

# An Adaptive Deep Learning Approach for PPG-Based Identification

V. Jindal, J. Birjandtalab, M. Baran Pouyan, M. Nourani

Quality of Life Technology Laboratory

The University of Texas at Dallas, Richardson, TX 75080

{vasu.jindal, birjandtalab, mxb112230, nourani}@utdallas.edu

**Abstract**—Wearable biosensors have become increasingly popular in healthcare due to their capabilities for low cost and long term biosignal monitoring. This paper presents a novel two-stage technique to offer biometric identification using these biosensors through Deep Belief Networks and Restricted Boltzman Machines. Our identification approach improves robustness in current monitoring procedures within clinical, e-health and fitness environments using Photoplethysmography (PPG) signals through deep learning classification models. The approach is tested on TROIKA dataset using 10-fold cross validation and achieved an accuracy of 96.1%.

## I. INTRODUCTION

### A. Motivation & Prior Work

In the emerging era of Internet of Things (IoT) wearable devices like smart bands and smart watches e.g. Fitbit, Samsung Gear [1], etc., have gained interest in monitoring individual health metrics. Recent developments in biometrics have focused on the usage of physiological signals in heart rate calculation, biometric authentication and blood pressure calculation. These applications make use of physiological signals like electrocardiograms (ECG) [2], electroencephalograms (EEG) and photoplethysmograms (PPG) [3], [4]. These physiological signals offer many advantages for personalization and identification of individuals due to their capability of providing insights into clinical conditions.

Additionally, physiological signals like PPG can be collected for prolonged periods without extra effort using a wrist or a finger-tip based sensor. Although ECG signals have been proven to be more accurate in source identification, most ECG sensors need to be positioned close to the heart to prevent motion artifacts [2]. This makes ECG signals inefficient in terms of cost and wearability convenience as compared to PPG signals. Therefore, in our approach, we choose PPG signals as the basis of building deep learning models for subject identification.

According to [5], factors like age, weight, gender as well as eating habits affect clinical metrics like heart rate, blood pressure etc. For example, a person who weighs over 80 kg may have a sharp increase in heart rate in a short interval of physical exercise while there is a much more gradual increase in a person weighing between 60 kg and 80 kg. This change in heart rate necessitates to create an adaptive model for identification of individuals into different groups and subsequently build personalized models to calculate individual health metrics.

Biometric identifications using PPG signals have been widely investigated in the literature. However, all prior works

in PPG source identification are limited to PPG signals obtained from controlled (clinical) settings and highly accurate sensors. These restrictions may make these models unsuitable to use in day to day environments due to motion artifacts.

Early efforts to use PPG as a form of human identification were made in [3] [4] using fuzzy logic based classification. Authors used four distinctive features of PPG as fuzzy inputs in classification models: the number of peaks, the upward slope, the downward slope, and the time interval from the bottom to the peak. However, these approaches limit themselves to four features and thereby, restrict them for accurate classifier training. Another technique presented in [6] investigated the feasibility of using PPG signals for biometric identification using majority voting classification. According to [6], researchers pre-processed the signal and extracted features using Linear Discriminant Analysis (LDA). The research concluded that signals should be collected in controlled environment with high precision sensors for source identification. This is a major shortcoming in existing approaches as in practical situations, subjects can be in various environments with different forms of motion.

### B. Main Contributions

In this paper we propose a two-stage technique to first cluster individuals into different groups. Then, we present a robust biometric subject identification technique using Restricted Boltzman Machines and Deep Belief Networks. Additionally, we present several new discriminative features of the PPG signal. Our key contribution is to personalize training procedure for classification of human's PPG signal and use this data for customized calculation of clinical outcomes. We choose deep learning as the core basis of our approach due to its various advantages including fast interference and fast unsupervised learning. Deep learning has been intensively used in image recognition and voice source identification. To the best of our knowledge, it has never been applied to PPG signal classification.

