electronics

MDPI

Article

A Hybrid GAN-Based DL Approach for the Automatic Detection of Shockable Rhythms in AED for Solving Imbalanced Data Problems

Kamana Dahal and Mohd. Hasan Ali *

Department of Electrical and Computer Engineering, The University of Memphis, Memphis, TN 38152, USA
* Correspondence: mhali@memphis.edu

Abstract: Sudden cardiac arrest (SCA) is one of the global health issues causing high mortality. Hence, timely and agile detection of such arrests and immediate defibrillation support to SCA victims is of the utmost importance. An automated external defibrillator (AED) is a medical device used to treat patients suffering from SCA by delivering an electric shock. An AED implements the machine learning (ML)- or deep learning (DL)-based approach to detect whether the patient needs an electric shock and then automates the shock if needed. However, the effectiveness of these models has relied on the availability of well-balanced data in class distribution. Due to privacy concerns, collecting sufficient data is more challenging in the medical domain. Generative adversarial networks (GAN) have been successfully used to create synthetic data and are far better than standard oversampling techniques in maintaining the original data's probability distribution. We, therefore, proposed a GAN-based DL approach, external classifier–Wasserstein conditional generative adversarial network (EC–WCGAN), to detect the shockable rhythms in an AED on an imbalanced ECG dataset. Our experiments demonstrate that the classifier trained with real and generated data via the EC–WCGAN significantly improves the performance metrics on the imbalanced dataset. Additionally, the WCGAN for generating synthetic data outperformed the standard oversampling technique, such as adaptive synthetic (ADASYN). In addition, our model achieved a high sensitivity, specificity, and F1-score (more than 99%) and a low balanced error rate (0.005) on the balanced 4-s segmented public Holter databases, meeting the American Health Association criteria for AEDs.

Keywords: AED; SCA; imbalanced data; tabular synthetic data generation; GAN; EC–GAN; WC-GAN; ADASYN

check for updates

Citation: Dahal, K.; Ali, M.H. A Hybrid GAN-Based DL Approach for the Automatic Detection of Shockable Rhythms in AED for Solving Imbalanced Data Problems. *Electronics* **2023**, *12*, 13. https://doi.org/10.3390/electronics12010013

Academic Editors: Antoni Morell and Amir Mosavi

Received: 6 October 2022
Revised: 15 December 2022
Accepted: 17 December 2022
Published: 20 December 2022

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

An AED is a medical device used for treating people suffering from SCA through defibrillation [1,2]. SCA is a serious public health issue in which the heart stops beating unexpectedly and abruptly, leading to patients' death if treatment is not provided immediately. Furthermore, more treatment delay means more chance of a patient's death. The recovery rate of victims decreases to 5–10% when defibrillation is delayed over 10 min [3]. Hence, the person suffering from SCA needs immediate treatment through an AED. An AED can treat only those SCA victims who have shockable heart rhythms. It first analyses SCA victims' electrocardiogram (ECG) or heart rhythms to detect whether they belong to shockable or non-shockable rhythms. Shockable rhythms are abnormal heart rhythms that can be treated with fibrillation. If the heart rhythm is shockable, only after that can an AED provide an electrical shock to restore the normal rhythm [4]. In contrast, non-shockable rhythms are those rhythms that have a minimal chance of defibrillation and hence cannot be treated with an AED. Additionally, if the patient has a non-shockable heart rhythm, such as asystole or pulseless electrical activity (PEA), in that case, the patient needs to be treated with a ventilator, chest compression, and some medications, so taking them to the hospital is essential [5]. As the response time is critical for SCA victims, the response time can be

reduced immediately using an AED, which is portable, easy to use, and readily available in public areas. In addition, immediate detection of rhythms that cause cardiac arrests is essential for an AED to decide whether to deliver an electric shock or not to the patients.

Even though AEDs have contributed to better survival of out-of-hospital cardiac arrest victims, there have been some reports of their malfunctioning [6]. SCA victims' unexpected deaths can be reduced by reducing AED's equipment errors and time delays. As the AED's malfunction occasionally occurs, the total number of AED malfunctions is smaller than the number of lives saved [7]. Additionally, the time needed to detect whether the patient's heart rhythm is shockable should be less than seconds. Hence, for using an AED, we need to implement the models that give the most minimal error or no error to detect the shockable rhythm within a short time (less than 1 s) and deliver defibrillation to the patient as soon as possible.

Before 2017, only machine learning (ML) algorithms were used for detecting abnormal rhythms in an AED. The conventional ML approaches to identifying shockable rhythms include preprocessing, extracting features, selecting features, and classification carried out by independent algorithms [8]. The deep learning (DL) approaches have been used since 2017 for discriminating between shockable and non-shockable rhythms. The DL approaches outperform the classical feature extraction algorithms [9]. One of the prime benefits of DL models is their ability to develop discriminating features while utilizing all of the information included in the ECG [10]. This saves time and, more significantly, enhances the quality of the extracted features by avoiding time-consuming feature extraction operations. Additionally, nowadays, a hybrid model (DL and ML/DL) is more famous for this task. In the hybrid model, usually DL is used for feature extraction, whereas DL/ML is used for classification. Acharya et al. [11] proposed a novel model for the automated classification of 2 s segmented ECG signals into shockable (Sh) and non-shockable (NSh) ventricular arrhythmias using a convolution neural network (CNN). The input signals were first preprocessed for noise removal and then fed to the 11-layer CNN model for classification. The proposed CNN model is not only used for classification but also feature extraction and feature selection. They used three databases, namely, the Creighton University Ventricular Tachyarrhythmia Database (CUDB), the MIT-BIH arrhythmia database (MITDB), and the MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB), consisting of a total of 54,096 ECG segments (6001(Sh) + 48,095(NSh)). The proposed system was 10-fold cross-validated. They claimed that their proposed model is highly sensitive for capturing shockable rhythms and has higher specificity, showing a maximum accuracy (Acc) of 93.18%, a sensitivity (Se) of 95.32%, and a specificity (Sp) of 91.04%. Nguyen et al. [12] proposed a novel algorithm for detecting SCA on electrocardiogram (ECG) signals applied to an AED by using a convolution neural network as a feature extractor (CNNE) and a boosting (BS) algorithm as a classifier. The CNNE combines a CNN with an RF classifier applied to extract in-depth features and then it is fed to the boosting classifier to validate its performance using 5-fold cross-validation. The datasets used were CUDB and VFDB as a part of training and validation. Furthermore, the 5-fold CV is applied to both the training and validation phases. The modified variational mode decomposition (MVMD) technique is used to reconstruct the ECG 8 s signals into the NSh signal and Sh signal. Hence, three input signals (preprocessed ECG, Sh, and NSh signals) are applied to the CNNE, and the extracted features were fed to the BS classifier for further classification of SH or NSH rhythms. Their proposed model observed an Acc of 96.26%, an Se of 97.07%, and an Sp of 99.44% for 8 s segments of training data. They claimed that CNNE is less complicated and time-consuming than conventional feature extraction methods. Nguyen et al. [9] proposed a novel feature extraction scheme and shock advice algorithm (SAA) to detect SCA on ECG signals. They applied CNN as a feature extractor, whereas SVM was a classifier for the VFDB and CUDB databases. The MVDB technique reconstructs the Sh and NSh signals from ECG signals of the user databases to produce the three input channels (ECG segments, Sh signal, and NSh signal) implemented to CNN to extract the features. Then. the 5-fold CV procedure is applied to the extracted features and fed to an SVM classifier

