# Utilizing Deep Neural Nets for an Embedded ECG-based Biometric Authentication System

Adam Page, Amey Kulkarni, and Tinoosh Mohsenin
Energy Efficient High Performance Computing Lab (EEHPC)
University of Maryland, Baltimore County (UMBC)

***Abstract*—This work presents a low-power, embedded ECG pattern recognition system for the purpose of biometric authentication. We believe that ECG coupled with a secondary biometric marker such as fingerprint will play a key role in wearable security as wearables' popularity continues to grow. The key objective of this work is to implement a system that is reliable, robust, and fast while maintaining a low area and power footprint. A streamlined approach was devised that utilized neural networks to both identify QRS complex segments of the ECG signal and then perform user authentication on these segments. When tested on 90 individuals, the system is able to achieve 99.54% accuracy for QRS complex identification, and, on average, 99.85% sensitivity, 99.96% specificity, and 0.0582% EER for user identification. When implemented on an Artix-7 FPGA, the entire design occupies 1,712 slices (5%) and 978.7 KB of memory and dissipates 31.75 mW of total chip dynamic power when running at 12.5 MHz.**

## I. Introduction

The massive influx in smart, wearable devices has sparked a new chapter in how we perceive and interact with the world. These smart devices coupled with the internet of things (IOT) infrastructure enables countless new capabilities, i.e., the ability to monitor your heart-rate, start your car, and set a reminder all from your watch in real-time is astounding. As ubiquitous computing grows, it is critical to further security measures that pair well with these personal devices. Biometric authentication is a clear choice and much success has already been made in the form of fingerprint, face, voice, and iris recognition [1]. Each technique has inherent trade-offs, from the ability to be compromised, ease of use, obtrusiveness, and implementation complexity. For instance, face and iris require the user's direct attention in addition to needing to perform complex video processing. Furthermore, as has become standard, two-factor authentication (2FA) may also be required. This work explores the use of ECG as another form of biometric authentication. Initial work by [2]–[9], have already demonstrated the ability for ECG to be used for user authentication. We envision a system such as a watch that can confirm user-identification by both fingerprint and ECG biomarkers. The remainder of the paper is divided into three main sections. The first briefly discusses related work on ECG pattern recognition. The second describes and evaluates the proposed system. Finally, the last section describes and evaluates the hardware implementation of the system.

## II. Related Work

There has been a plethora of research done on using ECG data for a variety of classification tasks. Many of these classification tasks, however, are for the purpose of identifying diseases and conditions, such as heart arrhythmia. Some work has been done successfully on using ECG data for the purpose of user authentication. This is primarily done using either fiducial or non-fiducial feature extraction followed by simple statistical analysis and/or machine learning algorithms. These options often present trade-offs between accuracy and computational complexity. Table I presents the results from related works that perform user authentication. [2] is one of the earliest works that used 12-lead ECG features to identify 20 users. Using proprietary equipment to extract 30 different fiducial features per lead, they obtained 100% accuracy. [9] uses the same ECG-ID database as used in this work. They employed a large variety of preprocessing techniques in addition to machine learning for classification.

| Reference | Method | Subjects | Accuracy |
|---|---|---|---|
| Beil [2] | PCA | 20 | 100% |
| Shen [3] | Template+DBNN | 20 | 100% |
| Israel [4] | LDA | 29 | 98% |
| Wang [5] | KNN+LDA | 13 | 96% |
| Chiu [6] | Wavelet+LDA | 35 | 100% |
| Chan [7] | Wavelet DM | 50 | 89% |
| Sriram [8] | Wavelet+LDA | 17 | 88% |
| Lugovaya [9] | Wavelet+LDA | 90 | 96% |
| This Work | DNN | 90 | 99.96% |

TABLE I: Results from related work for user authentication. Direct comparison is challenging due to different datasets, ECG setup, approaches, and design goals.

For a more in-depth comparison of the most recent ECG authentication systems, we refer the reader to [10]. [10] evaluates 20 systems on an in-house database consisting of 265 volunteers recorded while resting during 3 sessions with 1-week to 6-month span between sessions. For multi-day training experiments, two systems performed far superior. The best system, developed by [10] on the same database, achieved an equal error rate (EER) of 5.47%, the second best system by [11] achieved an EER of 6.38%, while the average EER of the rest was 24.04%. There are two observations we would like to highlight. First, the version of [10] that achieved the best EER requires dataset-specific feature reduction. Their similar non-feature reduction version is only able to achieve an EER of 14.73%. Second, the original version of [11] utilized wavelet transform followed by a 2 hidden layer neural network. However, in [10] the wavelet transform was substituted with a 1-40 Hz band-pass filter. In the proposed system, we achieve an EER of 0.0582% on the public ECG-ID database consisting of 90 volunteers recorded while resting during 2-20 sessions over a 1-day to 6-month time-span.

