# Comparative Analysis of ECG Data Augmentation Methods in Arrhythmia Classification

Jonas Mindaugas Rimšelis, Povilas Treigys, Zigmantas Kestutis Juškevičius, Jolita Bernatavičienė

Institute of Data Science and Digital Technologies, Vilnius University

## Introduction

An electrocardiogram (ECG) measures electrical signals from the heart to capture various cardiovascular conditions. Distinct patterns can be identified in ECG signals during the development of arrhythmias. The growing availability of single-lead Holter devices enables continuous cardiac monitoring and supports earlier arrhythmia detection using deep learning [1]. Since abnormal heartbeats occur rarely, even in patients diagnosed with arrhythmia, data used for training models are imbalanced, leading to poor generalization and robustness [2]. To circumvent this, data augmentation is utilized to mitigate label balancing issues. This study is split into two parts:

- Literature review of data augmentation techniques and preprocessing methods;
- Empirical study of Synthetic Minority Over-sampling Technique (SMOTE) and Denoising Diffusion Probabilistic Model (DDPM) models to improve the accuracy of arrhythmia beat classification.

## Literature Review

Table: Literature review results.

| Ref. | Preprocessing | Segmentation | Augmentation | Classifier | Accuracy |
|---|---|---|---|---|---|
| [3] | Butterworth BP | Beat seg., convert to image | Auxiliary Classifier GAN | LC-CNN | **99.22%** |
| [4] | Butterworth BP | Beat seg. | BC-GAN | MISEResNet-BiLSTM | **99.93%** |
| [5] | - | Beat seg., convert to image | Multimodality Data Matching-Based | Multimodality Feature Encoding, Fusing, Softmax | **98.83%** |
| [6] | - | Beat seg. | TCGAN | Multi-scale Conv1d + Bi-LSTM | **94.69%** |
| [7] | Wavelet denoising | 5-10s seg. | Stretching, scaling, Gaussian noise | dual path CNN-BiLSTM | **95.48%** |
| [8] | Butterworth BP, zscore normalization | Beat seg. | K-means undersampling + SMOTE | MB-MHA-TCN | **98.75%** |
| [9] | Min-Max normalization | 10s seg., norm., beat seg. | SMOTE, ADASYN | CNN | SMOTE: **94.75%**, ADASYN: **95.78%** |
| [10] | - | Beat seg. | TimeGAN, ECGAN, DDPM | t-SNE | TimeGAN: **79.2 ± 12.8**, ECGAN: **93.8 ± 5.1**, DDPM: **69.2 ± 18.6** |
| [11] | - | Beat seg., convert to image | DCGAN, CGAN, WGAN-GP | 2D-CNN | Normal class: **77.9%**; Abnormal class: **93.1%** |

## Insights

- **Preprocessing and segmentation:** Filtering and denoising steps are rarely used, but Butterworth bandpass filters followed by Z-score or Min-Max normalization coupled with beat-level segmentation are the standard.
- **Augmentation:** Generative Adversarial Networks (GANs) have replaced traditional methods (like SMOTE) as the most common approach for data augmentation.
- The highest accuracy of 99.83% was achieved using BC-GAN (which uses both BiLSTMs and CBAM modules) for data augmentation.
- An emerging trend of converting beats to images and augmenting data with an Auxiliary Classifier GAN resulted in an accuracy of 99.22%.

For further research, SMOTE was chosen as a baseline for augmentation and DDPM for its novelty in ECG signal augmentation and its capability to generate realistic waveforms.

## Pipeline

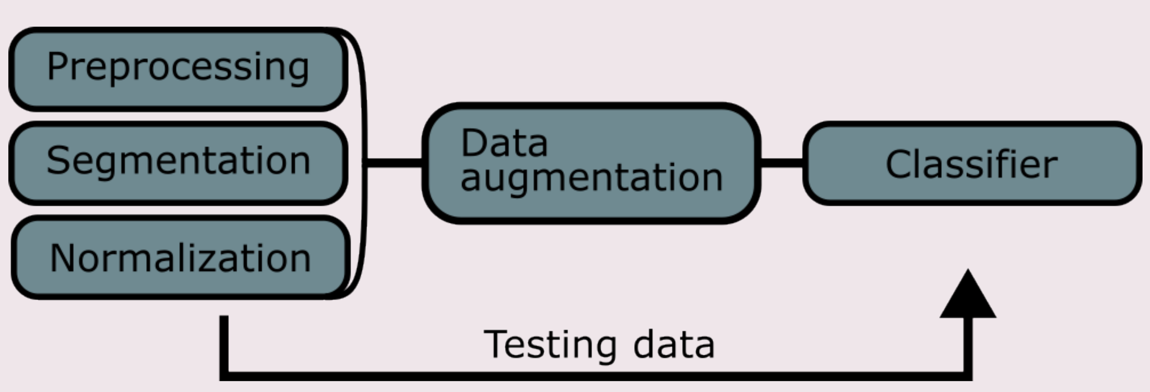


Figure: Overview of the ECG classification pipeline.