## II. PPG-BASED IDENTIFICATION METHODOLOGY

### A. Dataset

The TROIKA framework presented in [7] is used to verify the robustness of our novel approach. The TROIKA dataset consists of an ECG signal, two-channel PPG signals and three-axis acceleration signals from 12 male subjects with ages ranging from 18 to 35. For each subject, the PPG signals were recorded from wrist by two pulse oximeters with green LEDs of wavelength 609 nm. To make the data recordings

similar to practical world readings, the pulse oximeter and the accelerometer were embedded in a wristband and all signals were sampled at 125 Hz. During data recording, subjects walked or ran on a treadmill with the following speeds, in sequential order: 2 km/hour for 0.5 minute, 8 km/hour for 1 minute, 15 km/hour for 1 minute, 8 km/hour for 1 minutes, 15 km/hour for 1 minute, and 2 km/hour for 0.5 minute. Biometric identification using PPG should be possible even when the source is in heavy physical motion and henceforth, it is required to test learning models in practical day to day situations. Therefore, presence of different activities make TROIKA a suitable dataset for our model.

### B. Overview of the Approach

Figure 1 illustrates an overview of our two-stage method for identification of sources using PPG signals. Briefly, our technique involves Stage I.A: a pre-processing stage to extract noise reduced segments and features from the PPG signal. Consequently, in Stage I.B, the sources are categorized (clustered) into different subgroups. This clustering allows to make adaptive models for identification. Finally in Stage II.A and Stage II.B, we use Deep Belief Network (DBN) on each of these separate subgroups and identify individuals based on the extracted features. Each of these stages are discussed in the subsequent subsections.

### C. Stage I.A: Preprocessing

PPG signals obtained from embedded pulse oximeters in commercial products are quite noisy due to motion artifacts. This noise needs to be removed for accurate identification. Furthermore, distinctive feature should be extracted to input in deep learning models. Therefore, we perform a pre-processing stage including the following steps:

1. **Filtration:** PPG signals are highly susceptible to power line interference, respiration induced noise, muscle noise (EMG), and motion artifacts which may lead to inaccurate classification. We remove this noise in the PPG signal using a combination of normalization, moving average filter and Butterworth filter [8]. Firstly, the normalized PPG denoted by  $PPG_{norm}(t)$  is calculated as following:

$$PPG_{norm}(t) = \frac{PPG_{raw}(t) - \mu_{PPG}(t)}{\sigma_{PPG}(t)} \quad (1)$$

where  $\mu_{PPG}(t)$  and  $\sigma_{PPG}(t)$  represents the mean and standard deviation of the PPG signal respectively. This normalized PPG signal is further processed with a moving average filter of length  $F_c L$ , where  $F_c$  is the frequency of the signal and  $L$  is a variable user parameter. Similar to [8],  $L = 3$  was chosen to obtain the low-frequency trend. This low-trend is further removed using a 7th order Butterworth filter [8].

2. **Segmentation:** Segmentation of the PPG signal is necessary to extract discriminative features to input in classification models. We create segments each containing exactly one positive peak by using peak detection algorithms. The standard approach to find peaks by first deriving time series and identifying downward-going zero-crossing at peak

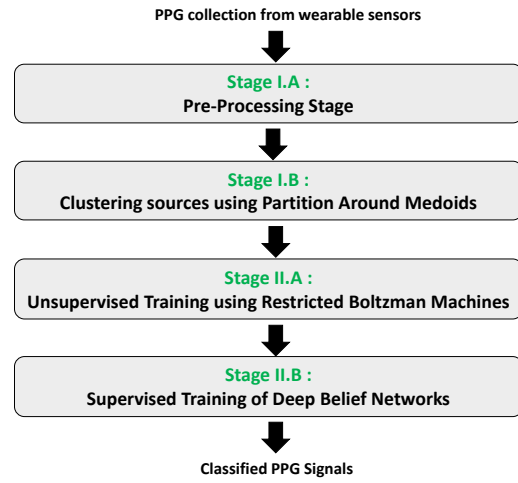


Fig. 1: The general view of our two-stage proposed method

maximums is incapable due to the random noise present in PPG. Therefore, we only consider zero crossings in the first derivative on two conditions: i) slopes exceeding the slope-threshold, ii) slope at locations where original signal exceeds the amplitude threshold. Essentially, each segment is a period of the signal.

