Transfer learning autoencoder used for compressing multimodal biosignal

Ithaya Rani Panneerselvam [1]

Abstract

Electrocardiogram, electromyogram, electroencephalogram are the foremost required vital signs for diagnosing chronic diseases like sleep disorder, mood disorder, epilepsy etc., which demands long-term monitoring. A sensor based wearable system which is enabled with internet technology, supports the continuous recordings of these vital signs without troubling the patient's daily activities. And the wearable hub is responsible for collecting the readings of biosignals from multiple micro-sensor nodes deployed around the body which creates the short range of communication and forward to the observer. These continuous monitoring increases the signal transmission cost and declines the battery life of wearables. So, the observed multiple biosignals can be compressed jointly than individually before sending, at an edge level. This paper proposes transfer learning based multimodal convolutional denoising auto encoder to perform multimodal compression and to reconstruct the data from its latent representation. Transfer learning helps the system to reuse the learned weights which may reconstruct the data with better quality score than by randomly initialized weights. The proposed work achieves compression ratio of 128 and it is proved that multimodal compression is better than unimodal compression in case of consuming multiple sensors. And the experimental result proves that the computation cost is low in multimodal compression than in unimodal compression.

Keywords Biosignal · Health monitoring system · Wearable IoT device · micro-sensors · micro-hub · Wireless communication · Multimodal compression · Transfer learning

1 Introduction

Feature extraction technology using CNN have been examined immense in the earlier [2, 8, 10, 25, 31]. Some of the most common chronic diseases like sleep disorder [16, 35], mood

✉ Ithaya Rani Panneerselvam
muhilrani@gmail.com

[1] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (KLEF),, Green Fields, Vaddeswaram, Guntur District, Vijayawada, Andhra Pradesh, India

disorder [20], epilepsy [19] requires multiple biosignals like Electrocardiogram (ECG), electromyogram (EMG), electroencephalogram (EEG) for the diagnosis and treatment. And it requires long-term recordings of these biosignals to prevent such chronic disease patients from the sudden death [21, 22] by predicting the symptoms earlier. E-health IoT scenario promotes long-term monitoring using micro-sensors and wearable micro-hub without affecting their daily activities. This sensor based wearable technology provides sophisticated environment to the patients and helps them to get treated at their door steps. These micro-sensor nodes are implanted or mounted on human body to collect the particular biosignal and forward to wearable micro-hub or sink thereby establishing the body area communications depicted in Fig. 1.

Though the sensor nodes can only communicate within a short distance due to the limited transmission range and power [37], the wearable micro-hub which is enabled with internet technology collects the readings from the sensor nodes and send to the observer. This micro-hub [4] is a resource constrained battery relayed device that collect, process and transmit the signal throughout the day. This prolonged transmission may increase the signal transmission cost and declines the battery life of wearables. This paper focuses on compressing the signal before transmission at sender side i.e., at edge level to reduce the transmission energy and thereby boosting battery life of wearables. The objective of this work is to jointly compress the multiple biosignals and reconstructing the original data from its compressed form with minimal loss. In case of consuming multiple sensors, compressing signal individually (unimodal compression) may consume more time and energy. This paper proves that multimodal compression consumes less energy than unimodal compression.

Most of the existing work has been concentrated on the unimodal compression of a biosignal using two emerging paradigms Compressive Sensing (CS) [12], and Auto encoders (AE) [13]. Recently, some of the researchers have concentrated on multimodal compression of multiple biosignals using CS algorithms and AE. CS is an efficient algorithm to jointly compress, but it is computationally intensive which is not suitable for lightweight devices and yields low compression ratio (CR) which cannot effectively reduce the transmission cost.
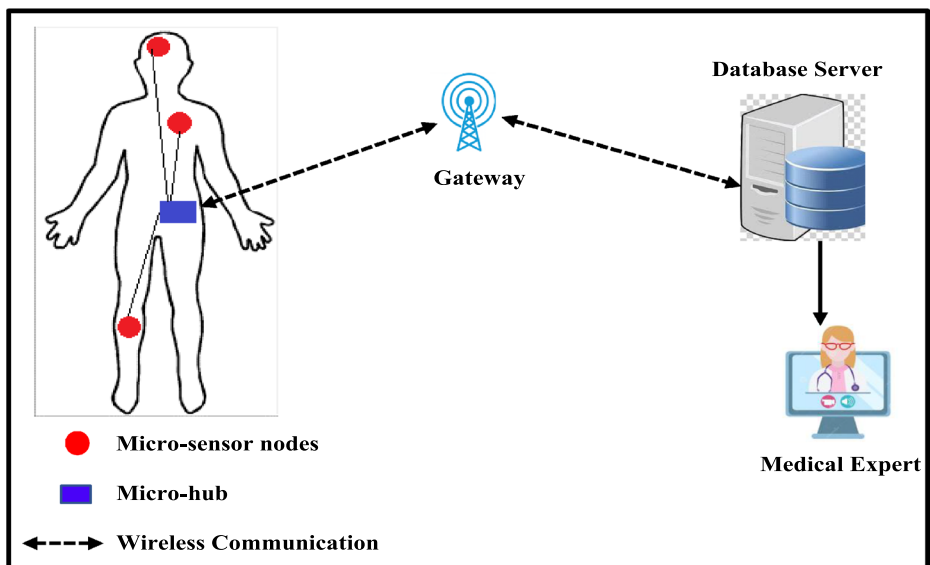

Fig. 1 Schematic Diagram of Sensor Based Wearable System acquiring multiple biosignals

Nevertheless, many of the past works [6, 33] have shown successive results with CS Singular value decomposition (SVD) and American standard code for information interchange (ASCII) character encoding-based algorithm [26, 30, 34] is a promising technique to jointly compress the multiple biosignals. But this CS and SVD-ASCII algorithms requires preprocessing of the biosignal to remove the noises as these biosignals are always prone to be contaminated with noises. Compression of multidimensional biomedical signals with spatial and temporal Codebook-Excited Linear Prediction [9] has achieved good CR of 89% with better reconstruction quality but the wearables are memory constrained device and it could not meet the growing dictionary size when it meets with new signals. However, the above discussed algorithms are good at compression and reconstruction quality, but they lack in focusing on light weight algorithm for these resource constrained devices.

Denoising Autoencoder (DAE) has been proved to be an efficient algorithm [27, 23] to compress the data with high CR at least computational cost and thereby boosting the battery life of the wearables. As DAE learns deep representation of the data and compress with high CR, multimodal stacked denoising autoencoder (SDAE) [28, 38] is proposed for joint compression and it is most suitable for light weight wearable devices. Because, the deep neural network is trained offline and not in the wearable devices. And once it achieves good performance, only the optimal configurations [1] called weights and biases are applied to the device for data processing. The trained network will not grow as in the case of codebook compression scheme when it meets new incoming data. Previously learned optimized configuration is much enough to reconstruct the unknown data and it is proven experimentally. Moreover, DAE has the ability to denoise signal simultaneously while reconstructing the data from its compressed form so DAE does not require any preprocessing step.

However, in case of deep neural network the weights are randomly initialized and optimized towards the target value while training, which may increase the training time and may result in poor performance. Transfer learning (TL) [29, 18] may address this issue, it learns the weights of the samples in source domain and transfers the learned parameters to train the target domains. So, combining TL with deep neural network may lessen the training time and improves the performance of the system as it uses the previously learned weights. And most of the researchers have succeeded with convolutional auto encoder (CAE) for unimodal compression and it is promising for signal reconstruction [17, 36]. So, the TL based multimodal convolutional denoising autoencoder (M-CDAE) is proposed in this work. Finally it is experimentally proved that to achieve better reconstruction quality, TL based M-CDAE model is better than stand-alone (SA) based (M-CDAE) model i.e. without TL. Also, it is clearly shown that it is better to do multimodal compression than unimodal compression in case of consuming multiple sensors to utilize the energy efficiently. The remainder of this paper is organized as follows. The basic framework of TL and M-CDAE is introduced in Section 2. The proposed methodology M-CDAE with TL is explained in Section 3. The experimental results obtained from appropriate database are presented in Section 4. The discussion of this experiment is carried out in Section 5. Finally, the work is concluded in Section 6.

