neural Networks for ECG biometric recognition. Pattern Recognit Lett (Internet). 2019;126:78–85. Available from: https://doi.org/10.1016/j.patrec.2018.03.028
[28] Pourbabaee B. Deep Convolutional Neural Network for ECG-Based Human Identification. C Proc (Internet). 2017;41:7–10. Available from: https://omsignal.com/wp-content/uploads/2017/12/OMsignal-Note-2017-009.pdf
[29] Lynn HM, Yeom S, Kim P. ECG-based biometric human identification based on backpropagation neural network. Proc 2018 Res Adapt Converg Syst RACS 2018. 2018;(1):6–10.
[30] Bogdanov MR, Dumchikova IN, Dokuchaev IS. Deep Learning Based Person Biometric Identification. Int Sci J 'Industry 40'. 2018;(4):219–22.
[31] Deshmane M, Madhe S. ECG Based Biometric Human Identification Using Convolutional Neural Network in Smart Health Applications. Proc - 2018 4th Int Conf Comput Commun Control Autom ICCUBEA 2018. 2018;1–6.
[32] Kim MG, Ko H, Pan SB. A study on user recognition using 2D ECG based on ensemble of deep convolutional neural networks. J Ambient Intell Humaniz Comput (Internet). 2019;0(0):0. Available from: http://dx.doi.org/10.1007/s12652-019-01195-4
[33] Kim JS, Kim SH, Pan SB. Personal recognition using convolutional

neural network with ECG coupling image. J Ambient Intell Humaniz Comput (Internet). 2019;(0123456789). Available from: https://doi.org/10.1007/s12652-019-01401-3

[34] Bento N, Belo D, Gamboa H. ECG Biometrics using Spectrograms and Deep Neural Networks. Int J Mach Learn Comput (Internet). 2020;Vol. 10, N(February). Available from: https://www.researchgate.net/publication/333311161

[35] Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. IEEE Eng Med Biol Mag. 2001;20(3):45–50.

[36] Bousseljot R, Kreiseler, D. A, Schnabel, A. N. EKG-Signaldatenbank CARDIODAT der PTB über das Internet. Biomed Tech / Biomed Eng. 1995;pp.317-318.

[37] Laguna P, Mark RG, Goldberg A, Moody GB. Database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. Comput Cardiol. 1997;(October 1997):673–6.

[38] Tatiana S. L. Biometric Human Identification Based on Ecg. 2015. p. 1–15.

[39] Nolle FM, Badura FK, Catlett JM, Bowser RW, Sketch MH. Crei-Gard, a New Concept in Computerized Arrhythmia Monitoring Systems. Computers in Cardiology. 1987. p. 515–8.

[40] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals. 2000;

[41] LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation applied to digit recognition. Vol. 1, Neural computation. 1989. p. 541–51.

[42] Rumelhart DE, Hinton GE, Williams RJ. Output Patterns. Parallel Distrib Process Explor Microstruct Cogn (Internet). 1987;567. Available from: https://web.stanford.edu/class/psych209a/ReadingsByDate/02_06/PDPVolIChapter8.pdf

[43] Wolpert DH, Nowlan SJ. Simplifying Neural Networks by Soft Weight Sharing. The Mathematics of Generalization. 2018. p. 373–94.

[44] Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging. 2018;9(4):611–29.

[45] Abdelwahab M, Busso C. Study of Dense Network Approaches for Speech Emotion Recognition. 2018 IEEE Int Conf Multimed Expo. 2018;5084–8.

[46] Wu J, Xie C-W, Luo J-H. Dense CNN Learning with Equivalent Mappings. 2016;(Nips). Available from: http://arxiv.org/abs/1605.07251

[47] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proc - 30th IEEE Conf Comput Vis Pattern Recognition, CVPR 2017. 2017;2017-Janua:2261–9.

[48] Hammad M, Pławiak P, Wang K, Acharya UR. ResNet-Attention model for human authentication using ECG signals. Expert Syst. 2020; (November 2019):1–17.