to identify a 5 s ECG segment as shockable or non-shockable rhythms. They observed a relatively high Acc of 99.02%, an Se of 95.21%, and an Sp of 99.31%. Picon et al. [13] proposed a model that combines CNN and long short-term memory (LSTM) networks to detect ventricular fibrillation (VF) in an AED shock decision algorithm. They used both public and out-of-hospital cardiac arrest (OHCA) datasets for testing their proposed model and claimed that for 4 s ECG segments, there was an observed Acc of 99.3%, an Se of 99.7%, and an Sp of 98.9% for the public data, and an Acc of 98.0%, an Se of 99.2%, and an Sp of 96.7% for the OHCA data. They claimed their proposed model (CNN–LSTM) is better than only CNN and SVM architecture. By adding the LSTM network, their model learned 20 features that provide higher Se and Sp values. To detect shockable or non-shockable rhythms in an AED, Krasteva et al. [14] proposed an optimized CNN with one to seven CNN layers and five to twenty-three hidden layers. First, their proposed model optimized CNN's hyperparameters and then validated the best hyperparameter setting for short and long (2–10 s) ECG segments. With the OCHA data, they observed an Acc of 99.5%, an Se of 99.6%, and an Sp of 99.4% for a 5 s analyses. With a tolerable drop in performance with a 2 s analysis, they observed an Acc of 98.2%, an Se of 97.6%, and an Sp of 98.7%.

Most public datasets studied in this area are unbalanced because samples of non-shockable rhythms are much larger than shockable rhythms, leading to the DL/ML models being less efficient [15]. DL/ML algorithms are often biased toward the majority class and treat minority examples as outliers of the majority class in the worst case [16]. Due to the minority classes being ignored, the learning algorithm generates a classifier that classifies every example as the majority class. Even though the detection accuracy is better in this condition, the generalization of these models will be poor. Additionally, insufficient data samples for different classes can hamper these models' learning. Hence, these algorithms require a sufficient sample of data to learn for each class. However, all the publications mentioned above had to face the class imbalance problem in the dataset. There are numerous classic methods for generating the minority classes; most researchers used oversampling methods to balance the data. In contrast, some focused on extracting and selecting better features to improve classification results. The standard oversampling methods artificially resample the data and randomly balance the class distribution by replicating minority class examples [17]. However, artificially resampling a large number of samples might cause overfitting. GAN has successfully created realistic synthetic data while maintaining the original data's probability distribution [18]. It generates artificial samples while still resembling real data [19]. These samples can be used to supplement real ones during training. GAN outperforms other oversampling approaches to the data imbalance problem that have been widely employed [20].

Based on the above background, this paper proposes a GAN-based DL approach, EC–WCGAN, for recognizing shockable rhythms in an AED with handling the imbalanced data problem. All the deep learning-based models are biased toward learning the distribution of the dominant class. That is, the performance is constrained for the low sample class. Most AED datasets have low sample shockable rhythms, so the DL model cannot perform well with the low sample data. Additionally, the traditional oversampling model fails miserably. The purpose of applying this model is to generate synthetic data samples for low-sample classes in parallel with classifier training. Further training is based on feedback from the classifier, which leads to improving classifier performance metrics. To compare our proposed model, we also applied the state-of-the-art DL approaches with the adaptive synthetic (ADASYN) oversampling technique to detect shockable rhythms in an AED. This comparison shows how the DL model with the GAN sampling technique (our proposed model) outperformed any DL models with the adaptive synthetic (ADASYN) sampling technique.

This paper has the following novelties and contributions.

- We proposed a generative model along with an external classifier to detect the shockable rhythms.
- We integrated WCGAN with the DL classifier to solve the low-sample class problem.

- We integrated WCGAN with the DL classifier in such a way that it trained the classifier together with the generation. Therefore, it eliminates the training overhead for the classifier.
- We improved shockable rhythms detection for an AED and a classification performance compared to the DL or combined DL and ML models.

The rest of the paper is organized as follows. Section 2 describes the proposed methodology with a detailed explanation of the datasets and state-of-the-art DL algorithms used to compare the proposed model. The results and discussions are provided in Section 3. Section 4 provides the conclusions. Section 5 describes future work.

2. Materials and Methods

We proposed a hybrid GAN-based DL architecture, external classifier–Wasserstein conditional generative adversarial network (EC–WCGAN), for detecting shockable rhythms in an AED on an unbalanced ECG dataset. Our proposed model used the EC–GAN model proposed by Haque [21] by replacing DC–GAN used by the author with the Wasserstein conditional GAN with a gradient penalty (WCGAN–GP), proposed by W. Manhar et al. [22]. The EC–WCGAN comprises a deep neural network (DNN) as an external classifier (EC) to classify heart rhythms into shockable or non-shockable rhythms and the WCGAN with a gradient penalty as a synthetic tabular data generator for the low-class to overcome the class imbalance problem. For GAN, WCGAN–GP, and the proposed EC–WCGAN, the operational algorithm and model topologies are shown in the following sections.

2.1. EC–GAN

Ayaan Haque [21] implemented the EC–GAN model to classify X-ray images on low-class data. EC–GAN is a generative adversarial network with an external classifier. It consists of three separate models: a generator (*G*), a discriminator (*D*), and a classifier (*C*), shown in Figure 1, where the generator takes a random vector (noise) as input and generates synthetic data samples, whereas the discriminator discriminates whether the generated data by the generator is fake or real. The author claimed that the EC–GAN method is unique, where the synthetic image is generated in parallel with classifier training, and feedback from the classifier is used for further training. In contrast, other methods employ the shared architecture paradigm, in which the discriminator serves as both a generator and a classifier. The author used ResNet18 (CNN) as a classifier and deep convolution GAN (DC–GAN) for generating X-ray image samples.

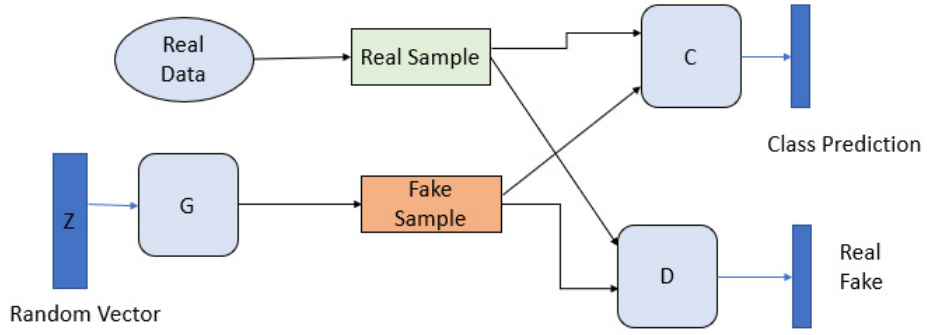


Figure 1. EC–GAN architecture.

The discriminator and generator loss of the standard GAN model is defined by [21]:

$$LD(x,z) = BCE(D(x),1) + BCE(D(G(z)),0) \quad (1)$$

$$LG(z) = BCE(D(G(z)),1) \quad (2)$$

where *BCE* is the binary cross-entropy, *D* is the discriminator, *G* is the generator, x is real data, and z is the random vector. For the discriminator, the first loss component in

Equation (1) trains the discriminator on real data, whereas the second component trains the discriminator on fake samples. The loss component in Equation (2) trains the generator to produce more realistic samples for the generator.

The author implemented an intuitive loss function for the classifier by utilizing both supervised and unsupervised methods, as defined below:

$$L_c(x,\ y,z) = CE(C(x),y)\ CE(C(G(z)),\ argmax(C(G(z))) > t) \quad (3)$$

where x represents real data, y represents real data labels, and z represents fake data. The $\searrow$ is the unsupervised loss weight (adversarial weight), CE means the cross-entropy loss, C is the classifier, $C(x)$ represents the outcome of real data classification, $C(G(z))$ represents the result of fake data classification, and t is the pseudo-labeled threshold. The first component in (3) is the typical cross-entropy loss with actual data and real labels. The second component is the cross-entropy generated between data and hypothesized labels, scaled with adversarial weight ($\searrow$). The pseudo-labeling threshold, t, assures predicted labels are over a given probability threshold.

