IoT

Article

Autoencoder-Based Neural Network Model for Anomaly Detection in Wireless Body Area Networks

Murad A. Rassam [1,2]

[1] Department of Information Technology, College of Computer, Qassim University, Buraidah 51452, Saudi Arabia; m.qasem@qu.edu.sa
[2] Faculty of Engineering and Information Technology, Taiz University, Taiz 6803, Yemen

Abstract: In medical healthcare services, Wireless Body Area Networks (WBANs) are enabler tools for tracking healthcare conditions by monitoring some critical vital signs of the human body. Healthcare providers and consultants use such collected data to assess the status of patients in intensive care units (ICU) at hospitals or elderly care facilities. However, the collected data are subject to anomalies caused by faulty sensor readings, malicious attacks, or severe health degradation situations that healthcare professionals should investigate further. As a result, anomaly detection plays a crucial role in maintaining data quality across various real-world applications, including healthcare, where it is vital for the early detection of abnormal health conditions. Numerous techniques for anomaly detection have been proposed in the literature, employing methods like statistical analysis and machine learning to identify anomalies in WBANs. However, the lack of normal datasets makes training supervised machine learning models difficult, highlighting the need for unsupervised approaches. In this paper, a novel, efficient, and effective unsupervised anomaly detection model for WBANs is developed using the autoencoder convolutional neural network (CNN) technique. Due to their ability to reconstruct data in a completely unsupervised manner using reconstruction error, autoencoders hold great potential. Real-world physiological data from the PhysioNet dataset evaluated the suggested model's performance. The experimental findings demonstrate the model's efficacy, which provides high detection accuracy, as reported F1-Score is 0.96 with a batch size of 256 along with a mean squared logarithmic error (MSLE) below 0.002. Compared to existing unsupervised models, the proposed model outperforms them in effectiveness and efficiency.

Keywords: wireless body area networks; anomaly detection; autoencoders; neural networks; machine learning; real-world dataset

check for updates

Citation: Rassam, M.A. Autoencoder-Based Neural Network Model for Anomaly Detection in Wireless Body Area Networks. *IoT* **2024**, *5*, 852–870. https://doi.org/10.3390/iot5040039

Academic Editor: Amiya Nayak

Received: 6 September 2024
Revised: 16 November 2024
Accepted: 19 November 2024
Published: 25 November 2024

1. Introduction

In developed countries, governments have plans to increase the average lifespan of their citizens, while the number of older adults requiring continuous monitoring is exponentially growing. This increase burdens the healthcare sector, emphasizing the need for pervasive systems capable of autonomously monitoring large numbers of patients. Hence, remote and ubiquitous vital sign monitoring has become essential. Furthermore, the increasing number of patients requiring admission and monitoring in Intensive Care Units (ICUs) necessitates automated systems to manage continuous monitoring of patients' critical conditions, facilitating decision-making by doctors and healthcare professionals.

The Internet of Medical Things (IoMT) refers to the collection, analysis, and storage of health-related data by tiny sensors comprising wireless body area networks. Vital signs such as blood pressure (BP), oxygen saturation (SPO2), blood pressure, body temperature, and pulse rate are collected and stored for the healthcare providers and practitioners to make informed decisions [1]. Figure 1 depicts various sensors implanted on or attached to the human body to measure vital signs and monitor patients' health at home or in the ICUs. The collected observations are sent from sensors at regular intervals to the sink node, which forwards them periodically to a base station where healthcare professionals can monitor patients' health conditions. The data may also be stored in the cloud, which will be available for further analysis.

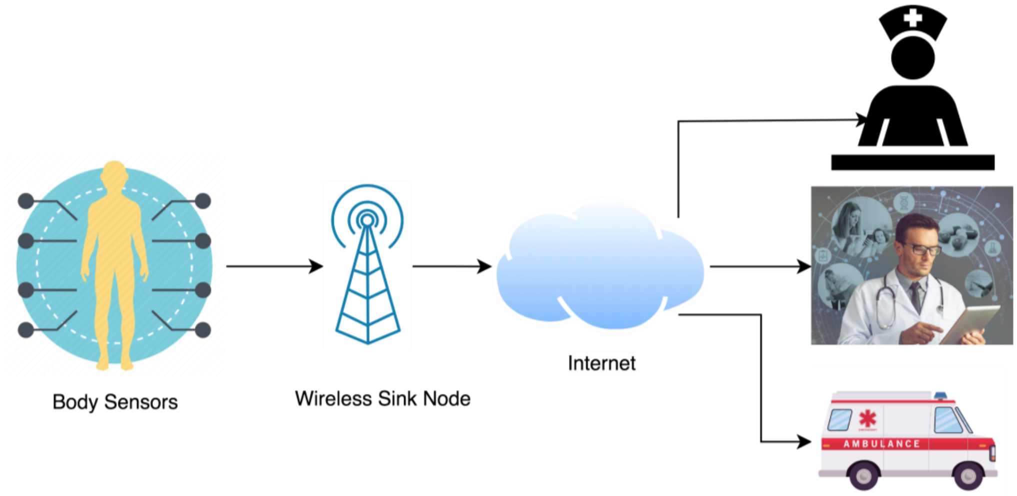


Figure 1. Healthcare monitoring via wireless body area networks.

The gathered data, however, are susceptible to several irregularities brought on by malicious attacks, inaccurate sensor readings, and other causes. Therefore, anomaly detection is a crucial procedure to guarantee data quality in many real-world applications, including healthcare monitoring.

Anomaly detection refers to identifying abnormal observations that occur for various reasons to ensure the quality of the data collected for healthcare monitoring applications, which is a significant research topic. Numerous statistical-based [2–5], machine learning-based [2,6–10], and other approaches to anomaly detection in WBANs have been introduced in the literature. However, most existing methods rely on computationally inefficient techniques and cannot be deployed online. Furthermore, statistical-based techniques depend on parameters that should be tuned for dynamic environments where each case requires individual consideration. Additionally, some current solutions do not account for multivariate sensor readings and consider only individual vital sign variables.

The absence of ground truth datasets for evaluating anomaly detection approaches in WBANs drives the research community toward unsupervised machine learning techniques. In these techniques, normal data observations are only needed for constructing the normal profile model that is used later to detect any deviations in the readings that indicate health degradation or signal events of interest that require immediate investigation by healthcare professionals.

To this end, this paper employs the concept of error reconstruction provided by autoencoder neural networks for unsupervised anomaly detection in WBAN. The autoencoder CNN technique is applied to multivariate healthcare data recorded at the ICU or from older adults to achieve this goal. According to [11], the autoencoder neural network can adapt to unexpected and new changes in a non-stationary environment due to the unsupervised learning feature.

The contributions of this paper are as follows:

- Proposing a novel anomaly detection model for WBAN based on the autoencoder neural network technique.
- Analyzing the performance of the proposed model in different settings to show its consistency and reliability in detecting anomalies in WBAN for various samples of real data streams.

The rest of this paper is outlined as follows: Section 2 critically reviews and analyzes existing approaches in anomaly detection for WBAN. The proposed model is introduced in Section 3, along with some background information on the autoencoder convolutional neural networks (CNN). Section 4 shows the findings of the experimental evaluation and compares the suggested model with previously published research. A discussion on the findings is provided in Section 5. This paper is concluded in Section 6.

2. Literature Review

There are several real-world application systems designed for remote patient monitoring based on WBANs, which gather and transmit patients' vital signs either at home or outdoors, such as CodeBlue [12], MEDISN [13], Vital Jacket [14], and Medical MoteCare [15]. A comprehensive survey of sensor-based medical applications can be found in [16,17]. Regarding research contributions, considering WBANs as a type of wireless sensor network (WSN), various anomaly detection strategies for WSNs, such as those in [18–23], can be applied to WBANs. However, most proposed models for WSNs need to be adapted due to the dynamic nature of vital sign readings.

Several anomaly detection models have been specifically proposed to help healthcare providers make accurate and timely decisions to identify anomalous readings received by WBANs. These models were developed based on various approaches, such as statistical [3,5,24,25], game theory [26], and machine learning [2,27] techniques. In addition, statistical-based techniques can be either parametric or non-parametric, and machine learning techniques can be supervised or unsupervised.

In [3], the authors presented a centralized anomaly and intrusion detection method in electrocardiogram (ECG) data. The proposed model utilized a simplified Markov model mechanism to detect deviations in ECG data. Several features were extracted from the ECG data and divided into sequences. The method then calculates each sequence's probability and identifies any deviations based on that. According to the authors, experimental results showed a low false positive rate, a comparable true positive rate, and a relatively short execution time. However, the Markov model is a parametric technique that relies on a threshold, which is challenging to set in context-aware environments.

The authors in [28] proposed a framework for event detection in biomedical sensor data based on the Kalman filter. The proposed mechanism forecasts the current observation and then derives the time series baseline of the collected data. In this approach, the distance between the predicted values is measured by the power divergence of the Kalman filter, which significantly deviates from its past values when a change occurs. Furthermore, this approach utilizes the spatial correlation between monitored observations to differentiate emergency events from faulty readings. In addition to the good detection accuracy reported by the authors, the proposed framework is efficient for real-world deployments. Similarly, the Kalman filter, like the Markov model, is based on a sliding window and requires several parameters to be set for each case.

