**DOI: http://doi.org/10.32792/utq.jceps.12.02.14**

# ECG Classification System based on time Domain Features with Least Square Support Vector Machine (LS-SVM)

**Russel R. Majeed & Sarmad K. D. AL Khafaji**

**Received 18/5/2022 Accepted 21/8/2022 Published 01/12/2022**



**Abstract:**

The development of authentication and identification procedures has become a critical requirement to ensure the integrity of device data. Although passwords provide enough control and authentication, they have been shown to have significant speed and security flaws, making biometrics the preferred authentication technique. Consequently, most authentication systems have given much consideration to an electrocardiogram (ECG) signal. However, ECG signals are distinct, making them difficult to imitate yet relatively frequent. We offer a novel ECG validation model that combines multi-domain characteristics with an LSSVM in this work. ECG signals examine two characteristics to determine each individual's ideal mix of characteristics. First, the best 3-band filter bank is used to extract time-domain properties from ECG data. The retrieved characteristics are examined for the most relevant ones, and the unimportant ones are removed. The Least Square Support Vector Machine(LS-SVM) classifier collects specified characteristics data. Our ECG authentication technology outperformed the competition, according to the findings. The suggested model was applied to categorize EKG signals, and it achieved an average accuracy of 92 %and 91 %when time features were used. The suggested model is evaluated based on its ability to interact with open data.

**Keywords:** ECG, Authentication, Time domains features, LS-SVM.

## Introduction:

Biometric identification technologies provide a highly secure means of identifying and confirming individuals based on their characteristics. [1], [2]. In most biometric systems, physiological and behavioral variables such as fingerprints, face, and voice are used [3], [4]. And although these items offer a high degree of security, various investigations have shown that they are vulnerable to fraud through attacks. Some biometric system attacks, for example, use latex to recreate features such as fingerprints and voices and use them as original features [2], [5]. EEG and (ECG), have recently become popular [6], [7]Most modern biometric systems support the use of ECG signals. To obtain information on myocardial activity from humans, the electrocardiogram is a practical, easy-to-understand, and low-cost procedure [5], [8] Clinical study suggests that each person's ECG has specific properties that may help protect them against falsifications. ECG features, on the other hand, have been difficult to duplicate and reproduce. [9], [10] Due to the time-varying nature of ECG signals, the irregularity of cardiac disorders, and the length of time required to collect ECG data from individuals, techniques-based ECG signals have a number of disadvantages. Therefore, the performance of a biometric identification model is the most important factor to consider while assessing it. Due to the inconsistent and recurrent ECG readings. The most problematic component of developing a functional ECG biometric system is the feature extraction technique, or how to choose the most important characteristics to characterize ECG signals. This research examines time and frequency factors to identify the optimal feature set for showing ECG signals. T wave, P wave, and QRS complex are the three basic ECG wave components. The landmark characteristics are the amplitudes of the P, R, and T waves. It shows the time difference between wave starting borders, offsets for all P, Q, R, S, and T waves, and regression data. [5]Several methods have been developed in recent years to use ECG signals and other biometric features such as face and fingerprints as a tool for human identification. For example, [11] Using a multimodal biometric property system, such as electrocardiograms and fingerprints, a group of researchers led by Arteaga-Falconi has suggested a way of preserving relationships between creatures and humans. The collected ECG and fingerprint were inspected and evaluated. According to many research and articles, ECG biometric technologies are frequently used in government commercial systems, as well as travel documents, health monitoring systems, and distributed systems such as smart cards. However, despite the advantages and dependability of ECG biometric systems, technology-based ECG signals have several limitations due to the fluctuating nature of ECG signals, the fickle nature of cardiac illnesses, and the time it takes to gather ECG data from individuals. Arteaga-Falconi, et al [12] suggested a concept that uses a multimodal biometric characteristics system with ECG and fingerprints to protect interactions between objects and humans. The ECG and Fingerprint characteristics that were retrieved were evaluated and explored. The MINDTCT algorithm was utilized as the minutiae extractor and the BOZORTH3 method as the minutiae matcher for fingerprints. Rabinezhadsadatmahaleh et al. [5] ECG heartbeats were examined to create an authentication model. Deep learning methods were used with SVM to create an ensemble classification model. They demonstrated the benefits of combining traditional machine learning models with deep neural networks. The average accuracy was 99.02, the FAR was 0.95, and the FRR was 1.02. Wang et al. [13]. suggested a unique ECG biometrics technique [13]. ECG signals were filtered to improve signal quality using a band pass Butterworth filter with cut-off frequencies ranging from 1 to 40 Hz. A multi-scale differential feature was used with one-dimensional multi-resolution local binary patterns to extract features from ECGs. In that study, four datasets were used to assess the proposed model. The detail matcher and extractor were built using the BOZORTH3 approach.

