This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3325540

LOGO

Deep Representation Learning with Sample Generation and Augmented Attention Module for Imbalanced ECG Classification

Muhammad Zubair*, Sungpil Woo*, Sunhwan Lim, and Daeyoung Kim

***Abstract*— Developing an efficient heartbeat monitoring system has become a focal point in numerous healthcare applications. Specifically, in the last few years, heartbeat classification for arrhythmia detection has gained considerable interest from researchers. This paper presents a novel deep representation learning method for the efficient detection of arrhythmic beats. To mitigate the issues associated with the imbalanced data distribution, a novel re-sampling strategy is introduced. Unlike the existing oversampling methods, the proposed technique transforms majority-class samples into minority-class samples with a novel translation loss function. This approach assists the model in learning a more generalized representation of crucially important minority class samples. Moreover, by exploiting an auxiliary feature, an augmented attention module is designed that focuses on the most relevant and target-specific information. We adopted an inter-patient classification paradigm to evaluate the proposed method. The experimental results of this study on the MIT-BIH arrhythmia database clearly indicate that the proposed model with augmented attention mechanism and over-sampling strategy significantly learns a balanced deep representation and improves the classification performance of vital heartbeats.**

***Index Terms*— Arrhythmia detection, Beat classification, Imbalanced learning, Remote health monitoring.**

I. Introduction

ELECTROCARDIOGRAM is an important tool for monitoring heart activity over time. It is broadly considered the centerpiece of diagnostic tools used in cardiac health assessment as it reflects the heart's electrical activity. The existence of cardiovascular diseases like myocardial infarction, ventricular tachycardia, or arrhythmia alters various fiducial points of heartbeat and thus can be diagnosed by examining the electrocardiogram waveform [1], [2]. Arrhythmia is the most common and notable cause of mortality among cardiovascular diseases [3]. Arrhythmia, primarily caused by a disturbance in the heart's electrical conduction system, distorts heart rate, rhythm, and key fiducial points of ECG signal [4], [5]. Therefore, arrhythmia can be detected by assessing the morphological pattern of ECG beats. However, due to arrhythmia's rare and infrequent appearance, a long ECG recording (24 hours) comprising hundreds of thousands of beats is required to identify arrhythmia [6]. As the visual interpretation of ECG beats on a large scale is nonviable; therefore, an automatic classification system is needed to alleviate the problems of manual inspection. Additionally, abundant ECG data is received continuously; thus, incorporating an automated beat classification system would greatly assist a remote health monitoring system.

The recent advancements in IoT technology [7], wearable sensors [8], and deep learning [9] have significantly boosted the development of intelligent healthcare applications. Researchers have made efforts over the past few years to develop an efficient and robust heartbeat classification system for arrhythmia detection. However, heartbeat classification is challenging, and potential issues remain to be addressed. For example, one of the most prominent problems of the heartbeat classification model is the lack of generalization capability. The classification model trained for a specific group of patients may fail to classify the heartbeats of other patients accurately. This poor generalization is mainly caused by inter-patient variations in the morphological characteristics of ECG signals [10]. In addition to inter-patient variation based on the individual's nature, the ongoing physiological processes triggered by external or internal stimuli may also alter the morphological pattern of beats via the autonomic nervous system [11]. Therefore, heartbeat classification models with conventional feature extraction methods give lower classification performance and do not apply to a large population.

The most crucial issue that has been neglected in arrhythmia detection is the imbalanced distribution of ECG data. In healthcare, rare and infrequent events are of great importance. However, such events exist in the minority, resulting in an imbalanced data distribution. Similarly, in arrhythmia-related ECG data, anomalous samples are fewer in numbers as arrhythmic beats appear seldom and infrequently [12], [13]. However, the classification model expects a balanced distribution of samples among different classes, and thus the existence of imbalanced

Manuscript received April 12, 2023; revised September 27, 2023*; accepted October 10, 2023. Date of publication xxxxxxx xx, 2023. This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (2020-0-00048, Development of 5G-IoT Trustworthy AI-Data Commons Framework, 50%) and (2022-0-01032, Development of Collective Collaboration Intelligence Framework for Internet of Autonomous Things, 50%) *(Corresponding author: Sungpil Woo)*

Muhammad Zubair, Sungpil Woo, Sunhwan Lim are with the Autonomous IoT Research Section, Electronics and Telecommunications Research Institute (ETRI), Daejeon, South Korea, (e-mail: zubair5608, woosungpil, shlim @etri.re.kr)

Sungpil Woo, Daeyoung Kim is with the Department of School of Computing, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, (e-mail: woosungpil, kimd@kaist.ac.kr)

* These authors contributed equally to this work

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3325540

TABLE I: MIT-BIH Arrhythmia database beat types mapping to AAMI beat classes

AAMI Heartbeat Classes Description					
	Normal beat (N)	Supraventricular ectopic beat (S)	Ventricular ectopicbeat (V)	Fusion beat (F)	Unknown beat (Q)
MIT-BIH heart beat types	Normal beat (N)	Atrial premature beat(A)	Premature ventricular contraction (V)	Fusion of ventricular and normal beat (F)	Paced beat (/)
	Left bundle branch block beat (L)	Aberrated atrial premature beat (a)	Ventricular escape beat (E)		Fusion of paced and normal beat (f)
	Right bundle branch block beat (R)	Nodal (junctional) premature beat (J)			Unclassified beat (Q)
	Atrial escape beat (e)	Supraventricular premature beat (S)			
	Nodal (junctional) escape beat (j)				

data significantly deteriorates the model performance [6], [14], [15]. During training, the scarce representation of minority classes leads the classification model to learn an imbalanced and biased representation and thus perform poorly on minority class samples [16], [17]. Therefore, in arrhythmia detection where rare events are of vital importance, imbalanced data problems should be considered to characterize minority-class samples adequately.

In addition, the Association for the Advancement of Medical Instrumentation (AAMI) provided recommendations for developing a heartbeat classification system [18], [19]. For instance, AAMI recommended that ECG beats should be categorized into five groups: normal beats(N), supra-ventricular ectopic beats(S), ventricular ectopic beats (V), fusion beats (F), and non-classifiable beats(Q) as depicted in Table I. However, most of the studies on arrhythmia do not follow the AAMI recommendations for cataloging heartbeats [1], [4], [20].

This paper presents an efficient heartbeat classification system using deep learning to alleviate the problems mentioned above. The proposed model architecture and training strategy assures a balanced and efficient representation learning while taking the issue of imbalanced data into account. The contributions of this paper can be summarized as follows:

- We proposed a deep learning model for an efficient deep representation learning of heartbeats by extracting significant morphological information using an augmented attention module. The attention module, assisted with an auxiliary feature, focuses on the relevant and target-specific information of the ECG waveform for effective deep representation learning.
- To mitigate the issue of imbalanced data, we proposed an over-sampling strategy. New minority class samples are generated by transforming suitable majority class samples using a pre-trained base model. For this purpose, we introduced a novel translation loss function that efficiently alters suitable majority-class samples to construct minority-class samples.

The rest of the paper is organized as follows. Section II provides a survey of related literature with an explanation of the basic concepts of ECG beat classification. Heartbeat classification methodology, including the proposed augmented attention mechanism and oversampling strategy, is presented in Section III. Data description and evaluation strategy are given in Section IV. The classification results of the proposed method and performance comparison are reported in Section V. The concluding remarks are presented in section VI.

II. Related Work

This section presents a summary of literature published on arrhythmia classification. For a better illustration of the previous work with regard to this study, this section is divided into three parts. The first part describes the conventional feature extraction and classification methods for arrhythmia detection. The second part summarizes the literature focusing on deep learning methods for arrhythmia detection. The third and final part summarises the approaches used for handling imbalanced data in arrhythmia detection.

