IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 6, 2018, accepted May 28, 2018, date of publication June 5, 2018, date of current version July 6, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2844308

A Deep Learning Approach for Vital Signs Compression and Energy Efficient Delivery in mhealth Systems

AHMED BEN SAID[1], (Member, IEEE), MOHAMMAD FATHI AL-SA'D[1], MOUNIRA TLILI[1], ALAA AWAD ABDELLATIF[1], AMR MOHAMED[1], (Senior Member, IEEE), TAREK ELFOULY[1], (Senior Member, IEEE), KHALED HARRAS[2], AND MARK DENNIS O'CONNOR[3]

[1]Department of Computer Science and Engineering, Qatar University, 2713 Doha, Qatar
[2]School of Computer Science, Carnegie Mellon University in Qatar, Education City, Doha, Qatar
[3]Mobile Healthcare Service, Ambulance Service, Hamad Medical Corporation, Doha, Qatar

Corresponding author: Amr Mohamed (amrm@ieee.org)

This work was supported by NPRP through the Qatar National Research Fund (a member of the Qatar Foundation) under Grant 7-684-1-127.

ABSTRACT Due to the increasing number of chronic disease patients, continuous health monitoring has become the top priority for health-care providers and has posed a major stimulus for the development of scalable and energy efficient mobile health systems. Collected data in such systems are highly critical and can be affected by wireless network conditions, which in return, motivates the need for a preprocessing stage that optimizes data delivery in an adaptive manner with respect to network dynamics. We present in this paper adaptive single and multiple modality data compression schemes based on deep learning approach, which consider acquired data characteristics and network dynamics for providing energy efficient data delivery. Results indicate that: 1) the proposed adaptive single modality compression scheme outperforms conventional compression methods by 13.24% and 43.75% reductions in distortion and processing time, respectively; 2) the proposed adaptive multiple modality compression further decreases the distortion by 3.71% and 72.37% when compared with the proposed single modality scheme and conventional methods through leveraging inter-modality correlations; and 3) adaptive multiple modality compression demonstrates its efficiency in terms of energy consumption, computational complexity, and responding to different network states. Hence, our approach is suitable for mobile health applications (mHealth), where the smart preprocessing of vital signs can enhance energy consumption, reduce storage, and cut down transmission delays to the mHealth cloud.

INDEX TERMS WBASN, multiple modality data, compression, deep learning, cross-layer optimization.

I. INTRODUCTION

The World Health Organization (WHO) classifies chronic diseases by having one or more of the following features: everlasting, cause disability, exist due to non-reversible pathological alteration, demand patient rehabilitation, or require a long period of supervision, observation, or care [1], [2]. Internationally, the number of Chronic Disease (CD) patients is increasing exponentially, more than 900 million individuals worldwide, and 80% of health-care spendings are reserved for CDs [3]. Chronic diseases, such as Epilepsy [4], Diabetes [5], Asthma [6], Alzheimer [7], etc., are widely spreading. According to WHO, approximately 50 million people worldwide have Epilepsy, making it one of the most common neurological diseases globally with nearly 80% of them living in low to middle-income countries [8]. People suffering from brain disorders require constant monitoring for their brain activities and a rapid notification system that is connected to a health-care provider in case of an emergency [9]. In order for those patients to be constantly monitored while living a normal life, acquisition, processing and transmission of their biomedical data such as Electroencephalography (EEG), Electromyogram (EMG) and Electrooculography (EOG) is required.

Biomedical data have become abundant, thanks to the development of wearable commercial devices, and variant by coming from multiple modalities. Even the most complicated medical follow-ups that require in-bed procedures are now possible with commercial devices. For example, EEG can be recorded using wearable, non-invasive, commercially available devices such as Emotiv headsets [10]. However, the processing and transmission of these signals, in particular, can be challenging, as high-quality monitoring can generate data at 1 Mbps, resulting in more than 80 GB of data everyday [11]. Hence, compression is a highly desirable pre-processing stage in order to optimize resources.

Motivated by the myriad of biomedical sensors, mobile phones, applications, and the rise of Internet of Things (IoT) [12] and Internet of Medical Things (IoMT) [13], scientific communities have standardized a system that focuses on the acquisition of vital signs via Wireless Body Area Sensor Networks (WBASN) under IEEE 802.15.6 [14]. A typical WBASN system consists of a wearable device that collects vital signs from the body, sends them to a Personal/Patient Data Aggregator (PDA) on which data can be processed and forwarded to a remote medical server to be stored for further analysis. Generally, data delivery is likely to be hindered due to mobile device and network resource constraints. In addition, network state is continuously changing due to wireless channel impairments, network congestion, patient mobility, etc. Therefore, for data compression to be efficient, it must be adaptive to network dynamics.

Following edge computing paradigm where we push the intelligence closer to the patient to optimize performance [15], we propose an adaptive data compression scheme exploiting Deep Learning (DL) approach for single and multiple modality data compression. This scheme is dynamically adapted to wireless network variations to optimize the total energy consumption, while maintaining application constraints. Hence, our contributions are summarized as follows:

1) We present an energy efficient framework for adaptive multiple modality data compression over multi-user mHealth system, where each user is able to collect multiple vital signs using wearable devices, compress, and transmit multiple modality data to the mHealth cloud (MHC).
2) We propose a Deep Learning (DL) approach for single and multiple modality data compression at the PDA level, where we exploit the intra and inter-correlation among multiple modalities for enhancing compression efficiency. Our approach is reversible in the sense that data are efficiently recovered at the MHC level.
3) We formulate an optimization framework for resource allocation in multi-user system that defines the optimal compression ratio for each user based on network dynamics, while minimizing the total energy consumption.
4) Our simulation results demonstrate the advantages of the proposed approach with respect to state-of-the-art techniques for increased compression ratio while maintaining signal distortion below a predefined threshold. Furthermore, the proposed optimization framework illustrates its efficiency for reducing the total energy consumption, while optimally allocating network resources between multiple users.

The rest of the paper is organized as follows: Section 2, is an overview of related work. We present in Section 3 our mHealth system, and describe our multiple modality compression algorithm in Section 4. In Section, 5 we formulate our optimization problem, and in Section 6, we present results. Finally, we conclude in Section 7 and summarize our remarks.

II. RELATED WORK

In this Section, we review previous work related to the following topics: conventional lossless and lossy compression methods, WBASN constraints and application requirements for efficient compression, neural networks and deep learning for data compression, and lastly, compression enhancements via multiple modality paradigms.

Many successful time series compression algorithms have been proposed which take into account the intra-modality correlation, and possibly the varying wireless network conditions to be able to adapt compression parameters for the purpose of minimizing transmission energy.

On one hand, some of these techniques are lossless such as Dictionary-based coding. In [16], a compression technique was proposed by quantizing data vectors. Compression was achieved through sending codebook indexes to the decompresser in place of the original time series. Although this technique showed low computational complexity and relatively high compression ratios; it is inefficient for EEG compression due to its non-stationarity. Hejrati *et al.* in [17] focused on reducing the intra and inter-channel redundancy and proposed a lossless multichannel EEG compression algorithm. First, intra-channel redundancy was removed using Differential Pulse Code Modulation. Then, a K-means clustering algorithm was used to remove the inter-channel correlation. Finally, arithmetic coding was applied to further compress the resultant signal. Furthermore, in [18], the concept of Correlation Dimension (CD) was used to assess alikeness. EEG signals were divided into segments, then, CD was calculated for each segment and blocks of EEG samples were arranged according to their CDs. After that, compression using neural networks was conducted [19].

On the other hand, lossy compression was also tested on EEG while simultaneously preserving diagnostic accuracy and complying with application requirements. Commonly used lossy compression techniques include Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) and Compressive Sensing (CS) techniques [20], [21]. Srinvasan *et al.* in [22] suggested to arrange EEG signals into 2D matrices. Specifically, each 1D EEG signal is cut into segments of the same length and arranged to form a matrix, where odd rows are filled directly whereas even rows are flipped. A smoothness analysis showed that EEG has a short-term memory which reflects inherent hidden correlation. This characteristic was exploited to enhance compression performance using classic image compression techniques.

In [23], electromyography signals were compressed with 75-90% compression ratios with an average distortion of 3.4 to 7% via DWT. In addition, adaptive data compression using DWT, based on the network condition, was proposed in [24] to optimize total energy consumption. Compressive Sensing has also been a very popular technique for EEG compression as it is lightweight on the transmitter side and most of the computation is done at the receiver side. Shukla and Majumdar [25] exploited the correlation of EEG signals and argued that they have a common sparse support in a transform domain. A sparse recovery problem was then formulated and solved. However, Majumdar *et al.* in [26] argued that the CS approach is inefficient since there is no sparsifying basis that fulfills the requirements of incoherence and sparsity. Instead, they formulated the recovery as a rank deficiency problem. Furthermore, they suggested using Schatten-p norm instead of the classic Nuclear norm as it showed better recovery results. The main problem was then divided into two sub-problems solved using Conjugate Gradient and p-shrinkage of singular values [27]. Xu *et al.* in [28] proposed a 1.5D EEG compression approach for WBASN where DWT was applied on the signal and its coefficients were arranged in a 2D matrix to concentrate energy and allow compression.