2 Methodology

AE is an ideal methodology for non-linear dimensionality reduction, which typically trains and test the model with dataset from same scenario. So that the knowledge learned by the model

will be limited to that same range of distribution. As this work is a real time scenario the data range may vary from patients to patients. So, it is necessary to learn different distribution of data which may increase the dataset size for learning. This may result in time consuming for training the model, also the randomly initialized weights while training the data may yield poor performance. TL can help the model to learn different domains by transferring the knowledge from the previously learned source domain to the target domain. Recent studies made on TL based AE has achieved outstanding performance by knowledge transferring, in learning a low dimensional representation for maintenance prediction in intelligent transportation system [14], learning a set of hierarchical non linear transformations for cross-domain visual recognition [18],real-time production state prediction in manufacturing industry [30, 32].To the best of our knowledge, this is the first study based on TL with CDAE for joint compression of multiple biosignals and to effectively reconstruct the signal from its compressed form.

2.1 Transfer learning

TL overcomes the problem of random weight initialization while training AE, instead the training will be initialized by previously learned weights from source domain. This may train the model for huge dataset without storing the source data [3]. And it promotes adaptation of trained model for data across various range of distribution. In general, TL is categorized into two types namely instance-based and feature-based [18]. In case of instance-based only the weights are learned and transferred from source domain to target domain. And in case of feature-based the information learned from the source domain is transferred to the target domain. Feature-based transformation are mostly appropriate for classification purpose [7, 11]. For data reconstruction it is better to apply instance-based learning i.e. only weight transfer. By doing so, the model can be quickly optimized towards target value which consumes less training time. The structure of AE with TL is given in Fig. 2. Three steps are involved in Instance-based TL, that are Pre-training Source Data, Parameters Transferring, Fine-Tuning Target Data. These three steps are described as follows:

1) Pre-Training Source Data: In this step, AE is trained for source dataset to determine the optimal weight W_n^{ts} and bias b_n^{ts}. Where n is a number of layers in the AE model, W_n^{ts} & b_n^{ts} are the weight and bias trained for source dataset.

2) Parameters Transferring: Following the pre-training step, the learned parameters weight and bias W_n^{ts} & b_n^{ts} are transferred to the new AE model as the initial weight.

3) Fine-Tuning model: The transferred weight W_n^{tt} and bias b_n^{tt} are used as an initial weight for new AE model or for fine-tuning the model. Where W_n^{tt} & b_n^{tt} are the weight and bias trained for target dataset.

2.2 Multimodal autoencoder (M-AE)

The objective of this work is to observe a long-term recording of multiple biosignals (ECG, EMG, EEG) acquired through multiple biosensors mounted on human body. Figure 3. depicts the architecture of deep M-AE. It uses different pathways to jointly compress the multiple inputs to low dimensional vector called single shared
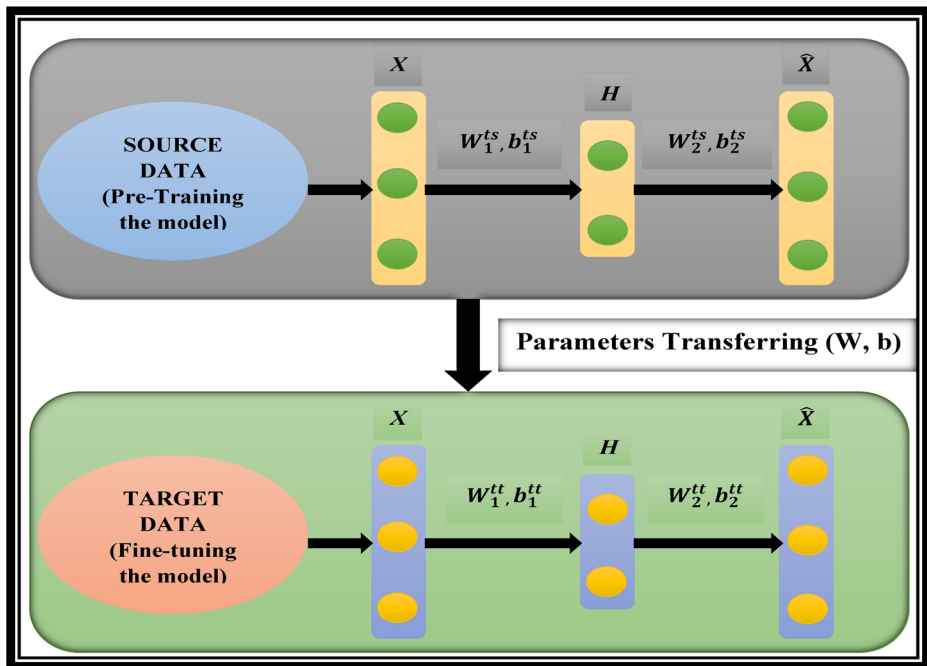
Fig. 2 Architecture of TL based AE

representations. This shared representation will be passed to the decoder (observer or Physician)in order to reconstruct the original signal, form its compressed form. Also, this single shared representation will be propagated to higher level features independently to reconstruct the original clean input. Once the optimal parameters are obtained through training, then it is configured [1] with wearable micro-hub for effective delivery of data. This multimodal compression is efficient than unimodal compression in terms of computation cost. Because deep M-AE simultaneously compress the multiple signals into a single shared representation before transmission which is processed separately in unimodal AE.

The deep M-AE is illustrated in Fig. 3. which obtains the multiple signals as input and compress to a single shared representation and then it is decompressed to original signal throughindependent pathways. The joint shared representation obtained by the encoder is as follows:

$$h = \sum_{i \in \{c,m,e\}} \tanh\left(x_i^{l-1} * W_i^l + b_i^l\right) \tag{1}$$

where c, m, e represents input ECG, EMG, EEG data respectively and l represents the no. of layers.$x_i^{l-1}$is the input from previous layer l. The decoded signals $\widehat{x}$ are represented as follows:

$$\widehat{x} = \left\{ \widehat{x}_{i \in c}^l = \tanh\left(h_i^{l-1} * \widehat{W}_i^l + \widehat{b}_i^l\right), \quad \widehat{x}_{i \in m}^l = \tanh\left(h_i^{l-1} * \widehat{W}_i^l + \widehat{b}_i^l\right), \quad \widehat{x}_{i \in e}^l = \tanh\left(h_i^{l-1} * \widehat{W}_i^l + \widehat{b}_i^l\right) \right\}$$
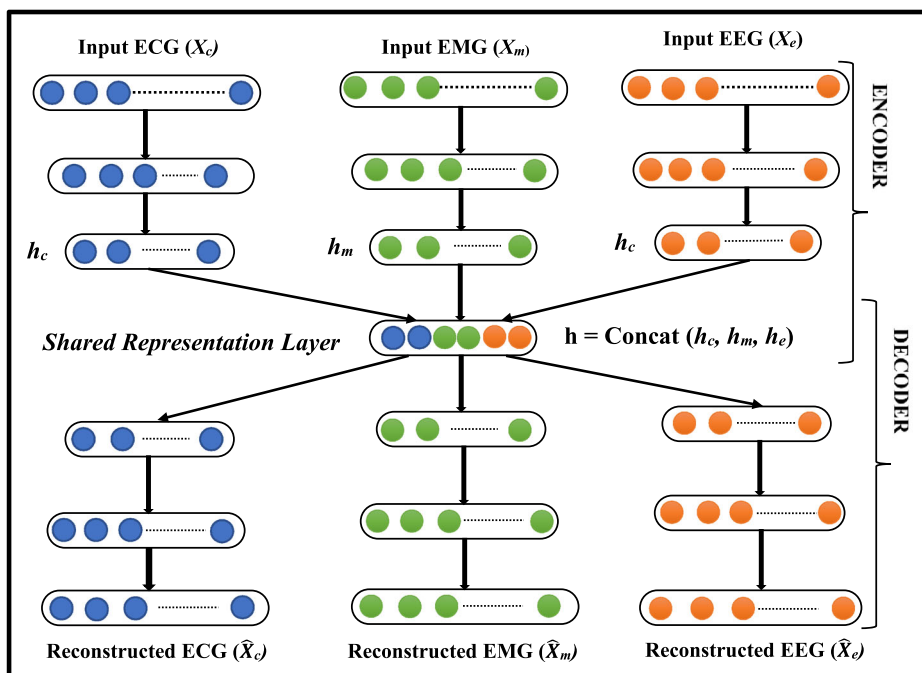
$$\tag{2}$$