## III. System Overview

The high-level system diagram is presented in Fig. 1. 1-lead ECG data is fed into the system and sampled at a rate of

978-1-4799-7234-0/15/$31.00 ©2015 IEEE

500 Hz. This data is buffered and passed to the main processing block in 250-sample segments. The main processing block consists of three major stages: 1. *Preprocessing*, 2. *QRS Identification*, and 3. *User Identification*.

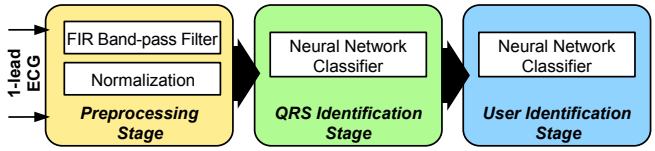



Fig. 1: High-level diagram of the ECG authentication system. The system consists of three main processing blocks: *Preprocessing, QRS Identification, and User Identification*.

*1) Preprocessing Stage:* The first stage consists of filtering and normalizing the input segment. The filter used is an FIR band-pass filter with a pass-band of 1-40 Hz and 60 dB attenuation in the lower and upper stop-bands. The filtering is done to help remove artifacts and noise such as power-line interference. Normalization is done by subtracting the mean and dividing by the standard deviation, in order to have segments with zero mean and unit variance.

*2) QRS Identification Stage:* In this stage, a neural network is used to identify segments that contain a QRS complex. Specifically, the segment must contain an R peak aligned approximately at the 80th sample. The reason for only performing user identification on QRS segments is because they provide the most relevant and unique information for a user. The alignment of the R peak around the 80th sample is also done to ensure the majority of the complex is captured in the segment to help improve user classification.

*3) User Identification Stage:* The last stage also uses a neural network to perform user authentication. Neural networks have shown great promise in biomedical signal classification and detection tasks [12]. Additionally, unlike other related biomedical signal processing works that require several pre-processing techniques and the use of feature engineering [13]–[15], the neural networks in this work are supplied only filtered raw ECG data for classification. The idea is to allow each layer in the neural network to abstract high-order features that perform better than hand-crafted features.

## IV. System Evaluation

The efficacy of the system is measured by its ability to securely identify a user with minimal effort and obtrusiveness. The most crucial goal is that the system is infallible at producing false matches. A longer processing time (allowing more false rejections) is acceptable if it enables the system to be more robust against false matches. Therefore, we want a system with a near perfect specificity while still maintaining an acceptable sensitivity. The second and third stages use neural networks to perform binary classification such that the output is the probability of a match. Typically, a threshold of 50% is used to obtain equal performance for both classes. By increasing the threshold, we can reduce the possibility of false matches at the expense of false rejections.

### A. Dataset

Data from the ECG-ID database is used in this study for both training and validation purposes [9]. The database consists of 310 1-lead ECG recording sessions obtained from 90 volunteers during a resting state. Each session produced 20 seconds of recorded data sampled at 500 Hz with 12-bit resolution. The number of sessions for each volunteer varied from 2 to 20 with a time span of 1-day to 6-months between the initial and last recordings. The specific time lapse between each session is not provided. There were 2 subjects with over 20 sessions, 20 subjects with at least 5 sessions, 41 subjects with at least 3 sessions, and 49 subjects with only 2 sessions.

### B. QRS Identification

*Setup:* To train and validate the neural network for generalized QRS complex identification all subjects were utilized. The users were divided into training, validation, and testing sets using a 70/15/15% split. This method of splitting by user should better approximate real world results by testing on individuals that were not part of the training process. For each ECG record, 250-sample segments were selected in two steps. In the first step, positive-labeled segments were strategically drawn that contained an R peak within 1-sample of the 80th index, which had been supplied with the database. In the second step, negative segments were drawn at random using an inverse-distance weighting based on R peak locations. This led to having more negative segments that coincided closely to QRS complexes. A 1:4 positive to negative instance ratio was also employed in order to reduce the large skew towards the negative class.

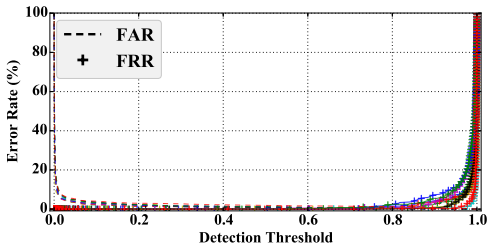



Fig. 2: DET curves for the 10 best neural network models used for QRS identification. Dashed and '+' lines correspond to false acceptance (FAR) and false rejection rate (FRR), respectively. The best model achieved a low EER of 0.4487% at threshold 71.8%.