A. Conventional methods

Conventional arrhythmia detection methods include basic sequential steps like data acquisition, pre-processing, feature extraction, and classification [21]. During the data acquisition, ECG data is deformed by various noise sources, such as motion artifacts and power line interference. Literature reveals different approaches for denoising ECG signals [22], [23]. ECG beat is composed of different fiducial points (P wave, R peak, QRS complex, T wave), segments (PR, ST), and intervals (RR, PR). Prior to feature extraction, these fiducial points are located with efficient peak detection algorithms [24]. The located points in ECG signals are further used for segmentation and feature extraction.

Feature extraction is the most crucial step in arrhythmia detection. Extraneous and unneeded features contribute nothing but deteriorate the performance of the classification model. Therefore, special attention should be paid while extracting features for arrhythmia detection. In literature, many feature extraction methods have been investigated for ECG signals [21], [25]. For instance, De Chazal [26] extracted morphological features from ECG signals to discriminate various patterns of arrhythmic beats. A similar approach of morphological feature extraction has also been adopted in [5], [10]. Moreover, wavelet transforms capture frequency and time domain

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3325540

information and are thus preferred by many researchers for promising results in arrhythmia detection [27]. In addition, statistical methods [28], [29], Hermite transform [28], and fractal dimension [30] based feature extraction methods have also been used in literature for improved classification performance. Similarly, different conventional machine learning algorithms have been used for arrhythmia detection in literature. For example, Linear Discriminate Analysis (LDA) [26], logistic regression(LR) [31], artificial neural networks (ANN) [32], support vector machine (SVM) [33], decision trees(DT) [34] and clustering [35] has been used by many researchers for arrhythmia classification. However, these methods are based on conventional signal processing and hand-crafted feature extraction methods and, therefore, show poor performance under the inter-patient classification paradigm due to inter-subject variability in morphological characteristics of ECG signals.

B. Deep learning methods

Besides computer vision, deep learning has also shown devastating performance in healthcare [9], [13]. The deep learning model aims to learn a generalized representation of ECG beats in arrhythmia detection. A deep learning model with advanced architecture and intelligent training strategy can potentially mitigate the issues associated with conventional machine learning methods. In this regard, many studies have been presented using deep learning for arrhythmia detection over the last few years. For instance, preliminary research on arrhythmia detection using deep learning is performed by Kiranyaz et al. [36]. Similarly, in [14], a 1D convolutional neural network (CNN) is trained for intra-patient beat classification of the heartbeat. Literature reveals that 1D CNN is suitable for extracting morphological information from 1D ECG beat. Therefore, in pursuit of improved classification performance, several authors have investigated 1D CNN for arrhythmia detection [12], [37]–[40].

In addition, deep neural networks(DNN) [20], [41], deep belief networks(DBN) [42], and generative adversarial networks(GAN) [40], [43] have also been used to learn a better deep representation of ECG beat for arrhythmia classification. Moreover, inspired by the performance of LSTM with time series data, many authors explored the potential of LSTM models in heartbeat classification for arrhythmia detection. For instance, Yildirim et al. [44] adopted a bidirectional LSTM model with an embedded wavelet sequence module for five-class beat classification. Similarly, in [6], the LSTM model is used in combination with the residual module for arrhythmia classification. Besides, the LSTM model has been used in combination with CNN for efficient representation learning of ECG beats for arrhythmia classification [45].

In deep learning, attention is a powerful mechanism that impressively enhances model performance by emphasizing more on target-specific information. In ECG signal analysis for arrhythmia detection, inter-patient discrimination of different beats is challenging due to the complex and visually identical patterns of the beats. Therefore, target-specific and complex morphological information must be extracted for improved classification performance. In [46], the author investigated the attention mechanism's impact on the beat classifier's overall performance. The author used Channel-wise attention in CNN model and reported an improved classification performance. In another study, a convolutional block attention module is employed to classify five classes of beats [47]. In this study, we also introduced a novel attention mechanism that emphasizes extracting target-specific information for an efficient and balanced representation learning of ECG beats.

C. Handling imbalanced data

In literature, numerous methods have been reported to alleviate the issue of imbalanced data. These methods include over-sampling, under-sampling, and cost-sensitive learning [48]–[50]. However, literature on arrhythmia detection reveals that the issue of imbalanced data is neglected, with few articles addressing this critical problem. For example, a batch-weighted loss function is introduced by Ali Sellami [12]. In this study, class weights are computed for each batch during training, and a weighted cross entropy loss function is used to train a CNN model for arrhythmia classification. In another study, the issue of imbalanced data in ECG classification is addressed by synthesizing minority class samples using a target-oriented augmentation method [16]. The data augmentation method has also been adopted by Yong Xia [51] to tackle the challenge of imbalanced data in arrhythmia detection. Similarly, over-sampling based on z-score normalization and SMOTE is also performed in [6] and [14], respectively. However, these re-balancing methods usually lead the model to overfit minority class samples with abbreviated generalization [52].

Despite the efforts made by researchers, overcoming the issues associated with arrhythmia classification is still challenging, especially under the inter-patient classification paradigm. The literature review also shows that most arrhythmia classification studies do not meet AAMI's recommended standards. Moreover, most studies have only focused on the intra-subject classification of beats, which is unsuitable for real-world application. In this study, we proposed a novel beat classification that significantly tackles the aforementioned issues.

III. Proposed Methodology

A. Framework Overview

The proposed heartbeat monitoring framework for remote health monitoring is given in Fig. 1. The first step of the framework includes the acquisition of ECG signals using wearable sensors. The proposed method is based on a single-lead ECG signal and can be acquired easily with a small and adaptable wearable device. After the signal acquisition, segmentation is performed. In this step, a pre-defined number of samples are extracted based on specific fiducial points of the ECG beat. The segment selection significantly influences the classifier's performance as the information content varies with the segment length. Therefore, segment selection should be performed wisely in order to incorporate all the significant morphological characteristics of ECG beat. The extracted beats are then classified using a deep model. The deep model

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3325540

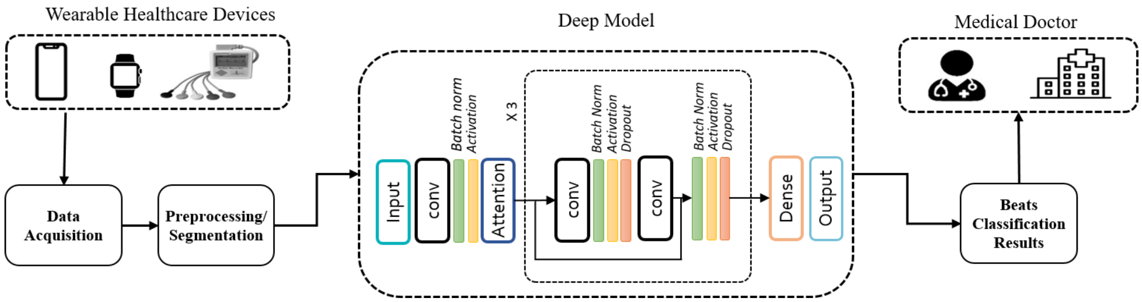

Fig. 1: Framework overview

significantly extracts the hidden patterns with 1D convolutional neural networks and predicts the class probabilities for each sample. The current study aims to develop an efficient deep model that learns to extract the most relevant and generalized features while handling the issue of imbalanced data. The proposed model can be used efficiently for on-site ECG classification as well as in remote health monitoring systems.

B. Proposed Over-sampling strategy

The proposed oversampling strategy aims to transform suitable majority-class samples to construct minority-class samples. In some cases, the minority classes have very few numbers of samples. Thus the regeneration of new samples from a limited number of samples leads to over-fitting and poor classification with degradation in the model's generalization capability [52]. Inspired by the re-sampling strategy adopted in [52], we introduced an efficient ECG beat re-sampling method to facilitate the balanced deep representation learning of ECG signals. The proposed over-sampling method boosts the model generalization performance by mitigating the over-fitting on minority samples caused by conventional oversampling strategies.