3. **Interpolation and Extrapolation:** Interpolation and extrapolation are critical as segments obtained after the segmentation step may be of different lengths leading to inaccurate calculations of features in subsequent steps. Therefore, the segments are interpolated and extrapolated to a common length of  $l = 125$  using Box-Jenkins AutoRegressive Moving Average (ARIMA) [9].

4. **Feature Extraction:** We used the following 11 features to be used as input for clustering and classification, a) Mean of PPG values in the segment b) Standard deviation of the values in the segment c) Average of the Dynamic Time Warp distance  $F_{dtw}$  between the chosen segment and all other segments, i.e.

$$F_{dtw} = \frac{\sum_{i=1}^n dtw(s_c, s_i)}{n}, i \neq c \quad (2)$$

where  $s_c$  represents the current chosen segment from  $n$  total segments and  $s_i$  for represents every other segment except the chosen segment, d) Maximum value of the segment, e) Minimum value of the segment, f) Position of maximum value within the segment, g) Position of minimum value within the segment, h) Maximum and minimum of computed discrete wavelet coefficients of the segment [10], j) Fisher-Pearson coefficient of skewness of the segment which we formulated as

$$g_1 = \frac{\sum_{i=1}^n (T_i - \bar{T})^3 / n}{s^3} \quad (3)$$

where  $T_1, T_2, \dots, T_n$  represents the PPG values in the segment,  $\bar{T}$  is the mean,  $s$  is the standard deviation, and  $n$  is the number of data points (in our case 125), k) kurtosis of the PPG segment.

#### D. Stage I.B: Clustering Using PAM

In Stage I.B, individuals are separated into subgroups (clusters) before the final identification stage. We used Partitioning around Medoids (PAM) clustering method to separate sources into different groups by clustering the features of the source obtained from pre-processing stage. If a new source (not belonging to the dataset) is added to the set, we find its nearest neighbor and use the identification model for its corresponding group. Separate groups based on different forms of motions as well as different types of sources (male, female, heavy-weight, under-weight) can be separated through this stage.

Given an initial set of sources it is often difficult to find the optimal number of clusters to create subgroups between the individuals. We address this shortcoming using GAP statistics [11] to estimate the initial number of clusters. GAP statistics resulted in 3 as the approximated optimal number of clusters for our dataset. The sources clustered into the three clusters  $C_1$ ,  $C_2$  and  $C_3$  are given in Table 1.

TABLE I: Clusters in TROIKA dataset

Clusters	TROIKA sources in Cluster
$C_1$	[1,3,8,11]
$C_2$	[2,4,6,7]
$C_3$	[5,9,10]

#### E. Stage II: Deep Learning to Identify Individuals

The core of our technique uses the state-of-the-art deep learning for identification of individuals. Neural Networks with more than two layers allow the network to capture more complex patterns. We train deep learning classification models separately for each cluster as the groups may have been separated based on different forms of motion and different clinical metrics of individuals as described in the clustering step. Nevertheless, a common challenge in deep learning is overfitting which occurs when few training samples are used to train large models. We address this shortcoming, using a grid search approach to estimate the optimal size and number of hidden layers [12]. Our deep network contains one input layer, three hidden layers and one output layer for  $C_1$  with hidden layers consisting of 40, 40 and 10 nodes respectively as shown in Figure 2. One input layer, three hidden layers and one output layer is utilized in models for both  $C_2$  and  $C_3$  with the hidden layers consisting of 30, 30 and 20 nodes.

The input of these classification models are the 11 feature vectors obtained after pre-processing stage. The activation function of hidden units is the sigmoid function which is traditionally used in nonlinear neural networks:

$$g(z) = \frac{1}{1 + \exp(-z)} \quad (4)$$

In the function  $z = \theta x$ , where  $x$  is input and  $\theta$  is the weight parameter. We further use a softmax classifier for output layer. This implies that the units in the output layer give the probability of the input features corresponding to the source. The output  $h(x)$  of the softmax classifier for an input vector  $x$  is given by:

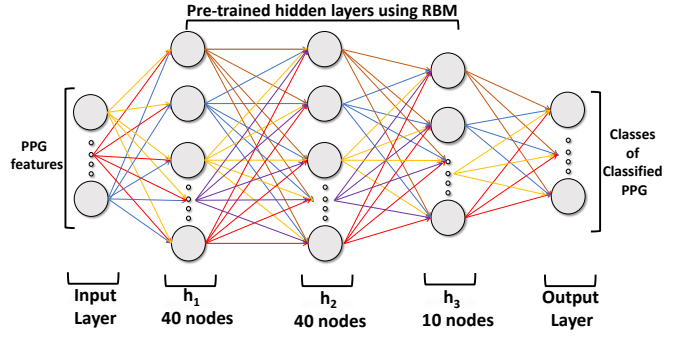


Fig. 2: Final Deep Belief Network for cluster  $C_1$

$$h_{\theta}(x) = \frac{1}{\sum_{j=1}^K \exp(\theta_j^T x)} \begin{bmatrix} \exp(\theta_1^T x) \\ \exp(\theta_2^T x) \\ \vdots \\ \exp(\theta_K^T x) \end{bmatrix} \quad (5)$$

where  $K$  equals the number of final classes and  $\theta$  is the weight of the parameter.

Another challenge in this technique is the higher number of parameters generally makes parameter estimation more difficult in neural networks. Therefore, it becomes extremely difficult to start the training of deep neural networks from random initial weight and bias values. We address this need and make our training faster and more feasible for day to day environments using a greedy unsupervised algorithm [13]. We further, incorporate Restricted Boltzmann Machines (RBM) to pre-train the network and find good initial weights for training deep belief networks. Briefly, RBMs are stochastic neural networks capable of predicting higher order and complex nonlinear input variables. Deep Belief Networks (DBN) are a class of Deep Neural Networks (DNN) with unsupervised pre-trained stacked Restricted Boltzmann Machines [14].

Let  $D_i$  be a DBN model for cluster  $C_i$ . In Stage II.A, the hidden layers of each  $D_i$  are first trained as RBMs using unlabeled inputs. We use Contrastive Divergence-1 (CD-1) [15] algorithm to obtain samples after 1-step Gibbs sampling. CD-1 allows accurate estimation of gradient's direction and minimize reconstruction error. This pre-training using unlabeled segmentations can be viewed as unsupervised learning and presents several advantages. First, due to this pre-training,  $D_i$  learns an identity function with same desired output as the original input. Secondly, it enhances the robustness of  $D_i$  by learning feature representations before the final supervised learning stage.

In Stage II.B, the pre-trained DBN network is fine-tuned by standard backpropagation with labeled segments as the input layer. Fine tuning is a critical step as it treats all pre-trained RBM layers as a single complete model and improves all weights in stacked RBMs in a single iteration. Figure 2 represents the final resulting  $D_1$  after tuning and is also used for identification for Cluster  $C_1$ .

TABLE II: Confusion Matrix of Clusters

(a) $C_1$					(b) $C_2$					(c) $C_3$			
Class/Predict	P1	P3	P8	P11	Class/Predict	P2	P4	P6	P7	Class/Predict	P5	P9	P10
P1	309	11	0	0	P2	311	1	0	6	P5	293	14	0
P3	12	251	4	0	P4	0	286	12	2	P9	18	294	2
P8	2	0	302	8	P6	23	0	318	0	P10	0	2	312
P11	6	0	1	319	P7	0	0	9	292				
Recall	93.9	95.8	98.3	97.5	Recall	93.1	99.5	93.8	97.3	Recall	94.2	94.8	99.3
Precision	96.5	94.0	97.4	98.0	Precision	98.1	95.3	93.2	97.0	Precision	95.4	93.6	99.3
Accuracy		96.4			Accuracy		95.7			Accuracy		96.1	

TABLE III: Comparing our algorithm against other methods

Method	Approach	# of Individuals	Setting of Collection	Personalization	Accuracy
[3]	Fuzzy Logic	17	Un-Controlled	No	82.3%
[6]	kNN	15	Controlled	No	74.0%
Ours	Clustering, Deep Learning	11	<b>Un-Controlled</b>	<b>Yes</b>	<b>96.1%</b>

### III. EXPERIMENTAL RESULTS

TROIKA [7] framework described earlier was used for all experiments to verify robustness of our model. A 10-fold cross validation is performed for each cluster to test the accuracy for source identification. Confusion matrix for each cluster with accuracy and recall rate is presented in Table II (a, b, c). We also compare the performance of our algorithm with two previous identification methods in Table III.