Although many lossless compression algorithms have been proposed in the literature; compression ratios obtained from these techniques remain limited. Consequently, in this work, we will focus on lossy compression techniques.

Few publications have focused on the network parameters and constraints when compressing EEG signals, particularly in WBASN system. Hussein *et al.* in [29] studied EEG compression in the context of WBASN with focus on encoding and transmission energy optimization in order to address power and distortion constraints. They proposed a scalable energy efficient EEG compression scheme based on DWT and CS. Several parameters have been considered to optimize the total energy consumption of the encoder and transmitter. The optimal configuration of these parameters was selected by limiting the total power consumption below a certain threshold. Awad *et al.* in [24] proposed an energy efficient cross-layer design where the optimal transmission rate and compression parameters were obtained based on the application requirements and network constraints.

Recently, data driven algorithms such as neural networks have emerged vastly. Artificial Neural Networks (ANN) proved to be state of the art in many cognitive tasks such as image recognition, speech detection and compression [30]. In [31], Google researchers have achieved state of the art image compression using ANNs. In [32], Kuleshov *et al.* introduced a state of the art ANN-based technique for audio prediction using Convolutional Neural Networks (CNN), and they were able to generate missing audio samples in low resolution audio signal. Furthermore, Sriraam and Eswaran [33] have compared different EEG near lossless compression techniques using: single layer perception, multilayer perception, log sigmoid activations, Elman networks and autoregressive models. They concluded that single layer perception with linear activations performs the best EEG near lossless compression.

Auto-Encoder (AE) is a particular type of ANN that seeks to approximate the identity function so that the output would be similar to the input [30]. Throughout this process, AE is able to learn important structures about data by constraining the neural network. Yan Ollivier in [32] proved that there is a strong relationship between minimizing the code length of data and minimizing reconstruction error that AE seeks. Tan and Eswaran [34] used a Stacked AE (SAE) for mammogram image compression where training was conducted on image patches instead of whole images. Testa and Rossi [35] applied AE for Electrocardiogram (ECG) compression. Comparisons with various classic compression methods showed that this special type of network is reliable for signal compression. Thus, in this paper we will focus on testing SAE for EEG compression.

Following the development of WBASN, vital signs data have become abundant. mHealth systems are now capable of collecting data from different modalities (EEG, EOG, etc). Although, they may seem totally unrelated, these data can give insights about the same phenomena. For example, in case of schizophrenic patients, when a stimulus is presented, a peak in EEG is observed and activations in the temporal lobe and the middle anterior cingulate region are depicted via functional Magnetic Resonance Imaging (fMRI) [36]. These findings confirm that both modalities are very likely to be correlated, as they present projections of the same underlying process.

In this paper, we hypothesize that deep learning is a good candidate to tackle this problem, due to its ability to efficiently exploit intra-modality correlation which allows extracting hierarchical representations of data. One of the well known advantages of deep neural networks is the ability to perform feature extraction. For example, when performing digit recognition using CNNs on the MNIST dataset [37], the extracted first order features represent image edges, the second order features correspond to patterns in the first features, such as corners and contours, and finally, higher orders tend to learn higher-level features. These characteristics have motivated us to investigate the application of deep learning on multiple modality data compression by exploiting not only intra-modality correlation but also inter-correlation among multiple modalities.

III. MULTI-USER mHEALTH SYSTEM

In this Section, we present the main components of our mHealth system, explain their individual and collective rules, and discuss their requirements to build an overall energy efficient vital signs delivery system.

A high-level description of our system is illustrated in Figure 1, where it consists of the following three major sub-systems/components:

- Edge network: multiple users equipped with PDAs to acquire vital signs using wearable devices. The PDA handles communication with the wearable device,

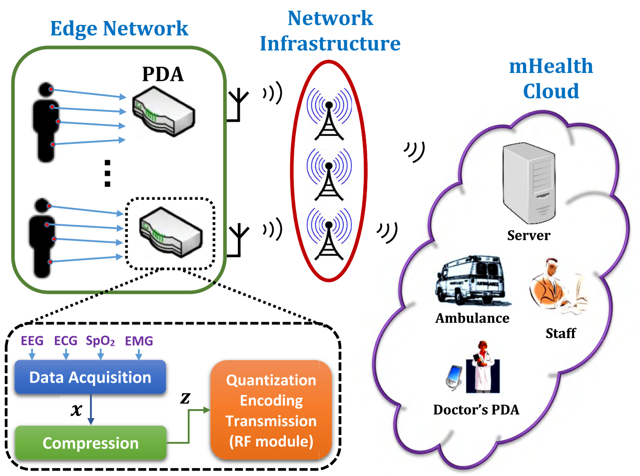

FIGURE 1. Overview of our system model consisting of edge network, network infrastructure, and mHealth cloud.

collects, pre-processes and transmits data to the MHC sub-system via the network infrastructure. Pre-processing consists of a compression algorithm that maps the original data to another representation. We propose to use a multiple modality deep learning compression approach that exploits the availability of multiple modality data and captures the inter-modality correlations to provide an adaptive compression technique. Specifically, we propose SAE-based compression schemes that compress acquired medical data before the transmission to the MHC taking into consideration network's state and application level Quality of Service (QoS).

- Network infrastructure: it allows the PDA to communicate with the MHC sub-system. The PDA is battery operated; thus, optimizing its transmission energy is essential. In this work, we formulate an optimization problem for multiple users to minimize energy consumption. We minimize a cost function that models the energy consumed by the different entities in the system, and their allocated resources based on the wireless channel state of each user. Furthermore, our model allows selecting an appropriate compression configuration given the current network dynamics.
- mHealth Cloud (MHC): a medical server that receives the compressed data from patients, decompresses and stores them for further analysis by the medical staff.

In what follows, we present the background and details of the proposed SAE-based compression schemes in Section 4. Then, to turn our proposed solution to be adaptive, we formulate and solve the optimization problem in Section 5 that aims to minimize the total energy consumption and select the optimal compression ratio for each user taking into consideration the network dynamics.

IV. PROPOSED DEEP LEARNING APPROACH FOR SINGLE AND MULTIPLE MODALITY DATA COMPRESSION

In this Section, we present the design requirements and methodology of the proposed compression scheme for vital signs. In particular, we first propose a single modality data compression technique using Stacked Auto-Encoders (SAE). Then, we extend the proposed approach for the multiple modality case discussing the obtained improvements in system performance. Finally, we present the complexity analysis of the proposed technique to assess its efficiency for maintaining low computational complexity.

A. DESIGN REQUIREMENTS

The following requirements guide the design of a compression approach used before transmitting vital signs as part of the mHealth system in accordance with the network and application constraints (see Section III):

1) Compressibility: It must be able to reduce input data dimensionality to a specific level that is required by the network state and capacity.
2) Reversibility: It needs to be able to reverse the compression operation (uncompress) at the receiver side within the performance criteria defined by the application.
3) Efficiency: It has to divide the required computational burdens between the edge node and MHC to reduce required network demands.

B. STACKED AUTO-ENCODER

SAE is a particular type of neural networks that can be used for supervised learning. It is comprised of an input layer, hidden layers and an output layer, see Figure 2. The output of each layer is fed as the input of the next layer until reaching the output layer. Within the hidden layers, a bottleneck is defined as the layer with the least number of neurons. SAE is designed to hierarchically capture abstractions of input data. It learns data representations; the first layer learns the first order features from the input data, the second layer learns the second-order features from the first order features and so on. SAE can be adopted for compression in the mHealth system as it complies with the design requirements defined in Section IV-A; it is able to compress data with specific ratios by adjusting the number of neurons in the bottleneck layer (compliance with Req. 1), reverse the compression operation by optimal decoding (compliance with Req. 2), and able to divide computational and resource burdens between the edge network and MHC by exclusively encoding in the first, while training and decoding in the latter (compliance with Req. 3). Let us consider an SAE with Q layers for encoding and

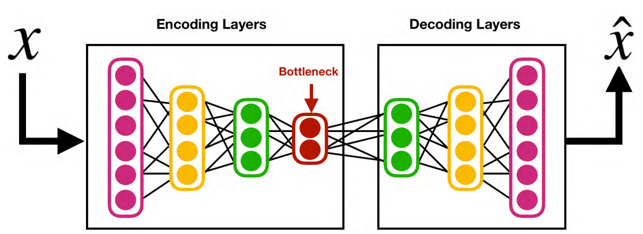

FIGURE 2. Stacked Auto-Encoder with three encoding and three decoding layers.