In this study, the sample generation from the majority class is performed in such a way that the translated sample does not deteriorate the model performance for majority class samples. We translated majority class samples by dragging them into the feature space of the targeted minority class. First, a base model $g(x, \theta)$ is trained on an imbalanced dataset. The higher sample representation of the majority class leads the trained base model $g(x, \theta)$ to be inclined towards the majority class. During the training of the target model, this trained base model is used to select suitable samples from the majority class. Samples of the majority class having higher similarity with minority class samples are selected for translation. The selected samples are then translated with an optimization approach using a novel translation loss function. In this approach, an objective function is optimized to increase the correlation of the generated samples with the target classes. As a result, a new balanced mini-batch is generated, which is used to train the proposed target model. Unlike [52], we adopted a more sophisticated and simple sample selection and

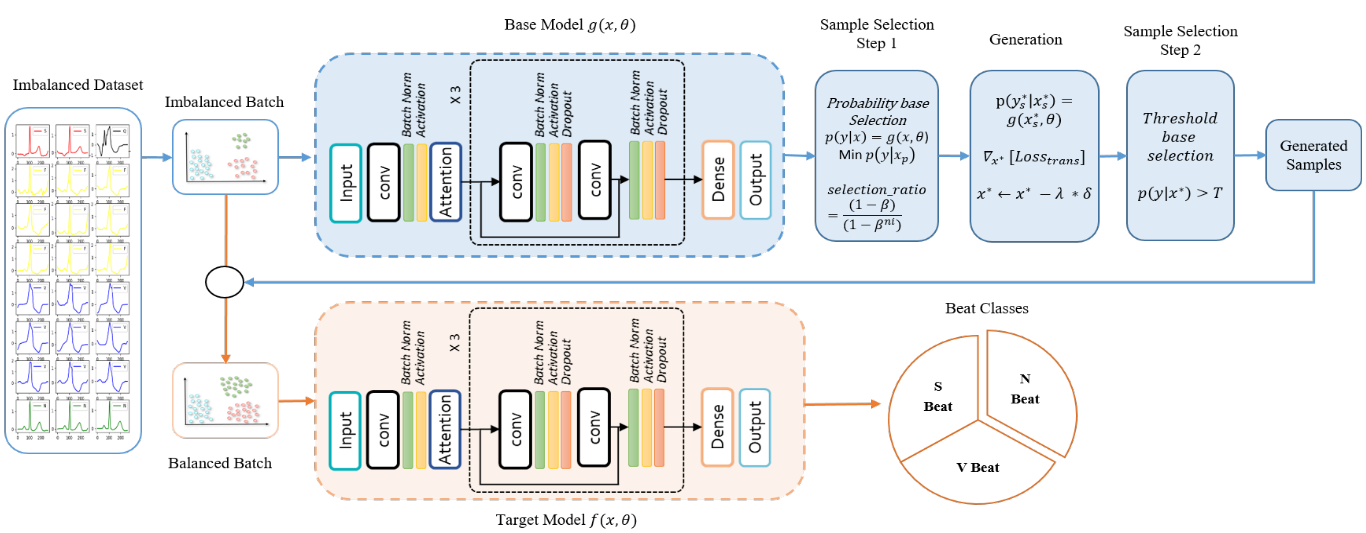

Fig. 2: The architecture of the proposed beat classification model consists of five components: (1) a base model $g(x, \theta)$ trained with an imbalanced dataset; (2) a probability-based sample selection step; (3) sample generation with translation loss function; (4) threshold based selection of the final set of generated samples; (5) a target model $f(x, \theta)$ trained with a balanced batch. Light blue blocks represent the oversampling strategy using the base model, while light orange blocks represent target model training with balanced data.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3325540

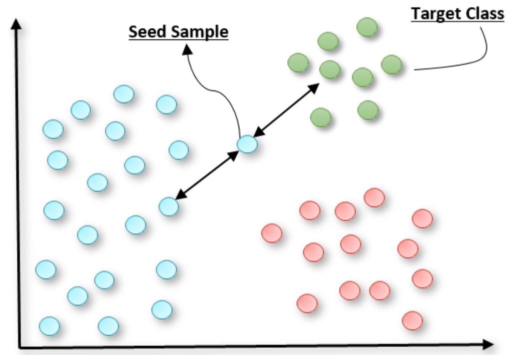

Fig. 3: Visualization of suitable sample for translation

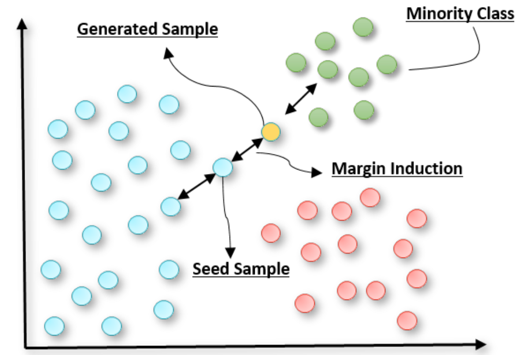

Fig. 4: Visualization of generated sample

translation method.

1) Sample Selection: For oversampling, we choose to use the most suitable samples from the majority class for transforming into minority class samples. Consider a base model $g(x, \theta)$ trained with an imbalanced dataset $D = \{(x_i, y_i)\}_{i=1}^{N}$, where $x \in \mathbb{R}$ and $y \in \{1, ...K\}$. Then, for a mini-batch $B = \{(x_i, y_i)\}_{i=1}^{m}$ of m samples, the probabilities can be expressed as

$$p(y_i|x_i) = g(x_i, \theta) \tag{1}$$

Now, for efficient translation, samples of the majority class having a lower probability of the parent class should be selected and samples having a strong correlation with the parent class should be ignored. For selecting the best samples for translation, we computed the distance of the sample with the minority class cluster as well as with the majority class cluster. Samples having a minimum distance with a minority class cluster are selected for translation. The visual interpretation of the suitable sample for translation is given in Fig 3. Moreover, all the samples of the majority class can not be selected for translation as it reduces the performance of the majority class. According to [53], as the number of sample increases as a result of over-sampling, the ability of the model to get adequate information diminish. Therefore an efficient number of samples is selected as given [53]. We selected an efficient number of samples from mini-batch with a specific selection ratio that is given as follows

$$N_s = \frac{1-\beta}{1-\beta^{n_i}} \tag{2}$$

where n_i is the number of samples in each class and $\beta = \frac{N-1}{N}$. The effective number of samples avoids the performance deterioration of the model for the majority class samples. These samples are then translated to minority class samples with the help of the translation loss function. The first six steps of Algorithm 1 represent the selection of suitable seed samples for translation.

2) Sample generation: For generating new samples of minority classes, our proposed method of oversampling solves an optimization task. The goal of this optimization problem is to reduce the loss of the majority sample for the minority class and induce some margin between the boundaries with other classes to ensure discrimination. This method will boost the generalization capability of the model. The proposed sample generation strategy aims to increase the similarity of the seed sample with the target minority class while decreasing its similarity with other classes. The visual interpretation of the sample generation is illustrated in Fig. 4.

In the process of sample generation, we used the base model $g(x, \theta)$ originally trained on an imbalanced dataset. The main objective behind generating samples for the minority class is to facilitate the training of a target model $f(x, \theta)$. This target model aims to acquire a balanced representation of the data and deliver high performance across all available classes. First, a small noise is added to the selected seed samples. Seed samples and the translated samples (with added noise) are then subjected to predictions using the previously trained base model. For an efficient translation, we proposed a translation loss function as follows.

$$Loss_{trans} = m + \|p(y_s|x_s) - p(y_s^*|x_s^*)\|^2 \tag{3}$$

Where $p(y_s|x_s)$ is the probability of the seed sample with parent class, $p(y_s^*|x_s^*)$ is the probability of the transformed seed sample with target minority class, and m is the margin induced to mitigate the overlap of seed sample and transformed sample. Steps seven to seventeen of Algorithm 1 show the translation of seed samples. Following the translation of appropriate seed samples, we proceed with a second round of sample selection using the predictions made by the base model (Algorithm 1, 17-21). The translated samples with prediction scores exceeding a specific threshold are selected for training the target model.

