International Journal of Electrical and Computer Engineering (IJECE)
Vol. 11, No. 1, February 2021, pp. 424~433
ISSN: 2088-8708, DOI: 10.11591/ijece.v11i1.pp424-433

Bio-signals compression using auto-encoder

Sunilkumar K. N., Shivashankar, Keshavamurthy
Sri Venkateshwara College of Engineering, India

Article Info

Article history:

Received Jul 18, 2019
Revised Jun 28, 2020
Accepted Aug 6, 2020

Keywords:

Auto-encoder
Bio-medical signals
Feature selection

ABSTRACT

Latest developments in wearable devices permits un-damageable and cheapest way for gathering of medical data such as bio-signals like ECG, Respiration, Blood pressure etc. Gathering and analysis of various biomarkers are considered to provide anticipatory healthcare through customized applications for medical purpose. Wearable devices will rely on size, resources and battery capacity; we need a novel algorithm to robustly control memory and the energy of the device. The rapid growth of the technology has led to numerous auto encoders that guarantee the results by extracting feature selection from time and frequency domain in an efficient way. The main aim is to train the hidden layer to reconstruct the data similar to that of input. In the previous works, to accomplish the compression all features were needed but in our proposed framework bio-signals compression using auto-encoder (BCAE) will perform task by taking only important features and compress it. By doing this it can reduce power consumption at the source end and hence increases battery life. The performance of the result comparison is done for the 3 parameters compression ratio, reconstruction error and power consumption. Our proposed work outperforms with respect to the SURF method.

Corresponding Author:

Sunilkumar K. N.,
Sri Venkateshwara College of Engineering,
Bangalore, India.
Email: sunilkumarkn26@gmail.com

1. INTRODUCTION

IoT technology offers devices to sense the practical surrounding and effortlessly combine the collected information into refined applications that permits for significant enhancements of activities of the person. The main aim of this paper is sensing the human's behavior [1] via IoT device that is wearable like chest straps, smart watches, wristbands, etc. that can be utilized to point-out the specific health and the required fitness of the clients [2]. Wearable devices can be combined into body sensor networks that are wirelessly connected to bring reports of the medical up-to-date though internet. Therefore, allowing anticipation, early examine and special care. Anyways, as they are needed to be tiny and less weight, they are also limited with the resources such as power, capability of the transition and memory.

In this paper, we implement new solution for data processing for the endurance monitoring of ECG signals. These medical records are basically simple to measure, but simultaneously, exceptionally valuable for the already mentioned determination. We assume the achievement of those kind of signals via small wearable machines [3] are worried about the extending the durability of the battery of those devices via compression of lossy signals. We assume the conditions where transmission of ECG signals wirelessly to some specific access point is needed, so that the signals can be preserved on the server and can be used by

the clinicians throughout the world. The advanced wearable IoT devices where the energy and memory are the main issues while collecting the high-dimensional data such as heart-rate, level of the oxygen, and other bio-information by the IoT device and those large data are transmitted to the units which are more capable of handling and picturing-(representing) those bio-informational data. As the data collected by the IoT devices are very large, we need an algorithm, which is very efficient in compressing the data and transmitting to the more capable units, which does the execution of handling and picturing of data accurately and efficiently as well. Although an auto-encoder and its variants are used to compress the data where the output should be almost similar to the input. But, there was a major limitation, it uses most of the high-level features that are redundant to use while representing the output in the training data which is not related with the pattern.

Therefore, to solve this problem, the proposed approach consists of both the parts of the algorithm to compress the data where simultaneously the selection of important data is performed to select the data from the learned hidden-layers in one group and the data, which are not relevant in another group. Concurrently, the execution of encoding is done to the hidden layer's selected data, which are important. We are proposing an algorithm where selection of important data from input is done in the hidden layer, which is correlated to each other, and neglecting the remaining data of the input, which are not much useful or non-correlated. Here we are taking the bio-signal data by a Wearable IoT device to perform the compression of the data, our algorithm has two parts:

a. Here, the necessary information (Data) from the overall input data is chosen for the operation of the auto-encoding process and the rest all data that are not necessary are ignored by which few of the unwanted data can be eliminated.
b. On the selected data, the non-linear data is performed with the unsupervised learning using auto-encoder. We demonstrate our work as follows: (i) First we convert the sensor data in to a hidden-layers by machine-learning process. (ii) Then our algorithm is used to separate the important and relevant data in one group and non-relevant data into another group. (iii) In parallel to the above point, encoding of the relevant data is performed which is in the hidden-layers.

The following paper is arranged as mentioned below. In section 2, describes studying about the related works based on the selection of features from the input data and performing the compression using auto-encoder technique, Section 3 provides an overview of the proposed work and section 4 provides the simulated results compared with the other compression approach SURF. Our proposed work and at last we will be proving that our results are better than other existing methods.

2. LITERATURE SURVEY

There is a compression algorithm for the signals of the ECG, which is divided into three important methods: Parameter extraction, Transformation and the direct methods. For our algorithm there are two important condition: 1) selecting necessary data from the hidden layers that are trained. 2) Processing of the selected important data using auto-encoder.

