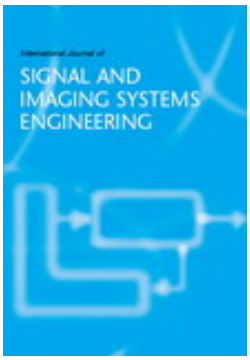


International Journal of Signal and Imaging Systems Engineering

ISSN online: 1748-0701 - ISSN print: 1748-0698
https://www.inderscience.com/ijsise

Involutional neural networks for ECG spectrogram classification and person identification

Hatem Zehir, Toufik Hafs, Sara Daas

DOI: 10.1504/IJSISE.2024.10064382

Article History:

Received:	03 March 2024
Last revised:	06 May 2024
Accepted:	07 May 2024
Published online:	15 July 2024

Involutional neural networks for ECG spectrogram classification and person identification

Hatem Zehir*, Toufik Hafs and Sara Daas

Faculty of Technology,
LERICA Laboratory,
Badji Mokhtar-Annaba University,
P.O. Box 12, Annaba, 23000, Algeria
Email: hatem.zehir@univ-annaba.dz
Email: toufik.hafs@univ-annaba.dz
Email: sara.daas@univ-annaba.dz
*Corresponding author

Abstract: Reliable personal identification is crucial. This study explores an ECG-based system using spectrograms and involutional neural networks (INNs) for accurate person identification. The system first preprocesses ECG signals with a Butterworth filter and segments them using the Pan-Tompkins++ algorithm. Each segment, representing a heartbeat, is then converted into a spectrogram using the short-time Fourier transform (STFT). Finally, an INN classifies the spectrograms for identification. Tested on the MIT-BIH Arrhythmia and Physikalisch-Technische Bundesanstalt (PTB) databases, the system achieved accuracies of 97.93% and 97.63%, respectively, surpassing conventional methods like convolutional neural networks (CNNs). This improvement is attributed to INNs' ability to better capture both local and temporal patterns in the spectrogram data.

Keywords: biometrics; deep learning; electrocardiogram; INNs; involutional neural networks; Person Identification.

Reference to this paper should be made as follows: Zehir, H., Hafs, T. and Daas, S. (2024) 'Involutional neural networks for ECG spectrogram classification and person identification', *Int. J. Signal and Imaging Systems Engineering*, Vol. 13, No. 1, pp.41–53.

Biographical notes: Hatem Zehir is a Master's graduate in instrumentation engineering from the University of Badji Mokhtar Annaba and is currently pursuing his PhD at the same institution. His research interests are signal processing, biometrics, biomedical engineering, and deep learning. He aims to develop secure and efficient biometric systems, exploring artificial intelligence to analyse biomedical signals.

Toufik Hafs is an Associate Professor at the University of Badji Mokhtar Annaba. He obtained his Bachelor's degree in Electronic Engineering in 2008 and continued his academic studies with a Master's degree in Instrumentation and Information Processing in 2010, both from the University of Badji Mokhtar Annaba. Building upon this educational background, he went on to earn his PhD in Instrumentation and Information Processing from the same institution in 2016. His research focuses on the areas of biometrics, signal processing, and digital image processing.

Sara Daas is an Associate Professor at the University of Badji Mokhtar Annaba. She obtained her Bachelor's degree in medical Electronic Engineering in 2011 and continued her academic studies with a Master's degree in Biomedical Engineering in 2013, both from the University of Constantine 1. Building upon this educational background, she went on to earn her PhD in image and signal processing from the University of Badji Mokhtar Annaba in 2021. Her research interests are biometrics and image processing.

1 Introduction

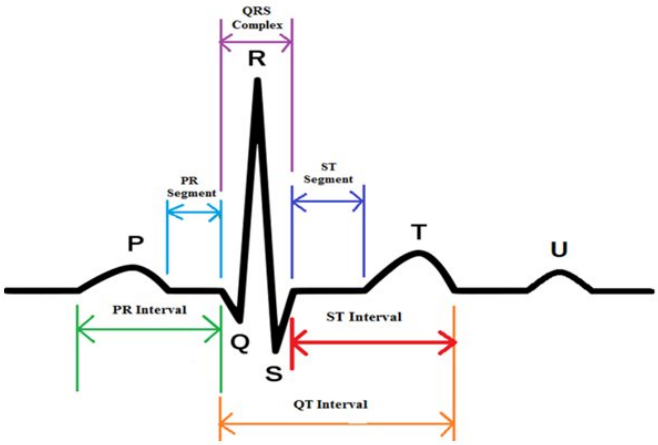
The need for secure and reliable person identification systems has been increasing in recent years. Traditional biometric modalities such as fingerprints (Yin et al., 2021), handwritten signatures (Jain et al., 2020), voice (Zouhir et al., 2020; Fredj et al., 2018), finger knuckle print (Malik et al., 2016) and facial recognition (Taskiran et al., 2020), have gained widespread adoption in various areas such as financial transaction, access control, and border security because of their robustness

and acceptance among users but they are susceptible to spoofing attacks. here is where electrocardiogram (ECG) signals emerge as a promising biometric trait because of their immunity to spoof in addition to their uniqueness and reliability.

The heart generates electrical signals that can be recorded as ECG signals, those signals exhibit complex variations that reflect an individual's unique physiological characteristics. These variations, influenced by various factors such as heart rate, rhythm, and morphology, provide a rich source of information for biometric identification. Compared to other biometric modalities, ECG-based biometrics possess several advantages. For example, they are unique, meaning every individual has a distinct ECG signal due to variations in many factors such as the heart anatomy. Unlike classical modalities such as facial recognition, ECG signals cannot be easily reproduced, which makes them difficult to forge. Another advantage is the continuous availability of ECG signals even during sleep or rest which enables continuous identification capabilities.

Figure 1 shows the four main waves of ECG: P, QRS, T, and U waves. Each wave has a unique amplitude and duration and can serve as a feature for biometric recognition. Additionally, three main intervals: PR, QT, and ST, are evident in the ECG. These intervals reflect various aspects of heart activity and can also be used as features for individual identification.

Figure 1 Graphical representation of a single heartbeat (see online version for colours)



Source: (Sinha, 2012)

ECG biometric systems address the limitations of traditional modalities, for example, fingerprint scanners can be easily tricked using prosthetics. Facial recognition algorithms can be fooled using a picture or a video which raises privacy concerns. On the other hand, ECG signals offer a more secure and immune approach due to their physiological origin which makes them harder to reproduce and less likely to spoof giving them an advantageous boost in applications where high security is required such as banking.