C. Attention-based Model

The deep model developed in this work for heartbeat classification is based on 1D convolutional neural network, as depicted in Fig.2. Multiple convolution layers were adopted, followed by a batch normalization, activation layer, and dropout layer. A skip connection is used to avoid the vanishing gradient problem and improve the model performance by mitigating the issue of information loss. The proposed attention module is used after the first convolution layer. The input (channels, segment length) to the attention modules includes the feature maps representing the features extracted for every

Algorithm 1 Minority class sample generation

Input: A mini-batch $\mathcal{B} = \{(x_i^s, x_i^a, y_i)\}_{i=1}^m$ with m samples of c different classes; A pre-trained model $g(x_s, x_a, \theta)$. Learning rate λ, acceptance threshold γ and number of iterations $T > 0$.

Output: A set of generated samples $\mathcal{B}^* = \{(x_i^*, ax_i^*, y_i^*)\}_{i=1}^m$

$P(\mathcal{B}) \leftarrow g(\mathcal{B}, \theta)$
$N_s \leftarrow (1-\beta)/(1-\beta^{n_i})$
$P_{maj} \leftarrow$ Probabilities of major class samples in $P(B)$
for $i = 2$ **to** c **do**
  $\hat{x}_s, \hat{x}_a \leftarrow P_{maj}[N_s] \in c_i$
end for
Initialize $x_s^* \leftarrow \hat{x}_s + \epsilon$ ▷ small noise ϵ
for $t = 0$ **to** T **do**
  $P(y) \leftarrow g(\hat{x}_s, \hat{x}_a, \theta)$
  $P(y^*) \leftarrow g(x_s^*, x_a^*, \theta)$
  $\mathcal{L}_{trans} \leftarrow m + \|p(y|\hat{x}_s, \hat{x}_a) - p(y^*|x_s^*, x_a^*)\|^2$
  $\mathcal{L}_{total} \leftarrow \mathcal{L}(g, x_s^*, x_a^*, y_s^*) - \mathcal{L}_{trans}$
  $\delta_s \leftarrow \nabla_{x_s^*}[\mathcal{L}_{total}]$
  $\delta_a \leftarrow \nabla_{x_a^*}[\mathcal{L}_{total}]$
  $x_s^* \leftarrow x_s^* - \lambda.\delta_s$
  $x_s^* \leftarrow x_a^* - \lambda.\delta_a$
end for
if $\mathcal{P}(y^*|x_s^*, x_a^*) \geq \gamma$ **then**
  $\{(x_s^*, x_a^*, y_s^*)\} \leftarrow x_s^*, x_a^*, y_s^*$
end if
$\mathcal{B}^* \leftarrow \mathcal{B} \cup \{(x_s^*, x_a^*, y_s^*)\}$

sample in the beat segment. The attention module assigns the weights to each feature based on its importance and feeds the masked input to fully connected layers for classification. For classification, we used a fully connected layer followed by a Sigmoid layer. The aim of using the attention module in the initial layers is to ensure the flow of relevant information for target-specific feature extraction.

1) Proposed Attention mechanism: In order to extract the target-specific features, we design an augmented attention module that focuses on the most relevant and target-specific information by exploiting an auxiliary feature. We used the RR interval as an auxiliary feature in this study. RR interval is known as the time difference between successive R peaks and has been widely used in conventional heartbeat classification methods [21], [54]. The morphological variations elicited by arrhythmia are efficiently reflected in the variations of RR interval width. Therefore, the RR interval feature is considered one of the most correlated and efficient handcrafted features to characterize temporal fluctuations provoked by arrhythmia [26], [55]–[57].

The architecture of the proposed attention module is designed to boost the representation power of the CNN model. The primary goal of the attention module is to refine the features associated with each time stamp of the ECG segment by emphasizing more on the target-specific features. The attention module takes the extracted deep feature maps and auxiliary feature (RR interval) as input to compute the

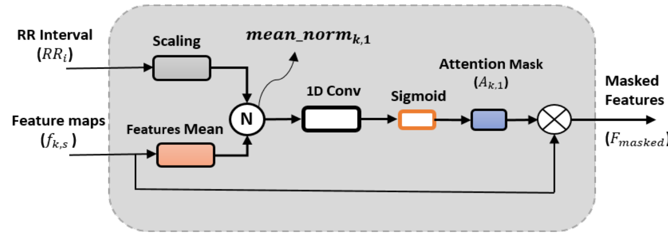

Fig. 5: Proposed attention module

attention mask. First, the mean of the feature maps is computed along the temporal axis(segment length). Literature reveals that global average pooling has been widely adopted in attention mechanisms for aggregating feature maps [58], [59]. Similarly, in [46], [47], average pooling has been employed to design attention-based deep models for arrhythmia detection. In order to further accentuate the performance of the attention module, the aggregated feature maps were normalized using the auxiliary feature (RR interval) as follows.

$$mean_{norm} = \frac{\frac{1}{S}\sum_i^s f_{k,i}}{\alpha.RR} \tag{4}$$

where s is the segment size and k is the number of kernels. α is a hyper-parameter to control and scale the effect of normalization. We used auxiliary feature-based normalization of feature maps for the following reasons.

- The auxiliary feature fuses the information related to morphological patterns and thus leads to the extraction of more refined, relevant, and target-specific features.
- It improves the discrimination performance of the classifier, especially for hard classes, by enriching feature refinement.
- It provides the flexibility to control the influence of external information on deep features for efficient feature extraction.

The normalized feature mean is then passed through a convolutional layer, and an attention mask $A_{k,s}$ is generated with the Sigmoid layer. The attention mask carries the weights associated with each feature map. In other words, this attention mask represents the importance of features linked with each segment of the ECG beat. The weighted feature maps are obtained as follows.

$$F_{masked} = f_{k,i} * A_{k,s} \tag{5}$$

The output of attention module F_{masked} holds the most refined and relevant information to the target class. Besides extracting target-specific features, the attention module also diminishes the effect of irrelevant information and thus mitigates performance deterioration due to redundant and irrelevant features. The architecture of the proposed attention module is given in Fig. 5.

D. Classification paradigm

The classification results of the previously published articles show low classification scores for S-type beats. These beats are hard to classify and are usually misclassified as N-class beats. Le Sun et al. [40] mentioned that N-class beats have nearly identical morphological characteristics and a shorter RR interval to that of the S class. Inspired by the hierarchical classification proposed in [40] and [60], we proposed a two-step classification strategy to overcome the aforementioned issue. Fig. 6 illustrates the proposed classification approach. The first model aims to discriminate N from S and V while S and V class beats are represented with the same label. The second model classifies S and V. This two-step classification approach alleviates the classification issue associated with N and S class beats. However, unlike [40], [60], the proposed two-step classification methods use the same input and deep model architecture in both steps, making the classification system more effective in terms of applicability and computation.

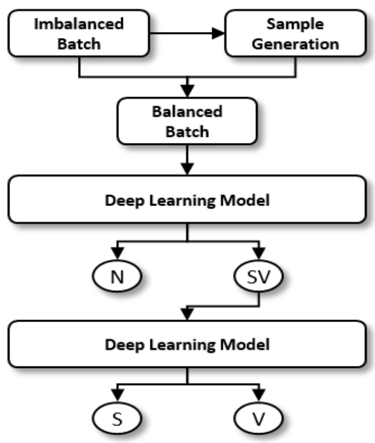

Fig. 6: Two-step classification of ECG beats

IV. Experiments

A. Dataset

MIT-BIH Arrhythmia database [61] has been widely used in previous studies for the evaluation of beat classification models [12], [26], [40], [62]. In this study, we also used MIT-BIH databases to evaluate the proposed model for a fair comparison with the previous studies. MIT-BIH database includes two channels' ECG recordings of 47 subjects. A total of 48 recordings (30 minutes long) were acquired. Each recording is digitized at 360 samples per second and annotated by two or more cardiologists independently. In this study, we used modified limb lead II signals to train and evaluate the proposed deep model. We selected 44 recordings out of 48 recordings for our experiment. Four recordings (102, 104, 107, 217) with low-quality signals and paced beats were excluded according to the recommendation of AAMI [12], [26], [40], [42].