A two-level anomaly detection approach was proposed in [26] based on game-theoretic techniques and the Mahalanobis distance measure. The proposed model was claimed to be lightweight and adaptive, such that it raises alarms only when the patients enter an emergency and discards false alarms caused by faulty observations. The first level constitutes the game-theoretic technique that exploits the spatiotemporal correlation of readings of the body sensor nodes and therefore detects local anomalous events according to the WBAN context. The Mahalanobis distance measure is employed at the second level for global multivariate analysis in the local processing unit attached to the body. Several numerical simulations were conducted on a real-world physiological data set to evaluate the efficacy of the proposed approach, which revealed a high detection accuracy and a low false alarm rate and energy consumption, according to the authors. However, the proposed approach is inefficient for big data cases due to the calculations required for the distance measure and therefore becomes impractical.

Another paper [2] contributed a model that can detect various types of anomalous observations in WBAN, such as simple, point, and contextual anomalies. The proposed model was designed based on the hybrid Convolutional Long Short-Term Memory (ConvLSTM) techniques, which can detect correlations between readings of WBAN sensors. According to the authors, experimental evaluations showed a high detection rate with a lower loss rate on different data subjects of a real-world physiological dataset. Furthermore, it was claimed that the hybrid ConvLSTM-based model achieved better results than CNN and LSTM separately. Deep learning models, as claimed, are helpful for big data streams generated by various sensors implanted in the patient's body. However, supervised approaches to anomaly detection problems suffer from the lack of ground truth datasets and depend on manual labeling, which may not be accurate.

In [29], a unified big sensor data framework for detecting anomalies in the WBAN environment was proposed. The proposed framework uses data compression and parallel fuzzy clustering based on the Hadoop MapReduce platform. The result clusters are further refined to provide better anomalous data detection accuracy. It is reported that the experimental results of the proposed framework using the real-world physionet dataset collected by sensors at ICU revealed the proposed framework's time efficiency and high classification accuracy. However, the proposed framework is of high cost and cannot be adopted for many WBAN applications, such as elderly monitoring at home.

A hybrid approach of isolation forest and K-means clustering was proposed in [30] to detect anomalies in elderly vital signs readings. A public dataset for vital signs of older adults was used to evaluate the proposed hybrid approach and compare it to the separated isolation forest approach. It is claimed that the proposed hybrid approach was more sensitive in detecting anomalous measurements than individual techniques in terms of low error rates for some labeled datasets. However, this hybrid approach was not successfully generalized for all dataset samples.

In [31], the authors proposed a model based on a dynamic sliding window and Weighted Moving Average (WMA) for predicting the vital signs and comparing them with the actual sign readings to reduce the number of false alarms and detect anomalous readings. Some statistical metrics were used to evaluate the performance of the proposed approach using publicly available datasets. Another statistical lightweight anomaly detection (LWAD) framework was proposed in [24] to detect anomalies in remote patient monitoring systems based on WBAN. The distance correlation of linear and nonlinear physiological parameters is used to design the proposed framework. Furthermore, a dynamic sliding window algorithm was utilized to predict the short range of vital sign parameters efficiently. It is claimed that the proposed LWAD framework outperforms existing methods according to the validation using three real-world datasets.

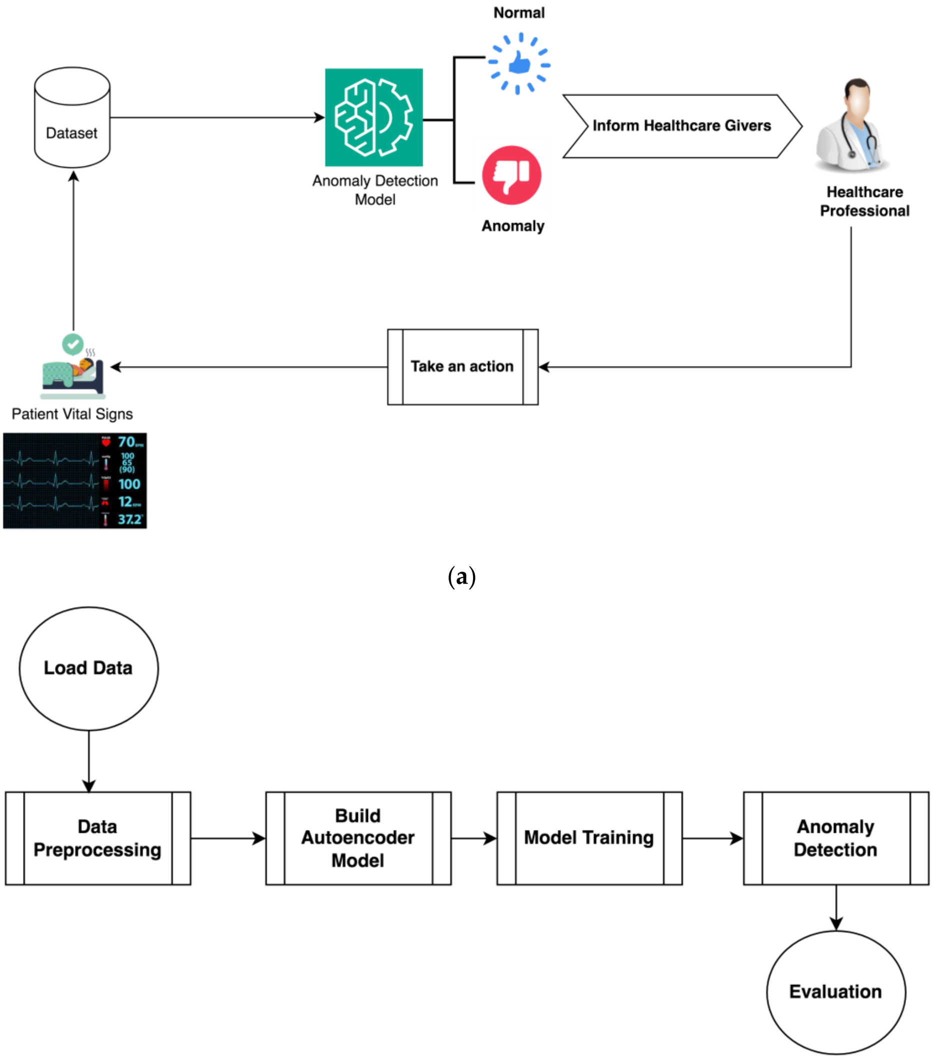
Autoencoder neural networks were used for the first time in [11] for anomaly detection in WSNs for IoT applications. A two-level approach was introduced, whereby the first-level algorithm resides on sensor nodes and the second level on the cloud. The detection mechanism is fully distributed so that each node can detect anomalies locally without communicating or cooperating with other nodes or the cloud. The computation-intensive learning task is carried out on the cloud level with a much lower frequency. The proposed design was claimed to be more efficient and reduce the power consumption and communication overhead. Experiments on real-world datasets collected by a real indoor WSN testbed revealed that the autoencoder-based anomaly detection approach was very effective regarding high detection rates and low false alarms.

Autoencoder CNN was further used in [32] to detect anomalies in multi-sensor time series in an unsupervised manner. Furthermore, the autoencoders have been utilized to identify device types for anomaly detection through network traffic analysis in IoT [33]. Additionally, in [34], the authors introduced three methods for better training of autoencoders for unsupervised anomaly detection, including cumulative error scoring (CES), percentile loss (PL), and early stopping via knee detection. The authors in [35] proposed an IoT-enabled WBAN and machine learning for recognizing patient speech emotions. In this study, authors developed a hybrid CNN and bidirectional long short-term memory (BiLSTM) techniques. More recently, the research in [36] applied Generative Adversarial Networks (GAN) for detecting anomalies in WBAN.

To conclude, existing anomaly detection approaches for WBAN have drawbacks, such as statistical-based techniques having many parameters to tune and becoming context-dependent. In contrast, supervised machine learning approaches require ground truth-labeled datasets, which are difficult to obtain as the context differs from patient to patient. Few unsupervised approaches exist in the literature; however, most of them cannot deal with massive data observations and cannot be scalable in real-time scenarios. This paper proposes an autoencoder convolutional neural network (CNN)-based approach. It has been claimed that autoencoders can adapt to novel and unseen changes in a dynamic environment and therefore fit the context scenarios. In addition, as a deep learning approach, autoencoders can learn the context of big data streams, a characteristic of WBAN vital sign readings.

3. Proposed Model

This section introduces the proposed autoencoder CNN-based anomaly detection model for WBANs (AUCNN-AD). The high-level diagram of the proposed model is depicted in Figure 2a. The detailed design steps of the anomaly detection engine are shown in Figure 2b and explained in the following paragraphs. Some preliminary details about the autoencoder CNN are also provided.