2.2. WCGAN–GP

Walia et al. [22] proposed the WCGAN–GP model for generating synthetic tabular data. The WCGAN–GP model is an extension of WGAN–GP by inputting the conditional vector or target labels. Further detail of WGAN–GP model is found in the paper [23], where the authors successfully implemented WGAN–GP on largescale image and language datasets. To limit the occurrence of failure modes associated with GANs, WCGAN–GP employs Wasserstein distance and the gradient penalty. In WCGAN–GP, the generator and discriminator/critic are conditioned on extra class label information, known as the conditional vector. Figure 2 shows the structure of WCGAN–GP where, unlike GAN, the conditional vector is provided to both the generator and critic. The input to the generator is noise and the label/conditional vector. The input to the critic is the synthetic data generated by the generator, real tabular data, and the label/conditional vector. Instead of distinguishing samples as real or fake, the critic predicts large values for true samples and small values for fake samples. The loss function used by the WCGAN–GP model is Wasserstein loss. Only the Wasserstein loss is used to update the weights of the generator, whereas the critic employs the Wasserstein loss and the gradient penalty loss for both real and generated samples. When using the Wasserstein loss function, a lower critic loss indicates a superior generator. The gradient penalty (GP) forced the norm of the gradients to be 1 and complied with the 1-Lipschitz constraint for overcoming the training instability of GANs.

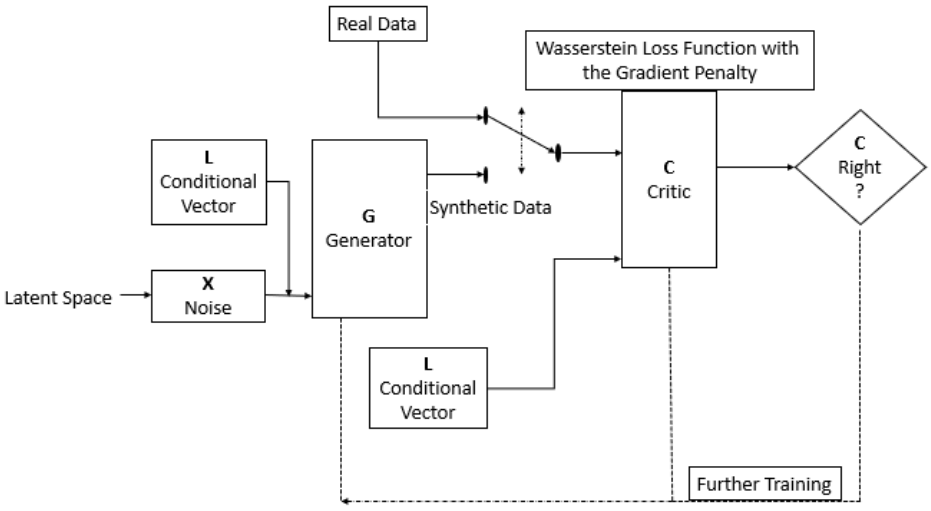


Figure 2. The structure of WCGAN–GP.

2.3. The Proposed EC–WCGAN Method

Figure 3 shows the steps involved in our proposed model. First, according to the common practice of splitting data into train–test, the real ECG dataset was divided into 70% training and 30% testing for our research. Pre-processing and normalization were applied to the raw data to convert the data into a uniform format. The random noise and target label (conditional vector) were provided as inputs to WCGAN–GP to generate synthetic tabular data. The proposed model's training processes are similar to the EC–GAN model, in which synthetic data samples are generated in parallel with classifier training, and further training is based on feedback from the classifier. Hence, the classifier of our model takes both the real training data and fake data from WCGAN–GP to train it. As the real data is unbalanced, the WCGAN–GP generates synthetic data to overcome the problem of unbalanced data. Hence, the classifier becomes more robust as it is trained with sufficient (real and generated/fake) data. For testing, we have applied two approaches. The first one used a trained classifier model on real test data, and another used a trained classifier model on real and generated test data. The classifier output should be either SHOCK or NO SHOCK. If it is SHOCK, then the AED provides the electric shock to the patients; otherwise, it will not.

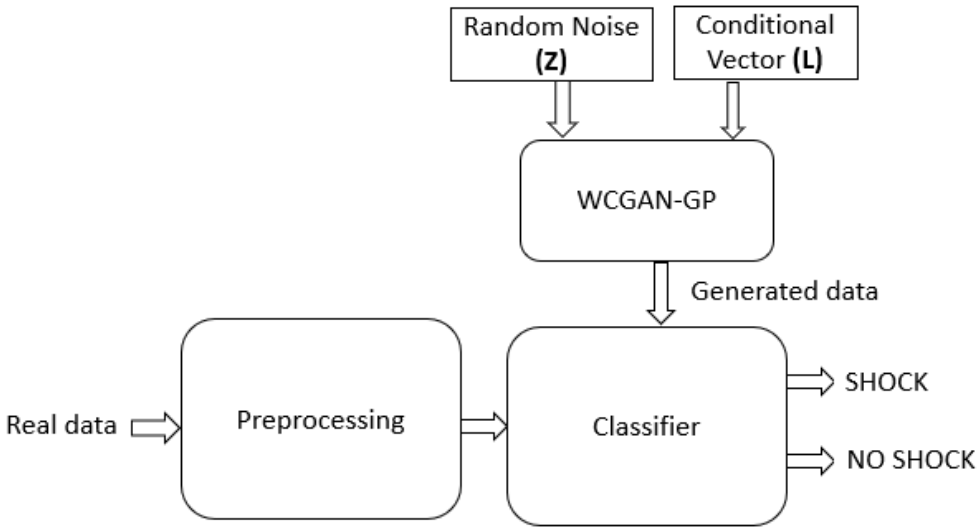


Figure 3. The steps involved in the proposed EC–WCGAN model.

Google collab with Python 3.7 version was used to run this experiment. The code was implemented using Python libraries such as TensorFlow, NumPy, Pandas, Matplotlib, and Keras. The computational specs for the algorithm are 32 GB RAM using Google Collaboratory Pro. The code sources for our work are found in [19,24]. The proposed model implemented the same hyperparameter tunning and network architecture used by WCGAN–GP [22]. However, we added an extra classifier model similar to EC–GAN [21]. The following network architecture is implemented in the proposed model.

All three neural networks (generator, critic, and classifier) have different architectures. Three hidden layers exist in these neural networks. The generator used hidden units (256, 512, and 1024) with 0.3 dropouts; the critic used hidden units (1024, 512, and 256); and the classifier used hidden units (128, 256, and 128) with 0.3 dropouts. The classifier for binary and multiclass classification uses the same neural architecture. However, the classifier employs the standard "ReLU" activation function, and the generator and critic use the "LeakyReLU" activation function with a negative slope coefficient (α) of 0.2 for hidden layers. The classifier employs "softmax" activation for its terminal layers and the generator and critic use the "tanh" activation. The parameters of the model have been optimized using the Adam optimizer. The model was trained with a batch size of 128, a number of

critic parameter updates per generator (n critic) of 5, a noise dimension of 30, a confidence threshold of 0.2, and an adversarial weight of 0.1.

The complexity of the proposed model is analyzed with training and testing time. The training time for the proposed model to achieve almost 100% detection accuracy is 90.202 s within 50 epochs. The real detection time, the test time for the proposed model, is 10.62 ms.

2.4. Dataset

The following three public databases were used in this research [25].