## 3. ECG Data:

ECG dataset was collected from the physio net organisation. The data is publicly available at https://physionet.org/content/mitdb/1.0.0/. The database was collected from 47 healthy. in this study, all ECG recordings were used in the performance evaluation. More detail in [18,17]. we considered a single ECG lead for collecting ECG signals for realistic scenarios. The sampling rate was 1000 Hz. A total of 54 healthy persons were involved in the recording ECG signals.

## 4. Materials and methods:

This research presents a novel ECG-based paradigm for person identification. The ECG signals were first utilizing a low pass filter (LPF) having a 128 Hz cut-off frequency. Min, Max, First Quartile (Q1), Second Quartile (Q2), Third Quartile (Q3), Quartile Range (IQR), mean, mode, median, range, variance (2), and standard deviation (IQR) were retrieved from the temporal domain to find and categorize the most robust feature set. For human identification, an ECG and an LS-SVM classifier were utilized. The suggested framework for human recognition is shown in Figure 1.

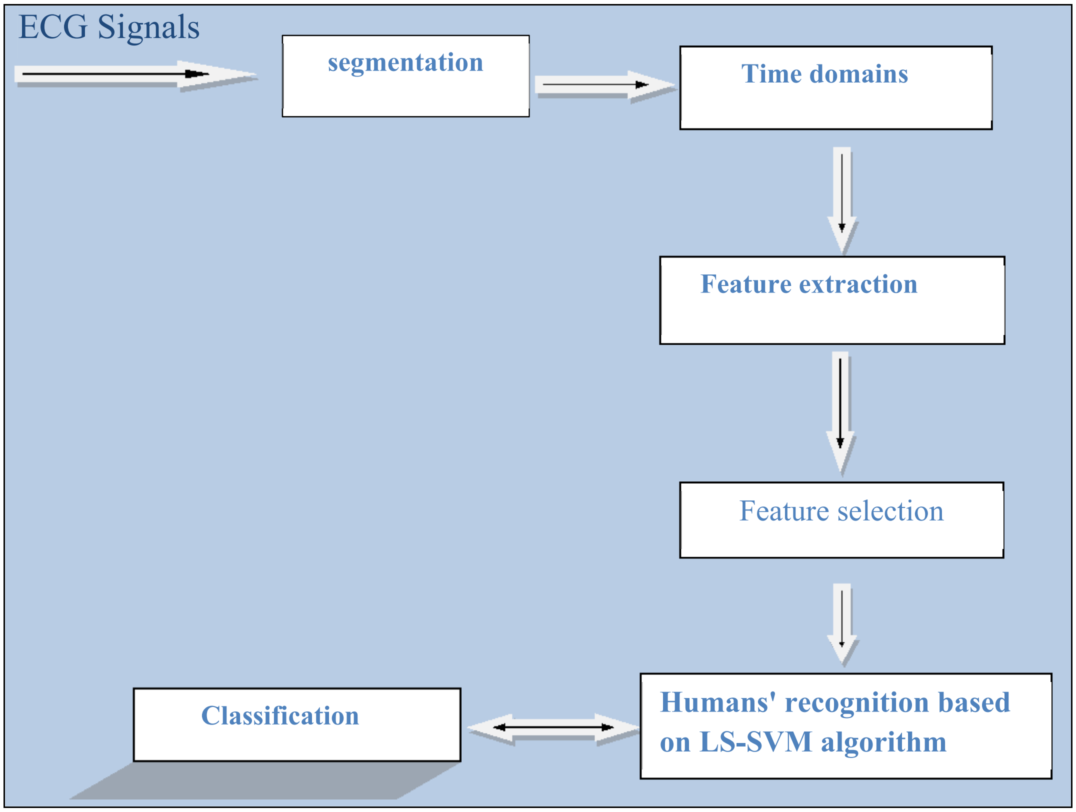


**Figure 1**: proposed system based on time domain

## 3. ECG PREPROCESSING:

### 3.1 ECG segment based on (pan Tompkins algorithms):

In this paper, ECG signals were separated into m intervals. As a reference, we employed Fiducial points to segment the ECG signals. The two fiducial points LP and TP are used to pinpoint the beginning and end of a heartbeat. To distinguish between LP and TP, the R peak must be recognized. Figure 2 of an ECG segment shows the LP and TP positions. The differentiation detection technique from [4] .was applied in this work. This method is useful because it is quick, requires no threshold, and has a small average time error. Consider the m data points in the ECG signal X=x 1,x 2,x 3,....,x m. The signal X was then segmented into n windows using LP and TP, with each n comprising k data points and stored in a vector for the next phase, X=n 1,n 2,n 3,....,n m. As a result, the LP and TP fiducial sites partition each heartbeat across the ECG signal. These heartbeats are recorded as a vector, which will be utilized in the next stage to create features.