Fig. 3 Schematic Representation of Deep M-AE

Where $\widehat{x}$ consist of set of multiple reconstructed signals and h_i^{l-1} represents the input from previous hidden layer l. where W_i^l denotes the weight for current layer l and b_i^l represents a bias for current layer l. Andtanh (•) is the Hyperbolic tangent ($\frac{2}{1+e^{-2z}}-1$) activation function applied for both encoder and decoder section.

3 Proposed methodology

This work proposes multimodal biosignal compression with TL based CDAE (M-CDAE). Recent work on TL based CAE in non-linear dimensionality reduction [24] for image processing in remote sensing, a cross-domain feature learning scheme for target objects recognition [7] in intelligent urban construction has been succeeded. In this work the idea behind to use TL is to reuse the learned weights from the pre-trained model (source data) for fine-tuning the target data. This may reduce the training and inference cost and increases the signal reconstruction quality which is proven experimentally.

And TL may increase the adaptability of the proposed model for the data across various distribution is proven by reusing the weight which is optimized for different dataset. Also, the comparisons are made between TL model and stand-alone (SA) model. The objective of this work is to efficiently compress the multiple biosignals (ECG, EMG, EEG) and to reconstruct from its compressed form. Moreover, these biosignals are prone to be contaminated with noise due to the body movement, so noise removal should be considered. As, DAE architecture is
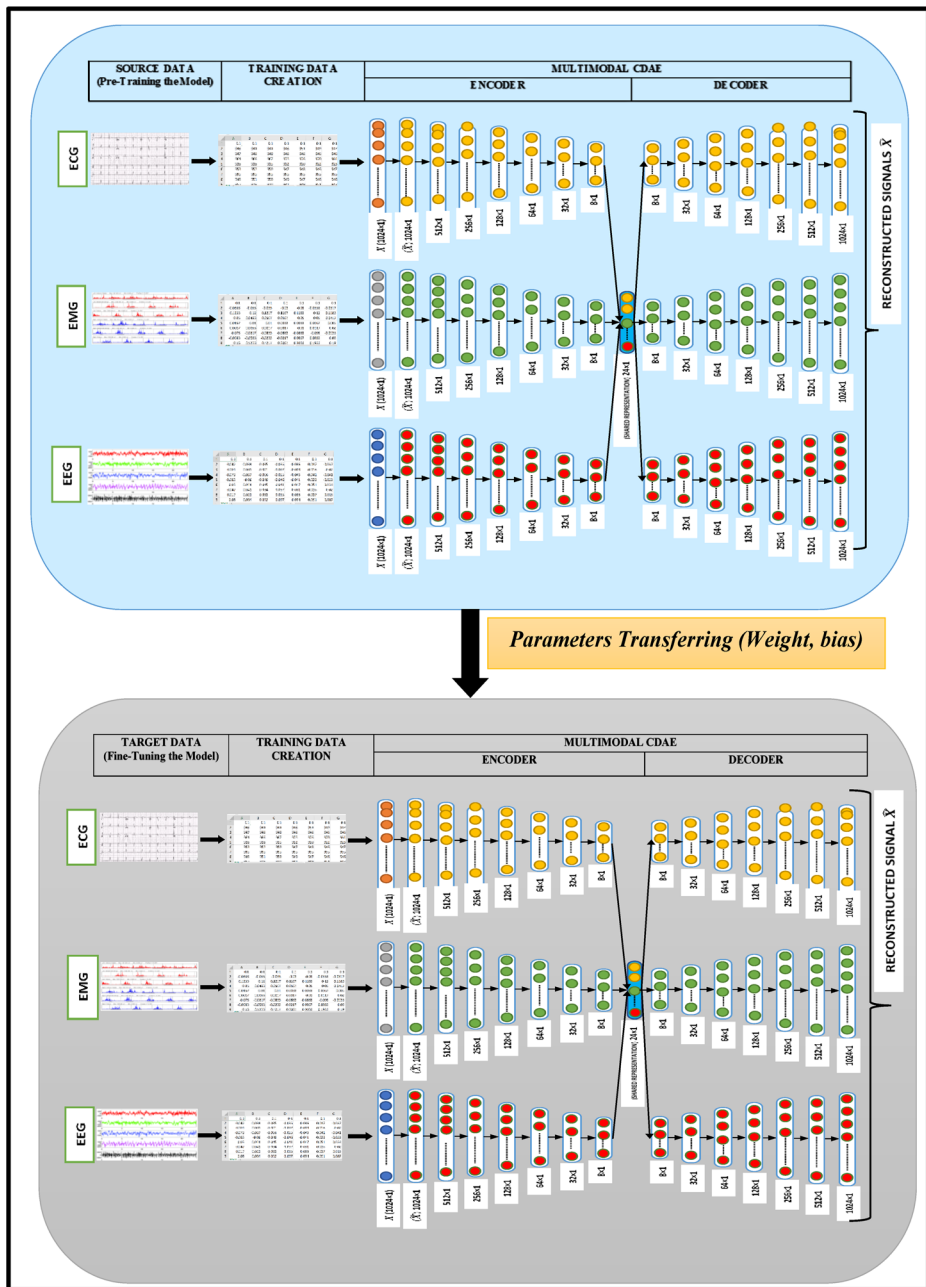
Fig. 4 Framework for the Proposed TL based M-CDAE model

employed in the proposed work the signals were simultaneously recovered from its noisy form in the decoder section.

The overall framework of proposed method is depicted schematically in the Fig. 4. and it consists of three pathways for input ECG, EMG, EEG signal. Before processing

the signals, they are additively corrupted by random gaussian white noise at specified SNR levels of −1, 0, 5 dB. Then they are passed into the model simultaneously each with the size of 1024 × 1. The proposed model employs CDAE architecture to reduce the redundant information and to learn the locally-spatial information efficiently by convolutional filters. And maxpool layers which helps to reduce the input size, hence the signal of size 1024 × 1 is reduced to the size of 8 × 1 and thereby achieving CR of 128. The CDAE architecture employed for these three signals will have same configurations to reveal its multi-modal property and the structure of CDAE is depicted clearly in Fig. 5 and its detailed parameters are given in Table 1. Thus, the encoder section efficiently compresses the signals independently but simultaneously. And the encoded data is concatenated finally using concatenate function before sending to the decoder section in the representation layer.