Our methodology is efficient in all three clusters  $C_1$ ,  $C_2$ ,  $C_3$  with accuracies 96.4%, 95.7%, 96.1% respectively, giving an average 96.1% accuracy for our dataset. Most misclassified cases in the confusion matrix may have occurred due to the noise present in the PPG signal. This can be solved using new enhanced filters in the pre-processing stage like spectral noise filter and adaptive noise filter. Table III illustrates that our algorithm achieves a better accuracy than other previous works. Furthermore, our model can handle activity independent data while previous approaches rely on controlled environments. We acknowledge that at present, the number of subject may be comparatively smaller for training purposes but we intend to evaluate our approach by collecting more data in the future.

### IV. CONCLUSIONS

This paper presents a novel two-stage technique for biometric identification involving clustering of individual PPG sources into different groups and using Deep Belief Networks as classification models. Our technique is robust against motion induced noise and thus suitable for usage in day to day settings. Furthermore, this procedure can be extended to build personalized models for clinical metrics calculation like heart rate, blood pressure using PPG signals in uncontrolled environments. Future work includes extracting more discriminative features for PPG signal using deep learning, and enhancing training of our model using larger datasets.

### REFERENCES

- [1] J. Takacs, C. L. Pollock, J. R. Guenther, M. Bahar, C. Napier, and M. A. Hunt, "Validation of the fitbit one activity monitor device during treadmill walking," *Journal of Science and Medicine in Sport*, vol. 17, no. 5, pp. 496–500, 2014.
- [2] A. Lourenço, H. Silva, and A. Fred, "Unveiling the biometric potential of finger-based ecg signals," *Computational intelligence and neuroscience*, vol. 2011, p. 5, 2011.
- [3] Y. Gu, Y. Zhang, and Y. Zhang, "A novel biometric approach in human verification by photoplethysmographic signals," in *4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine*, 2003, pp. 13–14.
- [4] Y. Gu and Y. Zhang, "Photoplethysmographic authentication through fuzzy logic," in *IEEE EMBS Asian-Pacific Conference on Biomedical Engineering*, 2003, pp. 136–137.
- [5] J. Hirsch, R. L. Leibel, R. Mackintosh, and A. Aguirre, "Heart rate variability as a measure of autonomic function during weight change in humans," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 261, no. 6, pp. R1418–R1423, 1991.
- [6] P. Spachos, J. Gao, and D. Hatzinakos, "Feasibility study of photoplethysmographic signals for biometric identification," in *17th IEEE International Conference on Digital Signal Processing (DSP)*, 2011, pp. 1–5.
- [7] Z. Zhang, Z. Pi, and B. Liu, "Troika: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 522–531, 2015.
- [8] E. Grisan, G. Cantisani, G. Tarroni, S. K. Yoon, and M. Rossi, "A supervised learning approach for the robust detection of heart beat in plethysmographic data," in *37th IEEE Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 5825–5828.
- [9] G. E. Box and D. A. Pierce, "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," *Journal of the American statistical Association*, vol. 65, no. 332, pp. 1509–1526, 1970.
- [10] D. Cvetkovic, E. D. Übeyli, and I. Cosic, "Wavelet transform feature extraction from human ppg, ecg, and eeg signal responses to elf pemf exposures: A pilot study," *Digital signal processing*, vol. 18, no. 5, pp. 861–874, 2008.
- [11] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 2, pp. 411–423, 2001.
- [12] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 281–305, 2012.
- [13] Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle *et al.*, "Greedy layer-wise training of deep networks," *Advances in Neural Information Processing Systems*, vol. 19, p. 153, 2007.
- [14] N. Le Roux and Y. Bengio, "Representational power of restricted boltzmann machines and deep belief networks," *Neural computation*, vol. 20, no. 6, pp. 1631–1649, 2008.
- [15] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural computation*, vol. 14, no. 8, pp. 1771–1800, 2002.