Q layers for decoding, and an input signal $\boldsymbol{x}$ with n samples such that $\boldsymbol{x} = [x(1), x(2), \cdots, x(n)]^T$. SAE seeks to reconstruct $\boldsymbol{x}$ using two operations; encoding and decoding, where the first can be performed in the edge network, and the latter can be performed in the MHC (compliance with Req. 3) First, the encoder gradually transforms $\boldsymbol{x}$ to a compressed representation z, at the bottleneck layer, such that: $z = [z(1), z(2), \cdots, z(m)]^T$ where $m < n$. Due to the encoding operation, an intermediate compressed signal z_q is produced at each layer q by the following expression:

$$z_q = f\left(W_q z_{q-1} + b_q\right), \tag{1}$$

where f is the activation function, $q = [1, 2, \cdots, Q]$, $z_0 = \boldsymbol{x}$, z_q has m_q samples such that: $m = m_Q < \cdots < m_1 < m_0 = n$, $z_Q = z$, W_q is a $m_q \times m_{q-1}$ weight matrix and b_q is a $m_q \times 1$ bias vector.

After that, the decoder gradually transforms back z to produce an estimate $\hat{\boldsymbol{x}}$. Due to the decoding operation, an intermediate estimate $\hat{\boldsymbol{x}}_q$ is produced at each layer q by the following expression:

$$\boldsymbol{x}_q = f\left(W'_q \boldsymbol{x}_{q-1} + b'_q\right), \tag{2}$$

where f is the activation function, $q = [1, \cdots, Q]$, $\boldsymbol{x}_1 = z, \boldsymbol{x}_q$ has n_q samples such that: $m = n_1 < \cdots < n_{Q-1} < n_Q = n$, $\boldsymbol{x}_Q = \hat{\boldsymbol{x}}$, W'_q is a $n_q \times n_{q-1}$ weight matrix and b'_q is a $n_q \times 1$ bias vector.

The former operations are optimized by training the SAE in a greedy layer-wise fashion. Each layer is trained such that the optimal set of parameters $\boldsymbol{\Theta} = [\Theta_1, \Theta_2, \cdots, \Theta_Q]$, containing each layer weight and bias, is obtained to minimize a reconstruction error $L_{\boldsymbol{\Theta}}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ (compliance with Req. 2). This error is generally modeled by the squared Euclidean distance, Eq. (3), or by cross-entropy, Eq. (4), and can be minimized using a Gradient descent algorithm [38].

$$L_{\boldsymbol{\Theta}}(\boldsymbol{x}, \hat{\boldsymbol{x}}) = ||\boldsymbol{x} - \hat{\boldsymbol{x}}||^2. \tag{3}$$

$$L_{\boldsymbol{\Theta}}(\boldsymbol{x}, \hat{\boldsymbol{x}}) = -\sum_i^n x_i \log(\hat{x}_i + (1 - x_i) \log(1 - \hat{x}_i)). \tag{4}$$

C. SINGLE MODALITY DATA COMPRESSION

We will refer to this technique as SAE-S. In this context, we compress every signal from any modality independently using a Stacked Auto-Encoder. Without loss of generality, let's consider a scenario where we have two modality signals $\boldsymbol{x}$ and $\boldsymbol{y}$ with number of samples n_x and n_y. Compression of $\boldsymbol{x}$ and $\boldsymbol{y}$ can be conducted using two separate SAEs, see Figure 3. SAE-S is a technique that uses a simple architecture of SAE and that yields good compression distortions. However, it comes with the following limitations:

- Separate SAE models need to be stored on each PDA for every modality.
- SAE-S exploits the modality's intra-correlation only, i.e. greedy solution.

Such restrictions force the mHealth system to operate in local optimal conditions that can quickly deteriorate when the number of modalities and/or the number of users increase.

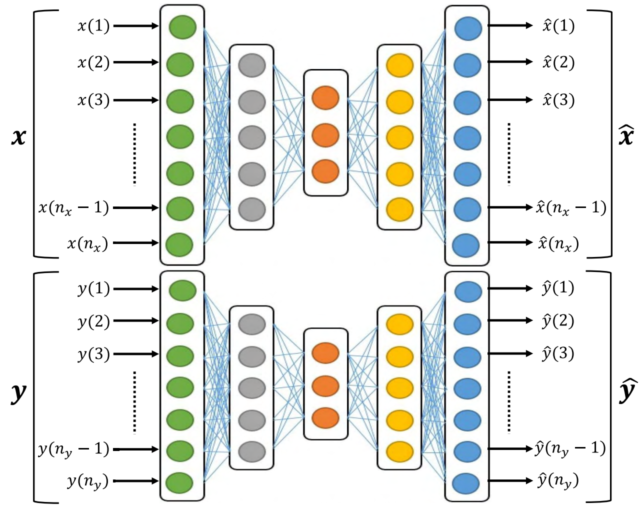

FIGURE 3. Single modality compression for two modalities x and y using SAE-S.

D. MULTIPLE MODALITY DATA COMPRESSION

We will refer to this technique as SAE-M. In this regard, we compress acquired signals from more than one modality jointly using one Stacked Auto-Encoder. SAE-M allows encoding multiple modalities in a single shared representation, which can in return, produce better compression results through joint optimization leveraging inter-modality correlations. Assuming the same two modality signals $\boldsymbol{x}$ and $\boldsymbol{y}$ as in Section IV-C, SAE-M is conducted by concatenating the two modalities into one vector $[\boldsymbol{x}, \boldsymbol{y}]$. After that, one SAE is applied to this vector as depicted in Figure 4. Multiple modality compression is a simple solution that liberates the mHealth system from the local optimal conditions dilemma as follows:

- Only one SAE-M model needs to be stored on the user's PDA for all modalities for a specific application.
- Data transmission is free of modality type headers.
- SAE-M performs joint optimization leveraging both intra and inter-modality correlations.

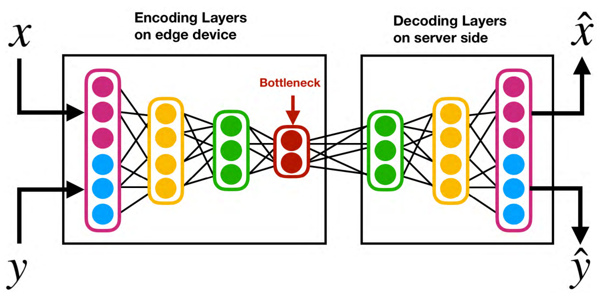

FIGURE 4. Multiple modality compression for two modalities x and y using SAE-M.

E. COMPLEXITY ANALYSIS

Training an SAE is an expensive task, however, it can be conducted off-line at the server side in the MHC to obtain an optimal configuration of weights and biases, which then can be stored on the patient PDA to be used in real-time. In what follows, we assess the computational complexity of SAE-S and SAE-M during test time (forward propagation) through formalizing the number of performed operations, multiplications and/or additions, in both the encoder and the decoder.

Let us re-write Eqs. (1) and (2), assuming linear activation function f, such that:

$$\boldsymbol{z} = \left(\prod_{q=1}^{Q} W_{Q-q+1}\right)\boldsymbol{x} + \sum_{p=1}^{Q}\left(\prod_{q=1}^{Q-p} W_{Q-q+1}\right) b_p. \quad (5)$$

$$\boldsymbol{x} = \left(\prod_{q=1}^{Q} W'_{Q-q+1}\right)\boldsymbol{z} + \sum_{p=1}^{Q}\left(\prod_{q=1}^{Q-p} W'_{Q-q+1}\right) b'_p. \quad (6)$$

The number of operations ζ (complexity) for a Q-layered SAE encoder or decoder can be deduced as:

$$\zeta = Q\, m_Q + \sum_{q=1}^{Q}\left(q\, m_{q-1}\, m_q\right), \quad (7)$$

where m_0 up to m_Q are decaying number of samples in case of encoding, or increasing number of samples in case of decoding.

V. ENERGY CONSUMPTION OPTIMIZATION

In this section, we propose an optimization framework for multi-user mHealth system taking into consideration the network and application requirements. Specifically, we adapt our SAE-M technique based on network dynamics and application demands to obtain optimal compression ratios that minimize total energy consumption through selecting the DL configuration that guarantees such compression ratio. First, we summarize the network/application constraints, and then, formalize the total energy consumption to be minimized. After that, we formulate the optimization problem and solve it using convex optimization tools.

A. NETWORK AND APPLICATION CONSTRAINTS

The aim of the proposed scheme is not only proposing a low-complexity compression technique but also to be adaptive to network dynamics and application demands. Our mHealth system is optimized such that the total energy consumption is minimized while optimally assigning transmission rates and bandwidths to all PDAs. Such assignment must be in accordance with the following application constraints: distortion threshold PRD_{thr}, and the following network requirements: delay deadline D_{thr} and total available bandwidth ω_t. In our model, we assume the network is interference free, and the main energy consuming components are the compression module, quantization, encoding, and transmission, as shown in Figure 1.