B. Evaluation Strategy

We performed inter-patient classification of ECG beats in order to develop an efficient arrhythmia detection model that can be used in real-world healthcare applications for a wide range of populations. We adopted the same procedure for splitting the data into a training set (DS1) and test set (DS2) as illustrated in [12], [26], [40]–[42], [62]. Both sets of data include recordings of 22 subjects. Unlike [40], we selected the segment between two successive T peaks as depicted in Fig. 7. The T-T segment includes all the key fiducial points of an ECG beat (P, QRS complex, T) and thus yields efficient classification performance. The summary of sample distribution among different classes in Table II shows an extreme imbalance of the samples across different classes. In this study, we only considered N, S, and V class beats for classification as performed in [40], [41].

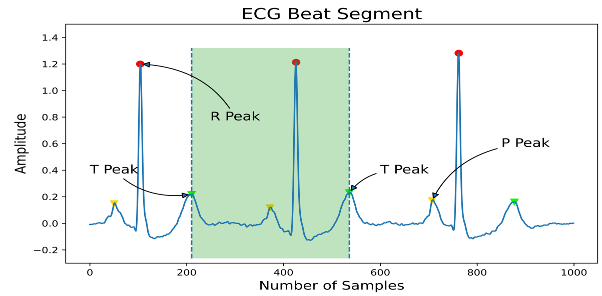

Fig. 7: ECG waveform with a TT interval segment(green shaded area). The red, yellow, and green dots represent R, P, and T peaks, respectively.

TABLE II: Summary of class distribution in MIT-BIH dataset

Types	Number of Beats
Normal beats (N)	90087
Supraventricular ectopic beats (S)	2781
Ventricular ectopic beats (V)	7008
Fusion beats (F)	802
Unknown beats (Q)	15
Total Beats	100,693

Additionally, the classification accuracy does not provide equitable information for the evaluation of a model trained on an imbalanced dataset as it treats all classes equally. Therefore, besides classification accuracy, we also computed three additional performance metrics, which include sensitivity, specificity, and positive productivity. These metrics have been employed widely in arrhythmia detection literature [12], [26], [40]–[42], [62]. Furthermore, the AAMI also recommended these metrics for model evaluation [12], [18], [19], [21].

C. Experimental Setup

In order to train the proposed deep learning model, we used weight decay (0.001) and dropout of 0.5 in addition to batch normalization layer for addressing the over-fitting problem. The initial learning rate is set to 0.001. During feature normalization in the attention module, we used a scaling score (α) of 0.01. Additionally, the training is performed on Nvidia Geforce RTX 3090.

V. Results and Discussion

A. Classification Performance

This study involves an investigation of the effectiveness of using attention and oversampling strategies in deep representation learning of ECG beats for arrhythmia detection.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3325540

To verify the validity of the proposed method, we carried out the evaluation under the inter-patient classification paradigm. The data from the MIT-BIH database is divided into training (DS1) and test set (DS2). Each of these sets includes data from 22 different patients. The same approach for data distribution has been employed in [12], [26], [40]–[42], [62]. We adopted the inter-patient paradigm for model evaluation as it can significantly reveal the applicability of the arrhythmia classification model for real-world application. Additionally, to boost the classification confidence of the model for discrimination between three classes of arrhythmia (N, S, and V), we used a two-step classification approach as reported by [40], [60].

First, we trained the model on an imbalanced dataset to highlight the effect of imbalanced data distribution. Table III presents the classification results of the proposed deep architecture trained with imbalanced data. The results obtained with imbalanced data clearly depict that the disproportion distribution of samples among three classes has significantly deteriorated the model performance. Although the sensitivity scores for all the three classes of beats (N, S, and V) reported in Table III are not adequate for a crucial healthcare application model, a notably low sensitivity score is achieved for the majority class (N) as compared to minority class (S) samples. These findings are related to the two-step classification approach adopted in this study. It has been reported that N class beats have nearly identical morphological characteristics to that of S class beats [40]. Therefore, the classification of N and SV leads to an increase in false negative samples. Interestingly, these observations are commending for the base model used in the proposed oversampling method. In other words, these findings reveal that the base model's tendency to over-fit minority class is mitigated partially with two-step classification, and thus it would be of great assistance in improving the seed selection in sample generation.

TABLE III: Classification results with imbalanced data

Classes	Accuracy	Sensitivity	Specificity	Positive productivity
N	87.35%	86.95%	90.68%	98.73 %
S	91.93%	93.90%	90.80%	85.35%
V	91.92%	90.08%	93.90%	96.31%
Average	90.40%	90.31%	91.79%	93.46%

Table IV presents the performance evaluation metrics obtained for the model trained with over-sampled data. These results of the inter-patient classification experiment illustrate that the proposed beat classification method significantly discriminated different types of heartbeats with improved performance metrics. In comparison to Table III, the results presented in Table IV clearly provide evidence that the proposed oversampling strategy significantly mitigates the issue of overfitting associated with imbalanced data. These findings are in complete agreement with the previous study [52], which concluded that the conventional re-balancing methods lead the model to over-fit on minority class samples with abbreviated generalization. The sensitivity score, which is considered to be the most suitable performance metric for imbalanced data, shows a significant increase for minority class samples. This boost in classification performance (especially in sensitivity), as compared to [12], [40], validates efficient translation of majority class samples to minority class samples.

Moreover, the performance metrics of majority class also ensure that the translation loss efficiently induced the adequate discrimination margin between seed and generated samples and thus achieved a higher score for majority class samples as well. In other words, the majority-class results depict that the translation loss function significantly eradicated parent class information from seed samples [52]. Similarly, these results also depict the contribution and efficacy of the proposed attention module that significantly refined the features linked with each time stamp of ECG segment.

TABLE IV: Classification results with over-sampling

Classes	Accuracy	Sensitivity	Specificity	Positive productivity
N	94.71%	95.63%	86.59%	98.42 %
S	96.93%	96.30%	97.29%	95.31%
V	96.94%	97.30%	96.30%	97.88%
Average	96.19%	96.41%	93.39%	97.20%

The findings of the current study have important implications for developing a robust arrhythmia detection model that can be used effectively for a large population. For instance, in cardiac health monitoring, arrhythmic beats (S and V) are rare and infrequent events and are greatly important to be detected accurately for early diagnosis of arrhythmia [21]. However, training with imbalanced data yields an over-fitted model on minority class beats and thus mostly fails to detect these crucially important beats, which is highly undesirable. Therefore, the proposed method with significant results provides further intuition of learning a balanced deep representation of ECG beats and thus would provide great assistance in designing a real-world arrhythmia detection model.