*Results:* A hyper-parameter optimization was used to help discover the best neural network model. 1,000 experiments were performed in which hyper-parameters were selected at random. These hyper-parameters included the number of hidden layers (1-4), the number of nodes per layer (250-1,000), the dropout rate per layer (40-80%), and the activation function (sigmoid, tanh, relu). The best network model, based on test results, required a single hidden layer with 307 nodes, 60% dropout, and a *tanh* activation function. From analyzing the results, it was found that deeper networks suffered from overfitting even when higher dropout rates were utilized. The best model obtained an overall accuracy of 99.54% with 99.49% sensitivity and 99.55% specificity. Figure 2 depicts the detection error trade-off (DET) with linear scales for the 10 best neural network models. The equal error rate (EER) of the best model is 0.4487% for a threshold of 71.8%. Figure 3 provides the histograms of the best model's predictions for non-QRS and QRS segments of the test set. This plot demonstrates the network's ability to separate the classes effectively. Using a threshold over 70% helps to remove the few occurrences of false matches.

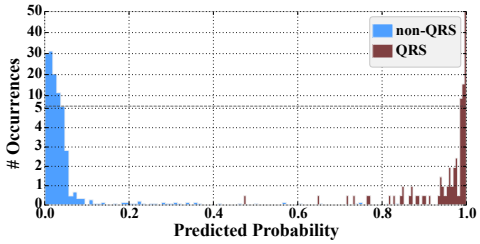



Fig. 3: Histograms of the probabilities produced by the best model for non-QRS (blue) and QRS (red) test segments.

### C. User Identification

*Setup:* The process of training and validating the models for user authentication was done using cross validation with a one-hot subject encoding technique. This means that for each cross validation iteration only one subject's QRS complex segments are given positive labels while all other subjects' QRS segments are given negative labels. This differs from other works that assign a unique label for each individual [9]. Assigning a unique label often leads to more complex classification models as it must learn to discern all subjects. Furthermore, the approach is more difficult to scale to larger databases containing more individuals. For the target subject, their recorded sessions were divided into training, validation, and test sets using a 70/15/15% split. For target subjects with only 2 records, one record was split between training and validation. For the non-target individuals, they were split into training, validation, and test sets using a 70/15/15% split. This means that the test and validation contain individuals not seen during training.

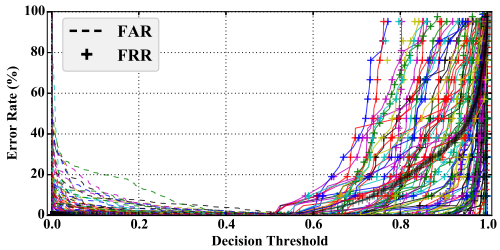



Fig. 4: DET curves of the best neural net model for user authentication for all 90 subjects. Dashed and '+' lines correspond to false acceptance (FAR) and false rejection rate (FRR), respectively. Utilizing all results, the model achieved a low EER of 0.0582% at 50.31% threshold.

*Results:* Similar to QRS identification, 1,000 experiments were performed using a hyper-parameter optimization. The same parameters and ranges were used in these experiments. For each experiment, 90 iterations were performed for cross validation across all subjects. The best neural network model, based on the test sets, required 2 hidden layers with 577 and 380 nodes, 76% and 78% dropout rates, and a *tanh* activation function. The model obtained an average accuracy of 99.93% with 99.85% sensitivity and 99.96% specificity. The DET curves for each subject using the best model are provided in Fig. 4. The model achieves an EER of 0.0582% at a threshold of 50.31%. Figure 5 demonstrates the model's effectiveness at performing user identification using only raw ECG data. Figure 6 provides box-plots of the ROCs area under the curve (AUC), accuracy, sensitivity, and specificity for the test sets. All metrics perform well with little variance. Specificity, the most important metric, has both the highest average and the lowest variance.

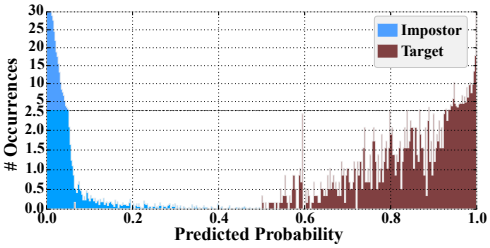



Fig. 5: Histograms of the probabilities produced by the best model for all subjects combined. The target (red) refers to the target subject, while impostor (blue) refers to all others.