(**a**)

(**b**)

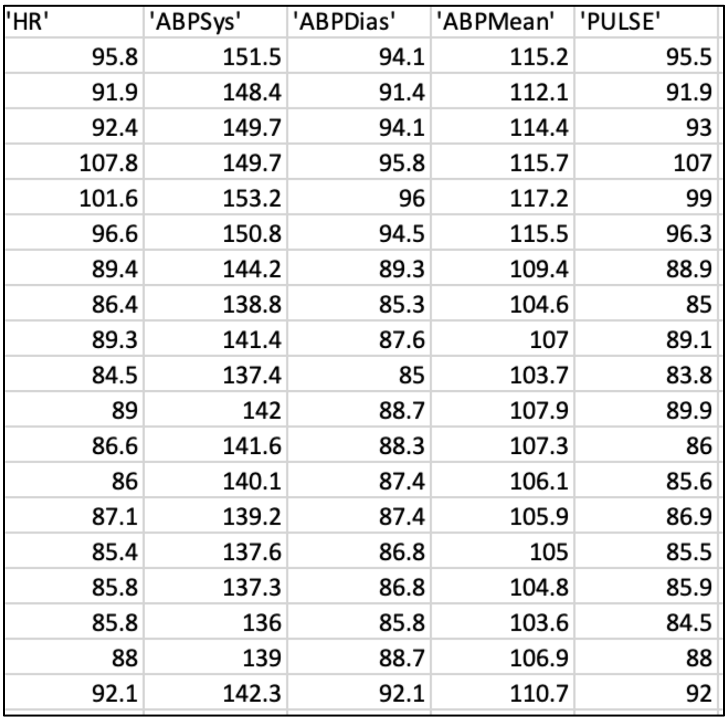
Figure 2. (**a**) High-level diagram of the proposed AUCNN-AD model. (**b**) The detailed design of the proposed AUCNN-AD model.

Figure 2a illustrates how data are collected by various sensor nodes implanted on the patient, which send their readings to a sink node to form a dataset used by the anomaly detection engine to classify them as normal or abnormal. Thus, the decisions made by the model will help healthcare professionals to follow up with the patient early and assess their situation accordingly. Based on the steps shown in Figure 2b, the proposed model is composed of the following main stages:

3.1. Data Collection and Preprocessing

3.1.1. Dataset Acquisition

This step collects vital sign readings from various sensors to form a dataset encompassing diverse physiological signals under multiple conditions. The collected dataset should include normal physiological patterns and instances of abnormal behavior. This ensures adherence to medical distinctions between what is considered a normal reading and what is considered abnormal. Figure 3 shows representative sample readings for various vital signs in the Multiple Intelligent Monitoring in Intensive Care (MIMIC-II) dataset used in this research to evaluate the proposed anomaly detection model.



'HR'	'ABPSys'	'ABPDias'	'ABPMean'	'PULSE'
95.8	151.5	94.1	115.2	95.5
91.9	148.4	91.4	112.1	91.9
92.4	149.7	94.1	114.4	93
107.8	149.7	95.8	115.7	107
101.6	153.2	96	117.2	99
96.6	150.8	94.5	115.5	96.3
89.4	144.2	89.3	109.4	88.9
86.4	138.8	85.3	104.6	85
89.3	141.4	87.6	107	89.1
84.5	137.4	85	103.7	83.8
89	142	88.7	107.9	89.9
86.6	141.6	88.3	107.3	86
86	140.1	87.4	106.1	85.6
87.1	139.2	87.4	105.9	86.9
85.4	137.6	86.8	105	85.5
85.8	137.3	86.8	104.8	85.9
85.8	136	85.8	103.6	84.5
88	139	88.7	106.9	88
92.1	142.3	92.1	110.7	92

Figure 3. Representative reading samples for various vital signs in the MIMIC-II dataset.

3.1.2. Data Preprocessing

In this subphase, the dataset is cleaned by removing noise artifacts and irrelevant information. After that, the physiological signals are normalized to maintain consistency across different sensors and to be suitable for use in the CNN model.

3.2. Autoencoder Architecture Design

An autoencoder neural network (AE) [37,38] is a feed-forward neural network where the output layer dimension equals the input layer, as shown in Figure 4. These types of neural networks are designed to operate unsupervised, where they are fed with training input vectors to reconstruct output vectors. As shown in Figure 4, AEs comprise two main components: an encoder and a decoder.

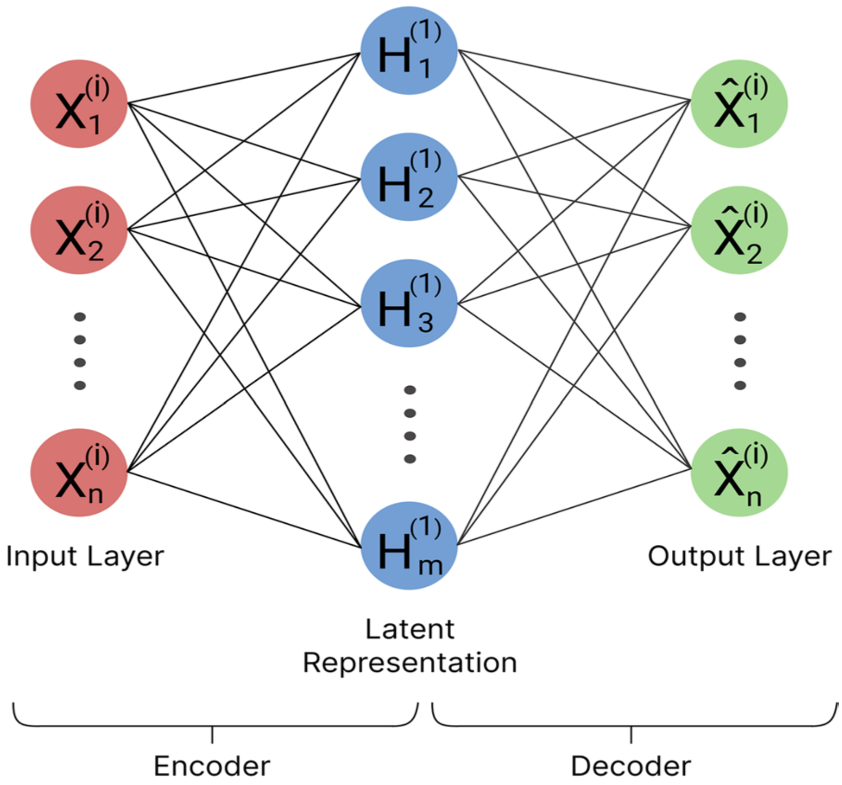


Figure 4. The basic autoencoder neural network architecture [39].

The AE works in a way such that the input vector X is transformed by the encoder to a hidden representation H, as presented in Equation (1).

$$H = \partial(W_{xh}X + \beta_{xh}) \tag{1}$$

W is a weight matrix, ∂ is any activation function, preferably a sigmoid function or rectified linear unit (RLU), and β is a bias vector.

The initial input space is reconstructed by transforming the hidden representation vector H using a decoder, as in Equation (2).

$$\hat{X} = \partial(W_{\hat{x}h}h + \beta_{\hat{x}h}) \tag{2}$$

The reconstruction error ε is then calculated by finding the difference between the reconstructed vector $\hat{X}$ and the original input vector, as in Equation (3).

$$\varepsilon = \|X - \hat{X}\| \tag{3}$$

The AE model is trained in an unsupervised approach to minimize the reconstruction error ε. An AE achieves this by learning the relationships between input features. If trained with data that resemble the training data, the AE should reproduce the input reasonably. If the input creates a high RE, it is considered an anomaly. The trained AE model can reconstruct normal input data with low RE but not anomaly data.

3.3. Model Training

Since labelled data are not required for autoencoder training, it is considered unsupervised. Optimizing a cost function remains the foundation of the training procedure. The cost function measures the error between the input x and its reconstruction at the output. The pseudocode algorithm for training the AE neural network is shown in Algorithm 1.

Algorithm 1. Pseudocode algorithm of the AE training process.

```
1.  Input: dataset samples X = ($x_1 \ldots x_n$) constitute normal and anomalous instances.
2.  Output: autoencoder model ($f_\varphi$, $g_{\theta.}$).
3.  Initialize the cost functions for encoder $\varphi$ and decoder $\theta$
4.  Calculate the sum of reconstruction errors E.
5.  Update parameters $\varphi$ and $\theta$ based on the gradient of E.
6.  If parameters $\varphi$ and $\theta$ still converge, go to 4.
    Else
    construct the encoder function $f_\varphi$ and the decoder function $g_\theta$.
7.  Stop.
```

As shown in Algorithm 1, the dataset serves as the process's input, and a trained autoencoder comprises an encoder function (f) and a decoder function (g) as its output. The encoder and decoder cost functions are initialized at the beginning of the algorithm. These cost functions are essential for measuring the difference between the autoencoder's reconstructed and original input data. The algorithm determines how well the autoencoder captures the patterns in the data by adding up the reconstruction errors after startup. Based on the error, gradient descent is used iteratively to update the encoder and decoder parameters. This procedure continues until the parameters converge when no appreciable advancements are possible. The training process ends when convergence is reached and the final encoder and decoder functions are built. After training, this autoencoder is prepared for usage in additional tasks like anomaly detection.