B. Impact of Augmented Attention Module

The augmented attention module is designed to accentuate the most relevant feature maps for efficient classification of beats. The three classes of beats, as shown in Fig. 9, have complex differences from one another in morphological characteristics [5], [26]. Therefore, eliminating the conflicting feature maps and focusing more on the target-specific areas would significantly improve the classifier performance by avoiding the deterioration caused due to redundant and irrelevant features. The efficacy of feature (Channel) attention in arrhythmia detection has been reported in literature [46], [47]. However, we designed augmented attention with an auxiliary feature to further boost the performance of the attention mechanism for highlighting the most relevant and target-specific features. A temporal attention module can also be used to emphasize the time stamps of beat segments. However, we choose to adopt feature attention to avoid any information loss and assure the extraction of more efficient features while considering the feature maps associated with all the samples in beat segments.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3325540

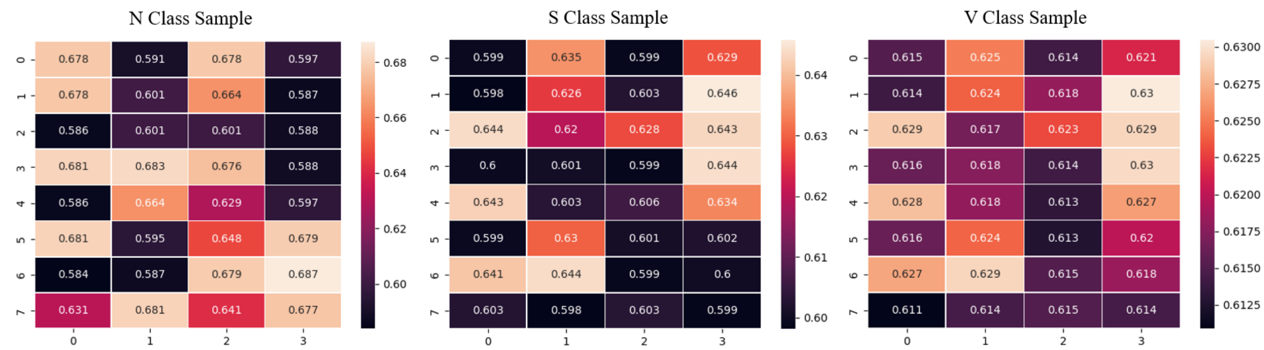

Fig. 8: Attention mask learned for three classes (N, S, V) of beats.

Figure 8 shows the attention mask (attention score) learned for three classes of beats. The attention mask for 32 feature maps is plotted on an 8x4 grid of heat maps. The attention masks clearly depict that a different attention score is assigned to each feature map in each category of beats. These findings corroborate that the most relevant and target-specific information propagates forward to learn an efficient representation of ECG beats. The proposed attention mechanism endorses the improved discrimination performance reported in Table IV. It is interesting to note that the feature maps assigned with maximum weight in Class N were assigned with minimum scaling weight in Class S, as depicted in Figure 8. A plausible explanation of this observation is that the S class has a limited representation in the training set and has nearly identical morphological characteristics with that of N class [40], which makes the classifier puzzling to discriminate efficiently between these classes. Therefore, the attention module has given more emphasis on refining the target-specific features clearly. Surprisingly, these observations also validate the fact that the parent class information (N class) has been eliminated precisely from newly generated samples of the minority class (S class).

Moreover, the improvement in the performance of the attention module is also supported by the insertion of an auxiliary feature. The position of R-peak in different classes clearly reveals the correlation of RR interval with arrhythmic beats as depicted in Fig. 9. Therefore, to put more emphasis on the relevant feature maps that contribute more toward the targeted class, we used an RR interval that reflects the presence of S and V beats. Thus incorporating the handcrafted feature to induce target-specific information assisted in the identification of feature maps in a more efficient way and achieved an improved classification performance.

C. Impact of Sample Generation

In this study, we introduced an oversampling method that translates majority-class samples to minority samples in order to mitigate the issues of imbalanced data in arrhythmia detection. The oversampling method aims to reduce the chances of learning a biased deep representation with imbalanced label distribution. Moreover, to overcome the over-fitting on minority class samples and deterioration of generalization capability caused by conventional over-sampling methods [52], we choose to translate suitable majority class samples into minority class samples. For this reason, the proposed oversampling strategy makes use of an optimization process with a novel translation loss function to generate a balanced batch.

The sample generation is performed for every batch to acquire the maximum possible number of samples for the minority class to efficiently mitigate the imbalanced data issue and improve the classification performance of minority class samples. Figure 10 compares the model performance trained with imbalanced data and over-sampled data. The proposed

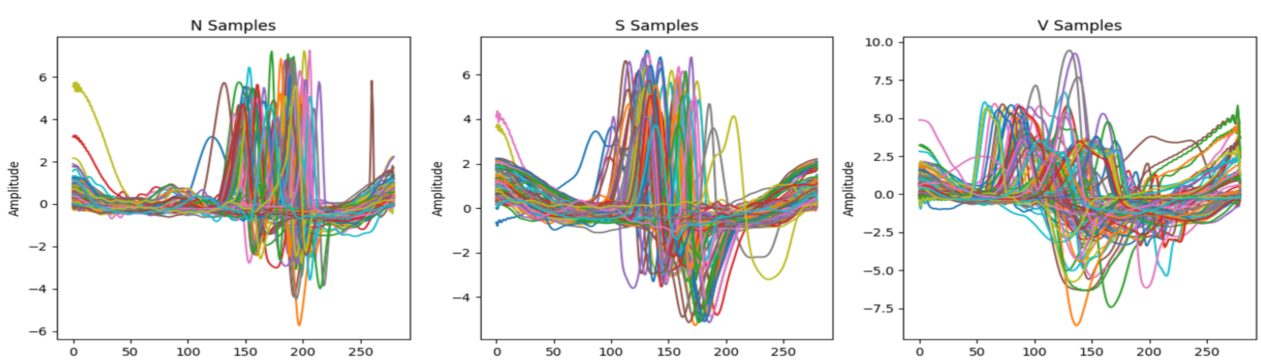

Fig. 9: ECG beat segments for three classes (N, S, V)

TABLE V: Classification comparison with previously published results on MIT-BIH dataset

Authors	Classifier	ACC(%)	N(%)			S(%)			V(%)		
			SEN	SPE	PPR	SEN	SPE	PPR	SEN	SPE	PPR
Dechazal et al. (2004) [26]	LDA	81.9	86.9	–	99.2	75.9	–	38.5	77.7	–	81.9
Huang et al. (2014) [62]	SVM + Threshold	93.8	99.2	–	95.2	91.1	–	42.2	93.9	–	90.9
Mathews et al. (2018) [42]	Deep belief networks	94.13	86.1	94.58	99.27	70.99	94.34	32.44	85.23	95.66	55.28
A. Sellami et. al.(2019) [12]	CNN	95.05	88.51	91.30	98.80	82.04	92.80	30.44	92.05	97.54	72.13
Wang et al. (2020) [41]	Dual fully connected neural networks	93.4	86.9	–	99.2	75.9	–	38.50	77.7	–	81.9
Le Sun et. al.(2022) [40]	Beatclass + MorphGAN	98.7	99.9	–	99.1	94.7	–	96.8	97.1	–	97.0
Yuan et. al.(2023) [16]	CNN + TAPDA	93.9	96.1	–	98.1	80.3	–	73.4	74.3	–	66.1
Xia et. al.(2023) [51]	CNN, DAE + Transformer	97.66	97.35	71.09	96.47	70.26	99.44	82.90	73.92	96.42	71.67
Proposed	**CNN + Attention**	96.19	95.63	86.59	98.42	**96.30**	**97.29**	95.31	**97.30**	96.30	**97.88**

over-sampling strategy significantly improved the classification performance. Specifically, the classification performance of the most crucial minority class (S) is improved substantially, which depicts the efficacy of the proposed training strategy. Comparing results presented in Table III and Table IV provides indisputable evidence that the model learned a deep balanced representation with the proposed oversampling strategy. Table IV is quite revealing in several ways for interpreting the significant impact of sample generation. For instance, sensitivity, specificity, and positive productivity for the minority class (S) improved with over-sampling, indicating an impressive reduction in false negative and false positive samples. This finding further validates that major-to-minor translation for sample generation alleviates the over-fitting of minority class and thus corroborates the previous work [52]. Similarly, Table IV also reveals that the model tendency towards majority class (N) is also attenuated adequately with the proposed oversampling method.