Prior research has explored various methods for extracting features and classifying ECG signals to achieve accurate and reliable biometric systems. The concept of ECG-based biometrics emerged in the work of Biel et al. (2001), the authors proved that ECG signals can be used as a biometric trait by collecting 12 lead recording from 20 persons ranging in age from 20 to 55 years from both genders. They extracted 30 features such as the durations of the different waves and their onsets from the 12 leads and an accuracy of 100% was achieved using the SIMCA (Wold, 1976) classification method. The initial research in ECG biometrics focused on extracting fiducial time-domain features from the signals. These features, such as heart rate, R-R interval, and QT duration, were used to represent an individual's ECG signature. Kyoso and Uchiyama (2001) used four features, the duration of the P wave and the intervals of the PQ, QRS, and QT waves extracted from the limb lead II of recordings collected from 9 different subjects ranging in age from 21 to 35 years. Shen et al. (2002) used recordings from 20 individuals obtained from MIT-BIH, they started by removing noises from the signals by applying preprocessing techniques, after that, they extracted seven QRST features from 20 heartbeats of each individual resulting in a total of 400 features, for the classification, two methods were used: template matching achieved an accuracy of 95% and a decision-based neural network (DBNN) that achieved an accuracy of 80%, an accuracy of 100% was achieved by combining the two methods. Israel et al. (2005) used signals with 1000 Hz as a sampling frequency from 29 subjects ranging in age from 22 to 48 over 41 sessions. The authors started by removing noises using a filter with [2–40] Hz bandpass. After that fiducial points were detected to extract the features that will be used to train discriminant functions for individual identification. However, these features were found to be susceptible to variations due to physiological factors and environmental conditions, limiting their accuracy in person identification. This is why researchers focused on developing more robust feature extraction techniques to capture more representative patterns embedded in ECG signals. Non-fiducial characteristics like frequency-domain features and techniques were introduced to provide a more comprehensive representation of ECG signals. In the research by Zehir et al. (2023b), the authors used single lead signals from 47 subjects obtained from MIT-BIH. They started by processing the signals to remove noises using a Butterworth filter of the 4th order with a frequency range for cutoff between 1 Hz and 40 Hz. After that, they segmented the signals around each R-peak into windows of around 700 ms. Four frequency domain features were extracted and are power spectral density (PSD) (Stoica and Moses, 2005), band power (Stoica and Moses, 2005), median frequency (Phinyomark et al., 2012), and mean frequency (Phinyomark et al., 2012). To classify individuals, three types of support vector machine (SVM) kernels (Cervantes et al., 2020) were used, linear SVM achieved an accuracy of 93.6%, quadratic SVM achieved 96.4%, and cubic SVM achieved 97.0%. In another paper, Pinto et al. (2017) used ECG signals recorded using a steering wheel. After the acquisition of the signals, they were denoised using Savitzky-Golay and moving average filters. Then, the signals were segmented into heartbeats around R-peaks detected using the Trahanias algorithm (Trahanias, 1993). The amplitude of each obtained segment was subject to a z-score normalisation. The discrete

cosine transform (DCT) (Othman and Zeebaree, 2020) and the Haar Wavelet transform (Guo et al., 2022) were used to extract features after detecting and removing outliers. The proposed method achieved an identification rate of 94.9% and an equal error rate (EER) of 2.66% for authentication. In Belgacem et al. (2013), the authors used ECG data obtained from four physionet databases and a fifth database consisting of volunteers studying at Paris Est University. To remove the noises from those signals, they are bandpass filtered and the frequency is limited in the range of [1–40] Hz. Following this, the Discrete Wavelet Transform (DWT) (Osadchiy et al., 2021) is applied to extract features. These features are subsequently input into a random forest classifier, resulting in a recognition rate of 100%. Researchers also investigated various classification algorithms. Conventional approaches to machine learning, like SVMs (Zehir et al., 2023b), and the method of k-nearest neighbours (k-NN) (Camara et al., 2018), were initially employed. However, with the advancement of deep learning techniques, convolutional neural networks (CNNs) (AlDuwaile and Islam, 2021), long short-term memory (LSTMs) (Jyotishi and Dandapat, 2020), and gated recurrent units (GRUs) (Zehir et al., 2023a) have emerged as powerful tools for ECG-based biometrics. However, emerging kernels such as involutional neural networks (INNs) (Li et al., 2021a) are yet to be explored for this specific type of biometrics.

The paper (Li et al., 2021a) introduces a novel operation for deep neural networks called involution, which is proposed as an alternative to the standard convolution operation used in vision tasks. The key idea behind involution is to invert the design principles of convolution, making it spatial-specific and channel-agnostic. This means that involution kernels are unique for each spatial location but shared across channels, allowing for dynamic parameterisation and efficient handling of variable feature resolutions. The authors demonstrate that involution-based models can achieve superior performance on various benchmarks, including ImageNet classification and COCO detection and segmentation while reducing computational costs. The paper also discusses the relationship between involution and self-attention, suggesting that self-attention can be seen as a complex form of involution. The proposed involution operation is shown to be versatile and effective across a range of vision tasks, offering a promising direction for future research in visual recognition.

By reviewing the existing literature, we can notice that despite the importance of ECG signals as a biometric trait, several algorithms such as the ability of INNs to improve identification accuracy are yet to be explored. This paper aims to address this gap and explore the combination of a technique of feature extraction that employs the short-time Fourier transform (STFT) and INNs as a classification algorithm. The main contributions of this paper are:

- investigating the integration of INN classification method and spectral characteristics of heartbeats for person identification using ECG signals
- comparing the performance of the STFT-INN model with a traditional CNN architecture, both trained using identical datasets
- the comparison of the proposed approach, to other state-of-the-art methods that were tested on MIT-BIH and PTB databases

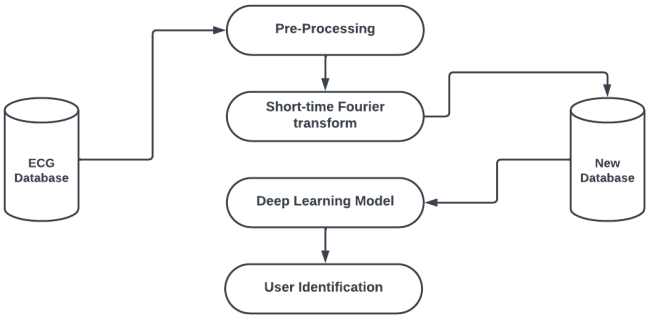
To the fullest extent of our knowledge, this is the first paper to explore those specific settings, as the combination of STFT and INNs is not been previously investigated in the literature.

To provide a clear and organised overview of the proposed approach, the following structure has been adopted: Section 2 presents the proposed system in detail, outlining the methods, techniques, and algorithms used. In Section 3, the experimental results of the study are discussed, with a focus on the analysis and interpretation. Finally, Section 4 concludes the paper, summarising the findings and outlining future directions for research in this area.

2 Proposed method

A graphical representation of the proposed approach is illustrated in Figure 2. ECG signals that are part of the MIT-BIH and PTB databases were employed for this study. The signals were preprocessed and segmented to enhance their quality and prepare them for subsequent analysis. STFT was utilised to extract features from the preprocessed ECG signals. These extracted features were then fed into an INN-based deep-learning model for biometric identification.

Figure 2 Schematic representation of the proposed system's architecture, illustrating the interconnected components



2.1 ECG databases

The MIT-BIH Database (Moody and Mark, 2001; Goldberger et al., 2000) served as the first ECG data source utilised in this paper. Data collection originally started in 1975 at the Beth Israel Hospital (BIH), long before ECG signals were proven as a biometric trait, however, it was proved to be valid for human identification in previous research (Hamza and Ayed, 2022; Fatimah et al., 2022; Lynn et al., 2019). This publicly available database is composed of 48 two channels recordings collected from 47 unique individuals, each recording is half-hour long.

There are two recordings from the same subject (signals 201 and 202), for this study the 202 signal was excluded. The recordings were collected from a diverse population, covering a broad age range from 23 to 89 years old and a balanced gender distribution, with 26 male and 25 female participants. Of these individuals, 23 were chosen randomly, and the remaining 24 were patients diagnosed with significant arrhythmias. All signals within MIT-BIH were recorded at a sampling frequency of 360 Hz.

The PTB Database (Goldberger et al., 2000; Bousseljot et al., 1995) is the second database used in this study. It is another publicly available database from physionet. It contains more participants compared to the MIT-BIH database, it is a collection of 549 recordings collected from 290 unique individuals. The participant demographic is also diverse, showcasing a broad age range spanning from 17 to 87 years old, but it has a non-balanced gender distribution, with 209 males and 81 females. The data incorporates individuals with and without heart conditions. While 216 subjects in the database had documented health conditions, 22 had no clinical reports, and 54 served as healthy controls. Each of the 549 recordings has 15 simultaneous signals, consisting of the standard 12-lead ECG configuration alongside the 3-Frank lead configuration. Additionally, the signals collection process adopted a higher sampling frequency of 1000 Hz compared to MIT-BIH.