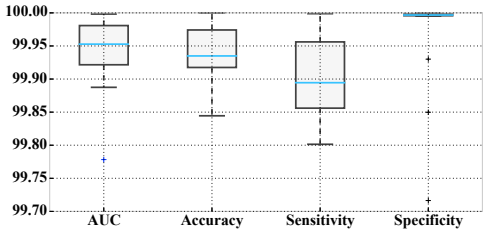
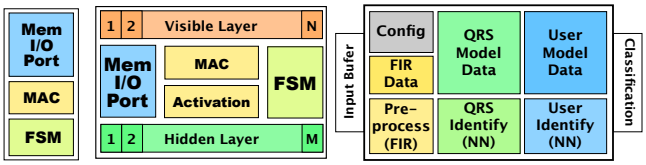



Fig. 6: AUC, accuracy, sensitivity, and specificity metrics for the best user identification model using all users' test sets.

## V. Implementation Overview

After demonstrating the system's ability to perform authentication, the next step is to implement the design in hardware. The main objective of the hardware implementation is to minimize the area and power footprint while meeting processing deadline and maintaining system accuracy. The two main hardware blocks required consist of the FIR filter (FIR) and the neural network (NN) classifier. These blocks along with the entire system block were implemented using Xilinx's Vivado HLS. The software not only allowed fast implementation but also enabled fixed-point precision experimentation, resource utilization management, and system verification of the RTL using C/C++ test-benches. The high-level system was evaluated using 32-bit floating-point precision. Due to the large amount of resources required for floating-point operations, the system was translated into fixed-point. 6.12 fixed-point format was selected as it maintained the same accuracy and fit nicely into 18-bit wide block RAMs (BRAM).




(a) FIR (b) NN (c) SYSTEM

Fig. 7: Block-diagrams of the filter (FIR), neural network classifier (NN), and system block.

*1) FIR Block:* The FIR filter is implemented in a serial manner using a single MAC operator as depicted in Fig. 7a. The filter coefficients are stored externally with I/O ports added to allow the block to load the coefficients as needed. The band-pass filter required for this design consisted of 500 symmetric coefficients. The input, output, and coefficients were all stored in 6.12 format with saturation and truncation employed.

*2) NN Block:* The neural network classifier (NN) block is implemented in a partially-parallel configuration. At RTL-

synthesis time, the number of parallel MAC operations can be specified. The network topology, weights, and biases are stored externally to enable run-time configuration.

*3) SYSTEM Block:* The high-level hardware design is depicted in Fig. 7c. The *Preprocessing Stage* consists of the FIR block and additional logic to perform normalization. This stage receives and produces 256-sample segments serially. The *QRS Identify* and *User Identify* stages are implemented using the NN blocks. Both stages are fed the same input from *Preprocessing* and produce a single classification output. *User Identify Stage* is only enabled when the *QRS Identify Stage* detects a QRS complex. For each of the three stages, there is a corresponding memory to store required configurations.

## VI. Implementation Results

The design is synthesized and implemented on an Artix-7 FPGA (xc7a200tlffq1156-2L). The resource utilization of the design is summarized in Table II. The full design occupies 1,712 slices of the total 33,650 available (5%). The total memory required by the system is just under 1 MB. When operating at 50 MHz, the FPGA consumes 256 mW of power. The authentication latency is dictated chiefly by the user's heart rate and false rejections. For instance, a user with a heart rate of 40 BPM could require 3 seconds if the first beat is misclassified. The minimum and maximum computational latency to process a segment takes 17.1 and 97.3 ms, respectively. The minimum consists of just the first two stages. Dynamic overlap of the segments is employed using a simple algorithm based on the probability output of the *QRS Identification Stage*. The higher the probability, the more overlap is used to ensure a peak is not missed. Conservatively, minimum and maximum processing must occur for every 50 and 200 samples received. This means that the minimum and maximum processing must occur within 100 ms and 400 ms deadlines given a sampling frequency of 500 Hz. The total chip dynamic power consumption is 31.75 mW when scaling the frequency to 12.5 MHz to meet the worst-case 400 ms deadline. A major contributor to the power is the use of the majority of the BRAMs available to store the network weights. Some possible ways to reduce this power include pruning the network, storing weights off-chip and locally caching, or translating into an ASIC [13]. Regardless, we envision that such a system will only need to be active for brief moments periodically throughout a given day. Shutting off the system while inactive will significantly reduce the effective power consumption.