B. ENERGY CONSUMPTION CALCULATION

Let us consider a multi-user mHealth system with N users, where each records signals $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N]^T$ with $\boldsymbol{n}$ samples, and compresses them to $\boldsymbol{Z}$ with $\boldsymbol{m}$ samples such that: $\boldsymbol{n} = [n_1, n_2, \cdots, n_N]$ and $\boldsymbol{m} = [m_1, m_2, \cdots, m_N]$, where n_i and m_i correspond to the original and compressed number of samples of user i data. The amount of transmitted bits in such system can be expressed as: $\boldsymbol{L} = \boldsymbol{\beta}\,\boldsymbol{m}$, where $\boldsymbol{L} = [l_1, l_2, \cdots, l_N]$ contains the number of transmitted bits of all users and $\boldsymbol{\beta} = [\beta_1, \beta_2, \cdots, \beta_N]$ contains the number of bits/sample of all modules. Note that in this Section, the same indexing scheme will be used for all vectors.

The total energy consumed by user i, denoted as $E^{(i)}$, can be expressed as follows:

$$E^{(i)} = E_t^{(i)} + E_c^{(i)} + E_q^{(i)}, \quad (8)$$

where $E_t^{(i)}$, $E_c^{(i)}$, and $E_q^{(i)}$ are the energies consumed for data transmission, compression and quantization/encoding in module i respectively. Following [39], $E_t^{(i)}$ can be computed using the following equations:

$$E_t^{(i)} = \frac{\omega_i\, l_i}{r_i\, g_i}\left(2^{\frac{r_i}{g_i}} - 1\right), \quad (9)$$

where r_i is the transmitted data rate of PDA i over bandwidth ω_i, and g_i is the channel gain. Furthermore, $E_c^{(i)}$ can be calculated using its proportionality to the compression algorithm complexity (see Section IV-E) such that:

$$E_c^{(i)} = \zeta_i\, E_p, \quad (10)$$

where ζ_i is the compression algorithm complexity of module i encoder and E_p is the energy consumed per computation [40]. Using Eq. (7), for a Q-layered SAE-M, modifies Eq. (10) to:

$$E_c^{(i)} = \left(Q\, m_i^{(Q)} + \sum_{q=1}^{Q}\left[q\, m_i^{(q-1)} m_i^{(q)}\right]\right) E_p. \quad (11)$$

Finally, $E_q^{(i)}$ depends on the number of conversion steps, which is proportional to m_i and the energy consumed at each step E_s [41]. Thus, $E_q^{(i)}$ can be computed by the following:

$$E_q^{(i)} = m_i\, E_s. \quad (12)$$

C. PROBLEM FORMULATION

Performance of compression is quantified using Compression Ratio (CR) and distortion via Percentage Root Mean Square Difference (PRD) such that:

$$\text{CR}^{(i)} = 100 \times \left(1 - \frac{m_i}{n_i}\right), \quad (13)$$

$$\text{PRD}^{(i)} = 100 \times \frac{||x_i - \hat{x}_i||}{||x_i||}, \quad (14)$$

where $||\ ||$ is the Frobenius norm. Using regression analysis, PRD can be estimated from CR by the exponential operator h such that:

$$\text{PRD}^{(i)} = h\left\{\text{CR}^{(i)}\right\} = a\, e^{b\,\text{CR}^{(i)}}, \quad (15)$$

where a and b are the regression parameters.

The objective of our optimization problem is to to minimize the total energy consumption of all users $\boldsymbol{E}$, given the application requirements and network constraints, i.e. maximum allowed distortion $\mathrm{PRD}_{thr}^{(i)}$, delay deadline $\mathrm{D}_{thr}^{(i)}$, and the total available bandwidth ω_t. Consequently, our problem can be formulated as follows:

$$\min_{\mathrm{CR}^{(i)},\, r_i,\, \omega_i} \left(\frac{\omega_i\, l_i}{r_i\, g_i} \left(2^{\frac{r_i}{g_i}} - 1 \right) + Q\, m_i^{(Q)} + \sum_{q=1}^{Q} \left[q\, m_i^{(q-1)} m_i^{(q)} \right] E_p + m_i\, E_s \right), \tag{16}$$

$$\text{subjected to}: \; a\, e^{b\, \mathrm{CR}^{(i)}} \leq \mathrm{PRD}_{thr}^{(i)}. \tag{17}$$

$$\frac{l_i}{r_i} \leq \mathrm{D}_{thr}^{(i)}. \tag{18}$$

$$\sum_{i=1}^{N} \omega_i \leq \omega_t. \tag{19}$$

Checking convexity shows that this form of problems is non-convex due to $\boldsymbol{E_t}$ [42]. Thus, as noted in [39], we leverage the Geometric Program transformation to put this optimization in a convex form and solve it using convex optimization tools, as follow:

1) In Eq. (9), substitute $2^{r_i/g_i}$ using Taylor's theorem by: $1 + \frac{r_i \log(2)}{g_i} + \frac{{r_i}^2 \log^2(2)}{2g_i^2} + \frac{{r_i}^3 \log^3(2)}{6g_i^3} + \cdots$.
2) Using change of variables, define $\widehat{\mathrm{CR}}^{(i)} = \log\big(\mathrm{CR}^{(i)}\big)$, $\widehat{g}_i = \log(g_i)$ and $\widehat{r}_i = \log(r_i)$.

The function $\log\left(\sum a e^{b\,s}\right)$ is convex if $a > 0$ [42]. Hence, the problem turns to be convex in $\widehat{\mathrm{CR}}^{(i)}$, $\widehat{g}_i$ and $\widehat{r}_i$, and is solved using convex optimization tools. Solving this optimization problem results in optimal compression ratios, as well as transmission rates and bandwidths allocated for each user. Thus, each PDA can adapt its compression configuration of the SAE-M based on the obtained optimal CRs and network state.

VI. RESULTS AND DISCUSSIONS

In this section, we describe the utilized dataset with the required pre-processing, and present the following results:

1) We conduct experiments for single modality adaptive data compression, where we apply SAE for EEG and EOG compression and compare it with Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) and Compressive Sensing (CS) methods [20], [21].
2) We conduct experiments for multiple modality adaptive data compression, where we apply SAE-S and SAE-M for EEG and EOG compression and compare their results.
3) We assess the computational complexity of SAE-S and SAE-M, in terms of processing time, and compare it with DCT, DWT and CS.
4) We assess the efficiency of the proposed adaptive compression techniques by comparing the total energy consumption under different distortion thresholds and bandwidths.

A. DATASET, RATIONALE AND PRE-PROCESSING

Data compression and analysis is performed using the multiple modality DEAP dataset described in [43]. EEG, EMG, EOG and other physiological signals are collected from 32 healthy participants (50% males) aged between 19 and 37 years old. In this work, we confine our scope by only using EEG and EOG from this dataset. EEG was recorded using 32 active AgCl electrodes, placed according to the international 10-20 system, at 32 Hz, while EOG was recorded using four electrodes placed around each patient eyes [43].

EEG and EOG are selected as testing signals due to their ability in revealing the true performance of a compression system thanks to the following properties:

1) EEG is a highly non-stationary signal; it has a time-varying spectrum [44]. Thus, time or frequency domain compression techniques are naturally limited due to EEG non-stationarity, making it challenging to our design along with conventional compression techniques. In addition, EOG is a stationary signal; it has a time-independent spectrum [44]. Therefore, its opposing nature with EEG can be used for testing.
2) EEG and EOG are intensive signals that hinder real-time analysis and immediate medical interventions [45], [46]. Using the sampling rate of our utilized dataset, 32 Hz, approximately 6 million EEG and EOG data samples are generated per channel per day. This creates a Big Data problem for any automated analysis system and any network infrastructure [47]. Thus, efficient EEG and EOG compression can push the development of automated remote sensing systems.
3) EEG has been utilized in various applications such as: seizure detection [48], abnormality localization [49]–[51], brain-computer interfaces [46], etc. In addition, EOG along with EEG has been successfully used for artifacts detection and removal [50], eye-writing recognition [52], wheelchair control [53], etc. Hence, the integration of these two modalities is interesting to validate the viability of our compression technique to capture properties of each signal and provide a joint optimization that leverages the inter-modality correlations to minimize overall distortion.

In this work, we pre-process the utilized EEG and EOG by segmentation to produce a total of 2250 segments for each modality. Each segment corresponds to 7 seconds of activities or 896 samples of data. Moreover, each segment is whitened to become a zero mean vector with unity variance. After that, segments are randomly divided into 75% for training and 25% for testing to assess SAE-S and SAE-M performances in compression.