## 4. Time-domain in feature extraction:

As seen in Figure 3, the database sends the first ECG signal (time-domain). From this point, the mean, standard deviation, and variance are calculated straight from the ECG signal. The signal's temporal domain describes how it evolves over time. Analyses mathematical functions, physiological actions, or time-series data sets. It examines mathematical functions, biological activities, or time-related data sets in a time series. In the time domain, the value of the signal or function is known as the absolute number of continuous-time or at different points in discrete time. For example, in an ECG signal, the time domain indicates how the amplitude of the signal changes over time [18]. Previous research has shown that when various abnormalities occur during ECG monitoring, they change the signal's shape and make it challenging to find wave edges This is why time-domain processing of an ECG signal is less beneficial than frequency-domain processing. [18].

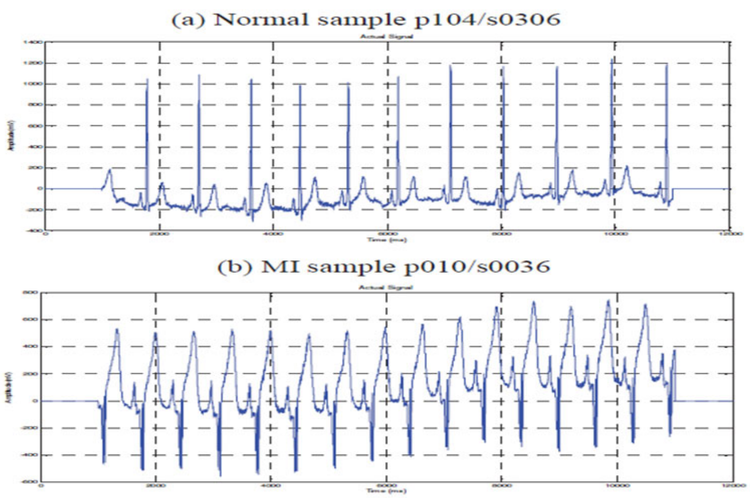


Figure 2: Different kinds of ECG signals in time domain

## 5. Feature extraction:

The goal of feature extraction is to convert the raw ECG signal into a feature vector. The ST shape, QRS width, QT intervals, R height, T height, and other features are calculated utilizing attribute vectors such as the QRS complex, S-wave, and J.[14].These things are called morphological features, viewed from the perspective of time-domain shapes. Some other methods have also been considered, such as wavelet transformations [15], [16].that change into a frequency domain. The first step for feature extraction techniques is to look at these two places. Statistical features can be pulled from frequency domain analysis (also called spectroscopy) using the fast Fourier transform. Much study has been conducted on extracting characteristics from the ECG signal, and it has become evident that frequency domain analysis outperforms time-domain analysis. [17].In the research, the discrete wavelet shunt is another method that has been used in the ECG signal.

## 6. Classification

In order to classify the classes of ECG signals, standard classifiers (Last Square Support Vector Machine (LS-SVM) were used). This classifier is among the most widely used classifiers for classifying EKG signals. Classification is of two types (classification by channels and range) for the algorithm LS-SVM used by the two systems and for the characteristics used, Wavelet features, the LS-SVM algorithm used a classification system based on channel selection, in addition, the statistical features were input into the LS-SVM algorithm.