2.1. Selection of the important data from the trained hidden layers

Here selection of the important data from the complete trained dataset hidden layers is grouped into one set and the data which are non-relevant is grouped into another dataset. We are using the feature selection method to group the relevant and non-relevant data in our algorithm. In [1], an "Automatic feature extraction (FE) and the selection for related datasets and the condition monitoring." Whereas, the combination of methods for FE and selection has been proposed suitable for extracting the huge related features for monitoring the machine condition and the relevant applications from the frequency domain, time domain, time-frequency domain and statistical sharing of the dimension values [1]. The approach is fully automated and suitable for multiple condition monitoring tasks such as vibration and process sensor-based analysis. This adaptability is explained by analyzing the two method of the monitoring datasets from the self-examination and the various freely present time series classification operations. Usually the condition monitoring is trained in three steps, namely feature extraction (FE) from raw data, feature selection (FS) of the most relevant features and fault classification using machine learning. FE and FS thereby represent a gradual data dimensionality reduction that is necessary to prevent the classifier to suffer from overfitting and the "curse of dimensionality" [2], which makes machine learning ineffective, if the number of features is high. Thus, high feature quality is the precondition for good prediction results. In the field of feature extraction, algorithms can be grouped into one of the following categories:

a. Piecewise approximation [3, 4].
b. Time domain transformations PCA [5])
c. Frequency domain transformations (e.g. Fourier transformation [6])
d. Time-frequency transformations (e.g. Wavelet transformation [7])

e. Statistical features (e.g. [8, 9])

Each of these approaches works under certain circumstances and can extract information unobtainable for the others. This and the impossibility to automatically (i.e. without detailed knowledge of the respective system) determine the best method in advance motivates the basic idea of the approach presented in this paper. In [10], they have introduced " A Novel merit for the selection of feature on the basis of rough set theory (RST)", which is motivated by the merit of the correlation-dependent and known as the rough set-based merit is implemented under rough set immediate reduces algorithm to choose a important subset of the features. In [11], the RST was defined for the analysis of data in the recognition pattern, machine learning and the data mining areas. RS is strong tool to identify a related relation amongst the class and the attributes under the dataset. Hence, utilization of RS as the FS step to identify a decrement of the features for a robust and fast categorization of the datasets. This paper giving attention on FS in DNA of the microarrays (also known as gene expression) datasets [12, 13]. The microarray datasets (MDs) had the features of real-value same as the real-world datasets. A MDs is a strong example of the large dimensionality. MDs consists of the several features and few numbers of the samples. Moreover, sharing of the classes in binary MDs is unequal and since they are the datasets of class-imbalanced.

2.2. Performing the operation of auto encoder on the selected data from the complete data

In paper [14], they introduced a "semi-supervised-stacked-label-consistent-auto-encoder" for reconstruction and investigation of Bio-signals. Old research has given reconstruction and categorizing as two different issues. For the purpose of tele-monitoring, the methodology of reconstruction of bio-signals are hugely depending on the compressed-sensing (CS); they are designed approaches where the formulation of the reconstruction is depended on some prediction about the signal. CS dependent project solutions, a part of the gathered signal (such as one second) on a matrix that is random (sparse binary, Gaussian and binomial) so that, the projected data's size is less when compared with the size of samples that was taken in one second [15, 16]. As CS needs only a product of matrix-vector, it costs less for computation. There are many research studies, which implement effective computation hardware and energy effective for same [17-21]. The data that is compressed wirelessly sent to base station. At base station, reconstruction with the help of CS methodology is done for later monitoring and examination [22, 23]. There can be many types of general CS methodology for re-construction; [17, 24, 25] use types of sparse-bayesian learning (SB) [26, 27, 28]. Many of the papers utilize the typical CS methods for recovery in few of the operations, the reconstruction of the signals was done with the help of intra and inter-channel correlations [29, 30].

In this article, they introduce a new argumentative graph, surrounding structure for data of the graph. The structure translates the topological framework and content of the node in a graph for compressed demonstration, on which the training of the decoder is done to reconstruct the structure of the graph [31, 32]. Graphs are important tools to recognize and build complex relations between the data. In different types of graph applications consisting protein-protein communication network, citation network and social media networks, data of the analyzing graph is essential in different data mining operations consisting classification of the nodes, prediction of the links and clustering of the nodes [32, 33]. Anyways, the parallelizability that is low, complexity of high computations, and incapability of machine-learning (ML) technique to the data of the graph made the operation of analytic of the graph more difficult. Algorithms that are based on factorization of the matrix. For example: HOPE, GraRep, M-NMF [15, 29] pre-process the framework of the graph into an adjacency matrix and obtain the embedding by disintegrating the adjacency matrix. Lately, it is represented that several algorithms of probabilistic are equal to factorization of matrix methods. Deep learning methods specifically auto-encoder dependent techniques, are also mainly researched for embedding of the graph. The algorithm of MGAE uses a relegated one-layer auto-encoder to learn clustering representation [34, 35].

An auto-encoder that is standardized by graph was proposed by Yu *et al*. aiming to select the graph to guide the operations of encoding and decoding. Anyways, it is still hard to learn with several non-related patterns in data, and present auto-encoder difference is still not included the units, which are hidden into 2 separate parts, first is work-related and the second one is work non-related.