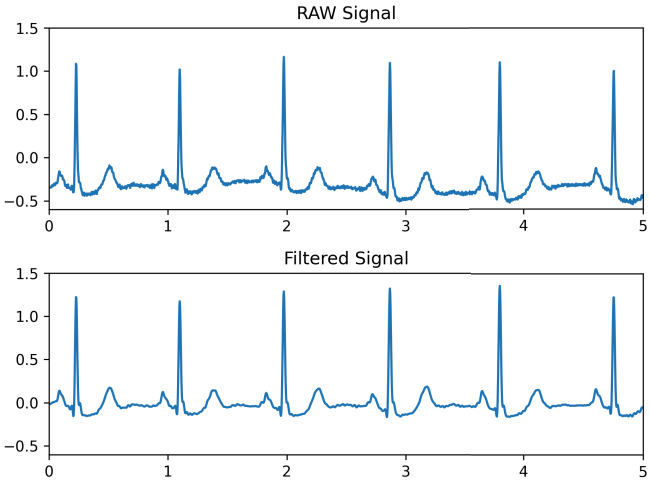
2.2 Filtering

Clean ECG signals are essential for getting better identification rates, but unfortunately, they are often contaminated by various types of noises that originate from various sources, such as baseline wander which is a low-frequency noise that can be caused by physiological factors related to the patient such as respiration. Muscle movements and contractions are another source of low-frequency noises and it is also known as electromyographic noise. Additionally, power line interference is a common type of noise, it occurs as sinusoidal waves of a frequency that could be either 50 Hz or 60 Hz depending on the country. When the electrodes attached to the patient's skin move due to factors such as sweat or loose leads it can cause changes in the recorded signal. Finally, ambient noise from nearby electronic devices can interfere with the ECG signal making it more noisy.

To ensure that the ECG signals that are used to extract features are reliable in biometric settings and to effectively remove noise from those signals, a Butterworth bandpass filter is employed. The filter parameters are as follows: the sampling frequency is set at 1000 Hz when filtering signals from the PTB datasets and 360 Hz when filtering signals from the MIT-BIH dataset. The lower and upper cutoff frequencies are specified as 1 Hz and 40 Hz, respectively. Additionally, a filter order of 4 is selected. These parameters are essential for designing the Butterworth filter, ensuring that it effectively attenuates noise and preserves the desired frequency components within the ECG signals while effectively suppressing unwanted noise and interference. The lower cutoff frequency eliminates baseline wander and low-frequency noise, while the upper cutoff frequency preserves the shape of the signals represented by the P, QRS, and T waves.

Figure 3 shows a 5-second segment from a randomly selected signal from MIT-BIH. The upper plot demonstrated the raw, unfiltered signal. In contrast, the lower plot shows the same signal post-filtering, where noise and unwanted artefacts have been reduced using a 4th-order Butterworth filter.

Figure 3 Noise reduction in a randomly selected ECG recording sourced from MIT-BIH using a Butterworth filter (see online version for colours)



2.3 Segmentation

Segmentation of ECG signals into individual heartbeats is an important step in ECG-based biometrics. To achieve this segmentation, the Pan-Tompkins++ algorithm (Imtiaz and Khan, 2022) was employed, it is a robust and adaptable technique for R-peak detection in ECG signals based on the well-known and reliable Pan-Tompkins algorithm (Pan and Tompkins, 1985) and offers enhanced performance in peaks detection in comparison with his older brother.

The algorithm starts by preprocessing the signal using a filter with [5–18] Hz as a bandpass to enhance its quality and reduce the presence of noise. After that, the signal is differentiated to get a more informative QRS slope. Then, the filtered signal is squared to make the QRS complexes more prominent. An integrating moving window of 150 ms is applied to extract information from the waveform after smoothing the signal using a 60ms flattop window.

For the decision-making stage, peaks that are at least 231ms apart are selected and thresholds are initialised. Peaks exceeding Threshold 1 are classified as R-peaks, with signal peak (SPK) and noise peak (NPK) adjustments using Rule-1 which is described by equation (1). If the number of detected beats exceeds 8, it calculates the Mean RR interval and classifies peaks based on RR interval conditions, adjusting SPK and NPK accordingly. For RR intervals below 360ms or 0.5 times the Mean RR interval, peaks are classified as T-wave or QRS complex based on slope, with additional adjustments. If the RR-interval surpasses 1s or 1.66 times the Mean RR-interval, it employs Threshold 3 and Rule-2 (this

rule is defined by equations (2) and (3)) adjustments for R-peak classification. If the RR interval exceeds 1.4s, peaks exceeding 0.2 times Threshold 2 are classified as R-peaks with Rule-2 adjustments. Threshold 1 and Threshold 2 are updated iteratively, providing a dynamic approach to peak classification and threshold adaptation.

$$SPK = 0.125PEAK + 0.875SPK \tag{1}$$

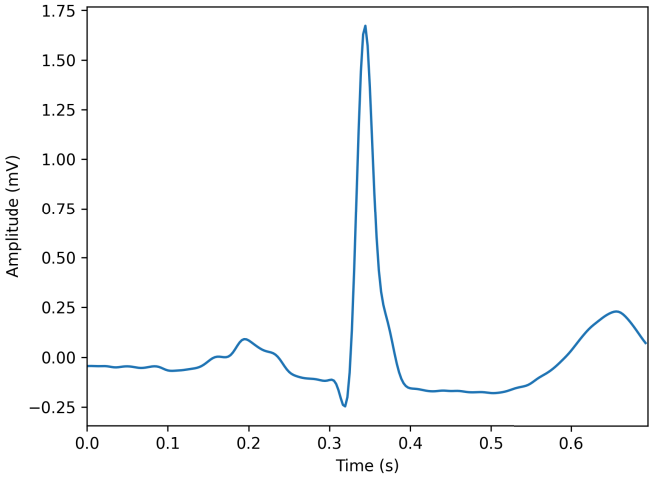
$$SPK = 0.75PEAK + 0.25SPK \tag{2}$$

$$NPK = 0.75PEAK + 0.25NPK \tag{3}$$

Following R-peak detection, a windowing approach to isolate individual heartbeats is implemented. Windows of 251 samples for MIT-BIH and 701 samples for PTB, centred around each detected R-peak, were created, The choice of different sample numbers for each database was necessitated by their different sampling frequencies. Each window spans approximately 700 milliseconds, representing a single heartbeat. This segmentation ensures that subsequent feature extraction only represents the characteristics of a single heartbeat.

Figure 4 represents a heartbeat segment extracted from a randomly chosen signal within MIT-BIH. The analysis of isolated heartbeat segments will help improve the feature extraction process and improve identification accuracy.

Figure 4 Segmented heartbeat from MIT-BIH prepared for feature extraction (see online version for colours)



2.4 Feature extraction

In signal processing applications, the Fourier transform (Gasmi, 2022), as described by equation (4), is a technique to get insights of a signal in terms of its frequency components rather than its time-varying amplitude. However, it has some limitations as we can't know exactly where a given frequency is located in time resulting in a complete loss of the temporal information.

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft}\mathrm{d}t \tag{4}$$

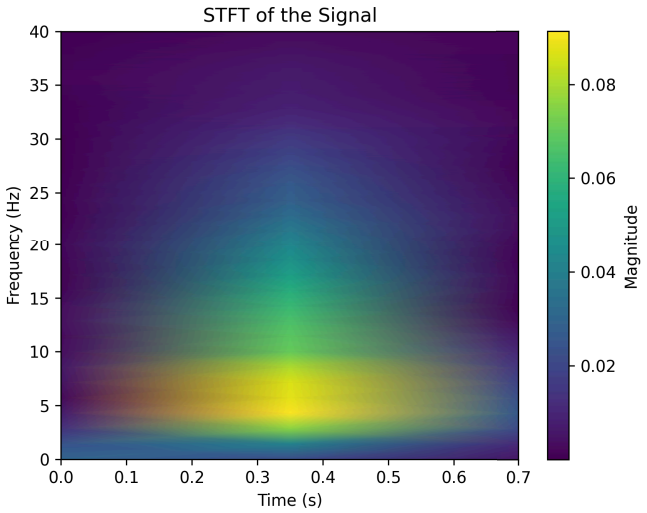
To overcome these problems and keep time information while transforming to the frequency domain, a new approach known as the STFT (Kumar et al., 2019; Mateo and Talavera, 2018) was proposed that allows the study of the signal in smaller windows. STFT is based on the use of a window, which is a regular function well localised in time or space (it is zero outside a certain area). The multiplication of the signal $f(t)$ by a sliding window $h(t - t_0)$ and the calculation of the Fourier transform of this product is given by the following mathematical relationship:

$$G_f(\nu, t_0) = \int_{-\infty}^{+\infty} f(t)h(t - t_0)e^{-2\pi i\nu t}\mathrm{d}t \tag{5}$$

Studying the signal after multiplication by this technique gives a local aspect, making it possible to determine where the different frequencies are located in time, which means that STFT is a technique that enables spectral analysis.