## MIT-BIH Arrhythmia Database

Table: Beat Counts of the MIT-BIH Arrhythmia Database by AAMI Class and DS1/DS2 split with adherence to AAMI labels and inter patient paradigm

| AAMI Class | DS1 (Training Set) Beats | DS2 (Testing Set) Beats | Total Beats |
|---|---|---|---|
| N (Normal) | 45,866 | 44,259 | 90,125 |
| S (Supraventricular) | 944 | 1,837 | 2,781 |
| V (Ventricular) | 3,788 | 3,221 | 7,009 |
| F (Fusion)* | 415 | 388 | 803 |
| Q (Unclassified)* | 8 | 7 | 15 |
| **Total** | **51,013** | **49,705** | **100,718** |

**Note: Q-class beats are **excluded** from training, while F beats from evaluation, allowing the analysis to focus on clinically relevant minority classes, particularly the S-class.*

## Classifier

2D-CNN-BiLSTM classifier was chosen for the classification task:

- Inputs: the raw beat segment, a QRS complex mask, and R–R interval features.
- A 3-layer 2D-CNN with a BiLSTM. Concurrently, R–R features are embedded from 3D to 16D.
- A 2-layer MLP is used for the class assignment.

Training process: 15 epochs with Focal Loss (No augmentation; DDPM weights: N–0.08, S–0.72, V–0.15, F–0.05; SMOTE weights: N–0.24, S–0.47, V–0.24, F–0.06). Because the F class received a very small effective weight, the model did not learn this class reliably; therefore, F-class performance is omitted from the results section.

## Data augmentation methods

- **SMOTE** with k neighbors set to 3 was utilized to augment the beat-level segment vector with concatenated RR features.
- **DDPM** architecture consists of a 1D U-Net with skip connections and ResBlock1D. Generated signals were filtered to have at least 0.3 correlation with their original signal. Example of the generated signals:

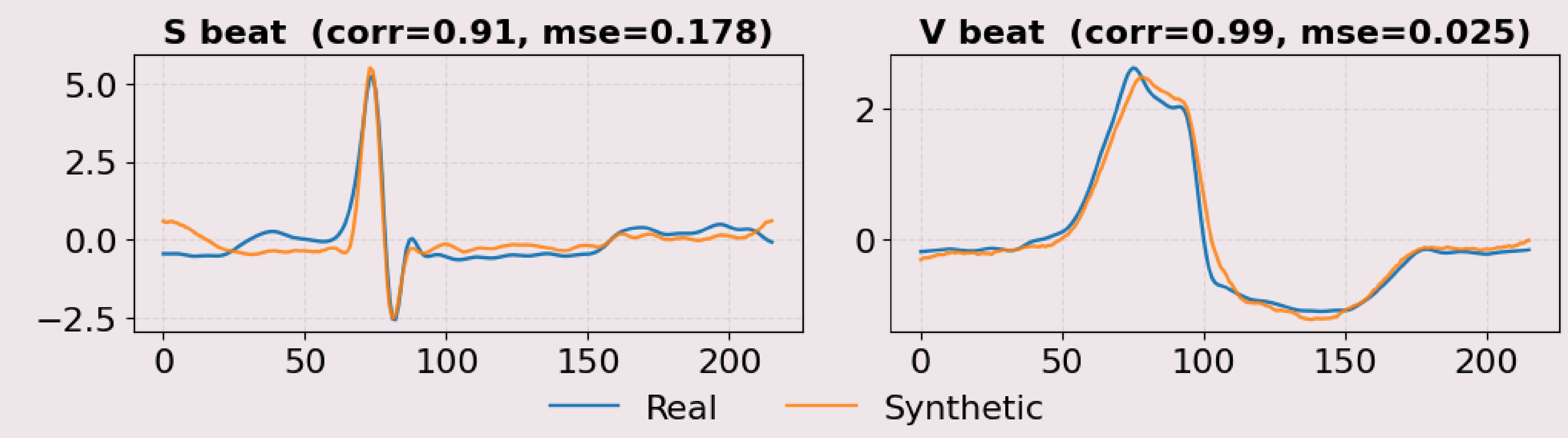


Figure: Generated S and V beats using DDPM

## Results

Table: Classification performance per class using different augmentation methods.

| Method | Class | Precision | Recall | F1-score | Overall Acc. |
|---|---|---|---|---|---|
| **No augmentation (Z-score)** | N | 0.96 | 0.98 | 0.97 | 0.93 |
| | S | 0.28 | 0.33 | 0.30 | |
| | V | 0.92 | 0.77 | 0.84 | |
| **SMOTE (Min-Max)** | N | 0.95 | 0.80 | 0.86 | 0.77 |
| | S | 0.11 | 0.33 | 0.16 | |
| | V | 0.75 | 0.77 | 0.76 | |
| **DDPM (Z-score)** | N | 0.96 | 0.97 | 0.97 | 0.94 |
| | S | 0.40 | 0.31 | 0.35 | |
| | V | 0.85 | 0.92 | 0.88 | |

## Conclusions and Future Work

- Classifier trained on SMOTE-augmented data performed worse (Overall Acc. – 77%) than the baseline - classifier with no augmentation (Overall Acc. – 93%).
- The DDPM method increased S-beat precision by 12% and F1-score by 5% compared to the baseline method.
- Focal loss with stronger class weighting can partially alleviate class imbalance.

Future work includes:

- Enhancing temporal feature extraction within the classifier architecture;
- Experimenting with GAN-based approaches for data augmentation;
- Incorporating longer ECG segments during training and augmentation to capture richer temporal dynamics.

## References

https://github.com/jonasmindaugasr/damms2025dataaugmentation

This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-ITP-25-9. Publication/Research is funded by Research Council of Lithuania under the Programme ‘University Excellence Initiatives’ of the Ministry of Education, Science and Sports of the Republic of Lithuania (Measure No. 12-001-01-01-01 ‘Improving the Research and Study Environment’). Project No.: S-A-UEI-23-11.