B. ADAPTIVE SINGLE MODALITY COMPRESSION

Figures 5 and 6 illustrate compression distortions for DCT, DWT, CS and SAE at different compression ratios for EEG and EOG respectively. Note that DCT is performed on a critically sampled version of the signals, i.e. zero DCT compression implies considering all frequencies that

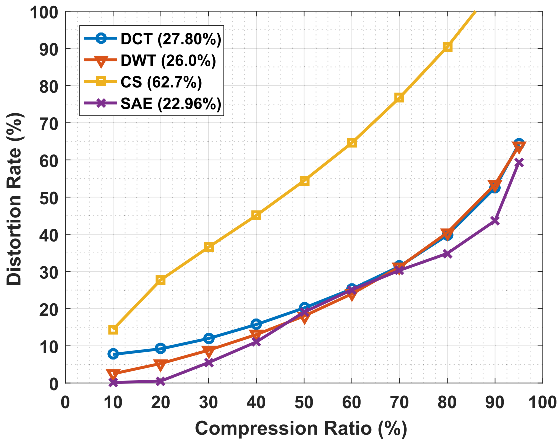

FIGURE 5. Performance of EEG compression using DCT, DWT, CS, and SAE.

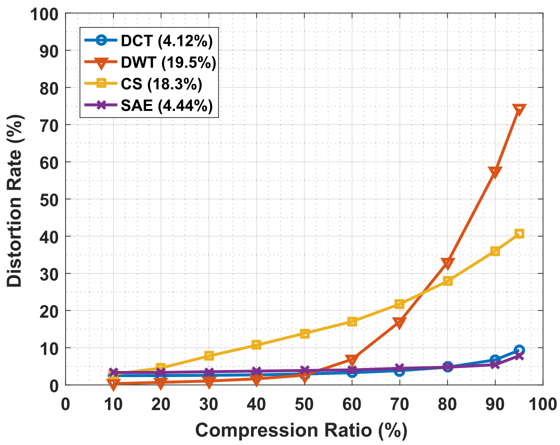

FIGURE 6. Performance of EOG compression using DCT, DWT, CS, and SAE.

cover 99.995% of the total energy. In addition, CS is applied on the Fourier transforms of the signals, assuming they are sparse, and samples are randomly selected for reconstruction [21]. Furthermore, DWT is optimized to minimize reconstruction distortion, where db29 and sym9 with single levels were found to be the best performing wavelets for EEG and EOG compression respectively [20]. In SAE, number of layers Q is selected to compress/decompress the signal by around 10% per layer, e.g. 40% SAE compression requires $Q = 4$, 4 for encoding and 4 for decoding. We use a linear activation function and the Squared Euclidean distance as our loss function regularized by a weight decay [54]. Finally, DCT and DWT compression results of the 2250 testing data segments are averaged via linear interpolation to ensure unbiased performance analysis.

On one hand, one can note in Figure 5 that DCT and DWT results are close to each other with no significant difference, and that SAE improves performance at high compression rates with an average distortion of 22.96% proving its ability to compress non-stationary signals. In addition, CS depicts a diverging distortion trend for high compression rates which agrees with Majumdar *et al.* conclusions in [26], EEG is not sparse. On the other hand, in Figure 6, DWT and CS results are opposing each other, as the first shows better performance up to 75% compression, and the second shows lower distortion above 75% compression. Furthermore, DCT and SAE depict higher performances when compared to the former with averaged distortions of 4.12% for DCT and 4.44% for SAE, proving SAE ability to compress stationary signals. Finally, Figures 7 and 8 demonstrate the compressed versus original EEG and EOG signals respectively using SAE for 20% and 80% compression ratios to give a visual assessment for the low distortions SAE can achieve.

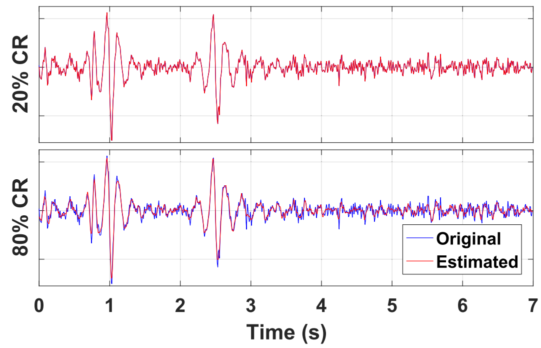

FIGURE 7. Comparison between original and compressed EEG using SAE at 20% and 80% compression ratios.

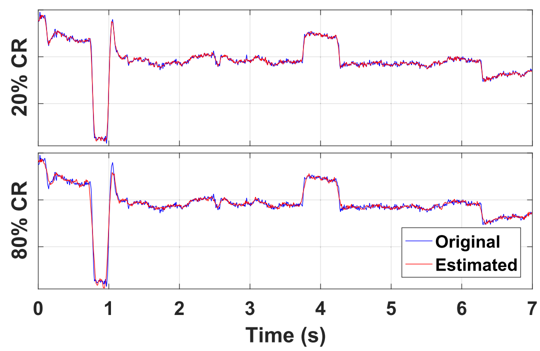

FIGURE 8. Comparison between original and compressed EOG using SAE at 20% and 80% compression ratios.

C. ADAPTIVE MULTIPLE MODALITY COMPRESSION

Figure 9 illustrates multiple modality compression distortions for DWT, SAE-S, and SAE-M at different compression ratios. Multiple modality results of DWT and SAE-S are computed by averaging the single modality outcomes using linear interpolation. The number of SAE-S and SAE-M layers Q is selected to compress/decompress the signal by around 10% per layer. In addition, we use linear activation functions and the Squared Euclidean distances as our loss functions regularized by a weight decay [54].

By examining the results, one can note that SAE-M shows an enhanced overall performance at all compression ratios with an average distortion of 13.21%. This proves the feasibility of SAE-M for multiple modality compression as it illustrates 72.37% and 3.71% relative increments in

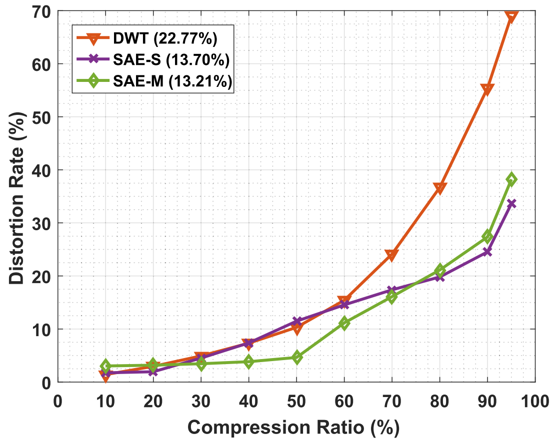

FIGURE 9. Performance of multiple modality compression using DWT, SAE-S and SAE-M.

performance when compared to DWT and SAE-S respectively. This illustrates SAE-M ability to compress both stationary and non-stationary signals as depicted in Figure 10.

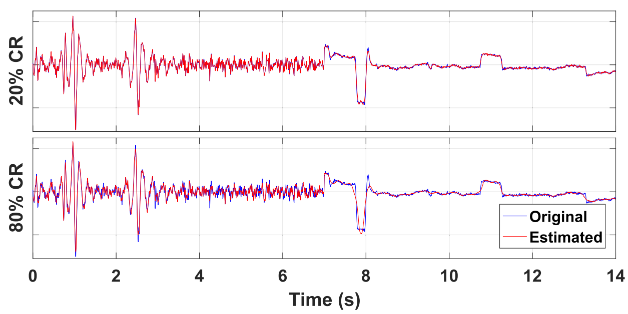

FIGURE 10. Comparison between original and compressed signals using SAE-M at 20% and 80% compression ratios.

D. COMPUTATIONAL COMPLEXITY FOR ADAPTIVE COMPRESSION

Assessment of computational complexity for the utilized adaptive compression techniques is conducted on a Lenovo Laptop (Y510P) with an Intel(R) Core(TM) i7-4700MQ x64-based processor, a 2.4 GHz CPU, 16 GB of memory and MATLAB R2016a. Results are produced by Monte-Carlo simulations where each single modality algorithm processing time is assessed by compressing 2250 segments of EEG or EOG ($\approx$ 4.4 hours of data per modality). Multiple modality results for DCT, DWT, CS, and SAE-S algorithms are produced by assuming a sequential paradigm; modalities are compressed independently in a sequential manner. Note that, DCT and DWT simulations are conducted with pre-defined thresholds using the optimal parameters summarized in Section VI-B, and that their results are averaged via linear interpolation to ensure integrity.

1) ADAPTIVE SINGLE MODALITY COMPRESSION

Figures 11 and 12 depict the required processing time for each single modality algorithm at different compression rates

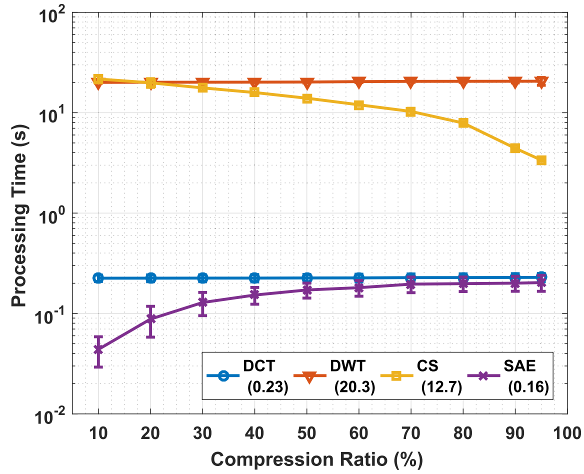

FIGURE 11. Computational complexity of EEG compression using DCT, DWT, CS, and SAE encompassed with 95% confidence intervals (CI). Note that some CIs are not visual on this plot due to estimation precision and logarithmic scale.