The decoder section receives the single shared representation of encoded data of three signals (ECG, EMG, EEG) with size of 24 × 1. This shared representation has been fragmented again to three inputs each with size of 8 × 1 for decoding. Then the fragmented input has its own pathway for decompression. The compressed input 8 × 1 is reconstructed to 1024 × 1 by Upsampling function which efficiently enhance the data and convolutional filters. Finally, the up sampled data is then fed to the fully connected dense layer which learns the high-levelfeatures to recover the original clean ECG signal. While training the model, the output produced from each dense unit is compared with original uncorrupted version. The gradient loss function optimizes the weight and bias towards original clean signal. Thus, the original clean signal is reconstructed independently but simultaneously. This simultaneous process may save computation time than separate unimodal compression and it is explained clearly in the experimental results. Hyperbolic tangent function is employed as an activation function for throughout the model. And the encoder section is equipped with Batch Normalization (BN) layer and dropouttechnique with the rate of 0.5 to avoid overfitting.

For training the proposed CDAE model, mean squared error (MSE) is employed to update the model's parameter $\theta = \left\{ W, b, \widehat{W}, \widehat{b} \right\}$ by optimizing the following objective function,

$$L(\theta) = \Sigma \left|\left| x - \widehat{x} \right|\right|_2^2 \tag{3}$$

where W, b are weight and bias in encoder section and $\widehat{W}, \widehat{b}$ are weight and bias in decoder section. It is important to note that the Eq. (3) propagate towards reconstructing the uncorrupted version of the corresponding corrupted input data.

4 Performance evaluation

4.1 Performance evaluation criteria

Quality Score (QS), Compression ratio (CR), root mean square error (RMSE), percentagerootmean square difference (PRD) and signal to noise ratio (SNR_{imp}) is an
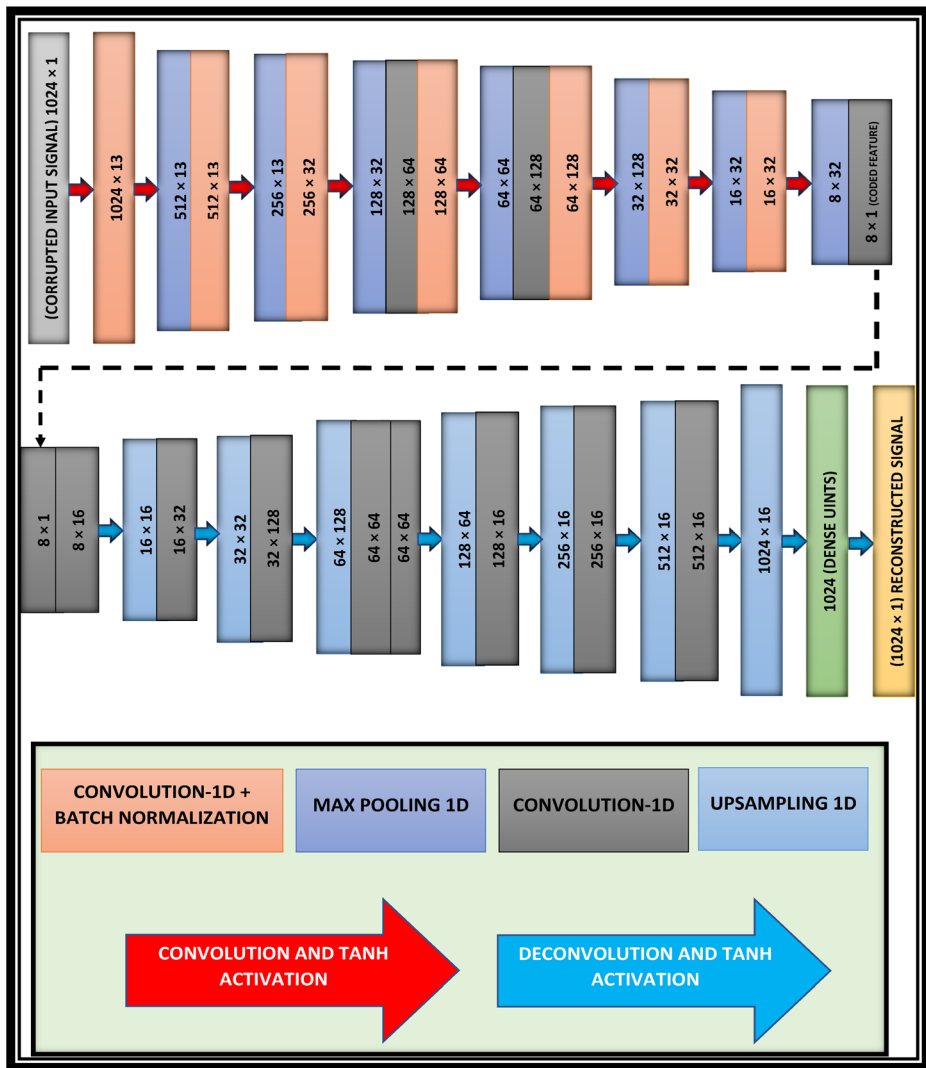
Fig. 5 Detailed Structure of Proposed CDAE Model

ideal criteria for performance evaluation for compression technique and it is defined as follows:

The algorithm's compression efficiency and its reconstruction quality are determined by QS and it is the ratio of CR and PRD which represents the reconstructed signal quality as given by (4).

$$QS = \frac{CR}{PRD} \qquad (4)$$

CR is defined as the ratio between size of original signal s_o and the size of compressed signal s_c. CR will be high for effective algorithm as given by 5.

Table 1 Structural Information of Proposed CDAE model

S. No.	Layer Name	No. of Filter × Kernel Size	Region / Unit Size	Activation Function	Output Size
ENCODER SECTION					
1	Input Layer	–	–	–	1024×1
2	Conv1D+BN	13×6	–	Tanh	1024×13
3	MaxPooling1D	–	2	–	512×13
4	Conv1D+BN	13×7	–	Tanh	512×13
5	MaxPooling1D	–	2	–	256×13
6	Conv1D+BN	32×7	–	Tanh	256×32
7	MaxPooling1D	–	2	–	128×32
8	Conv1D	64×7	–	Tanh	128×64
9	Conv1D+BN	64×7	–	Tanh	128×64
10	MaxPooling1D	–	2	–	64×64
11	Conv1D	128×13	–	Tanh	64×128
12	Conv1D+BN	128×13	–	Tanh	64×128
13	MaxPooling1D	–	2	–	32×128
14	Conv1D+BN	32×8	–	Tanh	32×32
15	MaxPooling1D	–	2	–	16×32
16	Conv1D+BN	32×8	–	Tanh	16×32
17	MaxPooling1D	–	2	–	8×32
18	Conv1D	1×8	–	Tanh	8×1
DECODER SECTION					
19	Conv1D	1×8	–	Tanh	8×1
20	Conv1D	16×8	–	Tanh	8×16
21	UpSampling1D	–	2	–	16×16
22	Conv1D	32×8	–	Tanh	16×32
23	UpSampling1D	–	2	–	32×32
24	Conv1D	128×13	–	Tanh	32×128
25	Conv1D	128×13	–	Tanh	32×128
26	UpSampling1D	–	2	–	64×128
27	Conv1D	64×7	–	Tanh	64×64
28	Conv1D	64×7	–	Tanh	64×64
29	UpSampling1D	–	2	–	128×64
30	Conv1D	16×8	–	Tanh	128×16
31	UpSampling1D	–	2	–	256×16
32	Conv1D	16×8	–	Tanh	256×16
33	UpSampling1D	–	2	–	512×16
34	Conv1D	16×8	–	Tanh	512×16
35	UpSampling1D	–	2	–	1024×16
36	Dense	–	1024	Tanh	1024×1

$$CR = \frac{s_o}{s_c} \tag{5}$$

The variance between the original signal x_i and the reconstructed signal $\widehat{x}_i$ is determined by RMSE. N is the total length of the signal. Smaller the value of RMSE represents the good performance and smaller variations. It is defined as (6):

$$RMSE = \sqrt{\frac{1}{N} \times \Sigma_{n=1}^{N} \left(x_i - \widehat{x}_i \right)^2} \tag{6}$$