To extract features from the preprocessed ECG segments, time-domain signals were transformed into time-frequency representations using STFT. As discussed earlier, this technique decomposes the signal into its constituent frequency components over time, generating a spectrogram that visualises the energy distribution of the signal across different frequencies and time intervals. Figure 5 illustrates a spectrogram of a randomly selected segment, displaying the heartbeat's frequency spectrum over time.

Figure 5 Signal's frequency components over time using a spectrogram extracted from a randomly selected segment from PTB (see online version for colours)



2.5 Deep learning

2.5.1 CNN model structure

CNN-based architectures (Li et al., 2021b; Gu et al., 2018) are a widely used type of neural network, especially in the

field of image processing. They are known for their ability to effectively extract features. In general, and as shown in Figure 6, a CNN block is composed of three layers, the first one is the convolution layer (CL), which is followed by an activation layer, a pooling layer (SL) to subsample data, and finally a flattening layer followed by fully connected layers (FC) for classification.

2.5.2 Convolution layer

The first and most important layer in a CNN block is the CL. To create new feature maps, the data forwarded to this layer is convolved using a collection of predetermined filters, also known as kernels, that are learned during training. The following equation describes the convolution operation:

$$X_j^{(l)} = f(\sum_{n=1}^{N} W_j^{(n,l)} \otimes X_{j-1}^n + B_j^l) \tag{6}$$

In the previous equation:

- $\otimes$ is the operator of convolution
- f is the activation function applied to the input
- $W_j^(n,l)$ is the nth channel filter from the lth convolution group at the layer indexed by j
- B_j^l stands for the bias term of the jth neuron in the lth group filter
- and finally, N represents the input feature maps' channel index number.

To further understand convolution, we consider **X**, an input tensor that represents input data which is in our case spectrograms extracted by STFT and has three dimensions, its height **H**, its width **W**, and its number of input channels denoted by C_{in} as shown in Figure 7. Each one of the convolutions Kernels C_{out} has a shape of **K**, **K**, and C_{in}, where **K** represents the kernel size; these kernels act as filters that extract features from input data. Using the multiplication operator between the kernel and corresponding input region and following it with a summation we obtain a single output tensor **Y** that also has three dimensions denoted by: **H**, **W**, C_{out}; the output tensor represents the extracted features. We can deduce from this process that the convolution has two particularities, the first one is the channel specificity which means that each output channel represents a unique feature extracted by a distinct kernel; the second one is location agnosticity which means that kernels capture spatial patterns within the input tensor regardless of their location.

2.5.3 LeakyReLU layer

By placing a layer operating a mathematical function on the output signals, also referred to as the activation function, between the processing layers of the neural network, it is frequently possible to increase the processing efficiency of the network. Since LeakyReLU has been shown to provide better results, it is one of the most widely used activation functions. The following equation serves as the definition of this function:

$$f(x) = \max(0, x) + \alpha * min(0, x) = \begin{cases} x & if x \geq 0 \\ \alpha * x & otherwise \end{cases} \tag{7}$$

where α is the negative slope.

2.5.4 Pooling layer

Pooling is a type of image subsampling that is another key idea in CNNs. The input image is divided into several non-overlapping rectangles with n pixels on each side during the pooling phase. You can think of each rectangle as a tile. Every tile's output signal can be expressed as a function of the values that each of its constituent pixels took. By pooling, the intermediate image's spatial size is decreased, which lowers the number of parameters and, consequently, the processing power required to run the network. This layer can be defined by:

$$x_j^l = f(\beta_j^l \mathbf{down}(x_j^{l-1}) + b_j^l) \tag{8}$$

where **down** represents the subsampling operation, β is the bias, and s is the additive bias.

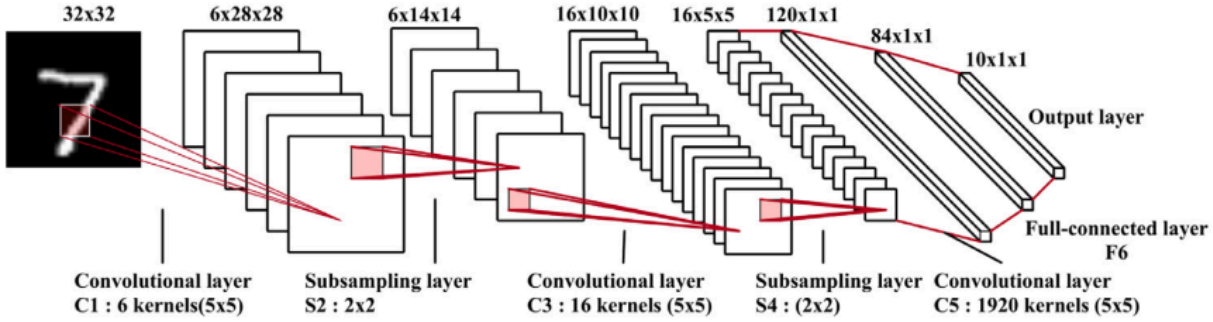
2.5.5 Proposed CNN model

To evaluate the effectiveness of the proposed system, a CNN-based model is trained in order to conduct a comparative analysis with the INN model proposed in Section 2.6. The CNN model's architecture is detailed in Table 1. It is composed of a two-dimensional convolutional layer with an input shape of (126, 3, 16) for the MIT-BIH and (351, 3, 16) for the PTB; followed by a LeakyReLU activation layer; after that, the output tensor is subsampled using a (2x1) pooling layer; then, the subsampled output tensor is converted into a one-dimensional vector using a flattening layer and fed into an FC layer with 256 hidden units; and finally, the classification task is performed using an FC layer that employs a softmax activation function, the number of subjects to be recognized and the number of hidden units are equal, so hidden units are set to 47 when training using MIT-BIH, and set to 290 when training using PTB.

Table 1 The architecture of the CNN model, showing the input layer and specific parameters of the subsequent layers for PTB

Layer	*Output shape*	*Param. #*
Conv2D	(None, 351, 3, 16)	160
LeakyReLU	(None, 351, 3, 16)	0
MaxPooling2D	(None, 176, 3, 16)	0
Flatten	(None, 8448)	0
Dense	(None, 256)	2162944
Dense	(None, 290)	74530

Figure 6 Standard topology of a CNN network (see online version for colours)

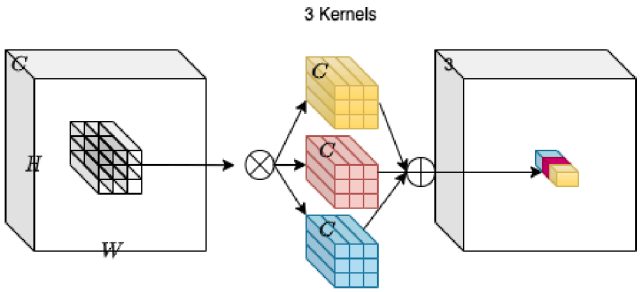


Source: Gu et al. (2018)

2.6 *Proposed involutional neural network*

To achieve a high identification accuracy, a deep learning model based on the INN architecture (Li et al., 2021a) is proposed. The model architecture is shown in Table 2. The proposed INN model is composed of an input of shape (126, 3, 1) for MIT-BIH and (351, 3, 16) for PTB followed by an involution layer; After that, a LeakyReLU activation layer is implemented; after that, the output tensor is subsampled with a (2x1) max pooling layer; then, the subsampled output tensor is converted into a one-dimensional vector using a flattening layer and fed into an FC layer with 256 hidden units; and finally, the classification task is performed using an FC layer that employs a softmax activation function; the number of hidden units is equal to the number of subjects to be identified and is set to 47 when training using signals from MIT-BIH and 290 when training using signals from PTB.