3.4. Anomaly Detection

After training the AE model, as described in Section 3.3, it can be used for anomaly detection by determining the threshold that will be used to classify the measurements as normal or abnormal. Algorithm 2 shows the pseudocode algorithm for the anomaly detection step. The threshold α is determined according to the pseudocode algorithm depicted in Algorithm 3.

Algorithm 2. Pseudocode algorithm of AE-based anomaly detection.

```
1.  Input: dataset samples X = ($x_1$,...$x_n$), constituting normal and anomalous instances,
    threshold $\alpha$.
2.  Output: classification of normal and abnormal data instances.
3.  Train the autoencoder with X to obtain $\varphi$ and $\theta$ as in Algorithm 1.
4.  Initialize counter i = 1.
5.  Calculate the sum of reconstruction errors $E_i$ for each i.
6.  If  ($E_i > \alpha$), $x_i$ is an anomaly
    Else
    $x_i$ is Normal
7.  If  (i<=n) go to 4
    Else Stop.
```

Algorithm 2 shows that the dataset's normal and abnormal instances and a threshold for identifying anomalies constitute the input for the anomaly detection phase. Each data instance is classified as normal or abnormal based on the algorithm's result. The encoder (f) and decoder (g) are first obtained by training the autoencoder on the dataset in Algorithm 1. Beginning with the first data instance, a counter is set to record the progress across the dataset. The algorithm determines the reconstruction error for each instance, quantifying the discrepancy between the input and the autoencoder's reconstruction. The event is classified as abnormal if this error is more than the predetermined threshold and as normal otherwise. For every data instance in the dataset, this procedure is repeated. The algorithm continues until every instance has been classified, after which it stops.

The threshold is a critical component in anomaly identification using autoencoders since it helps to identify abnormal data points. The procedure for determining the threshold utilized in anomaly detection (as seen in Algorithm 2) is explained by the pseudocode in Algorithm 3. Based on the reconstruction error, this threshold aids in deciding whether a data instance should be classified as normal or abnormal. A dataset, the number of samples (n_s), the number of features (n_f), and the autoencoder that has already been trained on the data are the inputs to the algorithm. The determined threshold value is the output. A threshold vector, which holds the threshold values for every sample, is initialized at the start of the process. After that, a loop is started that iterates through every sample in the dataset. The goal of the loop is to calculate the reconstruction error for every sample and modify the threshold appropriately. To properly distinguish normal data points from anomalies, this procedure seeks to establish a threshold value that achieves a balance between the sensitivity and specificity of the anomaly detection method.

Algorithm 3. Pseudocode algorithm of threshold calculation.

```
Input: dataset samples X = $(x_1 \ldots .. x_{n)})$, number of samples $n_s$, number of features $n_f$, trained autoencoder with X data AE.
Output: threshold $\alpha$.
Initialize the threshold vector $\alpha(0,0,\ldots..0_n)$.
For $i$ = 1 to $n_s$ do
    $AE(X_i) \xrightarrow{yields} \hat{X}_i RE(X_i, \hat{X}_i) \xrightarrow{yields} (r_1, r_2 \ldots ., r_{nf}) max((\alpha_1, \alpha_2, \ldots ., \alpha_{nf}), (r_1, r_2 \ldots ., r_{nf})) \xrightarrow{yields} \alpha$
End for
```

4. Experiments and Results

This section introduces the dataset used to evaluate the proposed model and presents the empirical results.

4.1. Datasets

The dataset used in this research is the Multiple Intelligent Monitoring in Intensive Care (MIMIC-I and II) [40], which includes physiological data records from more than 90 ICU patients, referred to as subjects. This dataset has been used in several research studies [2,25,28] as a benchmark to validate the viability of their proposed solutions. Two subjects, subject 330 and subject 441, were selected to test the proposed model in this paper. The data samples selected from these two subjects have seven features that describe the patient's current vital signs according to the time. The collected features include heart rate (HR), systolic arterial blood pressure (ABPsys), diastolic arterial blood pressure (ABPdias), mean arterial blood pressure (ABP-mean), pulse, temperature, respiration rate (RESP), and oxygen saturation (SPO_2), along with timesteps and dates. Sample sensor readings for heart rate (HR) and blood pressure features of Subject 441 are presented in Figure 5. MIMIC-II allows for time-dependent assessments of patient outcomes and the effects of interventions, which may be modeled using longitudinal statistical techniques and other machine learning and deep learning techniques because it records data over time. Like many healthcare datasets, MIMIC-II contains missing values, which are often handled statistically using techniques like removal or imputation, depending on the type of study.

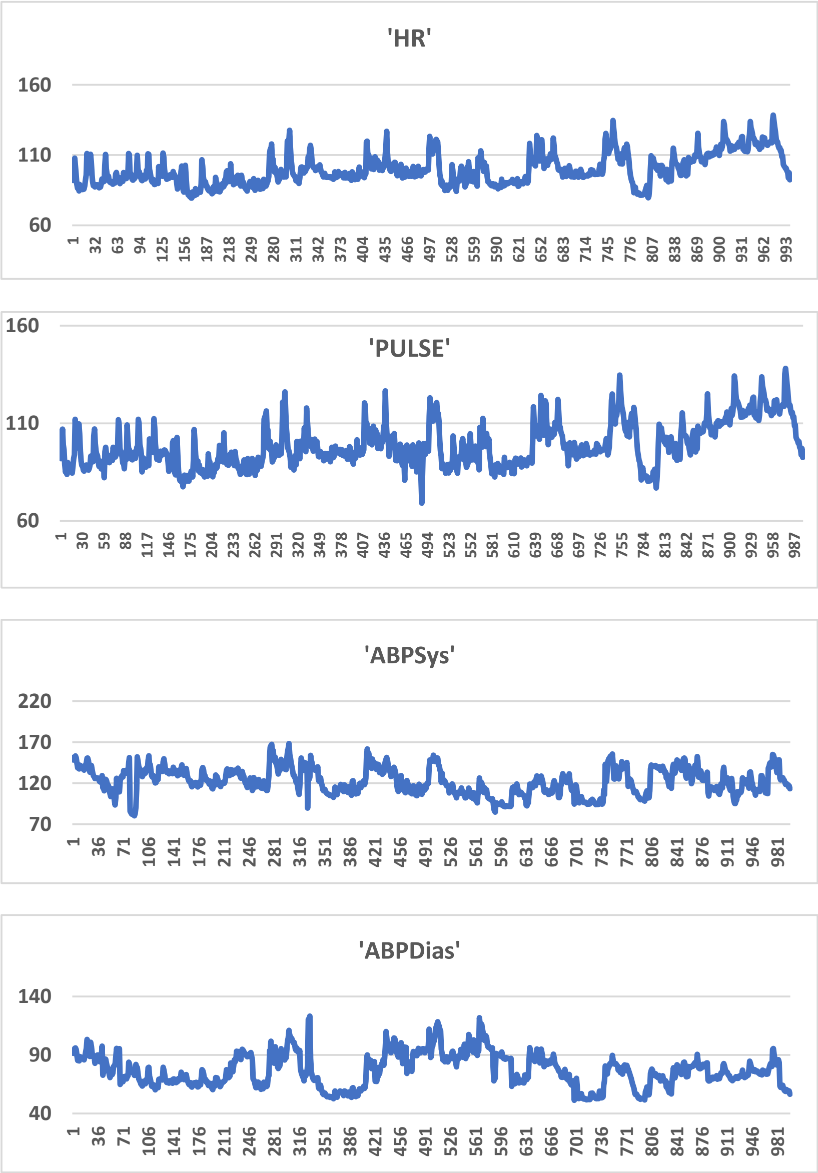


Figure 5. Sensor readings signals for four vital signs for subject 441.

After selection, the dataset samples are normalized to prepare them for deep learning processing. Normalization aims to use a standard scale to adjust the values of the dataset's numeric columns without distorting the range of values or losing data. The normalization procedure followed in this paper adjusts the data to a range between 0 and 1, using Equation (4).

$$x(i) = \frac{x(i) - \bar{x}}{S(x)} \tag{4}$$

where $x(i)$ is the dataset, $\bar{x}$ is one column in the dataset, and $S(x)$ is the number of the data samples.

4.2. Results

4.2.1. Model Setting and Evaluation Metrics

Before the actual implementation of the proposed model using the autoencoder neural network (AUN), the parameters are set as shown in Table 1. It is shown that the dataset samples are split into 80% for training and 20% for testing. The network architecture used contains 512 neurons in the input layer, 256 in the hidden layer, and 256 in the output layer. The number of epochs used to train the model is 20.

Table 1. Autoencoder neural network parameters.

Test Size	Network	Epochs
0.2	512-256-128	20

Four evaluation metrics are used to evaluate the proposed model: accuracy, recall, precision, and F1-score. These measures are commonly used in the literature to assess any classification-based task. The calculation of these metrics depends on the confusion matrix. Accuracy is a statistical metric that indicates how well the proposed model predicts outcomes, which, in our case, are normal or abnormal vital sign readings. Equation (5). shows how the accuracy metric is calculated.