PRD is given by (7) which shows the quality of recovered signal from compressed version by measuring the error between original signal and the reconstructed signal. A lower PRD value represents the better quality of the reconstructed signal.

$$PRD = \sqrt{\frac{\sum_{n=1}^{N}\left(x_i - \widehat{x_i}\right)^2}{\sum_{n=1}^{N} x_i^2}} \times 100 \tag{7}$$

SNR$_{imp}$is applied to observe how well the signal is denoised it is given by difference between SNR after noise reduction and the original input signal SNR. The greater the SNR$_{imp}$ is, the better denoising performance. SNR$_{imp}$ is formulated as follows:

$$SNR_{imp} = SNR_{out} - SNR_{in} \tag{8}$$

Where SNR_{in} and $SNR_{out}$is formulated as follows, here $\widetilde{x}_i$ is a corrupted input signal.

$$SNR_{in} = 10 \times log_{10}\left(\frac{\sum_{n=1}^{N}(x_i)^2}{\sum_{n=1}^{N}\left(\widetilde{x}_i - x_i\right)^2}\right) \tag{9}$$

$$SNR_{out} = 10 \times log_{10}\left(\frac{\sum_{n=1}^{N}(x_i)^2}{\sum_{n=1}^{N}\left(\widehat{x}_i - x_i\right)^2}\right) \tag{10}$$

4.2 Experimental data

This work implements three biosignals namely ECG, EMG, EEG which has been taken form standard databases and it is explained as follows. ECG signals used in this study is derived from the MIT Arrhythmia database [24] which are sampled at rate of 360 samples/s with 11-bit resolution over a 10-mV range. The MIT-BIH Arrhythmia Database contains 48 half-hour excerpts of two-channel (lead I & lead II) ambulatory ECG recordings, obtained from 47 subjects. Since the wearable sensors produces 1-D time-series data, this work focusses on the data of one of the electrodes lead II. The EEG data used in this study is taken from Bonn dataset [5] which is sampled at 173.6 Hz with 12-bit resolution. And it is composed of five sets of datawhich is denoted by Z, O, N, F & S, and each contains 100single-channel EEG segments. EMG signals recordings, which were recorded at 50 KHz and then down sampled to 4 KHz and exhibit sparsity in both the time and frequency domains, were collected from the Physiobank [15] for healthy, myopathy and neuropathy patients.

The idea behind the work is to transfer the weight learned in source data (pre-training model) to the target data (fine-tuning model) as initial weight instead of random weight initialization. This initial weight transfer may reduce the computation burden and make the model most adaptable for different distributions of data. For experiments, the dataset is divided into two parts source data and target data. And it is clearly charted in Table 2. different records are taken for source data (pre-training model) and target data (fine-tuning model) with equal data size of 100 fragments and each fragment with 1024 samples. The dataset (ECG, EMG, EEG) is split into 80% and 20% for training and testing respectively for both source and target dataset.

Table 2 Data Separation for the proposed TL based M-CDAE model

BIOSIGNAL	DATABASE	DATA SEPARATION	RECORD No. (or) Name	No. of Records × No. of Fragments	No. of Samples/ Fragment	TRAINING SIZE (%)	TESTING SIZE (%)
ECG	MIT Arrhythmia database	SOURCE DATA	100	(1×100)	1024	80	20
		TARGET DATA	117, 119	(2×50)	1024	80	20
EMG	Bonn dataset	SOURCE DATA	Myopathy	(1×100)	1024	80	20
		TARGET DATA	Healthy, Neuropathy	(2×50)	1024	80	20
EEG	Physiobank	SOURCE DATA	F	(1×100)	1024	80	20
		TARGET DATA	(Z, O, N, S)	(4×25)	1024	80	20

4.3 Experimental results

As shown in Table 2, the biosignals ECG, EMG, EEG are taken from particular dataset and split into training and testing set. Table 3 shows the quantitative analysis of the test scores of RMSE, PRD & SNR_{imp} for SA based M-CDAE model and TL based M-CDAE model overtarget dataset at specified input SNR noise level of (−1, 0, 5) dB. From the readings, it is observed that TL based M-CDAE model yields better result than SA based M-CDAE model for all the input SNR levels (−1, 0, 5) dB. Also, it is observed that as the input SNR level increases performance of the system degrades. The increased SNR levels increases the RMSE & PRD value and decreases the SNR_{imp} value.

Figure 6(a–c) shows the test QS for SL based M-CDAE model and proposed TL based M-CDAE model over target dataset at specified input SNR noise level of (−1, 0, 5) dB. For SL based M-CDAE model, the QS of ECG is 2226 when SNR is −1 dB and it is decreased to 1905 when SNR level is 5 dB. Similarly, for EMG & EEG when SNR is −1 dB the QS is 219& 124 respectively and when the SNR level increased to 5 dB the QS will be 116.4 & 124 respectively. These measurements are slightly improved through proposed TL based M-CDAE model but undergoing the same issue, as SNR level increases the performance degrades. For TL model, the QS of ECG is 2298 when SNR is −1 dB and it is decreased to 2370 when SNR level is 5 dB. Similarly, for EMG & EEG when SNR is −1 dB the QS is 305& 128 respectively and when the SNR level increased to 5 dB the QS will be 128& 116 respectively.

Figure 7 represents the overall graphical visualization of the waveform of ECG, EMG, EEG biosignals for proposed TL based M-CDAE model. And Fig. 8(a–f) clearly shows the comparison results of clean and reconstructed output of ECG, EMG, EEG biosignals produced by SA based M-CDAE model and TL based M-CDAE model over target dataset as given in Table 2. According to the tabulated results and graphical results, TL based M-CDAE model outperforms in all input SNR level of (−1, 0, 5) dB.