3. PROPOSED WORK

In this section we will be explaining what is the aim of our algorithm and how will that be performed, following by the combined model of selecting important data which are relevant and applying auto-encoder in the selected data of hidden layers to compress the data. After that we study about the two algorithms on which our proposed algorithm is based. And also, in detail our architecture is explained. The main concept of auto-encoder is to arrange the data from the non-linear encoder into the hidden layers and the hidden units are utilized as the new data arrangements:

$$HL_i = \sigma(W_{in}o_i + v_{in});\ \widehat{o_i} = \sigma(W_{out}HL_i + v_{out}) \quad (1)$$

Here, $HL_i \in \mathbb{S}r^{sv_{in}}$ is the arrangement of the hidden units and $\widehat{o_i} \in \mathbb{S}r^{sv_{out}}$ is assumed as a rebuilt of the given input $o_i \in \mathbb{S}r^{sv_{out}}$. The set of the parameters consists of weight matrices$W_{in} \in \mathbb{S}r^{sv_{in} \times sv_{out}}$, $W_{out} \in \mathbb{S}r^{sv_{out} \times sv_{in}}$ and the vectors which are offset are $v_{in} \in \mathbb{S}r^{sv_{in}}$ and $v_{out} \in \mathbb{S}r^{sv_{out}}$ whose sizes are sv_{in} and sv_{out} respectively. σ is defined as non linear activation-function. The Neural Network which has same input and target data is nothing but an auto-encoder who consist of only one hidden layer, such that

$$\min_{W_{in},W_{out},v_{in},v_{out}} \frac{1}{2n}\sum_{i=1}^{n}\|o_i - \widehat{o_i}\|_2^2 \quad (2)$$

Here, n is the data's size of the sample, o_i is the result after re-building it and $\widehat{o_i}$ is the target data to which the rebuilt data is compared. A better arrangement of the hidden layer data can be gained with the skills to rebuild the data.

As we studied before, all hidden units, which are at high-level, help in finding the essential data from the given input data while rebuilding the data or arranging the selected important data. Anyways, the importance of these units is not same as the importance of our classification work. For instance, some data are used to build the background image in the object image, which is not required. Those type of data are known as irrelevant data or less important data, which are unwanted for our features, which have been trained newly. In parallel, meanwhile the old method, which is not supervised, had a capacity with some limits to reconstruct the input distribution marginally for the aim-supervised processes. On the hidden units, few existing works made the label information with the help of Soft-Max Layer by "Socher *et al*." in the year 2011. Making count of the recent approximation regarding process units, which are not relevant, it is not okay or also to declare all the hidden units, as they are useless.

So we have two outcomes: 1) one is to select the important data required out of irrelevant data with the help of Feature selection approach, and 2) the selected important data only has to combined and sent by the auto-encoder. With the above study, we implement our combined selection of important data and compressing the selected data using auto-encoder algorithm in an integrated structure.

3.1. Bio-signals compression using auto-encoder (BCAE)

In this part, we implement our algorithm by combining both the concept of feature selection to select the important data and auto-encoder to compress the data within a single algorithm. Explicitly, we apply feature selection approach on the units of hidden layer data. Assume $P \in \mathbb{S}r^{sv2 \times n}$is the set of training elements, where $sv2$ is the size of the visual descriptor and n is the amount of samples of the data.

$$\min_{W_{in},W_{out},v_{in},v_{out},Q} \frac{1}{2}\left\|P - g\big(f(P)\big)\right\|_F^2 + \frac{10\lambda}{5}\mathcal{C}\big(Q, f(P)\big) \quad (3)$$

Where $\sigma(W_{in}P + V_{out})$ is assigned as $f(P)$ and $\sigma(W_{in}f(P) + V_{out})$ is assigned as $g\big(f(P)\big)$and V_{in} and V_{out} are each iteration's data reproduction of column of v_{in} and v_{out}respectively. The standardized form of selection of important data is given as $\mathcal{C}\big(Q, f(P)\big)$, along with Q i.e. matrix consisting trained feature selection data performing on $f(P)$ which is represented as hidden units. Explicitly, in Q, the vector of i^{th} column is represented as $q_i \in \mathbb{S}r^{sv_{in}}$ which is represented as,

$$q_i = [0, \ldots, 0, 1, 0, \ldots, 0]' \quad (4)$$

Here, the hidden units' amount is denoted as sv_{in} and j represents which column vector chooses the *j*-th unit from the sub collections of the features selected $np \in \mathbb{S}r^{eS}$. Then the method of feature selecting can be represented as the real set of features $f(P)$, looking for the Q-matrix to choose the new features and collect into a set $NP = Q'f(P)$ that optimizes the relatable criterion $\mathcal{C}\big(Q, f(P)\big)$.

Generally, operation for selection of data that are necessary can be takes place in 3 ways: unsupervised, supervised, semi-supervised. Coming to supervised methods, most of the times it consist of irrelevant information while selecting the necessary features such as fisher score that was introduced in the year of 2012, by Han, Li and Gu. In the year of 2005, Niyogi, He and Cai introduced laplacian score for unsupervised method. Unsupervised methods is assembles only the relevant and needed data while selecting main features. Trying to work with various cases in the actual world, we simplify the $\mathcal{C}\big(Q, f(P)\big)$ in simplest method. Explicitly, we introduce the regularizer for the selection of features in general as:

$$\mathcal{C}\big(Q, f(P)\big) = \frac{Tr\big(Q'f(P)Ml_w f'(P)Q\big)}{Tr\big(Q'f(P)Ml_b f'(P)Q\big)} \quad (5)$$