Figure 7 Visual representation of the involution operation's steps (see online version for colours)



Source: Keras (n.d.)

In contrast to conventional CNNs, which are location-agnostic and channel-specific, INNs exhibit inverse characteristics, meaning they are location-specific and channel-agnostic. This distinction results from the differences in their convolution kernel designs. In CNNs, the convolution kernel remains constant across the spatial dimension of the input tensor, while in INNs, the convolution kernel adapts to the spatial position of the input tensor. This adaptive nature allows INNs to capture local dependencies and spatial relationships within the input data, making them particularly well-suited for tasks involving spatial information, such as ECG signal analysis. The working of involution can be summarised in the following points:

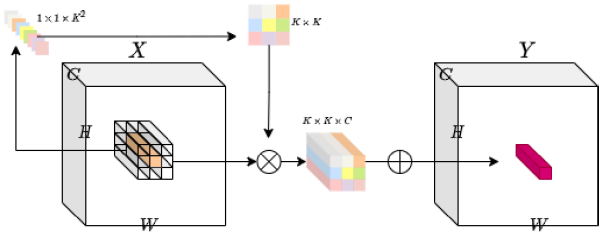
Table 2 The architecture of the proposed INN model, showing the input layer and specific parameters of the subsequent layers for PTB

Layer	*Output shape*	*Param. #*
Input layer	(None, 351, 3, 1)	0
Involution	(None, 351, 3, 1), (None, 351, 3, 9, 1, 1)	24
ReLU	(None, 351, 3, 1)	0
MaxPooling2D	(None, 175, 3, 1)	0
Flatten	(None, 525)	0
Dense	(None, 256)	134656
Dense	(None, 290)	74530

- **Kernel generation:** For each input feature map, an involution kernel is generated. This kernel is conditioned on the input tensor, meaning it changes based on the input.
- **Multiplication and addition:** The involution kernels are then multiplied with the input patches (small portions of the input image), similar to how convolution multiplies kernels with input patches. The results are then added together to produce the output feature map.

One of the key benefits of involution is that it has fewer parameters than convolution, making the network more efficient and less demanding in terms of computational power. The full involution operation is shown in Figure 8.

Figure 8 Visual representation of the involution operation's steps, including the transformation of input features through kernel generation and multiply-add operations (see online version for colours)



Source: Keras (n.d.)

3 Results and discussion

The proposed system was implemented using Python 3.9.18, a high-level, interpreted programming language. Python's open-source nature and extensive libraries make it a popular choice for deep-learning applications. For the deep learning part of the system, TensorFlow 2.10.1 is utilised alongside the Keras 2.10 library. TensorFlow is a free and open-source library created by Google Brain. Keras is a neural network library that works as an interface for the TensorFlow library. The codes were executed on a machine operating on Windows 11. In terms of hardware, the deep learning models were trained on a machine equipped with an Intel i3-10100F processor. This processor has 4 cores and 8 threads; it operates at a frequency of 2.30 GHz and maxes at a frequency of 4.30 GHz. The machine was also equipped with 8GB of DDR4-2666 memory (RAM). To speed up the neural network training and take advantage of the parallel computing abilities of GPUs, NVIDIA GTX 1650 GPU with 4GB of GDDR6 memory is used.

Four metrics were used to assess the system's performance:

- *Accuracy*: It evaluates how accurate a model is overall at making predictions.

$$Acc = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{9}$$

- *Precision*: it indicates the ability of a model to correctly identify labels or classes.

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{10}$$

- *Recall*: sometimes referred to as true positive rate or sensitivity, is a metric used to assess how well a model can accurately identify every positive case found in a given dataset.

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{11}$$

- *F1-Score*: It is a popular assessment that integrates recall and precision into one overall model performance metric.

$$F_1 = \frac{2 * \mathrm{P} * \mathrm{R}}{\mathrm{P} + \mathrm{R}} \tag{12}$$

TP (True positive) can defined as when the model accurately predicts the positive class; in other words, the model correctly identified the user and the actual value was positive. TN is true negative, it is when the negative class is correctly predicted by the model. That is, the actual value was negative, and the model also predicted a negative result; in other words, it is when the model correctly rejects an individual; for example, the system correctly identifies that the input heartbeat does not match the one it has on its database. FP is false positive, it is when the positive class is wrongly predicted by the model; So, the model predicted a positive result however the actual value was negative; it can also be defined by the biometric system incorrectly identifying the individual. Finally, FN is false negative, it is when the negative class is wrongly predicted by the model. In other words, the model predicted a negative outcome, while the actual value was positive; For example, the system fails to recognise the heartbeat and incorrectly rejects an individual even though it should accept him.

Both the CNN and INN models were trained using the same hyperparameters. The training process for both models was conducted using identical hyperparameters to ensure a fair comparison of their performances. The learning rate was set to 0.001. It defines the step size at each iteration while attempting to minimise a loss function; a smaller learning rate could improve learning performance but slow down the learning process and lead to overfitting. The batch size was set to 150 which is proven to be the best value for ECG biometrics (AlDuwaile and Islam, 2021; Lynn et al., 2019, 2021); this refers to the number of training samples utilised in one iteration; the batch size can significantly influence model performance and convergence of the learning process. The Adaptive Moment Estimation (Adam) optimiser was used which is an optimisation algorithm known for its efficiency and relatively low memory requirements. And The models were trained for 30 epochs. An epoch represents the number of times when all of the training data is used once to update the weights in the training process. More epochs could allow the model to learn better, up to a certain point but could also lead to overfitting.

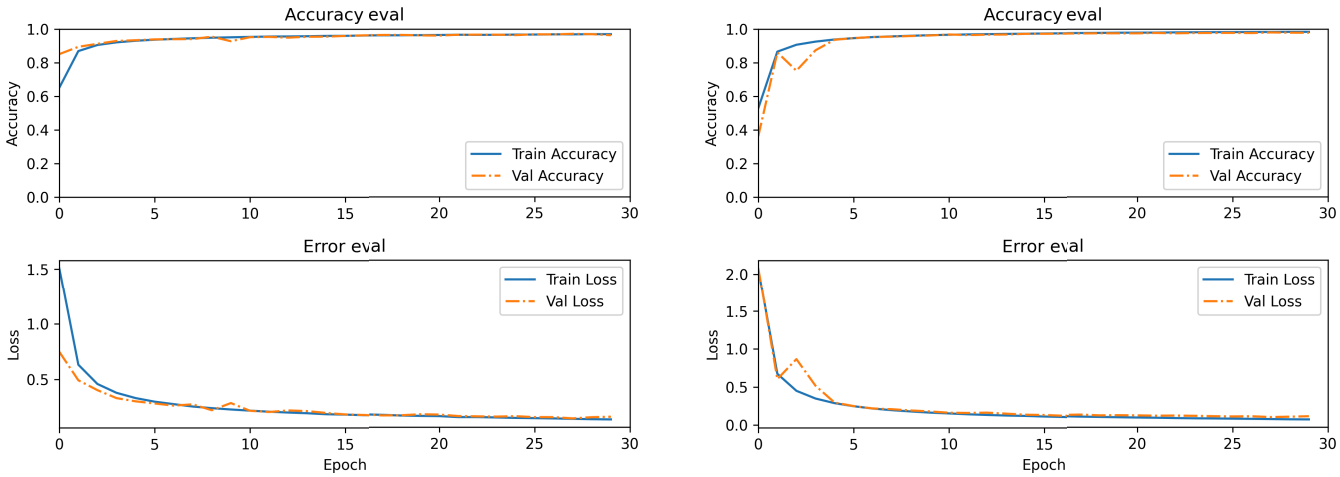
The datasets were separated into three subsets; namely, training, validation, and testing, which comprised 70%, 10%, and 20% of the signals, respectively. This is a common practice in deep learning applications, which helps in assessing the performance of the model. The validation set is used to fine-tune the hyperparameters during the training process, whereas the training set is used to train the model. Furthermore, the test set offers the optimal approximation of the model's performance.