Table 3 TL based M-CDAE vs SA based M-CDAE model with CR of 128

| MODEL TYPE	Input SNR (dB)	BIOSIGNALS	EVALUATION CRITERIA			
			RMSE	PRD	SNR_{imp}	QS
STANDALONE LEARNING MODEL	−1	ECG	0.0122	0.0575	25	2226
		EMG	0.1154	0.584	4.54	219
		EEG	0.167	1.03	4.7	124
	0	ECG	0.0126	0.06	24.7	2133
		EMG	0.234	1.1	2.1	116
		EEG	0.22	1.04	2.9	123
	5	ECG	0.013	0.0672	20.2	1905
		EMG	0.244	1.1	1.1	116
		EEG	0.159	1.4	1.4	91.4
TRANSFER LEARNING MODEL	−1	ECG	0.0119	0.05	26.3	2560
		EMG	0.0830	0.42	7.4	305
		EEG	0.153	1	5	128
	0	ECG	0.012	0.055	25.12	2327
		EMG	0.2	1	4.2	128
		EEG	0.158	1.01	4	127
	5	ECG	0.0125	0.063	25	2031
		EMG	0.23	1	2.14	128
		EEG	0.16	1.1	3.89	116
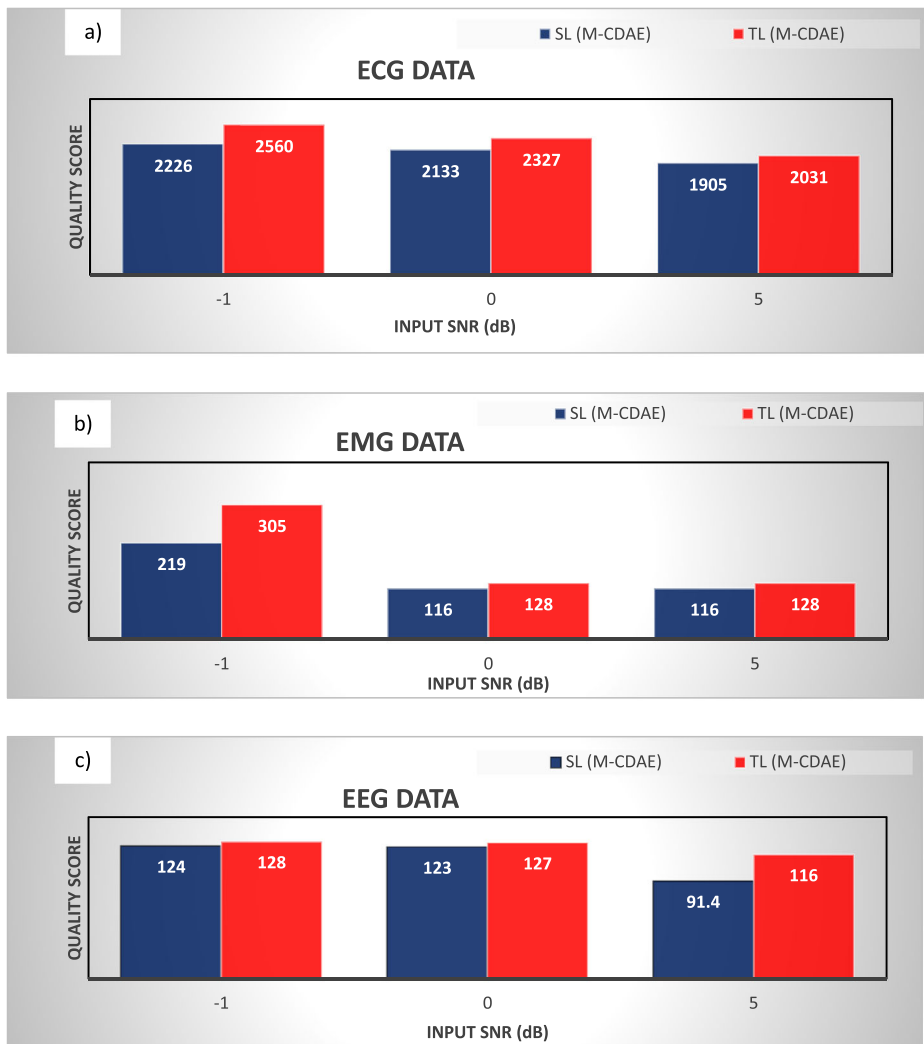
Fig. 6 Comparison of the Quality Scores for TL based M-CDAE and SA based M-CDAE model of all biosignals (ECG, EMG, EEG) at different input SNR levels of (−1, 0, 5) dB (**a**) Quality Scores of ECG (**b**) Quality Scores of EMG (**c**) Quality Scores of EEG

Energy efficiency is not a new issue but it is a major issue in wireless communication. Since the wearable hub which is responsible for acquiring and transmitting the multiple biosignal is miniaturized and battery relayed, the runtime of signal processing algorithm should be considered. This work focusses on jointly compressing the multiple biosignal before transmission, to reduce the transmission cost. In case of acquiring and transmitting multiple biosignals, it is better to do multimodal compression than unimodal compression. Because the training and testing (inference) time of multimodal compression is lower than unimodal compression as shown in Table 4 for highly corrupted input signal at SNR level of 5 dB.
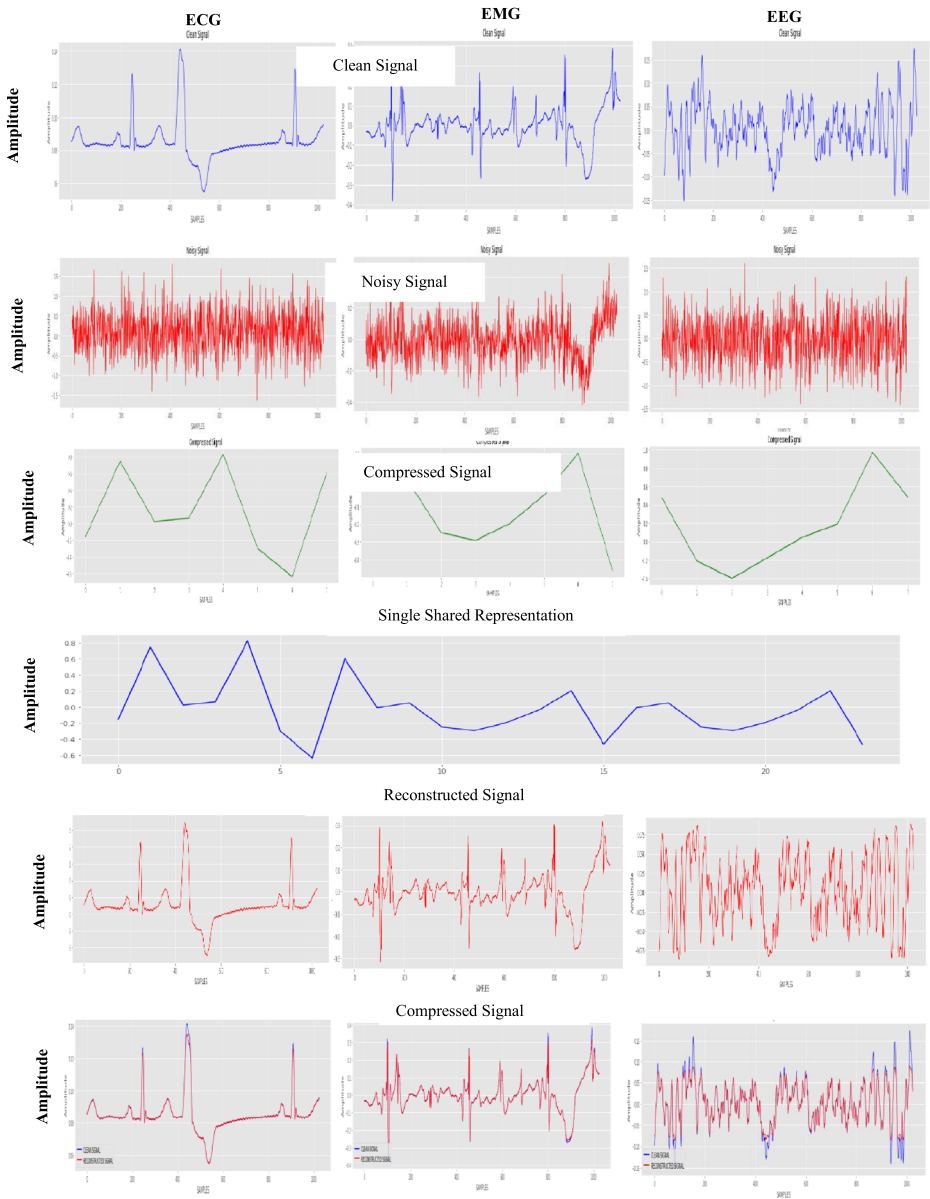
Fig. 7 Overall graphical representation of reconstructed results of ECG, EMG, EEG biosignals for proposed TL based M-CDAE model. X axis - Samples, Y axis - Amplitude