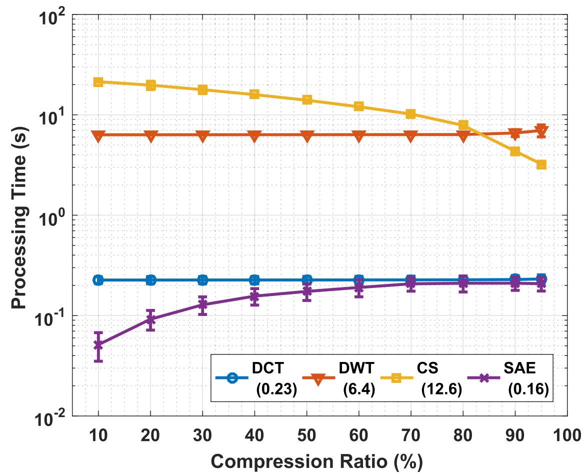

FIGURE 12. Computational complexity of EOG compression using DCT, DWT, CS, and SAE encompassed with 95% confidence intervals (CI). Note that some CIs are not visual on this plot due to estimation precision and logarithmic scale.

for EEG and EOG respectively. First, one can note that SAE requires the least processing time in both compression scenarios (EEG and EOG compression). Furthermore, DCT and DWT curves are independent of compression rate, while CS and SAE trends decrease and increase respectively with compression rate. In addition, mean processing times of DCT, CS, and SAE-S do not change, significantly, with the data type (EEG or EOG); however DWT does due to the different optimal parameters for each modality.

2) ADAPTIVE MULTIPLE MODALITY COMPRESSION

Figure 13 demonstrates the required processing time when compressing multiple modality data. SAE-S still shows the least required processing time; however, SAE-M illustrates a competing trend due to its lower distortion rates (Figure 9) with an average processing time of $\approx$ 1 second. Using

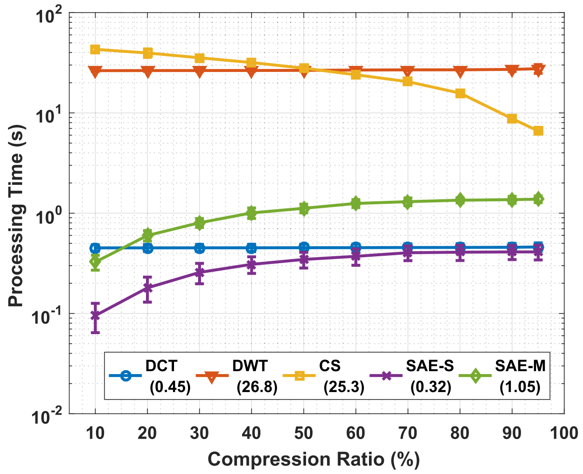

FIGURE 13. Computational complexity of multiple modality compression using DCT, DWT, CS, SAE-S, and SAE-M encompassed with 95% confidence intervals (CI). Note that some CIs are not visual on this plot due to estimation precision and logarithmic scale.

the averaged processing times, one can deduce the average number of compressed bytes per second as 67.80 MB/s for DCT, 1.15 MB/s for DWT, 1.21 MB/s for CS, 96.38 MB/s for SAE-S and 29.30 MB/s for SAE-M. By utilizing the former results, a pipeline for selecting an appropriate compression configuration depending on the network constraints can be created. For instance, SAE-S can be selected when the delay is the highest priority constraint, while SAE-M can be chosen when distortion is the most important concern.

E. ENERGY CONSUMPTION

We show in the following how the proposed SAE-M technique contributes in reducing the energy consumption under various network states. Based on the network topology in Figure 1, we analyze the energy consumption under different distortion thresholds and available bandwidths, given the simulation settings presented in Table 1. Figure 14 shows the total energy consumed for different values of distortion threshold. As shown, with more tolerance on the maximum allowed distortion, SAE-M can increase the compression ratio, hence, provides a significant reduction in the total energy consumption. Furthermore, using the proposed SAE-M technique, the achieved compression ratio has been increased compared to DWT and CS schemes, at the same distortion threshold. This enhancement in compression ratio leads to a significant reduction in energy consumption. Thus, we obtain about 45% and 20% reduction in energy consumption compared to DWT and CS, respectively.

TABLE 1. Simulation setting.

Parameter	Value	Parameter	Value
a	1.101	b	0.03682
N_0	-174 dBm	BER	10^{-4}
λ	12 m	E_{cs}	168 nJ
E_c	96 nJ	β	12 bps

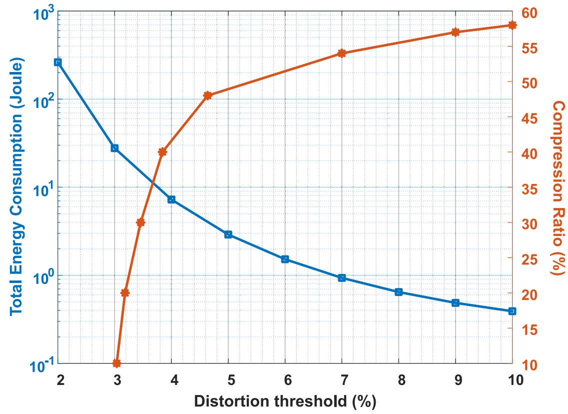

FIGURE 14. Total energy consumption and compression ratio of SAE-M with varying distortion threshold.

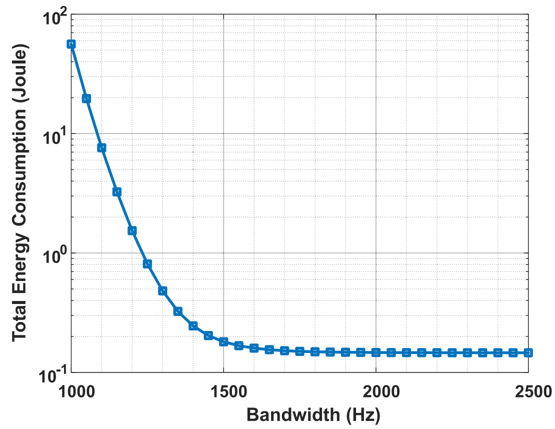

FIGURE 15. Total energy consumption of SAE-M for different total available bandwidths.

Regarding the effect of network dynamics, Figure 15 illustrates the resulting energy consumption with respect to the total available bandwidth, for $\text{PRD}_{thr} = 10\%$ and $\text{D}_{thr} = 1s$. At congested networks, the PDA will be allocated low bandwidth, hence will consume more energy to be able to transmit its data, while maintaining application's QoS requirements. On the contrary, at networks with large bandwidth, the PDA can consume less energy, hence obtaining less compression ratio, less distortion, and better energy efficiency. Findings show the importance of considering the variations of available network bandwidth and its impact on energy consumption. Thus, in the proposed adaptive SAE-M technique, we could optimize the total energy consumption, given the network dynamics, through adapting each user compression ratio, while maintaining the application QoS requirements.

VII. CONCLUSION

We proposed an adaptive biomedical data compression technique using Stacked Auto-Encoders for mHealth systems. We examined single and multiple modality data compression to exploit modalities intra and inter-correlation. We also proposed an energy and resource-aware framework for medical data delivery considering continuous changes in network dynamics. This was done by adapting the proposed algorithm to the following network constraints: delay deadline, the available bandwidth, and the application requirements, i.e. the maximum allowed distortion. We compared and analyzed our method with common compression approaches such as CS, DWT, and DCT and demonstrated that: 1) the proposed single and multiple modality compression techniques, SAE-S and SAE-M respectively, turn out to be adaptive to the network and application constraints. 2) Single modality SAE outperforms DWT by 13.24% reduction in distortion and DCT by 43.75% reduction in processing time. 3) SAE-M further decreases the distortion by 3.71% and 72.37% when compared to SAE-S and multiple modality DWT respectively. 4) SAE-S results depict its ability to disjointly compress stationary and non-stationary signals, while its extension, SAE-M, proves its ability to jointly leverage inter-signal correlations, making it essential for real-life applications. 5) SAE-S and SAE-M light computational complexities make them deployable on an edge device and suitable for real-time applications. 6) SAE-M technique was able to minimize the total energy consumption while adapting its compression ratio based on different network states.

ACKNOWLEDGMENT

(Mohammad F. Al-Sa'd and Mounira Tlili contributed equally to this work.) The statements made herein are solely the responsibility of the authors.