That offers a basic method to adjust with various cases by selecting Ml_b and Ml_w in various methods. Our proposed system follows the fisher score supervised approach in order to collect the information in supervised manner. For fisher score, two graphs G_b and G_w which are undirected are built using provided data (here, we take help of the input data P from the original one to keep the geometrical framework safe while selecting the features), which correspondingly effects the affinity relationship of the between-class and inside-class (introduced by- Yan *et al.* in the year-2007). In the same manner, two weighted matrices Mw_b and Mw_w are generated to categories both the graphs. Thus, we get the Laplacian matrices denoted as$Ml_w = Md_w - Mw_w$, here Md_w is the $Mw_w's$ diagonal matrices, in the same way for Ml_b andMw_b. By equating below issue of optimization, we can get the selection of feature matrix Q that generates the subset of the features with the low criterion score:

$$Q = arg\min_{Q} \frac{Tr\big(Q'f(P)Ml_w f'(P)Q\big)}{Tr\big(Q'f(P)Ml_b f'(P)Q\big)} \quad (6)$$

Unwantedly, there has not been a direct problem to resolve the given above trace-ratio issues, because of the absence of closed-form resolutions. Therefore, in place of straightforwardly dealing with trace-ratio issues. Many jobs desires to change to similar trace-various issues to gain a global resolution for optimization (Proposed by Nie *et al.* in year-2008). Assume standard score for subset-level is $\mathcal{C}\big(Q, f(P)\big)$ in the (5) goes up to the global lowest λ^*satisfying,

$$\lambda^* = arg\min \frac{Tr\big(Q'f(P)Ml_w f'(P)Q\big)}{Tr\big(Q'f(P)Ml_b f'(P)Q\big)} \quad (7)$$

This is to show that,

$$\Longrightarrow \frac{Tr\big(Q'f(P)Ml_w f'(P)Q\big)}{Tr\big(Q'f(P)Ml_b f'(P)Q\big)} \geq \lambda^*, \forall\, Q$$
$$\Longrightarrow Tr(Q'f(P)(Ml_w - \lambda^* Ml_b)f'(P)Q) \geq 0, \forall Q$$
$$\Longrightarrow \min_{Q} Tr(Q'f(P)(Ml_w - \lambda^* Ml_b)f'(P)Q) = 0$$

To this stage, we can represent the function of λ while using others as constant as:

$$R(\lambda) = arg\min_{Q} Tr(Q'f(P)(Ml_w - \lambda Ml_b)f'(P)Q) \quad (8)$$

Thus, looking for the global optimal λ can be transformed into equating the equation's root $R(\lambda) = 0$, which is trace-variance issue. See that $R(\lambda)$ is a function, which increases monotonically. By proposing the trace variance optimization issue, which is illustrated above in the (8) within the auto-encoder's hidden layer updating, we equate our final aim function as:

$$\min_{W_{in}, W_{out}, v_{in}, v_{out}, Q, \lambda} \mathfrak{m} = \frac{1}{2}\big\|P - g\big(f(P)\big)\big\|_F^2 + \frac{\gamma}{2} Tr(Q'f(P)(Ml_w - \lambda Ml_b)f'(P)Q) \quad (9)$$

Here, γ is the parameter used to balance which is placed between the selection of the feature and auto-encoder term. λ Denotes the ratio of trace's score after gaining the optimization Q and here Q is the trace ratio of optimization of recent issue.

3.2. Enhancement

Formula (9) is bit difficult to resolve because of its complex and non-linearity of the decoders and the encoders, therefore another way for enhancing is used to renew the parameters of the auto-encoder $W_{in}, W_{out}, v_{in}, v_{out}$ and variable of selection of the feature Q and also λ each time. Explicitly, we resolve the optimization with 2 sub-issues. First is learning the score of the selection of the feature and the second one is the optimization of the auto-encoder which is regularized.

3.3. Learning the score of the selection of the features

When the auto-encoder's parameters are made constant, we can enhance the selection of the features score λ and matrix of selection of the feature Q in an old method of trace ratio. Explicitly, we use trace variance equation

$$Q = arg\min_{Q} Tr(Q'f(P)(Ml_w - \lambda Ml_b)f'(P)Q) \quad (10)$$

Assume Q_k is the outcome after optimization in the k-th iteration of optimization, therefore λ_k is evaluated by

$$\lambda_k = \frac{Tr(Q_k'f(P)Ml_w f'(P)Q_k)}{Tr(Q_k'f(P)Ml_b f'(P)Q_k)} \quad (11)$$

Thus, we can get $R(\lambda_k)$ as

$$R(\lambda_k) = Tr(Q_{k+1}'f(P)(Ml_w - \lambda_k Ml_b)f'(P)Q_{k+1}) \quad (12)$$

where Q_{k+1} maybe evaluated proficiently with respect to score rank of every single feature. The equation's root $R(\lambda) = 0$ and the enhanced solution for (6) can be gained via its repeating process. See that λ is renewed as the enhanced score which is global for the criterion of the selection of the feature, and operates as a parameter in succeeding iteration of auto-encoder. Algorithm 1 explains the solution of enhancement.

Algorithm 1: Enhancement for the issue of Trace-Ratio.

Input: Trained features of the hidden layer $f(P)$, relevant data selected from the given input data eS, matrices Ml_w and Ml_b.

Initialize: assign Q with I, where I presents the identical matrix and it belongs to $\mathbb{S}r^{eS\times sv_{in}}$, λ with Equation (11), $expo = 10^{-9}$, $Iteration = 0$, $Max_{iteration} = 10^3$

IF $Iteration \leq Max_{iteration}$ and non-converge **do**

- Calculate score of every single feature of j-th iteration with the help of Equation(11) by assigning $Q = [0, \dots, 0, 1, 0, \dots, 0]'$
- Arrange the features according to their score in incremental order.
- Choose the highest eS feature to renew $Q \in \mathbb{S}r^{sv_{in}\times eS}$
- Evaluate λ with the help of Equation (11)
- Convergence constrains is checked: $\|\lambda_{old} - \lambda\| < expo$
- Start again from **IF** condition.