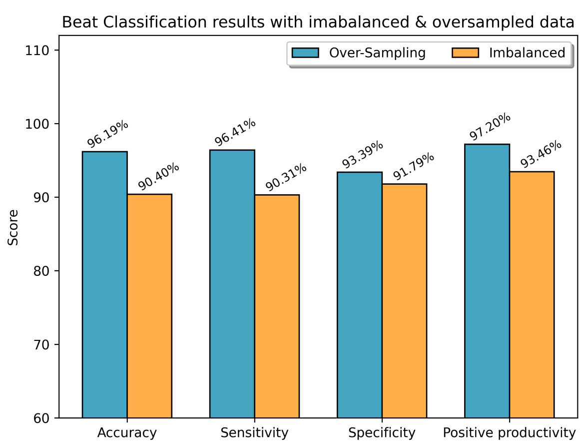

Fig. 10: Performance comparison between imbalanced data and over-sampled data

Furthermore, the over-sampling method assisted with a novel translation loss function that converts the majority class samples into target samples with higher confidence. The translation loss function induced an adequate discrimination margin between the seed sample (N class) and the newly generated sample (S class). It thus ensured the elimination of parent class information from seed samples. This observation can be validated clearly from the results of the majority class reported in Table IV. Similarly, it is also apparent from 8 that majority and minority class (parent class or translated) samples have learned significantly distinct features. The opposite feature mask is learned for corresponding feature maps in N and S class samples, resulting in improved classification scores for both classes. This finding validates the significant impact of the translation loss function and its intriguing correlation with attention masks.

D. Comparison

Finally, we compare our results with other studies on beat classification based on deep learning. For an honest and unbiased comparison, we selected only those studies that used the MIT-BIH dataset and adopted the same beat categorization and model evaluation methods. Table V illustrates that our proposed approach achieved better classification performance than other studies. The deep balanced representation learning with the proposed augmented attention module and translation loss function has significantly increased the classification rate of minority class samples, as depicted in Table IV. Le Sun et al. (2022) [40] adopted a hierarchical procedure for ECG beat classification and employed SMOTE to handle the issue of imbalanced data. The classification results of [40] presented in Table V depict that generating minority samples from minority class samples via SMOTE causes over-fitting on minority class samples [52] and, therefore, a lower sensitivity score is achieved for S and V class samples in [40]. Similarly, lower classification performance for S and V class samples is also reported in [16] and [51] where data augmentation methods are used to tackle imbalanced data issues. In comparison to [16], [40], [51], our proposed method achieved a better performance score for the most crucial class of S and V beats. These results substantiate previous findings in literature [52]. Comparatively high performance for N class samples is reported in [40] because fewer samples were used for the N class in train and test sets.

Similarly, a batch-weighted loss function is employed to counter the issue of imbalanced data in [12]. Compared with

[12], our proposed method with augmented attention module and over-sampling strategy achieved a higher classification score. Specifically, the efficacy of the proposed model can be justified by analyzing the performance of minority class samples (crucially important samples) compared to other studies. In addition, the T-T segment is also essential for efficient deep feature extraction as it takes all the vital fiducial points of ECG beats into account for deep feature extraction. In summary, the augmented attention mechanism and the sampling generation via majority-to-minority translation assured a deep balanced representation by extracting the most relevant and target-specific features, resulting in improved classification performance.

E. Challenges and Limitations

The most challenging step in the implementation of the proposed framework is eradicating parent class information from seed samples during transformation. The presence of parent class information in translated samples significantly deteriorates the classification accuracy of the majority class. In other words, the erroneous transformation of majority class samples reduces the discrimination performance of the classifier. Therefore considerable attention must be paid when transforming majority class samples into minority class samples. The most important limitation of this study is the computational complexity and convergence time due to multiple optimization steps.

VI. Conclusion

This paper proposes a heartbeat classification method for arrhythmia detection using deep learning. A novel re-sampling method is introduced to overcome the issues of imbalanced data distribution for an efficient representation learning of ECG. The proposed method also includes an augmented attention model for extracting the most relevant and target-specific features.

The experimental results reveal that the proposed augmented attention module and over-sampling method collectively improved the classification performance of crucially important beats and successfully mitigated the issues associated with imbalanced data. These results also indicate that the augmented attention module supported by an auxiliary feature significantly accentuates the target-specific features and alleviates the disputable features among the most identical beat classes. One of the most significant findings emerging from this study is that oversampling by translating suitable majority-class samples into minority-class samples reduces the over-fitting on minority samples and thus achieves higher classification performance for critically important arrhythmic beats. The findings of this study enhance our understanding of learning a balanced deep representation of ECG for arrhythmia detection and the role of advanced over-sampling methods in healthcare applications.

In the future, it would be of interest to assess the implications of transfer learning in addition to intelligent transformer algorithms on the classification performance of the arrhythmia detection model. More specifically, adversarial domain adaptation should be used to mitigate the inter-subject and intra-subject variability of ECG for better generalization of the arrhythmia detection model.

Acknowledgment

This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (2020-0-00048, Development of 5G-IoT Trustworthy AI-Data Commons Framework(50%), and 2022-0-01032, Development of Collective Collaboration Intelligence Framework for Internet of Autonomous Things(50%)).

[1] S. K. Berkaya, A. K. Uysal, E. S. Gunal, S. Ergin, S. Gunal, and M. B. Gulmezoglu, "A survey on ecg analysis," *Biomedical Signal Processing and Control*, vol. 43, pp. 216–235, 2018.
[2] O. Yildirim, U. B. Baloglu, R.-S. Tan, E. J. Ciaccio, and U. R. Acharya, "A new approach for arrhythmia classification using deep coded features and lstm networks," *Computer methods and programs in biomedicine*, vol. 176, pp. 121–133, 2019.
[3] W. H. Organization *et al.*, "Cardiovascular diseases (cvds): Fact sheet no. 317. 2015," 2016.
[4] J. Malik, O. C. Devecioglu, S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ecg classification by 1d self-operational neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 5, pp. 1788–1801, 2021.
[5] C. Ye, B. V. Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ecg signals," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2930–2941, 2012.
[6] Y. K. Kim, M. Lee, H. S. Song, and S.-W. Lee, "Automatic cardiac arrhythmia classification using residual network combined with long short-term memory," *IEEE Transactions on Instrumentation and Measurement*, 2022.
[7] L. Chettri and R. Bera, "A comprehensive survey on internet of things (iot) toward 5g wireless systems," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 16–32, 2019.
[8] J. Kim, A. S. Campbell, B. E.-F. de Ávila, and J. Wang, "Wearable biosensors for healthcare monitoring," *Nature biotechnology*, vol. 37, no. 4, pp. 389–406, 2019.
[9] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean, "A guide to deep learning in healthcare," *Nature medicine*, vol. 25, no. 1, pp. 24–29, 2019.
[10] P. De Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features," *IEEE transactions on biomedical engineering*, vol. 53, no. 12, pp. 2535–2543, 2006.
[11] M. Nardelli, G. Valenza, A. Greco, A. Lanata, and E. P. Scilingo, "Recognizing emotions induced by affective sounds through heart rate variability," *IEEE Transactions on Affective Computing*, vol. 6, no. 4, pp. 385–394, 2015.
[12] A. Sellami and H. Hwang, "A robust deep convolutional neural network with batch-weighted loss for heartbeat classification," *Expert Systems with Applications*, vol. 122, pp. 75–84, 2019.
[13] S. Hong, Y. Zhou, J. Shang, C. Xiao, and J. Sun, "Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review," *Computers in Biology and Medicine*, vol. 122, p. 103801, 2020.
[14] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. San Tan, "A deep convolutional neural network model to classify heartbeats," *Computers in biology and medicine*, vol. 89, pp. 389–396, 2017.
[15] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on knowledge and data engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.
[16] L. Yuan and M. Y. Siyal, "Target-oriented augmentation privacy-protection domain adaptation for imbalanced ecg beat classification," *Biomedical Signal Processing and Control*, vol. 86, p. 105308, 2023.
[17] Q. Dong, S. Gong, and X. Zhu, "Imbalanced deep learning by minority class incremental rectification," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 6, pp. 1367–1381, 2018.

[18] A. for the Advancement of Medical Instrumentation *et al.*, "Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms," *ANSI/AAMI EC38*, vol. 1998, 1998.

[19] A. ECAR, "Recommended practice for testing and reporting performance results of ventricular arrhythmia detection algorithms," *Association for the Advancement of Medical Instrumentation*, vol. 69, 1987.