REFERENCES

[1] N. Zwar *et al.*, "A systematic review of chronic disease management," Australian Primary Health Care Res. Inst., Australian Nat. Univ., Canberra, ACT, Australia, Tech. Rep., 2017.

[2] J. H. Hibbard, J. Greene, R. M. Sacks, V. Overton, and C. Parrotta, "Improving population health management strategies: Identifying patients who are more likely to be users of avoidable costly care and those more likely to develop a new chronic disease," *Health Services Res.*, vol. 52, no. 4, pp. 1297–1309, 2017.

[3] World Health Organization, Regional Office for the Eastern Mediterranean. (2006). *Country Cooperation Strategy for WHO and Qatar: 2005–2009*. [Online]. Available: http://www.who.int/iris/handle/10665/193975

[4] J. Halford, *Electrographic Seizures in Adults: Recognition and Examples*. Cham, Switzerland: Springer, 2017, pp. 87–113.

[5] W. H. Polonsky, D. Hessler, K. J. Ruedy, and R. W. Beck, "The impact of continuous glucose monitoring on markers of quality of life in adults with type 1 diabetes: Further findings from the DIAMOND randomized clinical trial," *Diabetes Care*, vol. 40, no. 6, pp. 736–741, 2017.

[6] M. H. Shamji *et al.*, "Biomarkers for monitoring clinical efficacy of allergen immunotherapy for allergic rhinoconjunctivitis and allergic asthma: An EAACI position paper," *Allergy*, vol. 72, no. 8, pp. 1156–1173, 2017.

[7] J. M. Corchado, J. Bajo, Y. de Paz, and D. I. Tapia, "Intelligent environment for monitoring alzheimer patients, agent technology for health care," *Decis. Support Syst.*, vol. 44, no. 2, pp. 382–396, 2008.

[8] World Health Organization. (Feb. 2017). *Epilepsy*. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs999/en/

[9] A. S. Detsky, "What patients really want from health care," *JAMA*, vol. 306, no. 22, pp. 2500–2501, 2011.

[10] Emotiv. *Emotiv Website*. [Online]. Available: https://www.emotiv.com/

[11] N. Memon, X. Kong, and J. Cinkler, "Context-based lossless and near-lossless compression of EEG signals," *IEEE Trans. Inf. Technol. Biomed.*, vol. 3, no. 3, pp. 231–238, Sep. 1999.

[12] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010.

[13] S. M. Riazul Islam, D. Kwak, M. Humaun Kabir, M. Hossain, and K.-S. Kwak, "The Internet of Things for health care: A comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, Jun. 2015.

[14] *IEEE Standard for Local and Metropolitan Area Networks—Part 15.6: Wireless Body Area Networks*. IEEE Std 802.15.6-2012, Feb. 2012, pp. 1–271.

[15] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016.

[16] M. Hooshmand, D. Zordan, D. Del Testa, E. Grisan, and M. Rossi, "Boosting the battery life of wearables for health monitoring through the compression of biosignals," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1647–1662, Oct. 2017.

[17] B. Hejrati, A. Fathi, and F. Abdali-Mohammadi, "Efficient lossless multi-channel EEG compression based on channel clustering," *Biomed. Signal Process. Control*, vol. 31, pp. 295–300, Jan. 2017.

[18] N. Sriraam, "Correlation dimension based lossless compression of EEG signals," *Biomed. Signal Process. Control*, vol. 7, no. 4, pp. 379–388, 2012.

[19] N. Sriraam, R. Kannan, and C. Eswaran, "Lossless compression of EEG data using neural network predictors," in *Proc. 9th Int. Conf. Neural Inf. Process. (ICONIP)*, vol. 4, Nov. 2002, pp. 2046–2048.

[20] S. Grgic, M. Grgic, and B. Zovko-Cihlar, "Performance analysis of image compression using wavelets," *IEEE Trans. Ind. Electron.*, vol. 48, no. 3, pp. 682–695, Jun. 2001.

[21] M. A. Qureshi and M. Deriche, "A new wavelet based efficient image compression algorithm using compressive sensing," *Multimedia Tools Appl.*, vol. 75, no. 12, pp. 6737–6754, Jun. 2016.

[22] K. Srinivasan, J. Dauwels, and M. R. Reddy, "A two-dimensional approach for lossless EEG compression," *Biomed. Signal Process. Control*, vol. 6, no. 4, pp. 387–394, 2011.

[23] M. V. C. Costa, P. D. A. Berger, A. F. D. Rocha, J. L. A. D. Carvalho, and F. A. D. O. Nascimento, "Compression of electromyographic signals using image compression techniques," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, Aug. 2008, pp. 2948–2951.

[24] A. Awad, A. Mohamed, and A. A. El-Sherif, "Energy efficient cross-layer design for wireless body area monitoring networks in healthcare applications," in *Proc. IEEE 24th Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 1484–1489.

[25] A. Shukla and A. Majumdar, "Row-sparse blind compressed sensing for reconstructing multi-channel EEG signals," *Biomed. Signal Process. Control*, vol. 18, pp. 174–178, Apr. 2015.

[26] A. Majumdar, A. Gogna, and R. Ward, "A low-rank matrix recovery approach for energy efficient EEG acquisition for a wireless body area network," *Sensors*, vol. 14, no. 9, pp. 15729–15748, 2014.

[27] A. Majumdar and R. Ward, "Some empirical advances in matrix completion," *Signal Process.*, vol. 91, no. 5, pp. 1334–1338, 2011.

[28] G. Xu, J. Han, Y. Zou, and X. Zeng, "A 1.5-D multi-channel EEG compression algorithm based on NLSPIHT," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1118–1122, Aug. 2015.

[29] R. Hussein, A. Mohamed, and M. Alghoniemy, "Scalable real-time energy-efficient EEG compression scheme for wireless body area sensor network," *Biomed. Signal Process. Control*, vol. 19, pp. 122–129, May 2015.

[30] Y. Bengio, "Learning deep architectures for AI," *Found. Trends Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.

[31] G. Toderici *et al.*, "Full resolution image compression with recurrent neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5435–5443.

[32] R. Dennis *et al.*, "A systematic review of chronic disease management interventions in primary care," *BMC Family Pract.*, vol. 19, no. 1, 2018, doi: 10.1186/s12875-017-0692-3.

[33] N. Sriraam and C. Eswaran, "Performance evaluation of neural network and linear predictors for near-lossless compression of EEG signals," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 1, pp. 87–93, Jan. 2008.

[34] C. C. Tan and C. Eswaran, "Using autoencoders for mammogram compression," *J. Med. Syst.*, vol. 35, no. 1, pp. 49–58, Feb. 2011.

[35] D. Del Testa and M. Rossi, "Lightweight lossy compression of biometric patterns via denoising autoencoders," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2304–2308, Dec. 2015.

[36] N. Correa, Y.-O. Li, T. Adalı, and V. D. Calhoun, "Examining associations between fMRI and EEG data using canonical correlation analysis," in *Proc. 5th IEEE Int. Symp. Biomed. Imag., Nano Macro (ISBI)*, May 2008, pp. 1251–1254.