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} \tag{5}$$

where *TP* is the true positive rate, *TN* is the true negative rate, *FP* is the false positive rate, and *FN* is the false negative rate. Similarly, Equations (6)–(8) show how other metrics are calculated.

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \tag{6}$$

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \tag{7}$$

$$\boldsymbol{F1-Score} = 2 \times \frac{\boldsymbol{Precison{*}Recall}}{\boldsymbol{Precison+Recall}} \tag{8}$$

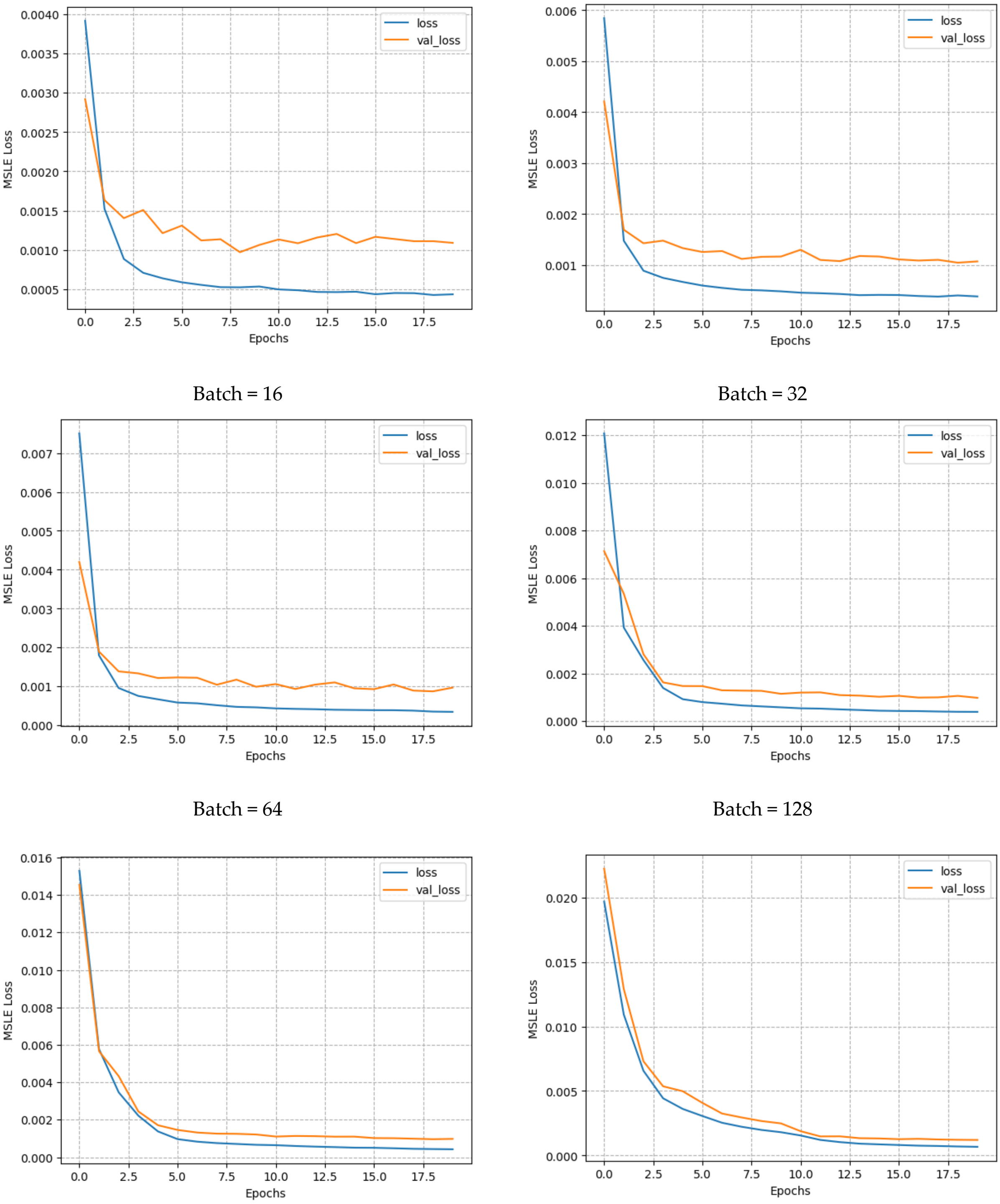
4.2.2. Performance Evaluation

Several experiments were conducted for each selected subject to train the proposed model. The experiments showed that the training batch size was the main parameter of the autoencoder technique that significantly influenced the classification of readings as normal or abnormal. The mean squared logarithmic error (MSLE) was used in our experiments to report the model's loss during training and validation. In the case of autoencoder neural networks, which are used to design the proposed model in this paper, the MSLE is a loss function used to evaluate the network's performance, particularly in regression tasks where the target variable is positive and can span several orders of magnitude. The *MSLE* can be mathematically defined, as shown in Equation (9).

$$\boldsymbol{MSLE} = \frac{1}{\boldsymbol{n}} + \sum_{i=1}^{\boldsymbol{n}} (\boldsymbol{log_e}(1+\boldsymbol{y_i}) - \boldsymbol{log_e}(1+\boldsymbol{\hat{y}_i}))^2 \tag{9}$$

where $\boldsymbol{y_i}$ is the true value, $\boldsymbol{\hat{y}_i}$ is the predicted value, and *n* is the number of readings.

By using the logarithm, MSLE reduces the effect of significant errors. This is beneficial when the goal is to focus the model on minimizing more minor errors, thereby reducing the impact of outliers in the dataset. In our case, the autoencoder is used for anomaly detection; therefore, it is trained to reconstruct normal data. Hence, the MSLE is crucial for identifying anomalies by detecting significant differences in reconstruction errors. As shown in Figure 6, the values of MSLE are minimum with a batch value of 16 and increase as the batch value gets larger. However, there is a significant difference between training loss and validation loss, which provides essential insights into the generalizability of the proposed model. This suggests that the model performs well during the training phase but does not generalize effectively to the validation phase. The possible reason for this outcome might be that the autoencoder has learned to reconstruct the training data very well, including noise and specific patterns, but fails to perform well on previously unseen data. This issue, known as overfitting, is clearly shown in Figure 6 for small batch sizes, such as 16, 32, and 64. This issue is gradually resolved by increasing the batch size, as the training loss curve becomes closer to the validation loss curve as the batch size increases. Despite the slight increase in MSLE values to around 0.014 for a batch size of 256, this batch size shows a better fit as the two curves converge. This outcome can be explained by increasing the batch size, which improves generalizability by reducing the effect of noise and other irrelevant patterns.