- The AHA fibrillation database (AHADB) [26], which includes 30 min ECG recordings from 10 patients.
- The Massachusetts Institute of Technology–Beth Israel Hospital (MIT–BIH) malignant ventricular ectopy database (VFDB) [27], which includes 22 half-hour ECG recordings of patients who experienced ventricular tachycardia, ventricular flutter, and ventricular fibrillation.
- The Creighton University (CU) ventricular tachyarrhythmia database (CUDB) [28], which includes 35 eight-minute ECG recordings of people who have undergone sustained ventricular tachycardia, ventricular flutter, and ventricular fibrillation episodes.

These three public datasets in the preprocessed form are available in [29]. A detailed description of dataset preprocessing, labeling, and segmentation can be found in the original paper by Figura et al. [30]. In this paper, the author computed 30 ECG features for each dataset for classification. The 30 features include the following four different types of features.

- Temporal features: to characterize the rhythm's amplitude, slope, sample distribution, or heart rate.
- Spectral features: to quantify spectral concentration, normalized spectral moments, or relative power content in distinct frequency bands
- Time-frequency features: based on the wavelet analysis of the ECG.
- Complexity features: include the Hilbert transform, sample entropy, complexity measure, covariance, etc.

Table 1 shows the number of samples for each class (Sh/NSh) for three public dataset datasets. We used these datasets because they are publicly available and also already used by many research articles, such as [9,11–13], for detecting shockable rhythms. The Sh rhythms are responsible for SCA. Due to low data samples for shockable rhythms in these datasets, only one dataset is insufficient for better performance of DL algorithms. Hence, we combined these three ECG datasets to increase the number of samples, especially for Sh rhythms. The combined dataset contains 3578 shockable (Sh) and 14,995 non-shockable (NSh) rhythm data samples.

Table 1. A description of the datasets used for classification.

Public Database	Patients	Sh	NSh
VFDB	22	1586	7761
CUDB	35	716	2986
AHADB	10	1276	3748
Total	67	3578	14,495

Thirteen different rhythms available in the combined dataset are shown in Table 2. This table shows the data distribution for each label of the combined dataset. In the combined dataset, labels 10, 6, 2, 11, 16, 13, 8, 5, 18, 15, and 2 belong to NSh rhythms or Class 0, whereas labels 19, 20, and 21 belong to Sh rhythms or Class 1. The normal sinus rhythm (NSR), asystole, atrial fibrillation, supraventricular tachycardia, sinus bradycardia, atrial flutter, and pulseless electrical activity are non-shockable rhythms. In contrast, coarse ventricular fibrillation and rapid ventricular tachycardia are examples of shockable rhythms [31]. Class

10 has a higher distribution (61.4%), followed by Class 19 (14.9%), whereas the other classes have very few data samples, with less than 5.8% data distribution. Hence, we need to generate synthetic data for the low class. To balance the dataset by generating more realistic synthetic data for minority class samples, the proposed model was applied to WCGAN–GP.

Table 2. Data description for each label.

Label	Count	Ratio (%)
10	11,098	61.40
19	2698	14.92
6	1028	5.68
2	855	4.73
21	780	4.31
11	433	2.39
16	373	2.06
13	364	2.01
8	203	1.12
5	125	0.69
20	100	0.55
18	10	0.05
15	6	0.03

2.5. The Deep Learning Models with Adaptive Synthetics (ADASYN)

This part of the work explains the deep learning algorithms implemented to compare with the proposed model EC–WCGAN. The reason for the comparison is to demonstrate the superiority of the proposed model to other state-of-the-art DL with ADASYN in terms of shockable rhythms detection and classification. Hence, for comparison, we applied the ADASYN sampling approach for generating synthetic data and then used deep learning approaches on a balanced dataset to detect shockable rhythms. The ADASYN is based on producing minority data samples in an adaptable manner based on their distributions [32]. It is used to generate synthetic data for minority-class examples that are more difficult to learn than minority-class examples that are simpler to learn. Some publications used the ADASYN sampling method successfully to deal with imbalanced datasets are [33–35]. After balancing the dataset using ADASYN, we applied the following deep learning algorithms to detect shockable rhythms in an AED.

2.5.1. Deep Neural Network (DNN)

DNN is based on the concept of artificial neural networks (ANN), which are designed to process large amounts of data through numerous layers of neurons to perform complicated analyses [36]. It is a neural network with more than one hidden layer [37]. Every node is linked to every other node and each link between two neurons has a weight. The output node (nodes) specifies the output of the output node for the input data provided, bypassing the output via the activation function.

2.5.2. Convolution Neural Network (CNN)

A convolution neural network (CNN) is a class of artificial neural networks that automatically and adaptively learns the hierarchical collection of features [38]. Indeed, it is an improvised neural network version consisting of multiple layers connected back-to-back in a feedforward manner. The main three layers used for extracting features are convolution, normalization, and pooling, whereas the fully-connected layer is used for the classification [39]. As the big data age progresses, a CNN with more hidden layers has a more complex network structure, more efficient feature learning, and feature speech abilities than conventional ML approaches [40].

2.5.3. Deep Convolution Neural Network (DCNN)

DCCNs are traditional artificial neural networks that use a three-dimensional neural pattern [36]. DCNNs' layering is what makes them so effective.

2.5.4. Long Short-Term Memory (LSTM)

Long short-term memory (LSTM) [41] is a recurrent neural network consisting of memory blocks or a set of recurrently connected blocks, considered as a differential version of a digital computer's memory chips [42]. It is accurate and faster than standard recurrent neural nets (RNNs) and time-windowed multilayer perceptions (MLPs).

2.5.5. Gated Recurrent Unit (GRU)

A GRU, a simpler version of LSTMs, is a recurrent neural network that relieves the problem of vanishing gradients in recurrent neural networks [43]. To avoid the vanishing gradient problem of a standard RNN, a GRU uses the so-called update gate (that controls the preservation of the last memory introduced) and reset gate (that adjusts the incorporation of new input with the previous memory).

2.5.6. Recurrent Neural Network (RNN)

Feedforward neural networks are unidirectional, where the outputs of one layer are transmitted to the following layer [37]. In these feedforward networks, past data cannot be stored. A RNN is a type of neural network that uses hidden layer loops to save information from previous time steps to predict the value of the current time step. The vanishing gradient problem is a critical flaw in the most basic RNN model, preventing it from being accurate [44].

3. Results and Discussion

In this section, we investigate and analyze the performance of the proposed model and several DL models to detect the shockable rhythms in an AED for different public datasets. This section also analyzed the confidence intervals for sensitivity and specificity using Wilson's method.

3.1. Plot-Based Responses

As there are three separate models (the critic/discriminator, generator, and classifier) in our proposed method, the loss for each model is shown in Figure 4. The plot shows that the generator and discriminator have a significantly stable loss after dropping in a few batches. The classifier, critic, and generator loss are stable around 0, −2, and −2.5, respectively.

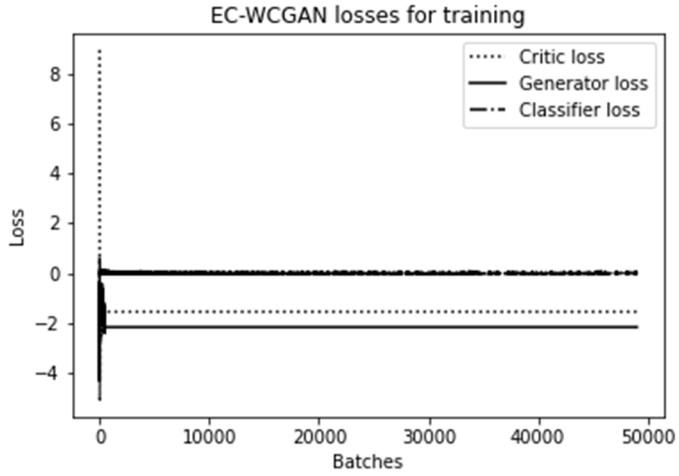


Figure 4. The critic, generator, and classifier loss for the EC–WCGAN model.