The initial evaluation tests were conducted using MIT-BIH, which provides a balanced length of ECG records. All the signals utilised in this study are 30 min long, as detailed in Section 2.1. However, it's important to note that after the segmentation process, the number of heartbeats obtained varies from one subject to another (Imtiaz and Khan, 2022). In this study, only one lead signal is used. This approach is based on previous research (Biel et al., 2001; Zehir et al., 2023a) which has shown that ECG biometric identification can be effectively achieved using just one lead signal.

The training progress and loss of the CNN model on the MIT-BIH are illustrated at the left of Figure 9. Similarly, the right side of Figure 9 presents the training progress and loss of the INN model on the same database.

The performance results of both the CNN and INN models on MIT-BIH are tabulated in Table 3. A comparison of the results reveals that the INN model outperforms the CNN model across all four metrics. Specifically, the INN model achieved an accuracy of 97.93%, compared to 96.58% for the CNN model; The precision, recall, and F1-score were : 96.56%, 96.56%, and 96.49% respectively for the CNN; While the INN achieved 97.92%, 97.94%, and 97.92% of precision, recall, and F1-score respectively.

Figure 9 Accuracy and loss of the INN and CNN models on MIT-BIH: (a) CNN model and (b) INN model (see online version for colours)



(a) (b)

Table 3 Performance comparison of the different classification metrics of the proposed INN model with classical CNN after training on MIT-BIH

DL model	*Accuracy*	*Precision*	*Recall*	*F1-score*
CNN	96.58%	96.56%	96.56%	96.49%
INN	97.93%	97.92%	97.94%	97.92%

The second phase of the evaluation tests were conducted using PTB, which is an unbalanced dataset. This is because the number of records for each subject varies from 1 to 5, as discussed in Section 2.1. Similar to the first tests, only one lead signal was used during these tests. The training progress and loss of the CNN and INN models on PTB are illustrated in Figure 10.

The performance results of both the CNN and INN models on PTB are presented in Table 4 and shown in Figure. A comparison of the results reveals that the INN model outperforms the CNN model across all four metrics. Specifically, the INN model achieved an accuracy of 97.63%, compared to 97.52% for the CNN model. The precision, recall, and F1-score for the CNN were 96.73%, 96.41%, and 96.38% respectively. On the other hand, the INN achieved 96.89%, 96.56%, and 96.54% for precision, recall, and F1-score respectively.

Table 4 Performance comparison of the different classification metrics of the proposed INN model with classical CNN After training on PTB

DL model	*Accuracy*	*Precision*	*Recall*	*F1-score*
CNN	97.52%	96.73%	96.41%	96.38%
INN	97.63%	96.89%	96.56%	96.54%

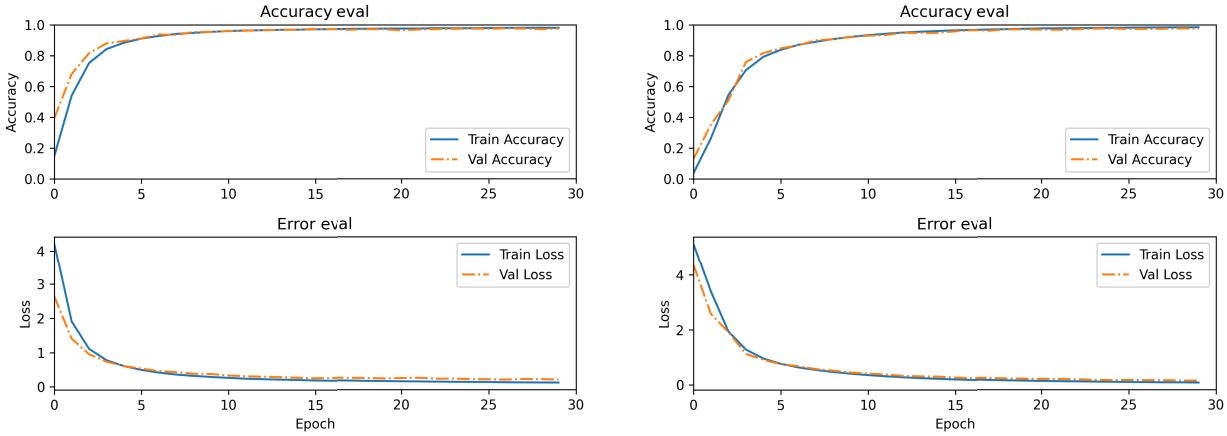
The performance of the model was explored using two different scenarios: one with all the subjects having 30 minutes of ECG recording using MIT-BIH, and another with a varying number of recordings from PTB, introducing a dataset imbalance which is a more effective approach to test the system's reliability. In both scenarios, our method achieved better results than the CNN model and other state-of-the-art ECG-based biometric systems, as shown in Table 5.

Table 5 Our introduced method in comparison With other State-of-the-art systems

Method	*Database*	*Classifier*	*Accuracy*
Hamza and Ben Ayed (2022)	PTB	SVM	95.40%
Fatimah et al. (2022)	MIT-BIH	SVM	97.92%
Zhang et al. (2017)	MIT-BIH	1D-CNN	91.10%
Jyotishi and Dandapat (2020)	PTB	LSTM	97.30%
	MIT-BIH		96.81%
Chu et al. (2019)	MIT-BIH	CNN	95.99%
Patro et al. (2022)	PTB	Random Forest	95.30%
Zehir et al. (2023b)	MIT-BIH	SVM	97.00%
Li et al. (2022)	PTB	XGBoost	95.77%
Zhao et al. (2013)	PTB	K-NN	96.00%
Fuster-Barceló et al. (2022)	MIT-BIH	CNN	89.80%
Deshmane and Madhe (2018)	MIT-BIH	CNN	81.33%
Byeon and Kwak (2019)	MIT-BIH	CNN	84.78%
Prakash et al. (2022)	PTB	CNN	96.66%
Proposed method	PTB	INN	97.63%
	MIT-BIH		97.93%

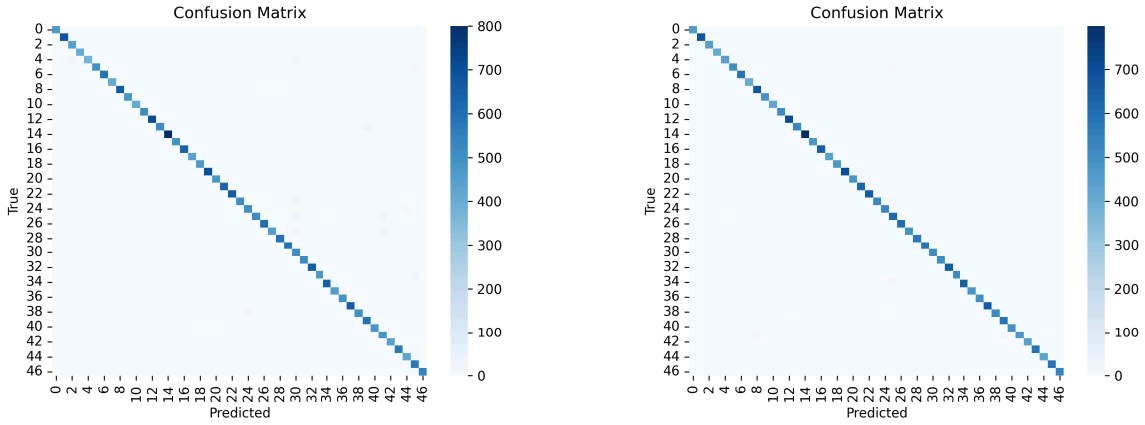
Upon examining the number of parameters in Tables 1 and 2, it becomes evident that the INN model demonstrates greater efficiency. The CNN model has a total of 2,237,634 parameters, whereas the INN model has a significantly lower number, with only 209,210 total parameters. This reduction in parameters indicates that the INN model is less computationally demanding. Consequently, it is faster to train than its CNN counterpart. This is a crucial aspect in deep learning, as models with fewer parameters typically require fewer computational resources, which can lead to significant time savings during the training phase. Moreover, the reduced parameter count in the INN model could also contribute to a decrease in overfitting. In deep learning, overfitting occurs when a model learns the training set too well, to the point where it underperforms on new data. This happens when the model has too many parameters, giving it the capacity to memorise