End of IF

Output: Q is the matrix of selection of the feature and λ is the enhancement score globally.

3.4. Enhancement of the auto-encoder which is regularized

When Q and λ are made constant, we can use stochastic-sub-gradient-descent technique to get $W_{in}, W_{out}, v_{in}, and\ v_{out}$. The slopes of the aiming function $\mathfrak{m}$ in (9) according to the parameters of the decoding is calculated as given below:

$$\frac{\partial \mathfrak{m}}{\partial W_{out}} = \Big(P - g\big(f(P)\big)\Big) \odot \frac{\partial g(f(P))}{\partial W_{out}} f'(P),$$

$$\frac{\partial \mathfrak{m}}{\partial B_{out}} = \Big(P - g\big(f(P)\big)\Big) \odot \frac{\partial g(f(P))}{\partial W_{out}} = \mathfrak{m}_{out},$$

$$\frac{\partial \mathfrak{m}}{\partial W_{in}} = \big(W_{out}'\mathfrak{m}_{out} + \gamma QQ'f(P)(\mathfrak{m}_w - \lambda \mathfrak{m}_b)\big) \odot \frac{\partial f(P)}{\partial W_{out}} P',$$

$$\frac{\partial \mathfrak{m}}{\partial V_{in}} = \big(W_{out}'\mathfrak{m}_{out} + \gamma QQ'f(P)(\mathfrak{m}_w - \lambda \mathfrak{m}_b)\big) \odot \frac{\partial f(P)}{\partial W_{out}}.$$

Then, $v_{in}, v_{out}, and\ W_{in}, W_{out}$ can be renewed with the help of the algorithm of gradient descent as given below:

$$W_{in} = W_{in} - \theta \frac{\partial \mathfrak{m}}{\partial W_{in}},\ v_{in} = v_{in} - \theta \frac{\partial \mathfrak{m}}{\partial v_{in}},$$
$$W_{out} = W_{out} - \theta \frac{\partial \mathfrak{m}}{\partial W_{out}},\ v_{out} = v_{out} - \theta \frac{\partial \mathfrak{m}}{\partial v_{out}},$$

Here, the rate of learning is denoted as θ.$\frac{\partial \mathfrak{m}}{\partial v_{in}}$ and $\frac{\partial \mathfrak{m}}{\partial v_{out}}$ are the column mean of $\frac{\partial \mathfrak{m}}{\partial V_{in}}$ and $\frac{\partial \mathfrak{m}}{\partial V_{out}}$ respectively. To add, the two part of the issues given above could be renewed recursively. Algorithm 2 explains the enhancement in detail.

Algorithm 2: Resolving the Issue in (9)

Input: parameters γ, Learning Data P, Layer_size, number of features selected $eS < sv_{in}$.

Initialize: $Max_{iteration} = 50, Iteration = 0, expo = 10^{-7}$.

While $Iteration \leq Max_{iteration}$ and non-converge **do**

- Keep others constant and renew Q and λ with the help of Equation (10).
- Make Q as constant and renew v_1, v_2 *and* W_1, W_2 with the help of Equation (13).
- Check for the constraints of convergence: $\|\mathfrak{m}_{latest} - \mathfrak{m}_{recent}\|_{\infty} < expo$.

End

Output: $Q, v_{in}, v_{out}, W_{in}, W_{out}$

$(NP = Q'\sigma(W_{in}P + V_{in}))$Can be utilized as an input next for our algorithm, to form the framework of stack.

Examine: the representation of the new features can be done as:

$$NP_{examine} = Q'\sigma(W_{in}P_{examine} + V_{out})$$

4. EXPERIMENTAL RESULT AND ANALYSIS

We represent the quantitative outcomes to estimate the efficiency of modified auto-encoders in the biomedical signal compression. MIMIC-II [21], the database can be available in 2 forms. In 1st form, the interested researchers can get the text version of flat-file of medical database and also the scheme of associated database which permits to recreate the database utilizing their technique of choice. In 2nd form, attentive researchers can obtain access to database via secure with password web services. The database researchers need the consumers to explain themselves with the help of database layout to the queries of program database utilizing the SQL (Query output can be exported to comma-separated files to be analyzed offline). The query output is transported to the files as comma-separated to be examined offline utilizing statistical or another software. The data is processing and accessing from the MIMIC-II, which is very complex. It can be highly suggested that the studies are based on the MIMIC-II database that can be directed as collaborative efforts, which contain statistical, relational and clinical database expertise. At the end, we utilized Physio net MIMIC II database [21] that consists information traces for the ECG. Additionally, we assumed our ECG traces, which developed via the Zephyr's Bio, harness 3 heart rate monitor. We represent the chosen outcomes which is gotten from the total of 16985 like sequences from 3 various Physio net of patients as well as 1400 for our method of ECG measurements and the test is picked so as to cover the important range of situations.

Figure 1 represents the actual signal graphs and recreated the waveform of ECG signal with trace 100. As we can see the graph, the actual signal was recreated successfully with the help of proposed model compression. The modified auto encoder (MAE) was worked in recreating all R in actual signal with the help of amplitude difference. 17 fragments with trace 100 were utilized in test phase.