The proposed model trains the classifier with real and generated or fake data. The average accuracy is the average of real and fake accuracy, and the average accuracy is stable after a few batches, as shown in Figure 5. On the other hand, the accuracy for DL algorithms, such as a DNN with ADASYN, shows overfitting, as shown in Figure 6.

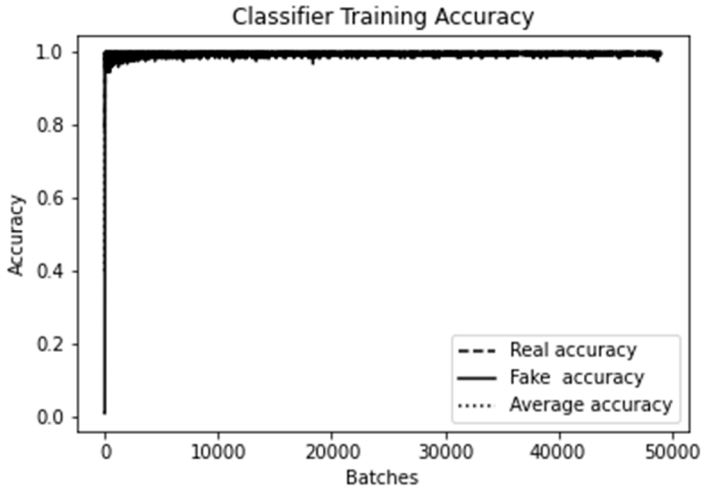


Figure 5. Classifier accuracy for the EC–WCGAN model.

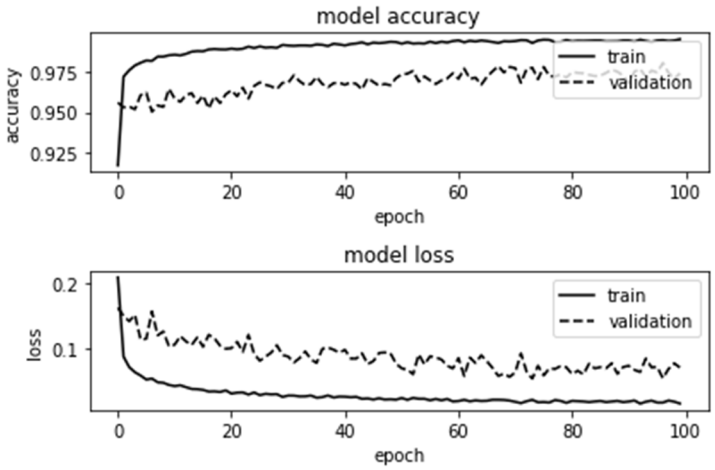


Figure 6. Classifier accuracy vs. loss using a DNN with ADASYN.

An AED's effectiveness depends on its ability to detect shockable (SH) rhythms on the test data and how correctly the operator can use them [45]. An AED device's accuracy is measured by quantifying the used model's performance to classify Sh or NSh rhythms.

Accuracy: It calculates the percentage of correctly categorized data across all datasets [46]. The better the DL model, the higher the accuracy. The accuracy is expressed in percentages (0–100%).

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (4)$$

For the evaluation of our proposed model, we tested the proposed model on real test data (30% of the original dataset). We achieved an accuracy of 99.43%, higher than other DL models with ADASYN, as shown in Figure 7. The DL algorithms with ADASYN achieved a maximum accuracy of 96.29% with a CNN.

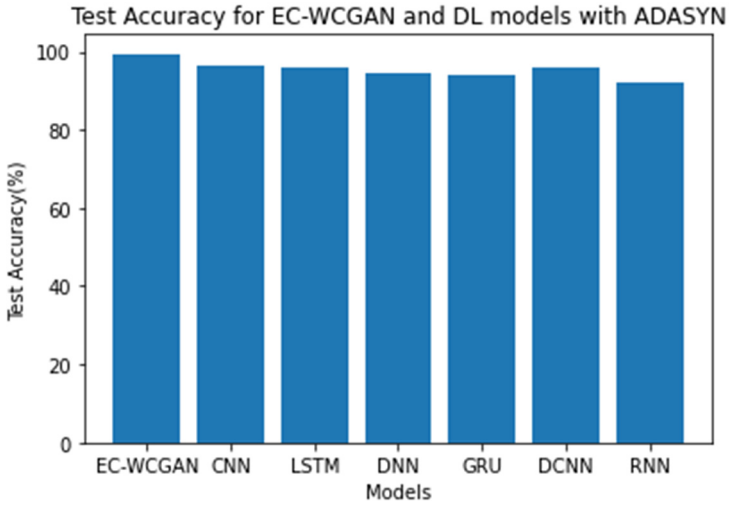


Figure 7. Testing accuracy of the EC–WCGAN and DL algorithms with ADASYN.

3.2. Performances Metrics

Specificity and sensitivity are the most widely used metrics among various performance metrics for an AED and are defined as follows [31].

- Sensitivity: Sensitivity is the probability of shock advised for patients who truly have shockable rhythms rhythm.
- Specificity: Specificity is the probability of no shock advised for patients with non-shockable rhythms.

"The American Health Association recommends a sensitivity (Se) higher than 90% for shockable rhythms, and a specificity higher than 95% for non-shockable rhythms, and above 99% in the case of normal sinus rhythms" [30].

- BER: BER represents the balanced error rate, a balanced statistic that considers shockable and non-shockable rhythm detection errors equally.

The confusion matrix in Table 3 measures how efficiently a specific algorithm classifies the actual data. The performance parameters TP, FP, FN, and TN, are explained as follows [31]. TP: a shock is correctly advised for a shockable rhythm, FP: a shock is incorrectly advised for a non-shockable rhythm, FN: no shock is advised for a shockable rhythm, and TN: no shock is advised for a non-shockable rhythm The performance metrics that have been considered in this paper are defined below in equation form [10,30]:

$$Sensitivity\ (Se) = \frac{TP}{TP + FN} \tag{5}$$

$$Specificity\ (Sp) = \frac{TN}{TN + FP} \tag{6}$$

$$BER = 1 - \frac{1}{2}(Se + Sp) \tag{7}$$

Table 3. A description of the datasets used for classification.

		Shockable (Sh)	**Non-Shockable (NSh)**
AED algorithm decision	Shock No Shock	True Positive (TP) False Negative (FN)	False Positive (FP) True Negative (TN)

Our article reports sensitivity, specificity, and BER in Table 4 to access the proposed model and DL algorithms, such as the CNN, DNN, DCNN, LSTM, GRU, and LSTM with

the ADASYN sampling technique, explained in Section 2.5. The purpose of the performance comparison of the proposed model with other deep learning algorithms is also to show how important it is to generate realistic data using GAN rather than standard sampling techniques for detecting shockable rhythms in an AED on an unbalanced ECG dataset. As our model is DL (DNN) with GAN, we can say that the DL performance with GAN is better than the DL performance with ADASYN in terms of sensitivity, specificity, and BER in Table 4. We achieved more than 99% for sensitivity, specificity, and a BER of 0.005 when using real and generated test data during testing. However, our model also tested on only real test data and achieved a higher specificity of more than 99% and a low BER of 0.01 compared to other DL models. On the top, the proposed model met AHA's criteria even without using generated data for testing. In addition, only the CNN meets the AHA's target on balanced datasets using ADASYN. Hence, each model can perform well only when the shockable and non-shockable rhythm classes are balanced using GAN-generated data samples.