| | | | | | |
|---|---|---|---|---|---|
| **Area** | Slice | 1,712 (5%) | **Clock** | Freq (MHz) | 50 |
| | FF | 4,335 (2%) | | Min Latency (ms) | 17.1 |
| | LUT | 3,734 (3%) | | Max Latency (ms) | 97.3 |
| **Memory** | FIR (KB) | 1.1 | **Power** | Static (mW) | 129 |
| | QRS NN (KB) | 172 | | Dynamic (mW) | 127 |
| | PT NN (KB) | 806 | | Total (mW) | 256 |

TABLE II: Resource utilization breakdown of the hardware implementation on an Artix-7 FPGA. The minimum latency consists of *Preprocessing* and *QRS Identification stages* while maximum also includes *User Identification Stage*

## VII. Conclusion

This paper presented a low-power, embedded system for the purpose of performing biometric authentication using ECG data. To streamline the design, the system used neural networks for both detecting QRS complex segments and performing user identification on these raw QRS segments. The only preprocessing performed is filtering and normalization. The QRS complex identification achieved 99.54% accuracy using a neural network model consisting of a single 307-node hidden layer with 60% dropout and a *tanh* activation function. The best model for user identification achieved, on average, 99.93% accuracy, 99.85% sensitivity, 99.96% specificity, and an EER of 0.0582%. The model required 2 hidden layers with sizes 577 and 380, dropout rates 76% and 78% and a *tanh* activation function. When implemented on an Artix-7 FPGA, the entire design occupied 1,712 slices (5%), ~1 MB of memory, and dissipates 31.75 mW of total chip dynamic power when operating at 12.5 MHz.

## References


[1] V. Bharadi, B. Pandya, and B. Nemade, "Multimodal biometric recognition using iris amp; fingerprint: By texture feature extraction using hybrid wavelets," in *Confluence The Next Generation Information Technology Summit, 2014 5th*, Sept 2014, pp. 697–702.

[2] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: a new approach in human identification," *Instrumentation and Measurement, IEEE Transactions on*, vol. 50, no. 3, pp. 808–812, 2001.

[3] T.-W. Shen, W. Tompkins, and Y. Hu, "One-lead ECG for identity verification," in *Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint*, vol. 1. IEEE, 2002, pp. 62–63.

[4] S. A. Israel *et al.*, "ECG to identify individuals," *Pattern recognition*, vol. 38, no. 1, pp. 133–142, 2005.

[5] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP journal on Advances in Signal Processing*, vol. 2008, p. 19, 2008.

[6] C.-C. Chiu, C.-M. Chuang, and C.-Y. Hsu, "A novel personal identity verification approach using a discrete wavelet transform of the ecg signal," in *Multimedia and Ubiquitous Engineering, 2008. MUE 2008. International Conference on*. IEEE, 2008, pp. 201–206.

[7] A. D. Chan, M. M. Hamdy *et al.*, "Wavelet distance measure for person identification using electrocardiograms," *Instrumentation and Measurement, IEEE Transactions on*, vol. 57, no. 2, pp. 248–253, 2008.

[8] J. C. Sriram *et al.*, "Activity-aware ECG-based patient authentication for remote health monitoring," in *Proceedings of the 2009 international conference on Multimodal interfaces*. ACM, 2009, pp. 297–304.

[9] L. T.S., "Biometric human identification based on electrocardiogram," June 2005.

[10] I. Odinaka, P.-H. Lai *et al.*, "Ecg biometric recognition: A comparative analysis," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 6, pp. 1812–1824, Dec 2012.

[11] Y. Wan, J. Yao *et al.*, "A neural network to identify human subjects with electrocardiogram signals," in *Proceedings of the world congress on engineering and computer science*. Citeseer, 2008, pp. 1–4.

[12] A. Page, J. Turner, T. Mohsenin, and T. Oates, "Comparing raw data and feature extraction for seizure detection with deep learning methods," in *The 27th International Conference of the Florida Artificial Intelligence Society (FLAIRS'27)*, 2014.

[13] S. Viseh, M. Ghovanloo, and T. Mohsenin, "Toward an ultra low-power onboard processor for tongue drive system," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 62, no. 2, pp. 174–178, Feb 2015.

[14] A. Page, C. Sagedy, E. Smith, N. Attaran, T. Oates, and T. Mohsenin, "A flexible multichannel EEG feature extractor and classifier for seizure detection," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 62, no. 2, pp. 109–113, Feb 2015.

[15] A. Page, S. Pramod, T. Oates, and T. Mohsenin, "An ultra low power feature extraction and classification system for wearable seizure detection," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, Oct 2015.