[37] Y. LeCun, C. Cortes, and C. J. C. Burges. *The MNIST Database of Handwritten Digits*. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[38] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. COMPSTAT*, Y. Lechevallier and G. Saporta, Eds. Heidelberg, Germany: Physica-Verlag, 2010, pp. 177–186.
[39] A. Awad, A. Mohamed, A. A. El-Sherif, and O. A. Nasr, "Interference-aware energy-efficient cross-layer design for healthcare monitoring applications," *Comput. Netw.*, vol. 74, pp. 64–77, Dec. 2014.
[40] E. Grochowski and M. Annavaram, "Energy per instruction trends in Intel microprocessors," *Technol. Intel Mag.*, vol. 4, no. 3, pp. 1–8, 2006.
[41] B. Murmann, "A/D converter trends: Power dissipation, scaling and digitally assisted architectures," in *Proc. IEEE Custom Intergr. Circuits Conf. (CICC)*, Sep. 2008, pp. 105–112.
[42] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[43] S. Koelstra *et al.*, "DEAP: A database for emotion analysis; Using physiological signals," *IEEE Trans. Affective Comput.*, vol. 3, no. 1, pp. 18–31, Jan./Mar. 2012.
[44] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing*, 2nd ed. Oxford, U.K.: Academic, 2015.
[45] M. Odabaee, W. J. Freeman, P. B. Colditz, C. Ramon, and S. Vanhatalo, "Spatial patterning of the neonatal EEG suggests a need for a high number of electrodes," *NeuroImage*, vol. 68, pp. 229–235, Mar. 2013.
[46] Y. Zhang *et al.*, "Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces," *Expert Syst. Appl.*, vol. 96, pp. 302–310, Apr. 2018.
[47] W. P. Fifer, P. G. Grieve, J. Grose-Fifer, J. R. Isler, and D. Byrd, "High-density electroencephalogram monitoring in the neonate," *Clin. Perinatol.*, vol. 33, no. 3, pp. 679–691, 2006.
[48] B. Boashash, H. Barki, and S. Ouelha, "Performance evaluation of time-frequency image feature sets for improved classification and analysis of non-stationary signals: Application to newborn EEG seizure detection," *Knowl.-Based Syst.*, vol. 132, pp. 188–203, Sep. 2017.
[49] B. Boashash and A. Aïssa-El-Bey, "Robust multisensor time-frequency signal processing: A tutorial review with illustrations of performance enhancement in selected application areas," *Digit. Signal Process.*, vol. 77, pp. 153–186, Jun. 2017.
[50] B. Boashash, A. Aïssa-El-Bey, and M. F. Al-Sa'd, "Multisensor time–frequency signal processing MATLAB package: An analysis tool for multichannel non-stationary data," *SoftwareX*, pp. 1–6, Dec. 2017.
[51] P. Nemtsas *et al.*, "Source localization of ictal epileptic activity based on high-density scalp EEG data," *Epilepsia*, vol. 58, no. 6, pp. 1027–1036, 2017.
[52] F. Fang and T. Shinozaki, "Electrooculography-based continuous eye-writing recognition system for efficient assistive communication systems," *PLoS ONE*, vol. 13, no. 2, p. e0192684, 2018.
[53] M. Djeha, F. Sbargoud, M. Guiatni, K. Fellah, and N. Ababou, "A combined EEG and EOG signals based wheelchair control in virtual environment," in *Proc. 5th Int. Conf. Electr. Eng.-Boumerdes (ICEE-B)*, Oct. 2017, pp. 1–6.
[54] W. Sun, S. Shaoa, R. Zhaob, R. Yana, X. Zhangc, and X. Chen, "A sparse auto-encoder-based deep neural network approach for induction motor faults classification," *Measurement*, vol. 89, pp. 171–178, Jul. 2016.

AHMED BEN SAID received the Ph.D. degree in computer Science from the University of Burgundy, France, in 2015. He was a Research Assistant with Qatar University on several projects, including the simulation of a surgical cutting operation using 3-D modeling, the usage of multispectral image for face recognition, and the development of reliable mHealth system for remote patient diagnosis. He currently holds a postdoctoral position at Qatar University. His research interests include machine learning and computer vision. He is also interested in crowd-sourced mobile applications and mobile health systems.

MOHAMMAD FATHI AL-SA'D received the B.Sc. and M.Sc. degrees in electrical engineering from Qatar University, Qatar, in 2012 and 2016, respectively. He is specialized in signal processing and graduated with honors under the supervision of Prof. B. Boashash. He has been with the College of Engineering, Qatar University, as a Research Assistant, since 2012. His research interests include multichannel EEG analysis and processing, time–frequency array processing, information flow and theory, statistical modeling and optimization, data mining, and machine learning. He has served as a Technical Reviewer for many journals, including *Biomedical Signal Processing and Control*, *Digital Signal Processing*, and so on.

MOUNIRA TLILI received the B.Sc. degree in computer science from Carnegie Mellon University in 2017. She is currently a Research Assistant with the Engineering Department, Qatar University. Her work focuses on data reduction and compression using deep learning for mHealth systems. The main objective of her work is creating a technology that senses the environment and reacts to it autonomously and seamlessly resulting in holistic experience for users with minimal interaction. Her research interests are in the fields of artificial intelligence, machine learning, and deep learning.

ALAA AWAD ABDELLATIF received the B.Sc. and M.Sc. degrees (Hons.) in electronics and electrical communications engineering from Cairo University, Egypt, in 2009 and 2012, respectively. He is currently pursuing the Ph.D. degree with the Department of Electronics and Telecommunications, Politecnico di Torino, Italy, and the Computer Science and Engineering Department, Qatar University. He was a Research Assistant with the Center for Wireless Studies, Cairo University, from 2009 to 2012, and with Qatar University from 2013 to 2014. His research interests include edge computing, in-network processing, heterogeneous networks, and resource optimization for efficient wireless sensor networks and smart health systems. He has served as a Technical Reviewer for many international IEEE journals and magazine, including the IEEE Systems, the IEEE Transactions on Multimedia, the *Journal of Biomedical and Health Informatics*, and *Communications Magazine*.

AMR MOHAMED (S'00–M'06–SM'14) received the M.S. and Ph.D. degrees in electrical and computer engineering from The University of British Columbia, Vancouver, Canada, in 2001 and 2006, respectively. He was an Advisory IT Specialist with the IBM Innovation Centre, Vancouver, from 1998 to 2007, taking a leadership role in systems development for vertical industries. He has over 25 years of experience in wireless networking research and industrial systems development. He is currently an Associate Professor with the College of Engineering, Qatar University, and the Director of the Cisco Regional Academy. His research interests include wireless networking and edge computing for IoT applications. He has authored or co-authored over 150 refereed journal and conference papers, textbook, and book chapters in reputed international journals and conferences. He holds three awards from IBM Canada for his achievements and leadership, and three best paper awards, latest from IEEE/IFIP NTMS 2015 in Paris. He has served as a Technical Program Committee (TPC) Co-Chair for workshops in IEEE WCNC 2016 and a Co-Chair for the technical symposia of international conferences, including GLOBECOM 2016, Crowncom 2015, AICCSA 2014, IEEE WLN 2011, and IEEE ICT 2010. He has served on the organization committee of many other international conferences as a TPC Member, including the IEEE ICC, GLOBECOM, WCNC, LCN, and PIMRC, and a technical reviewer for many international IEEE, ACM, Elsevier, Springer, and Wiley journals. He is serving as a Technical Editor for the *Journal of Internet Technology* and the *International Journal of Sensor Networks*.

TAREK ELFOULY has over 12 years of experience in computer network research. He published over 80 papers, most of them are related to wireless sensing and network security. He supervised many master's students and served as an examiner for many others. He has many projects under development related to computer networks and security. He has managed to secure over $3 million of funds for his research. His projects received many national and regional awards. His research work focuses on network security and protocols, physical layer security, and wireless sensor networks especially in the field of structural health monitoring and health applications. He has three provisional patent applications in USA, all of them are related to cyber security. He has 17 journal papers in reputable journals, such as the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE COMMUNICATION SURVEYS AND TUTORIALS. One of the projects that he secured a fund of 900K $ was related to network and service security using social networking, while another focused on information theory-enabled secure wireless networking by defining scaling laws, network control, and implementation. One of the projects that he lead with a total fund of $1million was in smart city structural health monitoring using wireless sensor networks for detecting the health of infrastructures in a city remotely and continuously while harvesting energy from the environment and optimizing the energy use. He has been involved in the deep learning field and has some publication to the applications of deep learning in health. He is currently with the Supreme Committee for Legacy and Heritage on a project to secure the evacuation of crowded venues, such as stadiums and shopping malls.

KHALED HARRAS is currently an Associate Professor with Carnegie Melon University in Qatar (CMUQ), where he is also the Director of the Computer Science Program. He is also the Founder and the Director of the Networking Systems Lab, CMUQ. He has been involved in or managing research grants that amount to over $ 4 million, and has supervised over 30 different personnel, including undergraduate and graduate students, post-doctoral researchers, and research engineers. He has over 100 refereed publications and holds four U.S. patents. Along with his research group in the past few years, he received the Best National Computing Research Award twice and two Best Paper Awards. His work has been featured online in various venues, such as *MIT Tech Review* and *Tech the Future*.

MARK DENNIS O'CONNOR was born in London, U.K., in 1962. He received the degree in physics from the St. Peter's College, Oxford University, in 1983, and the M.B.B.S. degree from the Charing Cross and Westminster Medical School, University of London. During his early postgraduate training, he became a member of the Royal College of Physicians (1992) and the Royal College of General Practitioners (1995) and an Associate Fellow of the Faculty of Occupational Medicine. He has spent a large part of his career specializing in the field of family medicine and occupational medicine, as a Senior Partner of the National Health Service General Medical Practice, Bristol, U.K. He has an interest in medical education. He was an Associate Lecturer in medicine with Bristol University, U.K. Since 2014, he has been the Associate Medical Director of mobile healthcare services, a division of the National Ambulance Service of Hamad Medical Corporation, Qatar. He is currently an Assistant Professor of clinical medicine with the Faculty of Clinical Medicine, Weill Cornell University, Qatar. He undertakes clinical work alongside holding a senior executive role within this National Organization. He leads a team of over 130 clinicians in establishing new community clinical services for the entire population of the country. He regularly contributes to the Ministry of Health National Health Strategy work groups and is a regularly invited speaker to International fora and other local organizations wishing to learn more about integrating primary and secondary care health systems to improve the care of long term conditions in the community. He sits on the National Ambulance Service's Group Research Oversight Committee and is currently collaborating, as a Principle Investigator, with the College of Engineering, Qatar University, on applying deep machine learning techniques to real-world clinical problems.