[20] M. M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Information Sciences*, vol. 345, pp. 340–354, 2016.

[21] E. J. d. S. Luz, W. R. Schwartz, G. Cámara-Chávez, and D. Menotti, "Ecg-based heartbeat classification for arrhythmia detection: A survey," *Computer methods and programs in biomedicine*, vol. 127, pp. 144–164, 2016.

[22] B. Chandrakar, O. Yadav, and V. Chandra, "A survey of noise removal techniques for ecg signals," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 3, pp. 1354–1357, 2013.

[23] M. AlMahamdy and H. B. Riley, "Performance study of different denoising methods for ecg signals," *Procedia Computer Science*, vol. 37, pp. 325–332, 2014.

[24] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE transactions on biomedical engineering*, no. 3, pp. 230–236, 1985.

[25] M. d. M. Elena, J. M. Quero, and I. Borrego, "Optimal selection of wavelet coefficients for electrocardiograph compression," *ETRI Journal*, vol. 29, no. 4, pp. 530–532, 2007. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.4218/etrij.07.0206.0213

[26] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *IEEE transactions on biomedical engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.

[27] P. S. Addison, "Wavelet transforms and the ecg: a review," *Physiological measurement*, vol. 26, no. 5, p. R155, 2005.

[28] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, and L. Sornmo, "Clustering ecg complexes using hermite functions and self-organizing maps," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 7, pp. 838–848, 2000.

[29] Y. Kutlu and D. Kuntalp, "Feature extraction for ecg heartbeats using higher order statistics of wpd coefficients," *Computer methods and programs in biomedicine*, vol. 105, no. 3, pp. 257–267, 2012.

[30] A. K. Mishra and S. Raghav, "Local fractal dimension based ecg arrhythmia classification," *Biomedical Signal Processing and Control*, vol. 5, no. 2, pp. 114–123, 2010.

[31] M. A. Escalona-Morán, M. C. Soriano, I. Fischer, and C. R. Mirasso, "Electrocardiogram classification using reservoir computing with logistic regression," *IEEE Journal of Biomedical and health Informatics*, vol. 19, no. 3, pp. 892–898, 2014.

[32] W. Jiang and S. G. Kong, "Block-based neural networks for personalized ecg signal classification," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1750–1761, 2007.

[33] S. Osowski, L. T. Hoai, and T. Markiewicz, "Support vector machine-based expert system for reliable heartbeat recognition," *IEEE transactions on biomedical engineering*, vol. 51, no. 4, pp. 582–589, 2004.

[34] J. Rodriguez, A. Goni, and A. Illarramendi, "Real-time classification of ecgs on a pda," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, no. 1, pp. 23–34, 2005.

[35] F. Sufi and I. Khalil, "Diagnosis of cardiovascular abnormalities from compressed ecg: a data mining-based approach," *IEEE transactions on information technology in biomedicine*, vol. 15, no. 1, pp. 33–39, 2010.

[36] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ecg classification by 1-d convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2015.

[37] H. Dang, M. Sun, G. Zhang, X. Zhou, Q. Chang, and X. Xu, "A novel deep convolutional neural network for arrhythmia classification," in *2019 International Conference on Advanced Mechatronic Systems (ICAMechS)*. IEEE, 2019, pp. 7–11.

[38] S. L. Oh, E. Y. Ng, R. San Tan, and U. R. Acharya, "Automated beat-wise arrhythmia diagnosis using modified u-net on extended electrocardiographic recordings with heterogeneous arrhythmia types," *Computers in biology and medicine*, vol. 105, pp. 92–101, 2019.

[39] H. Fujita and D. Cimr, "Computer aided detection for fibrillations and flutters using deep convolutional neural network," *Information Sciences*, vol. 486, pp. 231–239, 2019.

[40] L. Sun, Y. Wang, Z. Qu, and N. N. Xiong, "Beatclass: a sustainable ecg classification system in iot-based ehealth," *IEEE Internet of Things Journal*, vol. 9, no. 10, pp. 7178–7195, 2021.

[41] H. Wang, H. Shi, K. Lin, C. Qin, L. Zhao, Y. Huang, and C. Liu, "A high-precision arrhythmia classification method based on dual fully connected neural network," *Biomedical Signal Processing and Control*, vol. 58, p. 101874, 2020.

[42] S. M. Mathews, C. Kambhamettu, and K. E. Barner, "A novel application of deep learning for single-lead ecg classification," *Computers in biology and medicine*, vol. 99, pp. 53–62, 2018.

[43] A. M. Shaker, M. Tantawi, H. A. Shedeed, and M. F. Tolba, "Generalization of convolutional neural networks for ecg classification using generative adversarial networks," *IEEE Access*, vol. 8, pp. 35 592–35 605, 2020.

[44] Ö. Yildirim, "A novel wavelet sequence based on deep bidirectional lstm network model for ecg signal classification," *Computers in biology and medicine*, vol. 96, pp. 189–202, 2018.

[45] S. L. Oh, E. Y. Ng, R. San Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of cnn and lstm techniques with variable length heart beats," *Computers in biology and medicine*, vol. 102, pp. 278–287, 2018.

[46] F. Li, J. Wu, M. Jia, Z. Chen, and Y. Pu, "Automated heartbeat classification exploiting convolutional neural network with channel-wise attention," *IEEE Access*, vol. 7, pp. 122 955–122 963, 2019.

[47] J. Wang, X. Qiao, C. Liu, X. Wang, Y. Liu, L. Yao, and H. Zhang, "Automated ecg classification using a non-local convolutional block attention module," *Computer Methods and Programs in Biomedicine*, vol. 203, p. 106006, 2021.

[48] S. H. Khan, M. Hayat, M. Bennamoun, F. A. Sohel, and R. Togneri, "Cost-sensitive learning of deep feature representations from imbalanced data," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 8, pp. 3573–3587, 2017.

[49] C. Huang, Y. Li, C. C. Loy, and X. Tang, "Deep imbalanced learning for face recognition and attribute prediction," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 11, pp. 2781–2794, 2019.

[50] M. Zubair and C. Yoon, "Cost-sensitive learning for anomaly detection in imbalanced ecg data using convolutional neural networks," *Sensors*, vol. 22, no. 11, p. 4075, 2022.

[51] Y. Xia, Y. Xiong, and K. Wang, "A transformer model blended with cnn and denoising autoencoder for inter-patient ecg arrhythmia classification," *Biomedical Signal Processing and Control*, vol. 86, p. 105271, 2023.

[52] J. Kim, J. Jeong, and J. Shin, "M2m: Imbalanced classification via major-to-minor translation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 13 896–13 905.

[53] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-balanced loss based on effective number of samples," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 9268–9277.

[54] G. D. Clifford, F. Azuaje, P. McSharry *et al.*, *Advanced methods and tools for ECG data analysis*. Artech house Boston, 2006, vol. 10.

[55] G. K. Prasad and J. Sahambi, "Classification of ecg arrhythmias using multi-resolution analysis and neural networks," in *TENCON 2003. Conference on convergent technologies for Asia-Pacific region*, vol. 1. IEEE, 2003, pp. 227–231.

[56] W. Jiang and S. G. Kong, "Block-based neural networks for personalized ecg signal classification," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1750–1761, 2007.

[57] T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ecg signals," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1415–1426, 2009.

[58] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "Eca-net: Efficient channel attention for deep convolutional neural networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11 534–11 542.

[59] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 3–19.

[60] H. Shi, H. Wang, Y. Huang, L. Zhao, C. Qin, and C. Liu, "A hierarchical method based on weighted extreme gradient boosting in ecg heartbeat classification," *Computer methods and programs in biomedicine*, vol. 171, pp. 1–10, 2019.

[61] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[62] H. Huang, J. Liu, Q. Zhu, R. Wang, and G. Hu, "A new hierarchical method for inter-patient heartbeat classification using random projections and rr intervals," *Biomedical engineering online*, vol. 13, no. 1, pp. 1–26, 2014.