Batch = 16

Batch = 32

Batch = 64

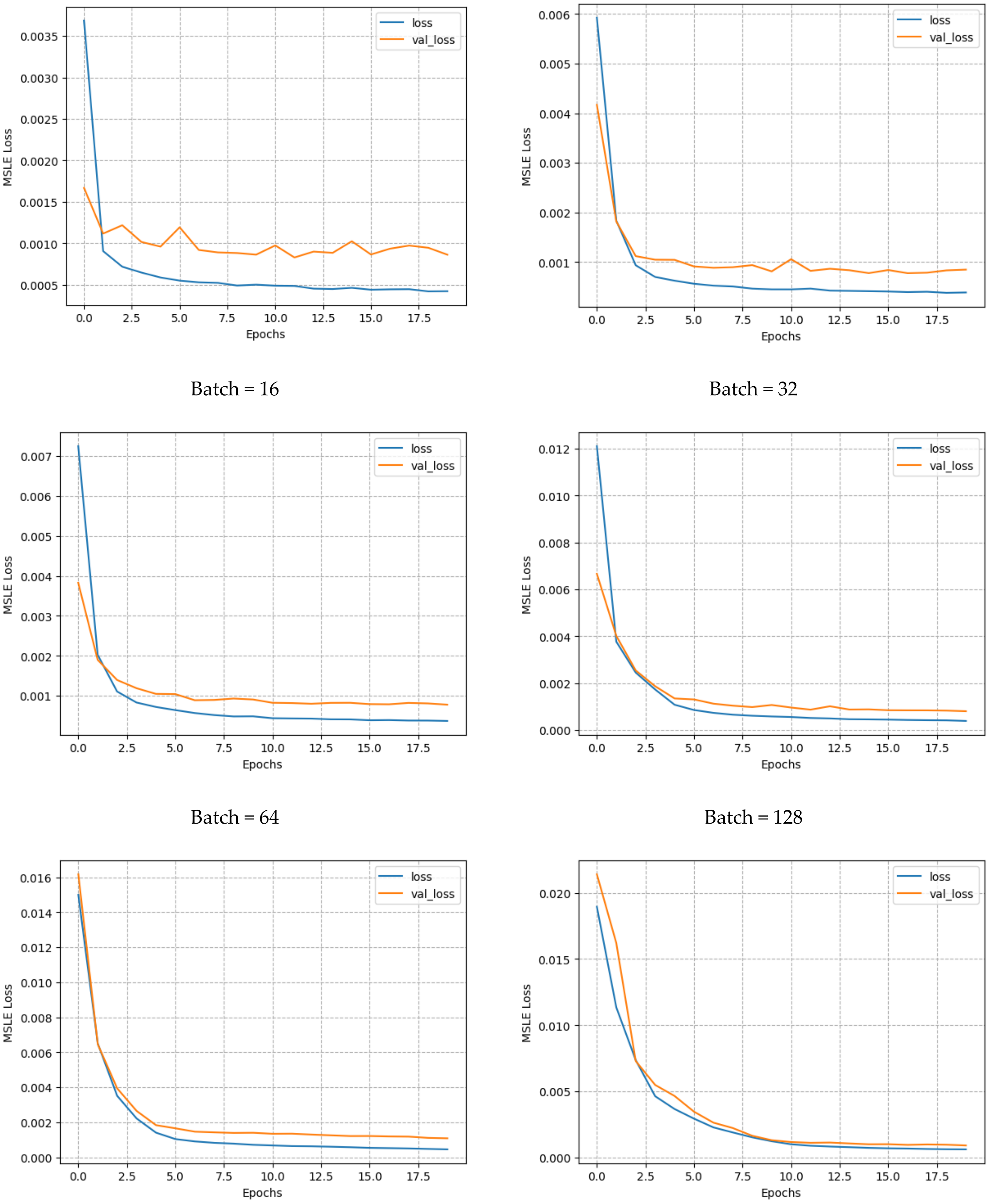
Batch = 128

Batch = 256

Batch = 512

Figure 6. Training and validation loss of the proposed model on subject 330.

Similarly, Figure 7 depicts the training and validation losses of the proposed mode on subject 441. The figure clearly shows that on a batch size = 256, the proposed model performs better than other batch sizes for the same justifications mentioned earlier for subject 330.



Batch = 16

Batch = 32

Batch = 64

Batch = 128

Batch = 256

Batch = 512

Figure 7. Training and Validation Loss of the Proposed Model on subject 441.

Figures 8 and 9 compare the performance of the proposed model using different batch sizes and various metrics, namely accuracy, precision, recall, and F1-score, for subjects 330 and 441, respectively. For subject 330, batch sizes of 128 and 256 yield the highest values for accuracy, recall, and F1-score. A similar trend is observed for subject 441, indicating consistency in the results across both subjects. An F1 score of 0.96 and an accuracy of 0.93 are reported for both subjects. These consistent results for the two subjects suggest the proposed model's performance stability. Figures 8 and 9 indicate that a batch size of 256 is the optimal choice for building the anomaly detection model for this application. Furthermore, the results suggest that other parameters have no significant effect on the performance of the proposed model for anomaly detection in WBANs.

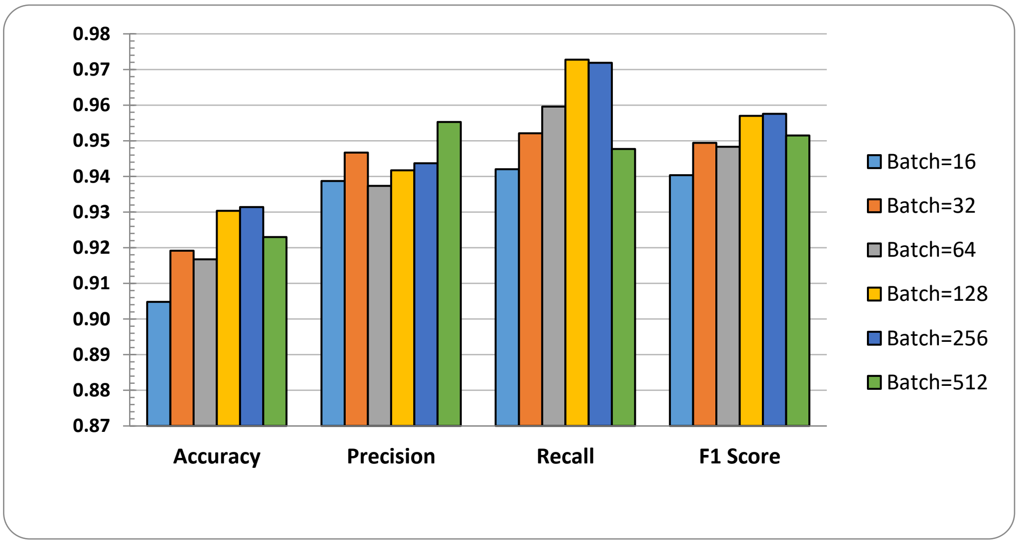


Figure 8. Performance evaluation metrics by different batch size for subject 330.

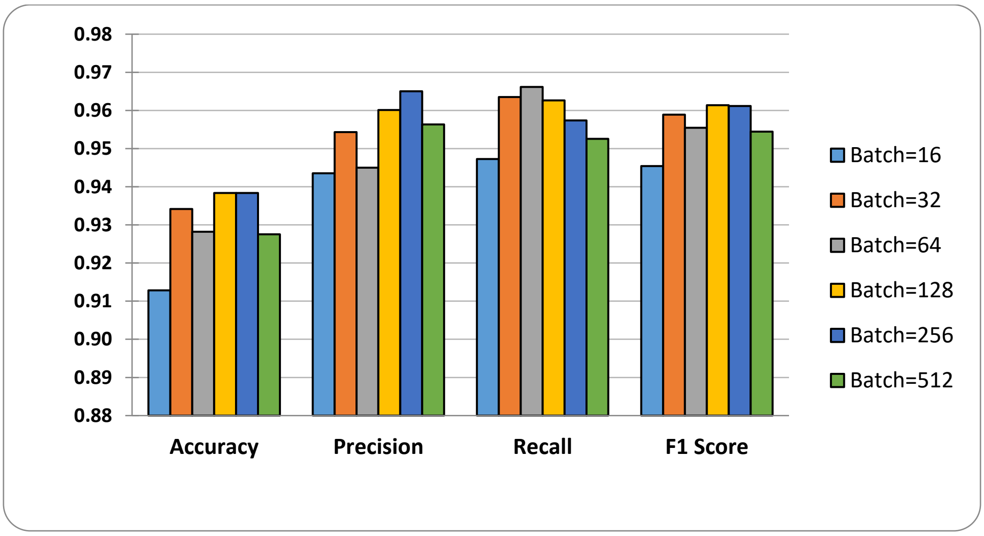


Figure 9. Performance evaluation metrics by different batch size for subject 441.

Figures 10 and 11 show the receiver operating characteristics (ROC) in different batch sizes for subject 330 and subject 441, respectively. They further show the area under the curve (AUC) values that indicate all values are above 50%.

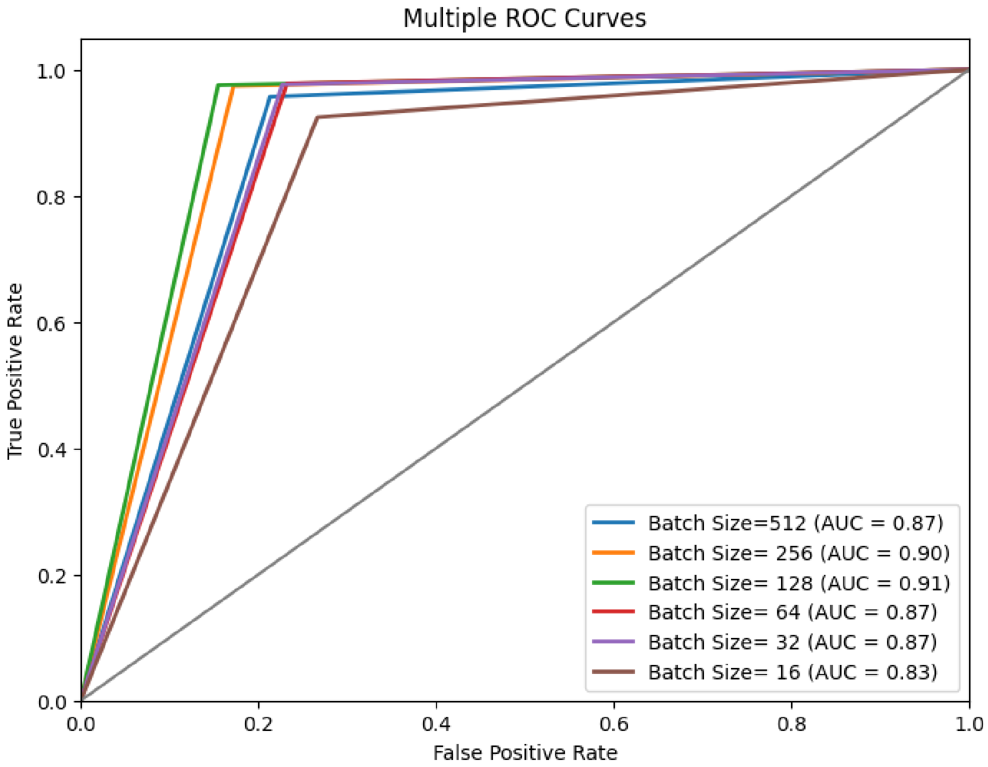


Figure 10. ROC plots by different batch size for subject 330.