Table 4. A performance comparison for EC–WCGAN and six DL algorithms with ADASYN.

Models	Se (>90%)	Sp (>95%)	BER
EC–WCGAN (Using real and generated test data)	99.16	99.67	0.005
EC–WCGAN (Using only real test data)	96.76	99.54	0.01
CNN	99.90	95.40	0.02
DCNN	99.72	94.98	0.02
DNN	99.72	93.30	0.03
LSTM	99.90	94.75	0.02
GRU	99.81	92.38	0.03
RNN	99.72	90.34	0.04

Other performance matrices from the confusion matrix to quantify the performance of the proposed detection method taken are precision, recall, and F1-score.

- Precision: It estimates the ratio of correctly classified rhythms to the number of all identified rhythms.
- Recall: It estimates the ratio of correctly classified non-shockable rhythms to all non-shockable rhythms.
- F1-Score: It is the harmonic mean of precision and recall. A higher value of the F1-score represents a good model, and its value lies between 0 and 1.

The formula of precision, recall, and F1-score is given in Equations (5), (6), and (7), respectively [46].

$$Precision\ (P) = \frac{TP}{TP + FP} \tag{8}$$

$$Recall\ (R) = \frac{TP}{TP + FN} \tag{9}$$

$$F1 - score = 2 \times \frac{P \times R}{P + R} \tag{10}$$

We also calculated the precision, recall, and F1-score for the proposed model and DL models, as listed in Table 5. A higher value of the precision, recall, and F1-score represents a good model, and these values lie between 0 and 1. We achieved higher precision, recall, and F1-score (more than 99%) using generated and real test data, whereas higher precision and F1-score (more than 99%) was achieved when using only real test data, as shown in Table 5. For used DL models with ADASYN, the precision and F1-score are below 85% and 92%, respectively.

Table 5. Classification matrices for EC–WCGAN and six DL algorithms with ADASYN.

Models	Precision	Recall	F1-Score
EC–WCGAN (using real and generated test data)	0.99	0.99	0.99
EC–WCGAN (using only real test data)	0.98	0.96	0.97
CNN	0.84	0.99	0.91
DCNN	0.83	0.99	0.91
DNN	0.79	0.99	0.88
RNN	0.71	0.99	0.83
LSTM	0.81	0.99	0.90
GRU	0.73	0.99	0.84

A binary classification means classifying the ECG signals either as shockable or non-shockable. Table 6 shows the precision, recall, and accuracy for each class. There are only two classes in the used dataset: shockable and non-shockable rhythms. As we have generated shockable rhythms using WCGAN, the shockable rhythm class's precision, recall, and accuracy are also high (more than 99%), similar to the non-shockable rhythm class.

Table 6. Classification metrics for EC–WCGAN for a binary classification.

Proposed Model	Data	Precision	Recall	F1-Score
EC–WCGAN	Shockable rhythm (1)	0.9965	0.9916	0.9940
	Non-shockable rhythm (0)	0.9922	0.9967	0.9945

Our proposed model also successfully distinguished 13 different kinds of rhythms available in the used dataset. The thirteen labels in Table 7 represent thirteen different shockable and non-shockable rhythms in the dataset. We successively applied the proposed model to generate synthetic data for any rhythms. We achieved the precision, recall, and F1-score for each rhythm or label between 0.63 and 0.99, except for label 2. Hence, our model can be applied to any field where binary or multiclass classification needs to be performed and the data has a highly unbalanced distribution.

Table 7. Classification matrices of EC–WCGAN for multiclass classification.

Proposed Model	Label	Precision	Recall	F1-Score
EC–WCGAN	10	0.89	0.97	0.93
	11	0.63	0.99	0.77
	13	0.87	0.95	0.91
	15	0.98	0.98	0.98
	16	0.96	0.99	0.97
	18	0.95	0.99	0.97
	19	0.95	0.95	0.95
	2	0.92	0.07	0.13
	20	0.99	0.99	0.99
	21	0.97	0.93	0.95
	5	0.97	0.98	0.98
	6	0.85	0.99	0.92
	8	0.98	0.97	0.98

3.3. Comparison of EC–WCGAN with Existing Hybrid Models

This section compares our work with existing models that were used in an AED to detect shockable rhythms [25]. We have compared most papers that used the hybrid model (a combination of DL and ML models) with our model.

Table 8 shows the value of Sp, Se, Acc, and BER for different ECG segments and databases for various algorithms (used by other researchers) and our proposed model. Our proposed model met the AHA's target with better performance for detecting shockable rhythms on the unbalanced dataset with or without using generated test data. Furthermore, our model outperformed other models with higher (more than 99%) Acc, Se, Sp, and a lower BER of 0.005 on a 4 s segmented public dataset (a combination of CUDB, VFDB, and AHADB). The DL and ML models suffer from the imbalanced dataset problem if insufficient datasets exist. However, we solved it using WCGAN to generate realistic synthetic data samples. Our proposed model performed better than other algorithms using sufficient data (real and generated).

Table 8. Classification metrics of EC–WCGAN for multiclass classification.

Ref	Type of Method	Approaches	Segments	Acc (%)	Se (%)	Sp (%)	BER	Database
	EC–WCGAN (Using real and generated test data)	DNN, GAN	4 s	99.45	99.18	99.70	0.005	AHADB, CUDB, VFDB
	EC–WCGAN (Using only real test data)	DNN, GAN	4 s	99.45	96.76	99.54	0.018	AHADB, CUDB, VFDB
[11]	DL	CNN	2 s	93.18	95.32	91.04	N/A	MITB, VFDB, CUDB
[12]	DL and ML	CNN, BS, MVMD	8 s	99.26	97.07	99.44	0.017	VFDB, CUDB
[9]	DL and ML	CNN, SVM, MVMD	5 s	99.02	95.21	99.31	0.027	VFDB, CUDB
[13]	DL and ML	CNN, LSTM	4 s 4 s 2 s 2 s	99.3 98.0 95.2 98.1	99.7 99.2 97.5 97.5	98.9 96.7 93.6 97.5	0.007 0.020 N/A N/A	Public OHCA OHCA Public
[14]	DL	DCNN, HP-optimization	5 s 2 s	99.5 98.2	99.6 97.6	99.4 98.7	0.005 0.018	OCHA

3.4. Calculation of Confidence Intervals for Sensitivity and Specificity

For model validation, confidence intervals were used. We computed confidence intervals for sensitivity and specificity using Wilson's method [47,48]. This method avoids a normal approximation and yields precise confidence intervals even for small sample sets. In this method, for binary outcomes, if r is the observed number of subjects with some feature in a sample size of n, the estimated proportion of having the feature is $p = \frac{r}{n}$. The proportion that does not have features is

$q = 1 - p$. Additionally, $z_{1-\frac{\alpha}{2}}$ is the $100\left(1-\frac{\alpha}{2}\right)$ percentile from the standard normal distribution. The three quantities needed to calculate the confidence interval are defined as:

$$A = 2r + z^2 \tag{11}$$

$$B = z\sqrt{z^2 + 4rq} \tag{12}$$

$$C = 2(n + z^2) \tag{13}$$

Now, the confidence interval for a population proportion is given by:

$$\frac{A - B}{C} \; to \; \frac{A + B}{C}$$

With a desired confidence of 95%, we achieved a confidence interval of 98.9161% to 99.4604% for sensitivity and 99.5497% to 99.8593% for specificity when using both real and generated data for testing. While using only real test data for testing, we observed a confidence interval of 95.5117% to 97.6781% for sensitivity and 99.2941% to 99.7035% for specificity. The confidence interval for sensitivity and specificity successfully met the AHA's criteria, leading to the EC–WCGAN-based AED being safe for SCA victims.