## 7. Performance Evaluation Measurements

When developing a specialized system, it is vital to analyze its performance using unique measurements in order to identify the system's efficacy. A confusion matrix was utilized in this study to determine some of the performance assessment metrics, which include (Accuracy, Sensitivity, Specificity, and false-positive rate (FPR)). In machine learning, a confusion matrix is also known as an error matrix [19] is a two-dimensional (actual*predicted) matrix that helps you to examine the performance of a classifier; each row of the matrix represents examples in a predicted class, while each column represents occurrences in an actual class [20]. Correct classifications are designated as 'true positives (TP)' or 'true negatives (TN)', while incorrect classifications are labeled as 'false positives (FP)' or 'false negatives (FN)'. Accuracy, sensitivity, specificity, and other metrics may be computed as well. This is particularly important in single-test diagnoses for the presence or absence of a specific disease or sickness [21]. Figures 2 show the confusion matrix for two and several classes, respectively.

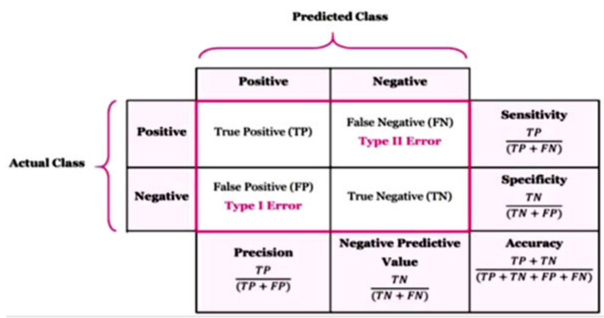


**Figure 2:** Confusion Matrix [**1**].

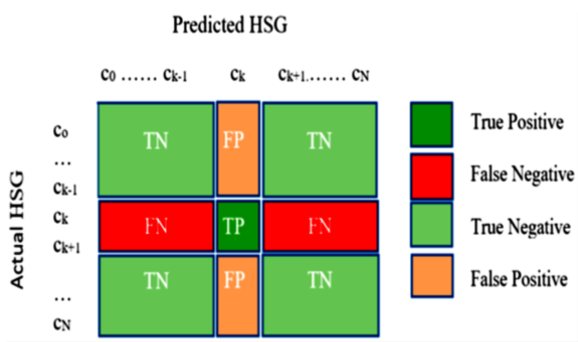


**Figure 3.** Confusion matrix, shown with totals for positive and negative tuples **[2]**

To calculate measures that used in this work, the following equations are used [22].

- Accuracy $\frac{TP+TN}{TP+FN+TN+FP}$ (1)
- Sensitivity, recall, hit rate, or true positive rate (TPR)

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (2)$$

- Specificity, selectivity or true negative rate (TNR)

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (3)$$

- False Positive Rate (FPR) or fall-out

$$\text{FPR} = \frac{FP}{FP + TN}$$

## 8. Results:

This section evaluates the suggested model's performance on various scales. The retrieved properties were also examined to determine which temporal and frequency feature sets were most advantageous. The study aims to create a set of feature extracts that may be utilized to verify ECG biometrics and evaluate classification algorithm performance. The ECG template was used to generate this template (cut off). We gathered ten outcomes from each patient's ECG recording to create a validation form. The database determines the number of EKG beats during the verification procedure. The accuracy, sensitivity, and specificity of the proposed model were evaluated. Several feature extraction and classification techniques were tested. We looked into various ECG properties using time domains and different classification methods. Several metrics are utilized to evaluate the effectiveness of the proposed model. The retrieved features were analysed to see whether feature sets obtained from temporal domains were the most effective. The research aims to find the best location of feature extracts to verify ECG biometrics and put classification algorithms to the test. The template was created using ECG (segment) beats. We used ten results from each subject's ECG recording to create a validation model. The database counts the number of ECG beats required during the verification step. The proposed model's accuracy, sensitivity, and specificity were used to assess its success. We performed experiments using various feature extraction and classification algorithms. In addition, we specifically investigated several components of the ECG using time domains with the classification algorithm and achieved the best results The ECG fragment was supplied using signal segmentation in this experiment. Then, the retrieved features were fed into the LS-SVM, and Table 1 illustrates the discovered components based on the time domain to get the best outcomes.

| Classification algorithm | Accuracy | Sensitivity | Specificity |
|---|---|---|---|
| LS-SVM | 92% | 91% | 100% |

## 6. Conclusions:

Despite massive efforts to build an electrocardiogram-based biometric system, several challenges remain unanswered, including the relationship between selection criteria and detection rate, as well as the impact of feature ECG biometric system type and feature extraction. As a consequence, several experiments were conducted to assess the suggested model's performance with various features. This study used temporal characteristics to extract ECG data from random ECG recordings, and validation system tests were conducted. To improve the robustness of the suggested model, further testing with a vast amount of data is required, according to the findings. New classes in ECG-based biometric systems should be evaluated as well.