5 Discussion

Multimodal compression has been carried out by most of the studies [4, 12, 22, 37] but this is the first study to propose TL based multimodal compression formultiple biosignals ECG, EMG, EEG. And each signal is taken from standard databases. This work performs four type of experimentations with the proposed CDAE model are SA based unimodal compression, SA based multimodal compression, TL based unimodal compression and TL based multimodal
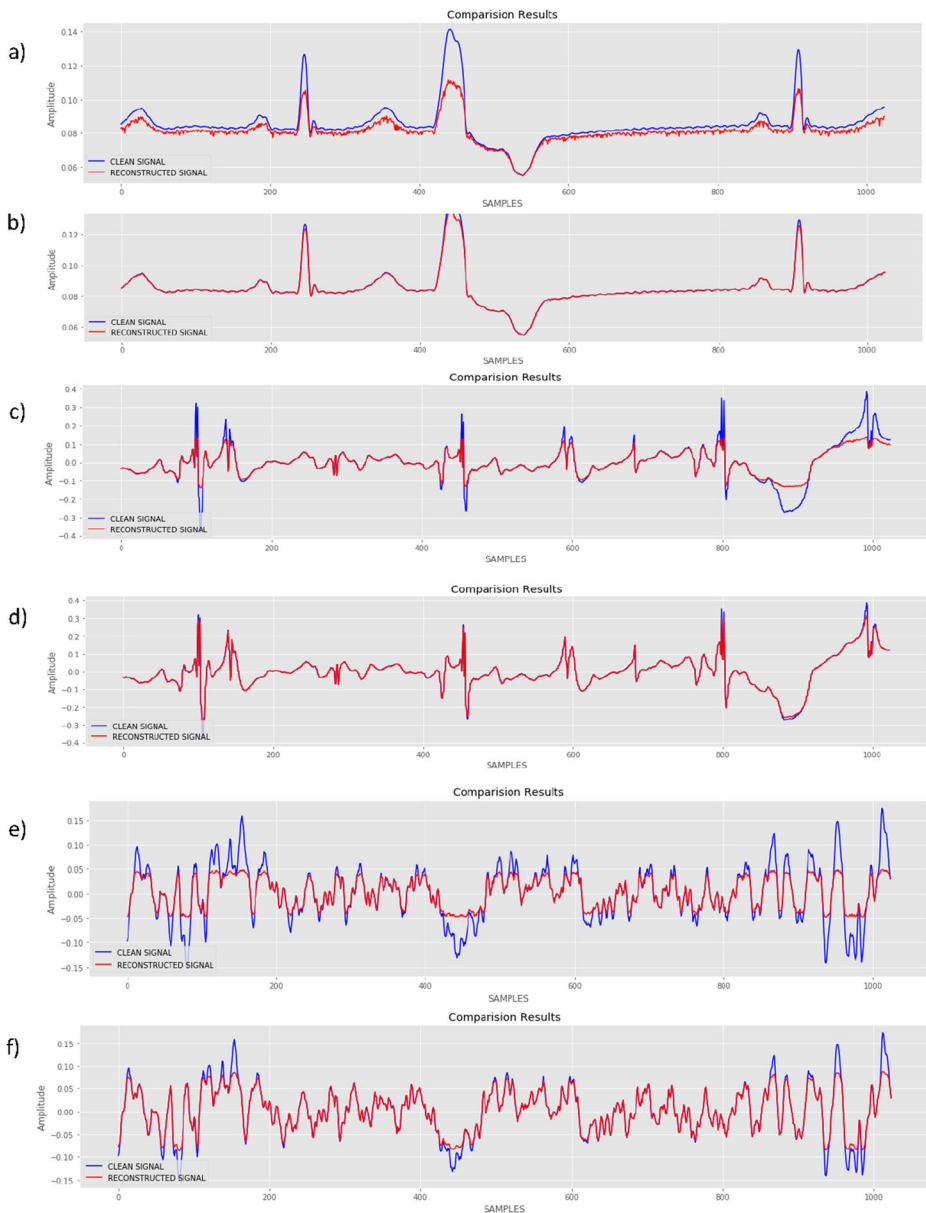
Fig. 8 Comparison results of clean and reconstructed biosignals over target dataset with and without TL. **a**) ECG signal reconstructed by SA based M-CDAE model **b**) ECG signal reconstructed by TA based M-CDAE model **c**) EMG signal reconstructed by SA based M-CDAE model **d**) EMG signal reconstructed by TA based M-CDAE model **e**) EEG signal reconstructed by SA based M-CDAE model **f**) EEG signal reconstructed by TA based M-CDAE model

compression. The experimental result proves that TL based multimodal compression will be better in terms of computation cost and QS. The reason behind the successive result is, in transfer learning model the learned weight is reused so that the system can soon propagates towards the optimal configurations (weights and bias). Also, while compressing the multiple

Table 4 Stand-Alone based CDAE model vs. Transfer Learning based CDAE model for unimodal and multimodal compression in terms of computation cost at the input SNR level of 5 dB

COMPRESSION TYPE	BIOSIGNAL	STAND-ALONE CDAE Model		TRANSFER LEARNING CDAE Model	
		COMPUTATION TIME (seconds)		COMPUTATION TIME (seconds)	
		Training cost	Inference cost	Training cost	Inference cost
UNIMODAL	ECG	5482.35	1.15	4961.86	0.0418
	EMG	6158.89	1.4	5117.86	0.323
	EEG	3923.38	0.7319	3614.98	0.455
TOTAL		15,564.62	3.28	13,694.7	0.8198
MULTIMODAL	ECG-EMG-EEG	8363.3	0.7	**8211.92**	**0.1007**

signals simultaneously the training time and inference cost is reduced further as shown Table 4.Most of the work achieved great success in signal reconstruction using CDAE so this work employs improved CDAE architectureand achieves CR of 128.

The dataset used for experiments has been separated into source and target dataset and the proposed model is pre-trained on source dataset. The learned weights and bias from the source dataset are reused for training target dataset. In neural network, rather than initializing weights randomly it is better to reuse thelearned weights which may reduce the training cost and inference cost. Also,by transferring the parameters like weight and biasthe model is trained for large data without storing the source data. The experimental results produced here arethe test score over target dataset.

Moreover, this biosignals are always prone to be contaminated with noise due to the body movement. So, the input signal is additively corrupted by random gaussian white noise at various noise levels (−1, 0, 5) dB. DAE architecture is utilized with the proposed work which efficiently denoises the signal simultaneously while reconstructing from its compressed form. This is done by propagating loss function of the model towards original clean signal of the corrupted version while training. Therefore, the proposed work can simultaneously denoise signal while reconstructing from its compressed form.

6 Conclusion

Wearable technology enriches the long-term monitoring system through wearables and wireless communication. This wireless communication always suffers from signal transmission cost due to acquiring and sending multiple signals throughout the day.So, the signals can be compressed before sending at edge level i.e. where they are acquired. The wearable hub which is responsible for acquiring and transmitting signals is miniaturized in nature and resource constrained. So, the designed algorithm should be light weight at the same it should be efficient to reconstruct the signal from its compressed form. This paper employs TL based M-CDAE to compress the signal efficiently. In neural network, rather than initializing the weight randomly it is better to reuse the learned weights so that the model performs well for data across various range of distribution. The experimental result shows that the proposed model consumes less training and inference cost when performing multimodal compression and gives better QS with CR of 128 with TL based model. Moreover, the proposed system

reveals its multimodality property by employing same CDAE architecture for reconstructing three biosignals ECG, EMG, EEG. As this work focusing on real time application the future work can be carried out for on-device compression so that the system can perform well for any new incoming unknown data without offline training.

Declarations

Conflict of Interest and Authorship Conformation Form.

Please check the following as appropriate:

All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.

This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.