Figure 10 Accuracy and loss of the INN and CNN models on MIT-BIH: (a) CNN model and (b) INN model (see online version for colours)



(a) (b)

Figure 11 Confusion matrices of the INN and CNN models on MIT-BIH: (a) CNN model and (b) INN model (see online version for colours)



(a) (b)

the training data instead of learning to generalise from it. By having fewer parameters, the INN model reduces the risk of this issue, potentially leading to better performance on unseen data. Figure 11 presents the confusion matrices for the INN and CNN models on the MIT-BIH dataset, with subfigure (a) illustrating the performance of the CNN model and subfigure (b) showcasing the INN model.

The training plots from MIT-BIH and PTB demonstrated by Figures 9 and 10 indicate that the CNN model converges faster than the INN model. suggesting that CNNs may require a lower number of epochs for training for ECG-based biometric systems. However, when comparing the performance metrics in Tables 3 and 4, the INN models demonstrate superior performance over the CNN models across all the metrics indicating that the INN model is more effective at capturing both spatial and temporal information in the spectrogram data of heartbeats. Moreover, INNs are characterised by having fewer parameters than CNNs, which can be advantageous in terms of model complexity and computational resource requirements.

The experimental results indicate that the INN architecture achieves a high degree of performance in identifying individuals, making it a promising choice for ECG biometrics. This performance can be attributed to many factors:

1 Representing ECG signals as spectrograms offers a more comprehensive view compared to traditional time-domain or frequency-domain features. Spectrograms capture both time and frequency information, providing a richer and more detailed representation of ECG signals which allows the INN to extract more meaningful features from the ECG data leading to an improvement in the accuracy of identification.

2 INNs outperform conventional CNNs in ECG identification tasks, suggesting that have a higher ability to recognise patterns.

3 Because of the lower parameters of the INNs in comparison to CNNs, the proposed system does require fewer computational resources.

However, despite these advantages, the proposed system does have some limitations. The primary limitation is that the system was evaluated on relatively small datasets. This raises questions about the system's ability to maintain its high performance when scaled up to larger datasets. Additionally, the datasets used for evaluation included individuals with different heart diseases. This could potentially introduce

bias into the system's performance, as the ECG patterns of individuals with heart diseases could differ significantly from those of healthy individuals.

4 Conclusion

This work proposed a biometric system based on spectrograms extracted by STFT, INN, and ECG signals. Two databases have been used for testing: MIT-BIH and PTB. The results achieved in terms of accuracies were 97.93% and 97.63% on these databases, respectively. Which outperforms traditional CNNs on the same task and also other state-of-the-art methods. These findings suggest that the proposed system, combining spectrograms and INNs, offers a promising approach for person identification with high accuracy and robustness. It has the potential to be applied in various fields and domains, including access control, financial transactions, and healthcare decision-making.

There is still potential for improvement in terms of accuracy and performance for ECG-based biometrics employing deep learning algorithms. Future research will aim to refine the proposed method by conducting further tests on larger and more diverse datasets. This approach will aid in verifying the system's generalisability. A possible improvement could involve the integration of multimodal biometric data. This addition could potentially boost the system's robustness by offering extra layers of identification and verification. In line with this, we plan to incorporate the Phonocardiogram (PCG) as a secondary biometric modality. Moreover, we are contemplating the implementation of the system on field programmable gate arrays (FPGAs). Given their high-speed processing capabilities and flexibility, FPGAs could serve as an efficient platform for deploying our biometric system.

Conflict of interest

The authors have no competing interests to declare.

Generative AI

The authors utilised Bard and ChatGPT to edit and refine the text's phrasing while preparing this work. The authors reviewed and revised the generated text as necessary.

References

AlDuwaile, D.A. and Islam, M.S. (2021) 'Using convolutional neural network and a single heartbeat for ecg biometric recognition', *Entropy*, Vol. 23, No, 6, p.733.

Belgacem, N., Nait-Ali, A., Fournier, R. and Bereksi-Reguig, F. (2013) 'ECG based human identification using random forests', *Proc. Int. Conf. E-Technol. Bus. Web (EBW)*, Jeju Island, Republic of Korea.

Biel, L., Pettersson, O., Philipson, L. and Wide, P. (2001) 'Ecg analysis: a new approach in human identification', *IEEE transactions on instrumentation and measurement*, Vol. 50, No, 3, pp.808–812.

Bousseljot, R., Kreiseler, D. and Schnabel, A. (1995) 'Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet', *Biomedical Engineering/Biomedizinische Technik*, Vol. 40, No. s1, pp.317–318.

Byeon, Y-H. and Kwak, K-C. (2019) 'Pre-configured deep convolutional neural networks with various time-frequency representations for biometrics from ECG signals', *Applied Sciences*, Vol. 9, No, 22, p.4810.

Camara, C., Peris-Lopez, P., Gonzalez-Manzano, L. and Tapiador, J. (2018) "Real-time electrocardiogram streams for continuous authentication', *Applied Soft Computing*, Vol. 68, pp.784–794.

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L. and Lopez, A. (2020) 'A comprehensive survey on support vector machine classification: Applications, challenges and trends', *Neurocomputing*, Vol. 408, pp.189–215.

Chu, Y., Shen, H. and Huang, K. (2019) 'ECG authentication method based on parallel multi-scale one-dimensional residual network with center and margin loss', *IEEE Access*, Vol. 7, pp.51598–51607.

Deshmane, M. and Madhe, S. (2018) 'ECG based biometric human identification using convolutional neural network in smart health applications', *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, IEEE, Pune, India, pp.1–6.

Fatimah, B., Singh, P., Singhal, A. and Pachori, R.B. (2022) 'Biometric identification from ecg signals using fourier decomposition and machine learning', *IEEE Transactions on Instrumentation and Measurement*, Vol. 71, pp.1–9.

Fredj, I.B., Zouhir, Y. and Ouni, K. (2018) 'Fusion features for robust speaker identification', *International Journal of Signal and Imaging Systems Engineering*, Vol. 11, No, 2, pp.65–72.

Fuster-Barceló, C., Peris-Lopez, P. and Camara, C. (2022) 'Elektra: Elektrokardiomatrix application to biometric identification with convolutional neural networks', *Neurocomputing*, Vol. 506, pp.37–49.

Gasmi, A. (2022) *What is Fast Fourier Transform?*, PhD Thesis, Société Francophone de Nutrithérapie et de Nutrigénétique Appliquée.

Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C-K. and Stanley, H.E. (2000) 'Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals', *circulation*, Vol. 101, No, 23, pp.e215–e220.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. and Chen, T. (2018) 'Recent advances in convolutional neural networks', *Pattern recognition*, Vol. 77, pp.354–377.

Guo, T., Zhang, T., Lim, E., Lopez-Benitez, M., Ma, F. and Yu, L. (2022) 'A review of wavelet analysis and its applications: challenges and opportunities', *IEEE Access*, Vol. 10, pp.58869–58903.

Hamza, S. and Ayed, Y.B. (2022) 'Toward improving person identification using the electrocardiogram (ECG) signal based on non-fiducial features', *Multimedia Tools and Applications*, Vol. 81, No, 13, pp.18543–18561.