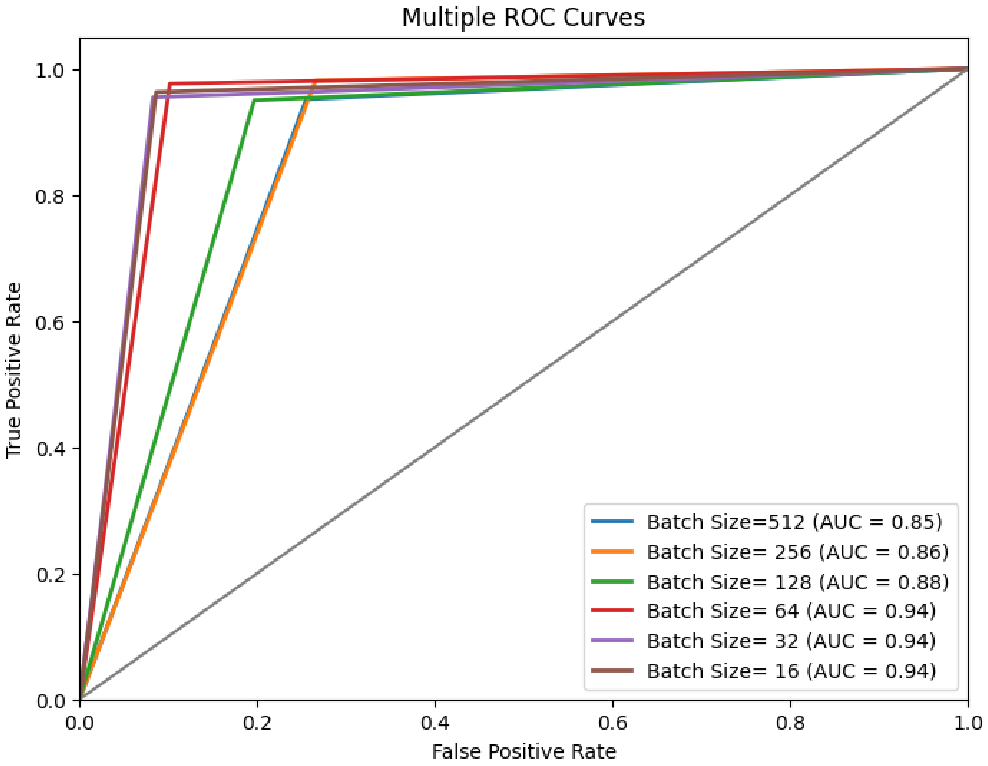


Figure 11. ROC plots by different batch size for subject 441.

Comparisons of performance measures (accuracy, recall, and F1-score) between the proposed model and two existing unsupervised models, namely the One-Class Support Vector Machine (OCSVM) and the Isolation Forest, applied to both subjects 330 and 441 are shown in Tables 2 and 3, respectively.

Table 2. Comparison with existing models on subject 330.

Model	Accuracy	Recall	F1-Score
OCSVM	0.79	0.81	0.79
Isolation Forest	0.81	0.82	0.77
Proposed Model	0.92	0.96	0.95

Table 3. Comparison with existing model on subject 441.

Model	Accuracy	Recall	F1-Score
OCSVM	0.76	0.80	0.76
Isolation Forest	0.81	0.82	0.79
Proposed Model	0.93	0.96	0.97

Table 2 reports that the proposed model outperforms the OCSVM and isolation forest models for subject 330 by a considerable margin. The proposed model achieves an accuracy of 0.92, a recall of 0.96, and an F1-score of 0.95. However, the OCSVM model performs worse, achieving lower accuracy and recall than the isolation forest. The isolation forest's accuracy of 0.81 and recall of 0.82 are marginally higher than OCSVM's, but its F1-score of 0.77 is marginally lower than OCSVM's. This implies that, compared to the other models, the proposed model performs better overall and achieves a more balanced trade-off between precision and recall for subject 330, yielding improved results across all measures.

Similar results are observed for subject 441 in Table 3, where the proposed model once again outperforms the existing models. It achieves the highest F1-score of 0.97, accuracy of 0.93, and recall of 0.96. The isolation forest again outperforms OCSVM for the other models, while it is still far worse than the proposed model.

5. Discussion

Wireless Body Area Networks (WBANs) have enabled numerous applications in human–computer interaction, sports monitoring, and healthcare. However, ensuring the security and reliability of data transmitted over WBANs remains a significant challenge, particularly in identifying abnormal activity or behavior. In this study, we specifically designed and evaluated an autoencoder neural network-based anomaly detection system for WBANs.

Our findings demonstrate that autoencoder neural networks are an effective tool for detecting anomalies in WBAN data streams. The results shown in Figures 6–9 prove the effectiveness of the proposed model in detecting anomalous observations, with the F1-score reaching 0.97 for subject 441 and 0.96 for subject 330. Such scores for an unsupervised approach without prior labeled data indicate the efficacy and novelty of this model. In addition, the comparisons reported in Tables 2 and 3 demonstrate how well the proposed model generalizes to different data samples, which makes it a more reliable choice for these datasets. They further indicate how the proposed model far outperformed the most common unsupervised models used for detecting anomalies in the same domain in the literature.

By leveraging the unsupervised learning capabilities of autoencoders, our model learns to reconstruct typical physiological signals and identify anomalies that indicate unusual events or conditions. This approach is beneficial for real-world deployments where obtaining labeled data for training is either prohibitively expensive or impractical, as the model can learn meaningful representations of WBAN data without requiring labeled anomalous examples. Utilizing a generic autoencoder architecture, our model captures intricate patterns and dependencies found in various physiological signal sources, providing a flexible solution for anomaly detection across a wide range of monitoring and healthcare applications.

Moreover, our proposed model can be easily adapted to support different levels of anomaly detection granularity, from single-sensor anomalies to higher-level event-based abnormalities involving multiple sensors. This adaptability is crucial for addressing the diverse needs of WBAN applications, as anomalies may occur at different temporal and spatial scales.

6. Conclusions

Our work presents a viable method for improving the security and reliability of WBAN-enabled applications in healthcare and beyond, utilizing autoencoder neural networks. These networks form the foundation of our innovative approach to anomaly detection in wireless body area networks. By addressing critical challenges in anomaly detection and WBAN data processing, our proposed model enhances the capabilities and usability of WBAN technology in promoting human health and well-being.

Although our study yielded promising results, several research avenues remain to be explored. First and foremost, testing our framework's resilience to hostile attacks and noisy environments is crucial for ensuring its dependability in real-world scenarios. Additionally, investigating semi-supervised and transfer learning techniques to leverage pre-trained models from related domains or labeled anomalous data could further improve the performance of our anomaly detection framework. Moreover, validating our proposed framework's practical usefulness and efficacy in clinical settings will require its integration into WBAN-enabled healthcare systems and conducting real-world validation studies involving human participants. Finally, continuous monitoring and updating with new data streams and evolving anomaly patterns will be necessary to ensure that our model remains effective and adaptable over time.

Funding: This research received no external funding.

Data Availability Statement: The original data used in the study are openly available in [Physionet] at [https://www.physionet.org/content/mimicdb/1.0.0]. The implementation code will be available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Santos, M.A.G.; Munoz, R.; Olivares, R.; Rebouças Filho, P.P.; Del Ser, J.; de Albuquerque, V.H.C. Online Heart Monitoring Systems on the Internet of Health Things Environments: A Survey, a Reference Model and an Outlook. *Inf. Fusion* **2020**, *53*, 222–239. [CrossRef]
2. Albattah, A.; Rassam, M.A. A Correlation-Based Anomaly Detection Model for Wireless Body Area Networks Using Convolutional Long Short-Term Memory Neural Network. *Sensors* **2022**, *22*, 1951. [CrossRef] [PubMed]
3. Khan, F.A.; Haldar, N.A.H.; Ali, A.; Iftikhar, M.; Zia, T.A.; Zomaya, A.Y. A Continuous Change Detection Mechanism to Identify Anomalies in ECG Signals for WBAN-Based Healthcare Environments. *IEEE Access* **2017**, *5*, 13531–13544. [CrossRef]
4. Lau, B.C.P.; Ma, E.W.M.; Chow, T.W.S. Probabilistic Fault Detector for Wireless Sensor Network. *Expert. Syst. Appl.* **2014**, *41*, 3703–3711. [CrossRef]
5. Mohamed, M.B.; Makhlouf, A.M.; Fakhfakh, A. Correlation for Efficient Anomaly Detection in Medical Environment. In Proceedings of the 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC), Limassol, Cyprus, 25–29 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 548–553.
6. Al-Mishmish, H.; Alkhayyat, A.; Rahim, H.A.; Hammood, D.A.; Ahmad, R.B.; Abbasi, Q.H. Critical Data-Based Incremental Cooperative Communication for Wireless Body Area Network. *Sensors* **2018**, *18*, 3661. [CrossRef]
7. Haque, S.A.; Rahman, M.; Aziz, S.M. Sensor Anomaly Detection in Wireless Sensor Networks for Healthcare. *Sensors* **2015**, *15*, 8764–8786. [CrossRef] [PubMed]
8. Pachauri, G.; Sharma, S. Anomaly Detection in Medical Wireless Sensor Networks Using Machine Learning Algorithms. *Procedia Comput. Sci.* **2015**, *70*, 325–333. [CrossRef]
9. Qu, H.; Lei, L.; Tang, X.; Wang, P. A Lightweight Intrusion Detection Method Based on Fuzzy Clustering Algorithm for Wireless Sensor Networks. *Adv. Fuzzy Syst.* **2018**, *2018*, 4071851. [CrossRef]
10. Zhang, H.; Liu, J.; Pang, A.-C. A Bayesian Network Model for Data Losses and Faults in Medical Body Sensor Networks. *Comput. Netw.* **2018**, *143*, 166–175. [CrossRef]
11. Luo, T.; Nagarajan, S.G. Distributed Anomaly Detection Using Autoencoder Neural Networks in WSN for IoT. In Proceedings of the 2018 IEEE International Conference on Communications (icc), Kansas, MI, USA, 20–24 May 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.