4. Conclusions

In this work, we proposed a GAN-based DL model, the EC–WCGAN for an AED, to detect shockable rhythms on imbalanced ECG datasets. We implemented the DNN as an EC to classify shockable and non-shockable rhythms and the WCGAN with a gradient penalty to generate synthetic data for the low class to handle the data imbalance problem. Based on the outcome of the experiments, the proposed model met the American Health Association's criteria for a sensitivity higher than 90% and a specificity higher than 95%. The proposed model outperformed the sampling-based state-of-the-art algorithms, such as ADASYN, for generating more realistic ECG samples. It improved the performance over the widely used DL-based algorithms such as the DNN, DCNN, CNN, LSTM, RNN, and GRU with ADASYN and achieved a BER of 0.005, a sensitivity of 99.16%, and a specificity of 99.67%. Furthermore, the model outperformed existing hybrid algorithms used by other researchers to detect shockable rhythms in an AED in terms of BER with 0.005 on a 4 s segmented public Holter dataset (CUDB, VFDB, and AHADB).

5. Future Work

In our future work, the same detection model (EC–WCGAN) will be applied to detect shockable rhythms in SCA patients without interrupting cardiopulmonary resuscitation (CPR). Since chest compressions produce artifacts in the ECG, CPR must be stopped for a reliable automated rhythm analysis [49]. However, interrupting CPR has a detrimental effect on survival. Indeed, interruption of chest compressions decreases the chances of adequate resuscitation by up to 50%. According to the AHA, immediate CPR can double or triple SCA patients' chance of survival [50]. Hence, we will also apply artifact detection algorithms and removal strategies to handle artifacts due to the person's movement while providing CPR. Additionally, the research should focus on fewer data-hungry unsupervised and semi-supervised algorithms that would eliminate extensive data labeling. Additionally, self-supervised learning is another paradigm worth exploring for shockable rhythm detection.

Author Contributions: Conceptualization, K.D. and M.H.A.; methodology, K.D.; software, K.D.; validation, K.D. and M.H.A.; investigation, K.D. and M.H.A.; writing—original draft preparation, K.D.; writing—review and editing, M.H.A.; visualization, K.D.; supervision, M.H.A.; project administration, M.H.A.; funding acquisition, M.H.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are pleased to acknowledge the partial financial support from the Dept. of Electrical and Computer Engineering at the University of Memphis, USA, to complete this work.

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

Abbreviations

Abbreviation	Meaning
AED	Automated External Defibrillator
AHA	American Health Association
BER	Balanced Error Rate
CNN	Convolution Neural Network
CNNE	Convolution Neural Network as a Feature Extractor
CPR	Cardiopulmonary Resuscitation
CUDB	Creighton University Ventricular Tachyarrhythmia Database
DCNN	Deep Convolution Neural Network
DNN	Deep Neural Network
DL	Deep Learning
ECG	Electrocardiogram
EC	External Classifier
GAN	Generative Adversarial Network
GP	Gradient Penalty
GRU	Gated Recurrent Unit
OHCA	Out-of-Hospital Cardiac Arrest
LSTM	Long Short-Term Memory
ML	Machine Learning
MVDB	Modified Variational Mode Decomposition
PEA	Pulseless Electrical Activity
NSH	Non-Shockable
NSR	Normal Sinus Rhythm
RNN	Recurrent Neural Network
SAA	Shock Advise Algorithm
SCA	Sudden Cardiac Arrest
SVM	Support Vector Machine
VF	Ventricular Fibrillation
VT	Ventricular Fibrillation
WCGAN	Wasserstein Conditional Generative Adversarial Network

References

1. Ferretti, J.; Di Pietro, L.; De Maria, C. Open-source automated external defibrillator. *HardwareX* **2017**, *2*, 61–70. [CrossRef]
2. Patil Kaustubha, D.; Halperin Henry, R.; Becker Lance, B. Cardiac Arrest. *Circ. Res.* **2015**, *116*, 2041–2049. [CrossRef]
3. Fabbri, A.; Marchesini, G.; Spada, M.; Iervese, T.; Dente, M.; Galvani, M.; Vandelli, A. Monitoring intervention programmes for out-of-hospital cardiac arrest in a mixed urban and rural setting. *Resuscitation* **2006**, *71*, 180–187. [CrossRef]
4. Kennedy Space Center Automated External Defibrillator (AED) Program—Home Page. Available online: https://aed.ksc.nasa.gov/FAQs/AED%20FAQs (accessed on 12 January 2021).
5. Sleightholm, K. Non-Shockable Rhythms. The First Aid Show. 2019. Available online: https://www.firstaidshow.com/non-shockable-rhythms/ (accessed on 12 January 2021).
6. The Shocking Truth About Defibrillators—IEEE Spectrum. IEEE Spectrum: Technology, Engineering, and Science News. Available online: https://spectrum.ieee.org/biomedical/devices/the-shocking-truth-about-defibrillators (accessed on 12 January 2021).
7. Js, S.; Wh, M. Recalls and safety alerts affecting automated external defibrillators. *JAMA* **2006**, *296*, 655–660. [CrossRef]
8. Amarappa, S.; Sathyanarayana, D.S.V. Data classification using Support vector Machine (SVM), a simplified approach. *Int. J. Electron. Comput. Sci. Eng.* **2014**, *3*, 435–445.
9. Nguyen, M.T.; Kiseon, K. Feature Learning Using Convolutional Neural Network for Cardiac Arrest Detection. In Proceedings of the 2018 International Conference on Smart Green Technology in Electrical and Information Systems (ICSGTEIS), Bali, Indonesia, 25–27 October 2018; pp. 39–42.
10. Isasi, I.; Irusta, U.; Aramendi, E.; Eftestøl, T.; Kramer-Johansen, J.; Wik, L. Rhythm Analysis during Cardiopulmonary Resuscitation Using Convolutional Neural Networks. *Entropy* **2020**, *22*, 595. [CrossRef]
11. Acharya, U.R.; Fujita, H.; Oh, S.L.; Raghavendra, U.; Tan, J.H.; Adam, M.; Gertych, A.; Hagiwara, Y. Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network. *Future Gener. Comput. Syst.* **2018**, *79*, 952–959. [CrossRef]
12. Deep Feature Learning for Sudden Cardiac Arrest Detection in Automated External Defibrillators. Abstract—Europe PMC. Available online: http://europepmc.org/article/PMC/6249221 (accessed on 12 January 2021).