The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

The following authors have affiliations with organizations with direct or indirect financial interest in the subject matter discussed in the manuscript:

References

1. Ahmed B, Amr M, Khaled H, Jane W (2017) Multimodal deep learning approach for joint EEG-EMG data compression and classification. IEEE Wirel Commun Netw Conf (WCNC). https://doi.org/10.1109/WCNC.2017.7925709
2. Akila T, Kumaragurubaran G (2021) An improved feature selection based classifier for prediction of different regions in Sar images. Multimed Tools Appl 80:33641–33662
3. Ali A, Mohammad A, Erik S, Kevin E (2020) A deep transfer learning approach to reducing the effect of electrode shift in EMG pattern recognition-based control. IEEE Trans Neural Syst Rehab Eng 28:370–379. https://doi.org/10.1109/TNSRE.2019.2962189
4. Amit S, Samaresh B, Sudip M (2015) Link quality-aware resource allocation with load balance in wireless body area networks 74–81. https://doi.org/10.1109/JSYST.2015.2458586
5. Andrzejak RG, Lehnertz K, Mormann F, Rieke C (2001) Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. Phys Rev E Stat Nonlin Soft Matter Phys 64:061907. https://doi.org/10.1103/physreve.64.061907
6. Anna MR, Dixon G, Allstot G (2012) Compressed sensing system considerations for ECG and EMG wireless biosensors. IEEE Trans Biomed Circ Syst 6:156–166. https://doi.org/10.1109/TBCAS.2012.2193668
7. Bing X, Zhemin D, Bin Z (2019) Research on target object recognition based on transfer-learning convolutional SAE in intelligent urban construction. IEEE Access 7:125357–125368. https://doi.org/10.1109/ACCESS.2019.2939284
8. Bowen D, Ronggui W, Juan Y, Lixia X (2021) Multi-scale feature self-enhancement network for few-shot learning. Multimed Tools Appl 80:33865–33883 57 access
9. Carotti ESG, Juan C (2019) Compression of multidimensional biomedical signals with spatial and temporal codebook-excited linear prediction. IEEE Trans Biomed Eng 56:2604–2610. https://doi.org/10.1109/TBME.2009.2027691
10. Chao S, Wen M, Zhang K, Ping M (2021) Traffic sign detection algorithm based on feature expression enhancement. Multimed Tools Appl 80:33593–33614
11. Chuang S, Meng M (2018) Deep transfer learning based on sparse autoencoder for remaining useful life prediction of tool in manufacturing. IEEE Trans Ind Inf 15:2416–2425. https://doi.org/10.1109/TII.2018.2881543
12. Craven D, Brian M, Liam K (2014) Compressed sensing for bioelectric signals: a review. IEEE J Biomed Health Inf 529–540. https://doi.org/10.1109/JBHI.2014.2327194
13. Del Testa D, Michele R (2015) Lightweight lossy compression of biometric patterns via denoising autoencoders. IEEE Signal Proc Lett 22:2304–2308. https://doi.org/10.1109/LSP.2015.2476667
14. Evaldas V, Matej U (2018) Learning low-dimensional representation of bivariate histogram data. IEEE Trans Intell Transp Syst 19:3723–3735. https://doi.org/10.1109/TITS.2018.2865103
15. Goldberger A (2000) Physiobank, PhysioToolkit, and PhysioNet:Components of a new research resource for complex physiological signals. Circulation 101:1–6. https://doi.org/10.1161/01.cir.101.23.e215

16. Hannah R, Kiri L, Wagstaff D, Patrick G, James F (2019) Toward generalized change detection on planetary surfaces with convolutional auto encoders and transfer learning. IEEE J Sel Top Appl Earth Observ Remote Sens 12:3900–3918. https://doi.org/10.1109/JSTARS.2019.2936771
17. Hsin-Tien C, Hsieh Y (2020) Noise reduction in ECG signals using fully convolutional denoisingautoencoders. IEEE Access 7:60806–60813. https://doi.org/10.1109/ACCESS.2019.2912036
18. Junlin H, Jiwen L (2016) Deep transfer metric learning. IEEE Trans Image Process 25:5576–5588. https://doi.org/10.1109/TIP.2016.2612827
19. Klosch G, Kemp B, Penzel A, Schlogl P (2001) The SIESTA project poly graphic and clinical database. IEEE Eng Med Biol 20:51–57. https://doi.org/10.1109/51.932725
20. Lang J, Ying Z, Xiao L, Wen J, Yong-Hong L, Zhao J (2017) Postictal apnea as an important mechanism for SUDEP: a near-SUDEP with continuous EEG-ECG-EMG recording. Elsevier J Clin Neurosci 43:1–3. https://doi.org/10.1016/j.jocn.2017.04.035
21. Liu S, Mengjing Y, Li M, Qingzhen X (2019) The research of virtual face based on deep convolutional generative adversarial networks using tensor flow. Phys A Stat Mech Appl 521(9):667–680
22. Marco T, Simone B (2015) Power line interference removal for high quality continuous bio-signal monitoring with low-power wearable devices. IEEE Sensors J 16:3887–3895. https://doi.org/10.1109/JSEN.2016.2536363
23. Mohsen H, Davide Z (2019) Boosting the battery life of wearables for health monitoring through the compressionof biosignals. IEEE Internet Things J 4:1647–1662. https://doi.org/10.1109/JIOT.2017.2689164
24. Moody GB, Mark RG, Goldberger AL (2001) Physionet: a web-based resource for the study of physiologic signals. IEEE Eng Med Biol Mag 20:70–75. https://doi.org/10.1109/51.932728
25. Neha S, Reecha S, Neeru J (2021) Prediction of face age progression with generative adversarial networks. Multimed Tools Appl 80:33911–33935
26. Qingzhen X (2013) A novel machine learning strategy based on two-dimensional numerical models in financial engineering. Math Probl Eng. https://doi.org/10.1155/2013/659809
27. Qingzhen X, Gvuangyi H, Mengjing Y, Yanliang G (2019) A novel edge-oriented framework for saliency detection enhancement, image and vision. Computing 87(1):1–12
28. Qingzhen X, Gvuangyi H, Mengjing Y, Yanliang G (2020) Fall prediction based on key points of human bones. Phys A Stat Mech Appl 532(5):123205
29. Qingzhen X, Jiayong W, Qiang C (2014) A novel mobile personalized recommended method based on money flow model for stock exchange. Math Probl Eng. https://doi.org/10.1155/2014/353910
30. Qingzhen X, Wang Z, Wang F, Yongyi G (2019) Multi-feature fusion CNNs for Drosophila embryo of interest detection. Phys A: Stat Mech Appl 531(19):121808
31. Senlin Y, Xin C (2021) Study on feature extraction technology of real-time video acquisition based on deep CNN. Multimed Tools Appl 80:33937–33950
32. Shaohua H, Yu G, Daoyuan L, Shanshan Z, Weiguang F (2019) A two-stage transfer learning-based deep learning approach for production Progress prediction in IoT-enabled manufacturing. IEEE Internet Things J 6:10627–10638. https://doi.org/10.1109/JIOT.2019.2940131
33. Singh A, Samarendra D (2017) Block sparsity-based joint compressed sensing recovery of multi-channel ECG signals. Healthc Technol Lett 4:50–56. https://doi.org/10.1049/htl.2016.0049
34. Sourav Kumar Mukhopadhyay M (2018) SVD and ASCII character encoding-based compression of multiple biosignals for remote healthcare systems. IEEE Trans Biomed Circ Syst 12:137–150. https://doi.org/10.1109/TBCAS.2017.2760298
35. Stanislas C, Galtier MN, Arnal PJ, Gilles W, Alexandre G (2018) A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. IEEE Trans Neural Syst Rehab Eng 26:758–769. https://doi.org/10.1109/TNSRE.2018.2813138
36. Wang F, Qiming M, Wenhan L (2020) A novel ECG signal compression method using spindle convolutional auto-encoder. ELSEVIER Comput Methods Programs Biomed 175:139–150. https://doi.org/10.1016/j.cmpb.2019.03.019
37. Xu X, Ning Z, Houbing S, Anfeng L, Ming Z (2018) Adaptive beaconing based MAC protocol for sensor based wearable system, 29700-29714. https://doi.org/10.1109/ACCESS.2018.2843762
38. Youshen C, Hanzhi Z (2020) Hybrid deep learning model assisted data compression and classification for efficient data delivery in mobile health applications. IEEE Access 8:94757–94766. https://doi.org/10.1109/ACCESS.2020.2995442