Figure 2 represents the actual signal graph and recreated the ECG signal waveform with trace 100. Whereas, we can cee that the graph, that actual signal was recreated successfully at receiver with the help of proposed model compression. The MAE was successful in recreating the actual signal with the help of little difference and as the hidden nodes were maximizing the reconstruction accuracy also increases.

Figure 3 represents the actual signal graph and recreated the ECG signal waveform with trace 100. From the graph, we can see the development in the reconstructed signal with the actual signal. With maximizes the hidden nodes the accuracy of reconstructing has improved when compared with hidden nodes 2 and 4. Figure 4 shows the improvement in the reconstructed signal of ECG signal at receiver with that of actual signal. In the above figure uses 24 hidden nodes for training and reconstructing the signal. The reconstructed signal is having negligible amplitude difference with respect to original signal.

Figure 5 represents the actual signal graph and the recreated the waveform of an ECG signal with trace 100. As we can see from the graph, the actual signal was successfully recreated with the proposed model of compression for hidden nodes 48. The modified auto encoder model succeeded in reconstructing with that of actual signal with zero amplitude difference.

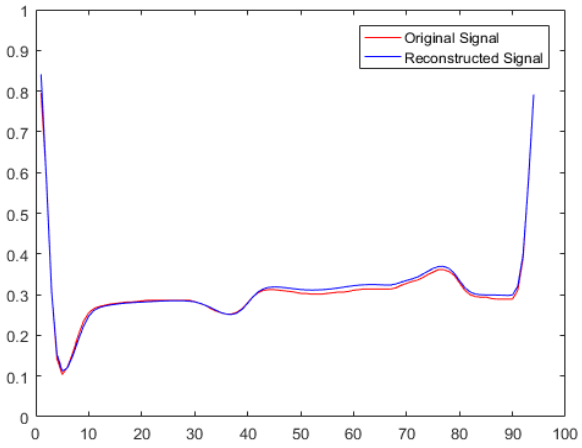

Figure 1. Feature graph at hidden nodes = 2

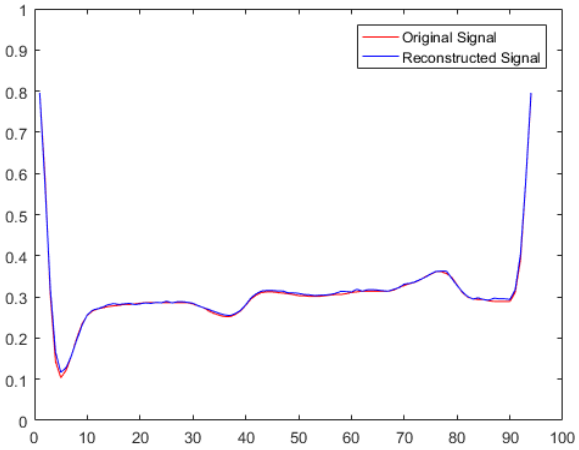

Figure 2. Feature graph at hidden nodes = 4

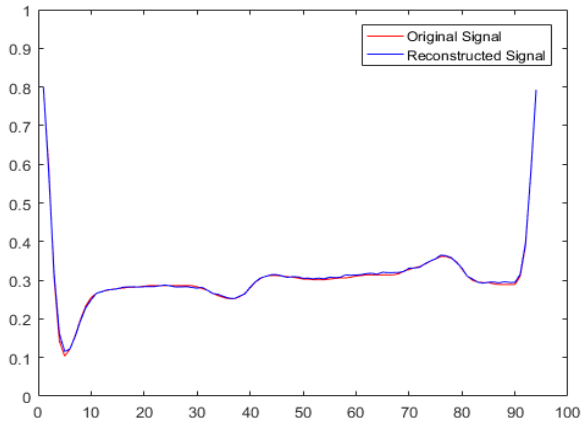

Figure 3. Feature graph at hidden nodes = 12

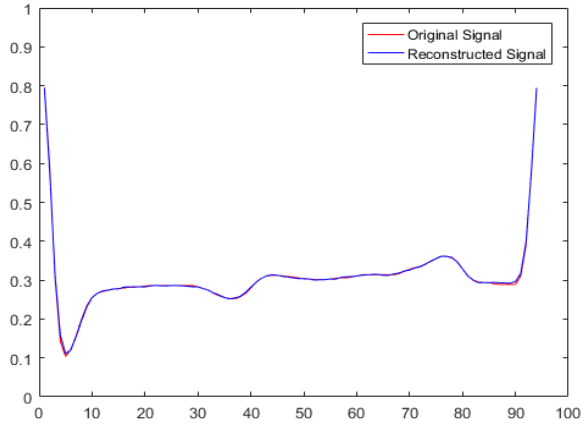

Figure 4. Feature graph at hidden nodes = 24

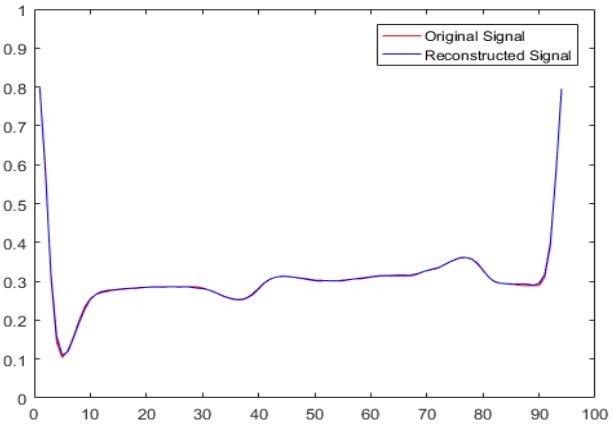

Figure 5. Feature graph at hidden nodes = 48

Figure 6 shows the root mean square error vs energy consumption (joule per bit) using our proposed model and SURF [24], where our modified auto-encoder got less RMSE w.r.t SURF [24] approach. Figure 7 shows the RMSE (%) with respect to Compression Efficiency where our model has performed significantly well w.r.t SURF [24] approach. Figure 8 represents the RMSE with regards to the various number of the Hidden Nodes, where our model has performed significantly well w.r.t SURF [24] approach.

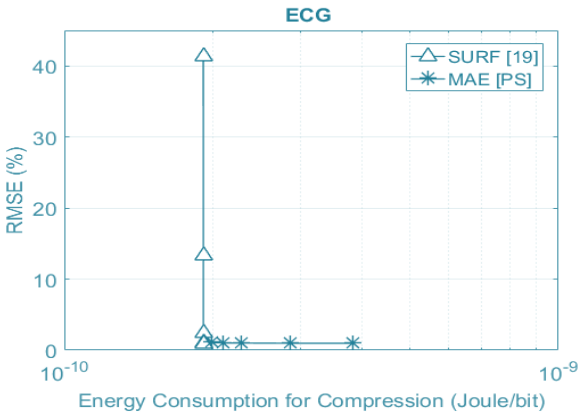