13. Picon, A.; Irusta, U.; Álvarez-Gila, A.; Aramendi, E.; Alonso-Atienza, F.; Figuera, C.; Ayala, U.; Garrote, E.; Wik, L.; Kramer-Johansen, J.; et al. Mixed convolutional and long short-term memory network for the detection of lethal ventricular arrhythmia. *PLoS ONE* **2019**, *14*, e0216756. [CrossRef]
14. Krasteva, V.; Ménétré, S.; Didon, J.-P.; Jekova, I. Fully Convolutional Deep Neural Networks with Optimized Hyperparameters for Detection of Shockable and Non-Shockable Rhythms. *Sensors* **2020**, *20*, 2875. [CrossRef]
15. Hammad, M.; Kandala, R.N.V.P.S.; Abdelatey, A.; Abdar, M.; Zomorodi-Moghadam, M.; Tan, R.S.; Acharya, U.R.; Pławiak, J.; Tadeusiewicz, R.; Makarenkov, V.; et al. Automated detection of shockable ECG signals: A review. *Inf. Sci.* **2021**, *571*, 580–604. [CrossRef]
16. Sannino, G.; De Pietro, G. A deep learning approach for ECG-based heartbeat classification for arrhythmia detection. *Future Gener. Comput. Syst.* **2018**, *86*, 446–455. [CrossRef]
17. Kanellopoulos, D. Handling imbalanced datasets: A review. *Int. Trans.* **2006**. Available online: https://www.academia.edu/16613336/Handling_imbalanced_datasets_A_review (accessed on 18 May 2022).
18. Shafqat, W.; Byun, Y.-C. A Hybrid GAN-Based Approach to Solve Imbalanced Data Problem in Recommendation Systems. *IEEE Access* **2022**, *10*, 11036–11047. [CrossRef]
19. Haque, A. EC-GAN. 2022. Available online: https://github.com/ayaanzhaque/EC-GAN (accessed on 27 April 2022).
20. Lee, J.; Park, K. GAN-based imbalanced data intrusion detection system. *Pers. Ubiquitous Comput.* **2021**, *25*, 121–128. [CrossRef]
21. Haque, A. EC-GAN: Low-Sample Classification using Semi-Supervised Algorithms and GANs. *arXiv* **2021**, arXiv:2012.15864. Available online: http://arxiv.org/abs/2012.15864 (accessed on 25 February 2022).
22. Walia, M.; Tierney, B.; Mckeever, S. Synthesising Tabular Data using Wasserstein Conditional GANs with Gradient Penalty (WCGAN-GP). In Proceedings of the AICS, Dublin, Ireland, 7–8 December 2020. [CrossRef]
23. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A. Improved Training of Wasserstein GANs. *arXiv* **2017**, arXiv:1704.00028. Available online: http://arxiv.org/abs/1704.00028 (accessed on 8 March 2022).
24. Zekan, M. WCGAN-GP. 2022. Available online: https://github.com/marzekan/WCGAN-GP (accessed on 27 October 2022).
25. Dahal, K.; Ali, M.H. Overview of Machine Learning and Deep Learning Approaches for Detecting Shockable Rhythms in AED in the Absence or Presence of CPR. *Electronics* **2022**, *11*, 3593. [CrossRef]
26. AHA. AHA Database Sample Excluded Record. 2003. Available online: https://physionet.org/content/ahadb/1.0.0/ (accessed on 16 December 2022).
27. Greenwald, S.D. The MIT-BIH Malignant Ventricular Arrhythmia Database. 1992. Available online: https://physionet.org/content/vfdb/1.0.0/ (accessed on 16 December 2022).
28. Nolle, F.M.; Bowser, R.W. Creighton University Ventricular Tachyarrhythmia Database. 1992. Available online: https://physionet.org/content/cudb/1.0.0/ (accessed on 16 December 2022).
29. Felipe Alonso—Home Page. Available online: https://tsc.urjc.es/~{}falonso/ohca_data.html (accessed on 5 May 2022).
30. Figuera, C.; Irusta, U.; Morgado, E.; Aramendi, E.; Ayala, U.; Wik, L.; Kramer-Johansen, J.; Eftestøl, T.; Alonso-Atienza, F. Machine Learning Techniques for the Detection of Shockable Rhythms in Automated External Defibrillators. *PLoS ONE* **2016**, *11*, e0159654. [CrossRef]
31. Performance Aspects of Automated Rhythm Detection Capabilities for AEDs. HMP Global Learning Network. 2010. Available online: https://www.hmpgloballearningnetwork.com/site/eplab/articles/Performance-Aspects-Automated-Rhythm-Detection-Capabilities-AEDs (accessed on 11 January 2022).
32. He, H.; Bai, Y.; Garcia, E.A.; Li, S. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–8 June 2008; pp. 1322–1328.
33. Cao, Y.; Zhao, X.; Zhou, Z.; Chen, Y.; Liu, X.; Lang, Y. MIAC: Mutual-Information Classifier with ADASYN for Imbalanced Classification. In Proceedings of the 2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), Jinan, China, 14–17 December 2018; pp. 494–498.
34. Shaker, A.M.; Tantawi, M.; Shedeed, H.A.; Tolba, M.F. Deep Convolutional Neural Networks for ECG Heartbeat Classification Using Two-Stage Hierarchical Method. In *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2020*; Hassanien, A.E., Slowik, A., Snášel, V., El-Deeb, H., Tolba, F.M., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 137–147.
35. Pham, T.-H.; Sree, V.; Mapes, J.; Dua, S.; Lih, O.S.; Koh, J.E.W.; Ciaccio, E.J.; Acharya, U.R. A novel machine learning framework for automated detection of arrhythmias in ECG segments. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 10145–10162. [CrossRef]
36. Deep Convolutional Neural Networks—Run:AI. Available online: https://www.run.ai/guides/deep-learning-for-computer-vision/deep-convolutional-neural-networks/ (accessed on 3 January 2022).
37. Shewalkar, A.; Nyavanandi, D.; Ludwig, S.A. Performance Evaluation of Deep Neural Networks Applied to Speech Recognition: RNN, LSTM and GRU. *J. Artif. Intell. Soft Comput. Res.* **2019**, *9*, 235–245. [CrossRef]
38. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629. [CrossRef]
39. Ebrahimi, Z.; Loni, M.; Daneshtalab, M.; Gharehbaghi, A. A review on deep learning methods for ECG arrhythmia classification. *Expert Syst. Appl. X* **2020**, *7*, 100033. [CrossRef]

40. Al-Saffar, A.A.M.; Tao, H.; Talab, M.A. Review of deep convolution neural network in image classification. In Proceedings of the 2017 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET), Jakarta, Indonesia, 23–24 October 2017; pp. 26–31.
41. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
42. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* **2005**, *18*, 602–610. [CrossRef] [PubMed]
43. Zhao, R.; Wang, D.; Yan, R.; Mao, K.; Shen, F.; Wang, J. Machine Health Monitoring Using Local Feature-Based Gated Recurrent Unit Networks. *IEEE Trans. Ind. Electron.* **2018**, *65*, 1539–1548. [CrossRef]
44. Kostadinov, S. How Recurrent Neural Networks Work. Available online: https://towardsdatascience.com/learn-how-recurrent-neural-networks-work-84e975feaaf7 (accessed on 4 December 2021).
45. Nishiyama, T.; Nishiyama, A.; Negishi, M.; Kashimura, S.; Katsumata, Y.; Kimura, T.; Nishiyama, N.; Tanimoto, Y.; Aizawa, Y.; Mitamura, H.; et al. Diagnostic Accuracy of Commercially Available Automated External Defibrillators. *J. Am. Heart Assoc.* **2015**, *4*, e002465. [CrossRef]
46. Basnet, M.; Ali, M.H. Deep Learning-based Intrusion Detection System for Electric Vehicle Charging Station. In Proceedings of the 2020 2nd International Conference on Smart Power Internet Energy Systems (SPIES), Bangkok, Thailand, 15–18 September 2020; pp. 408–413.
47. *Statistics with Confidence: Confidence Intervals and Statistical Guidelines*, 2nd ed.; Altman, D.G. (Ed.) BMJ Books: London, UK, 2011.
48. Confidence Intervals for Sensitivity and Specificity in Python. Available online: https://gist.github.com/maidens/29939b3383a5e57935491303cf0d8e0b (accessed on 20 April 2022).
49. De Gauna, S.R.; Irusta, U.; Ruiz, J.; Ayala, U.; Aramendi, E.; Eftestøl, T. Rhythm Analysis during Cardiopulmonary Resuscitation: Past, Present, and Future. *BioMed Res. Int.* **2014**, *2014*, e386010. [CrossRef]
50. American Heart Association. Compression-Only CPR Increases Survival of Out-Of-Hospital Cardiac Arrest. Available online: https://medicalxpress.com/news/2019-04-compression-only-cpr-survival-out-of-hospital-cardiac.html (accessed on 16 March 2021).

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.