Hamza, S. and Ben Ayed, Y. (2022) 'Recognition of person using ECG signals based on single heartbeat', *Intelligent Systems Design and Applications: 21st International Conference on Intelligent Systems Design and Applications (ISDA 2021)*, 13–15 December, 2021, Springer, Seattle, WA, USA, pp.452–460.

Imtiaz, M.N. and Khan, N. (2022) 'Pan-tompkins++: A robust approach to detect r-peaks in ECG signals', *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, Las Vegas, NV, USA, pp.2905–2912.

Israel, S.A., Irvine, J.M., Cheng, A., Wiederhold, M.D. and Wiederhold, B.K. (2005) 'ECG to identify individuals', *Pattern recognition*, Vol. 38, No, 1, pp.133–142.

Jain, A., Singh, S.K. and Singh, K.P. (2020) 'Handwritten signature verification using shallow convolutional neural network', *Multimedia Tools and Applications*, Vol. 79, pp.19993–20018.

Jyotishi, D. and Dandapat, S. (2020) 'An LSTM-based model for person identification using ECG signal', *IEEE Sensors Letters*, Vol. 4, No, 8, pp.1–4.

Keras (n.d.) *Keras Documentation: Involutional Neural Networks*, https://keras.io/examples/vision/involution/ (Accessed 28 November, 2023).

Kumar, A., Ranganatham, R., Komaragiri, R. and Kumar, M. (2019) 'Efficient qrs complex detection algorithm based on fast fourier transform', *Biomedical Engineering Letters*, Vol. 9, pp.145–151.

Kyoso, M. and Uchiyama, A. (2001) 'Development of an ECG identification system', *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, Istanbul, Turkey, Vol. 4, pp.3721–3723.

Li, D., Hu, J., Wang, C., Li, X., She, Q., Zhu, L., Zhang, T. and Chen, Q. (2021a) 'Involution: Inverting the inherence of convolution for visual recognition', *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, TN, USA, pp.12321–12330.

Li, Z., Liu, F., Yang, W., Peng, S. and Zhou, J. (2021b) 'A survey of convolutional neural networks: analysis, applications, and prospects', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, No. 12, pp.6999–7019.

Li, M., Si, Y., Yang, W. and Yu, Y. (2022) 'ET-UMAP integration feature for ECG biometrics using stacking', *Biomedical Signal Processing and Control*, Vol. 71, p.103159.

Lynn, H.M., Kim, P. and Pan, S.B. (2021) 'Data independent acquisition based bi-directional deep networks for biometric ECG authentication', *Applied Sciences*, Vol. 11, No, 3, p.1125.

Lynn, H.M., Pan, S.B. and Kim, P. (2019) 'A deep bidirectional GRU network model for biometric electrocardiogram classification based on recurrent neural networks', *IEEE Access*, Vol. 7, pp.145395–145405.

Malik, J., Dahiya, R., Girdhar, D. and Sainarayanan, G. (2016) 'Finger knuckle print authentication using canny edge detection method', *International Journal of Signal and Imaging Systems Engineering*, Vol. 9, No, 6, pp.333–341.

Mateo, C. and Talavera, J.A. (2018) 'Short-time Fourier transform with the window size fixed in the frequency domain', *Digital Signal Processing*, Vol. 77, pp.13–21.

Moody, G.B. and Mark, R.G. (2001) 'The impact of the MIT-BIH arrhythmia database', *IEEE Engineering in Medicine and Biology Magazine*, Vol. 20, No, 3, pp.45–50.

Osadchiy, A., Kamenev, A., Saharov, V. and Chernyi, S. (2021) 'Signal processing algorithm based on discrete wavelet transform', *Designs*, Vol. 5, No, 3, p.41.

Othman, G. and Zeebaree, D.Q. (2020) 'The applications of discrete wavelet transform in image processing: a review', *Journal of Soft Computing and Data Mining*, Vol. 1, No, 2, pp.31–43.

Pan, J. and Tompkins, W.J. (1985) 'A real-time QRS detection algorithm', *IEEE Transactions on Biomedical Engineering*, Vol. BME-32, No. 3, pp.230–236.

Patro, K.K., Jaya Prakash, A., Jayamanmadha Rao, M. and Rajesh Kumar, P. (2022) 'An efficient optimized feature selection with machine learning approach for ECG biometric recognition', *IETE Journal of Research*, Vol. 68, No, 4, pp.2743–2754.

Phinyomark, A., Thongpanja, S., Hu, H., Phukpattaranont, P. and Limsakul, C. (2012) 'The usefulness of mean and median frequencies in electromyography analysis', *Computational intelligence in electromyography analysis-A perspective on current applications and future challenges*, Vol. 23, pp.195–220.

Pinto, J.R., Cardoso, J.S., Lourenço, A. and Carreiras, C. (2017) "Towards a continuous biometric system based on ECG signals acquired on the steering wheel', *Sensors*, Vol. 17, No, 10, p.2228.

Prakash, A.J., Patro, K.K., Hammad, M., Tadeusiewicz, R. and Pławiak, P. (2022) 'Baed: a secured biometric authentication system using ECG signal based on deep learning techniques', *Biocybernetics and Biomedical Engineering*, Vol. 42, No, 4, pp.1081–1093.

Shen, T-W., Tompkins, W. and Hu, Y. (2002) 'One-lead ECG for identity verification', *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society, Engineering in Medicine and Biology*, Vol. 1, IEEE, Houston, TX, USA, pp.62–63.

Sinha, R. (2012) *An Approach for Classifying ECG Arrhythmia based on Features Extracted from EMD and Wavelet Packet Domains*, Master's Thesis, Bangladesh University of Engineering and Technology, Available at: http://lib.buet.ac.bd:8080/xmlui/handle/123456789/114 (Accessed 21 May, 2024).

Stoica, P. and Moses, R.L. (2005) *Spectral Analysis of Signals*, Vol. 452, Pearson Prentice Hall Upper Saddle River, NJ.

Taskiran, M., Kahraman, N. and Erdem, C.E. (2020) 'Face recognition: past, present and future (a review)', *Digital Signal Processing*, Vol. 106, p.102809.

Trahanias, P. (1993) 'An approach to QRS complex detection using mathematical morphology', *IEEE Transactions on Biomedical Engineering*, Vol. 40, No, 2, pp.201–205.

Wold, S. (1976) 'Pattern recognition by means of disjoint principal components models', *Pattern recognition*, Vol. 8, No, 3, pp.127–139.

Yin, X., Zhu, Y. and Hu, J. (2021) 'A survey on 2d and 3d contactless fingerprint biometrics: a taxonomy, review, and future directions', *IEEE Open Journal of the Computer Society*, Vol. 2, pp.370–381.

Zehir, H., Hafs, T., Daas, S. and Nait-Ali, A. (2023a) 'An ECG biometric system based on empirical mode decomposition and Hilbert-Huang transform for improved feature extraction', *2023 5th International Conference on Bio-engineering for Smart Technologies (BioSMART)*, IEEE, Paris, France, pp.1–4.

Zehir, H., Hafs, T., Daas, S. and Nait-Ali, A. (2023b) 'Support vector machine for human identification based on non-fiducial features of the ECG', *Journal of Engineering Studies and Research*, Vol. 29, No, 1, pp.61–69.

Zhang, Q., Zhou, D. and Zeng, X. (2017) 'Heartid: a multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications', *IEEE Access*, Vol. 5, pp.11805–11816.

Zhao, Z., Yang, L., Chen, D. and Luo, Y. (2013) 'A human ECG identification system based on ensemble empirical mode decomposition', *Sensors*, Vol. 13, No, 5, pp.6832–6864.

Zouhir, Y., Fredj, I.B., Ouni, K. and Zarka, M. (2020) 'Robust speaker recognition based on biologically inspired features', *International Journal of Signal and Imaging Systems Engineering*, Vol. 12, Nos. 1–2, pp.19–27.