## 7. REFRENCES:

[1] W. Wu, S. Pirbhulal, A. K. Sangaiah, S. C. Mukhopadhyay, and G. Li, "Optimization of signal quality over comfortability of textile electrodes for ECG monitoring in fog computing based medical applications," *Futur. Gener. Comput. Syst.*, vol. 86, pp. 515–526, 2018.

[2] M. Hammad and K. Wang, "Parallel score fusion of ECG and fingerprint for human authentication based on convolution neural network," *Comput. Secur.*, vol. 81, pp. 107–122, 2019.

[3] T. Bai *et al.*, "A lightweight method of data encryption in BANs using electrocardiogram signal," *Futur. Gener. Comput. Syst.*, vol. 92, pp. 800–811, 2019.

[4] P. Peris-Lopez, L. González-Manzano, C. Camara, and J. M. de Fuentes, "Effect of attacker characterization in ECG-based continuous authentication mechanisms for Internet of Things," *Futur. Gener. Comput. Syst.*, vol. 81, pp. 67–77, 2018.

[5] N. Rabinezhadsadatmahaleh and T. Khatibi, "A novel noise-robust stacked ensemble of deep and conventional machine learning classifiers (NRSE-DCML) for human biometric identification from electrocardiogram signals," *Informatics Med. Unlocked*, vol. 21, p. 100469, 2020.

[6] S. Hamza and Y. Ben Ayed, "Svm for human identification using the ecg signal," *Procedia Comput. Sci.*, vol. 176, pp. 430–439, 2020.

[7] S. Hong, Y. Zhou, J. Shang, C. Xiao, and J. Sun, "Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review," *Comput. Biol. Med.*, vol. 122, p. 103801, 2020.

[8] W. Klonowski, "Fractal analysis of electroencephalographic time series (EEG Signals)," in *The fractal geometry of the brain*, Springer, 2016, pp. 413–429.

[9] H. Namazi and S. Jafari, "Age-based variations of fractal structure of EEG signal in patients with epilepsy," *Fractals*, vol. 26, no. 04, p. 1850051, 2018.

[10] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, 1970.

[11] M. Hammad, S. Zhang, and K. Wang, "A novel two-dimensional ECG feature extraction and classification algorithm based on convolution neural network for human authentication," *Futur. Gener. Comput. Syst.*, vol. 101, pp. 180–196, 2019.

[12] J. S. Arteaga Falconi and A. El Saddik, "Security with ECG Biometrics," in *Advanced Methods for Human Biometrics*, Springer, 2021, pp. 65–79.

[13] J. S. Arteaga Falconi, "Towards an Accurate ECG Biometric Authentication System with Low Acquisition Time." Université d'Ottawa/University of Ottawa, 2020.

[14] J. S. Arteaga-Falconi, H. Al Osman, and A. El Saddik, "ECG and fingerprint bimodal authentication," *Sustain. Cities Soc.*, vol. 40, no. August 2017, pp. 274–283, 2018, doi: 10.1016/j.scs.2017.12.023.

[15] D. Jha *et al.*, “Resunet++: An advanced architecture for medical image segmentation,” in *2019 IEEE International Symposium on Multimedia (ISM)*, 2019, pp. 225–2255.

[16] S. G. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, 1989.

[17] D. Stoller, S. Ewert, and S. Dixon, “Wave-u-net: A multi-scale neural network for end-to-end audio source separation,” *arXiv Prepr. arXiv1806.03185*, 2018.

[18] T. Nakamura and H. Saruwatari, “Time-Domain Audio Source Separation Based on Wave-U-Net Combined with Discrete Wavelet Transform,” *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2020-May, pp. 386–390, 2020, doi: 10.1109/ICASSP40776.2020.9053934.

[19] S. V Stehman, “Selecting and interpreting measures of thematic classification accuracy,” *Remote Sens. Environ.*, vol. 62, no. 1, pp. 77–89, 1997.

[20] D. M. W. Powers, “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation,” *arXiv Prepr. arXiv2010.16061*, 2020.

[21] C. Sammut and G. I. Webb, *Encyclopedia of machine learning*. Springer Science & Business Media, 2011.

[22] H. Brooks *et al.*, “WWRP/WGNE Joint Working Group on Forecast Verification Research,” *Collab. Aust. Weather Clim. Res. World Meteorol. Organ.*, 2015.