Figure 6. RMSE vs energy consumption

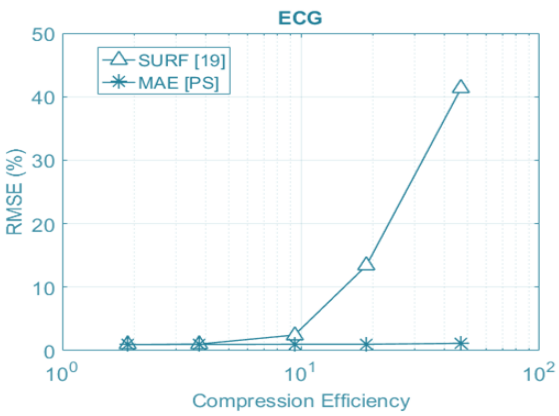

Figure 7. RMSE vs compression efficiency

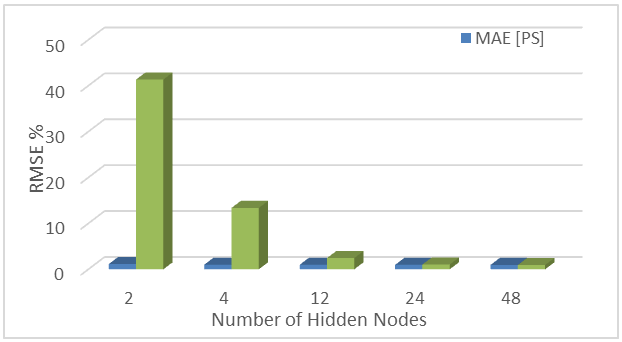

Figure 8. RMSE with respect to different number of hidden nodes

5. CONCLUSION

In this research work, we have presented *Bio signal compression using auto encoders (BCAE)* compression algorithm for wearable fitness monitors using IoT. We proposed the ECG-based biometric system whereas the lesser dimensional of non-linear representations of the heartbeat templates are learned through the deep modified Auto encoder. This study also represents the efficient signal of ECG compression method which is utilized in biomedical field i.e. e-health applications, Holter systems and telemetry. The proposed method of compression integrates the architecture of MAE, which have recently become very popular in the machine learning (ML) field. Therefore, it can be possible to get the data in the deep layers about the actual input signal, consisting lower level to higher level of features. In this study, the deep structure of MAE is proposed, 2000×1 ECG signals were recreated successfully by showing only 62×1 dimension. The application of comprehensive was made on the 4800 fragments of 2000 samples from the 48 patients in the dataset of MIT-BIH arrhythmia. With the help of proposed model, the important compression outcomes were gained with the rate of CR 32.25 and an average rate of PRD 2.73%. The proposed methodology also has the structure which can be utilized safely to transfer the biomedical information to remote locations securely.

REFERENCES

[1] P. Patel and A. Thakkar, "The upsurge of deep learning for computer vision applications," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, pp. 538-548, 2020.
[2] S. Narejo, et al., "EEG Based Eye State Classification using Deep Belief Network and Stacked AutoEncoder," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 6, no. 6, pp. 3131-3141, 2016.
[3] P. Reimann and A. Schütze, "Fire detection in coal mines based on semiconductor gas sensors," *Sensor Review*, vol. 32, no. 1, pp. 47-58, 2012.
[4] F. Yasmeen, et al., "Detection of Real Time QRSComplex Using WaveletTransform," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 5, pp. 2857-2863, 2018.
[5] S. Wold, et al., "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1-3, pp. 37-52, 1987.