12. Malan, D.J.; Fulford-Jones, T.; Welsh, M.; Moulton, S. Codeblue: An Ad Hoc Sensor Network Infrastructure for Emergency Medical Care. In Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks, London, UK, 6–7 April 2004.
13. Ko, J.; Lim, J.H.; Chen, Y.; Musvaloiu-E, R.; Terzis, A.; Masson, G.M.; Gao, T.; Destler, W.; Selavo, L.; Dutton, R.P. MEDiSN: Medical Emergency Detection in Sensor Networks. *ACM Trans. Embed. Comput. Syst.* **2010**, *10*, 1–29. [CrossRef]
14. Cunha, J.P.S.; Cunha, B.; Pereira, A.S.; Xavier, W.; Ferreira, N.; Meireles, L. Vital-Jacket$^{®}$: A Wearable Wireless Vital Signs Monitor for Patients' Mobility in Cardiology and Sports. In Proceedings of the 2010 4th International Conference on Pervasive Computing Technologies for Healthcare, Munchen Germany, 22–25 March 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 1–2.
15. Navarro, K.F.; Lawrence, E.; Lim, B. Medical MoteCare: A Distributed Personal Healthcare Monitoring System. In Proceedings of the 2009 International Conference on eHealth, Telemedicine, and Social Medicine, Cancun, Mexico, 1–7 February 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 25–30.
16. Alemdar, H.; Ersoy, C. Wireless Sensor Networks for Healthcare: A Survey. *Comput. Netw.* **2010**, *54*, 2688–2710. [CrossRef]
17. Grgić, K.; Žagar, D.; Križanović, V. Medical Applications of Wireless Sensor Networks-Current Status and Future Directions. *Med. Glas.* **2012**, *9*, 23–31.
18. Bettencourt, L.M.A.; Hagberg, A.A.; Larkey, L.B. Separating the Wheat from the Chaff: Practical Anomaly Detection Schemes in Ecological Applications of Distributed Sensor Networks. In Proceedings of the Distributed Computing in Sensor Systems: Third IEEE International Conference, DCOSS 2007, Santa Fe, NM, USA, 18–20 June 2007; Proceedings 3. Springer: Berlin/Heidelberg, Germany, 2007; pp. 223–239.
19. O'Reilly, C.; Gluhak, A.; Imran, M.A.; Rajasegarar, S. Anomaly Detection in Wireless Sensor Networks in a Non-Stationary Environment. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 1413–1432. [CrossRef]
20. Shahid, N.; Naqvi, I.H.; Qaisar, S. Bin One-Class Support Vector Machines: Analysis of Outlier Detection for Wireless Sensor Networks in Harsh Environments. *Artif. Intell. Rev.* **2015**, *43*, 515–563. [CrossRef]
21. Subramaniam, S.; Palpanas, T.; Papadopoulos, D.; Kalogeraki, V.; Gunopulos, D. Online Outlier Detection in Sensor Data Using Non-Parametric Models. In Proceedings of the 32nd International Conference on Very Large Data Bases, Seoul, Korea, 12–15 September 2006; pp. 187–198.
22. Zhang, Y.; Hamm, N.A.S.; Meratnia, N.; Stein, A.; Van de Voort, M.; Havinga, P.J.M. Statistics-Based Outlier Detection for Wireless Sensor Networks. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 1373–1392. [CrossRef]
23. Zhang, Y.; Meratnia, N.; Havinga, P. Adaptive and Online One-Class Support Vector Machine-Based Outlier Detection Techniques for Wireless Sensor Networks. In Proceedings of the 2009 International Conference on Advanced Information Networking and Applications Workshops, Bradford, UK, 26–29 May 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 990–995.
24. GS, S.; Balakrishnan, R. A Statistical-Based Lightweight Anomaly Detection Framework for Wireless Body Area Networks. *Comput. J.* **2022**, *65*, 1752–1759.
25. Salem, O.; Alsubhi, K.; Mehaoua, A.; Boutaba, R. Markov Models for Anomaly Detection in Wireless Body Area Networks for Secure Health Monitoring. *IEEE J. Sel. Areas Commun.* **2020**, *39*, 526–540. [CrossRef]
26. Arfaoui, A.; Kribeche, A.; Senouci, S.M.; Hamdi, M. Game-Based Adaptive Anomaly Detection in Wireless Body Area Networks. *Comput. Netw.* **2019**, *163*, 106870. [CrossRef]
27. Arfaoui, A.; Kribeche, A.; Senouci, S.-M. Context-Aware Anonymous Authentication Protocols in the Internet of Things Dedicated to e-Health Applications. *Comput. Netw.* **2019**, *159*, 23–36. [CrossRef]
28. Salem, O.; Serhrouchni, A.; Mehaoua, A.; Boutaba, R. Event Detection in Wireless Body Area Networks Using Kalman Filter and Power Divergence. *IEEE Trans. Netw. Serv. Manag.* **2018**, *15*, 1018–1034. [CrossRef]
29. Saneja, B.; Rani, R. An Integrated Framework for Anomaly Detection in Big Data of Medical Wireless Sensors. *Mod. Phys. Lett. B* **2018**, *32*, 1850283. [CrossRef]
30. Nugroho, L.E.; Lazuardi, L.; Prabuwono, A.S. Detection of Anomalous Vital Sign of Elderly Using Hybrid K-Means Clustering and Isolation Forest. In Proceedings of the TENCON 2018—2018 IEEE Region 10 Conference, Jeju, South Korea, 28–31 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 913–918.
31. Smrithy, G.S.; Balakrishnan, R.; Sivakumar, N. Anomaly Detection Using Dynamic Sliding Window in Wireless Body Area Networks. In *Proceedings of the Data Science and Big Data Analytics: ACM-WIR 2018*; Springer: Singapore, 2018; pp. 99–108.
32. Zhang, Y.; Chen, Y.; Wang, J.; Pan, Z. Unsupervised Deep Anomaly Detection for Multi-Sensor Time-Series Signals. *IEEE Trans. Knowl. Data Eng.* **2021**, *35*, 2118–2132. [CrossRef]
33. Tien, C.-W.; Huang, T.-Y.; Chen, P.-C.; Wang, J.-H. Using Autoencoders for Anomaly Detection and Transfer Learning in IoT. *Computers* **2021**, *10*, 88. [CrossRef]
34. Merrill, N.; Eskandarian, A. Modified Autoencoder Training and Scoring for Robust Unsupervised Anomaly Detection in Deep Learning. *IEEE Access* **2020**, *8*, 101824–101833. [CrossRef]
35. Olatinwo, D.D.; Abu-Mahfouz, A.; Hancke, G.; Myburgh, H. IoT-Enabled WBAN and Machine Learning for Speech Emotion Recognition in Patients. *Sensors* **2023**, *23*, 2948. [CrossRef]
36. Rao, V.A.; Rao, R.; Hota, C. Anomaly Detection in Wireless Body Area Networks Using Generative Adversarial Networks. In Proceedings of the 2024 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), Bali, Indonesia, 4–6 July 2024; pp. 60–65.

37. Hinton, G.E.; Zemel, R.S. Autoencoders, Minimum Description Length and Helmholtz Free Energy. *Adv. Neural Inf. Process. Syst.* **1993**, *6*, 3–10.
38. Kramer, M.A. Nonlinear Principal Component Analysis Using Autoassociative Neural Networks. *AIChE J.* **1991**, *37*, 233–243. [CrossRef]
39. Torabi, H.; Mirtaheri, S.L.; Greco, S. Practical Autoencoder Based Anomaly Detection by Using Vector Reconstruction Error. *Cybersecurity* **2023**, *6*, 1. [CrossRef]
40. MIMIC Datasets. Available online: https://www.physionet.org/content/mimicdb/1.0.0/ (accessed on 7 January 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.