[6] F. Morchen, "Time series feature extraction for data mining using DWT and DFT," Tech. Rep. No. 33, Dep. Mathmatics Comput. Sci. Philipps-University Marburg, pp. 1-31, 2003.
[7] R. Yan, et al., "Wavelets for fault diagnosis of rotary machines: A review with applications," *Signal Processing*, vol. 96, pp. 1-15, 2014.
[8] N. Helwig, et al., "Condition monitoring of a complex hydraulic system using multivariate statistics," in *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, pp. 210-215, 2015.
[9] S. S. Kohli, et al., "Hilbert Transform Based Adaptive ECG R-Peak Detection Technique," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 2, no. 5, pp. 639-643, 2012.
[10] M. Mohtashami and M. Eftekhari, "Using a novel merit for feature selection based on rough set theory," *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, Kerman, pp. 68-70, 2018.
[11] Z. Pawlak, "Rough sets," *International Journal of Parallel Programming*, vol. 11, no. 5, pp. 341-356, 1982.
[12] V. Bolon-Canedo, et al., "A review of microarray datasets and appliedfeature selection methods," *Information Sciences*, vol. 282, pp. 111-135, 2014.
[13] Y. Ma, Y. Wang, J. Yang, Y. Miao and W. Li, "Big Health Application System based on Health Internet of Things and Big Data," *in IEEE Access*, vol. 5, pp. 7885-7897, 2017.
[14] Lakshminarayana M. and M. Sarvagya, "OFCS: Optimized Framework of Compressive Sensing for Medical Images in Bottleneck Network Condition," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 5, pp. 2829-2838, 2018.
[15] S. Fauvel and R. K. Ward, "An Energy Efficient Compressed Sensing Framework for the Compression of Electroencephalogram Signals," *Sensors*, vol. 14, no. 1, pp. 1474-1496, 2014.
[16] M. Mohsina and A. Majumdar, "Gabor Based Analysis Prior Formulation for EEG Signal Reconstruction," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 951-955, 2013.
[17] F. Pareschi, et al., "Hardware-Algorithms Co-Design and Implementation of an Analog-to-Information Converter for Biosignals Based on Compressed Sensing," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 1, pp. 149-162, 2015.
[18] F. Chen, et al., "Design and Analysis of a Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 3, pp. 744-756, 2012.
[19] N. Verma and A. Chandrakasan, "An ultra low energy 12-bit rate-resolution scalable SAR ADC for wireless sensor node," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 6, pp. 1196-1205, 2007.
[20] R. Harrison and C. Charles, "A low-power low-noise CMOS amplifier for neural recording applications," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 6, pp. 958-965, 2003.
[21] Z. Zhang, et al., "Compressed sensing for energy-efficient wireless telemonitoring of noninvasive fetal ECG via block sparse Bayesian learning," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 2, pp. 300-309, 2013.
[22] M. H. Kamal, et al., "Compressive Multichannel Cortical Signal Recording," *2013 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2013.
[23] A. Majumdar, et al., "Low-rank Matrix Recovery Approach for Energy Efficient EEG Acquisition for Wireless Body Area Network," *Sensors*, vol. 14, no. 9, pp. 15729-15748, 2014.
[24] Z. Zhang, et al., "Spatiotemporal Sparse Bayesian Learning with Applications to Compressed Sensing of Multichannel Physiological Signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 6, pp. 1186-1197, 2014.
[25] Z. Zhang and B. D. Rao, "Extension of SBL Algorithms for the Recovery of Block Sparse Signals with Intra-Block Correlation," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2009-2015, 2013.
[26] S. Pan, et al., "Adversarially Regularized Graph Autoencoder for Graph Embedding," *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*, pp. 2609-2615, 2018.
[27] Z. Zhang, T. P. Jung, S. Makeig, Z. Pi and B. D. Rao, "Spatiotemporal Sparse Bayesian Learning with Applications to Compressed Sensing of Multichannel Physiological Signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 6, pp. 1186-1197, 2014.
[28] Jinkwon K., Hang S. S., Kwangsoo S., Myoungho L., "Robust algorithm for arrhythmia classification in ECG using extreme learning machine," 2011. [Online]. *Available: https://link.springer.com/article/10.1186/1475-925X-8-31*.
[29] Jing N., Victoria G., Shirley G., Gaelle L., Chee T. P., "A novel approach to detect atrial fibrillation efficiently and accurately from 48 hours of ECG data," *Journal of Electronics and Communication Engineering Research*, vol. 2, no. 1, pp. 01-06, 2014.
[30] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: Useful tips and tricks*," IEEE Commun. Surveys Tuts.,* vol. 19, no. 3, pp. 1527–1550, 2017.
[31] X. Wang, et al., "Community preserving network embedding," in *The 31st AAAI Conference on Artificial Intelligence*, pp. 203-209, 2017.
[32] S. Aviyente, "Compressed Sensing Framework for EEG Compression," *IEEE Workshop on Statistical Signal Processing*, pp. 181-184, 2007.
[33] C. Wang, et al., "MGAE: Marginalized graph autoencoder for graph clustering," in *CIKM'17*, pp. 889-898, 2017.
[34] J. Chiang and R. K. Ward, "Energy-Efficient Data Reduction Techniques for Wireless Seizure Detection Systems," *Sensors*, vol. 14, no. 2, pp. 2036-2051, 2014.
[35] W. Ding, Y. Lu, F. Yang, W. Dai, P. Li, S. Liu, and J. Song, "Spectrally efficient CSI acquisition for power line communications: A Bayesian compressive sensing perspective*," IEEE J. Sel. Areas Commun.,* vol. 34, no. 7, pp. 2022-2032, jul. 2014.