Preprints.org

Article

Not peer-reviewed version

End-to-End Premature Ventricular Contraction Detection using Deep Neural Networks

Dimitri Kraft , Gerald Bieber , Peter Jokisch , Peter Rumm *

Posted Date: 25 September 2023

doi: 10.20944/preprints202309.1588.v1

Keywords: Ventricular Premature Contractions (PVC) detection; 1D U-Net neural network; Holter monitoring

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

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.

Article

End-to-End Premature Ventricular Contraction Detection Using Deep Neural Networks

Dimitri Kraft [1], Gerald Bieber [1], Peter Jokisch [2] and Peter Rumm [2]

[1] Fraunhofer IGD Rostock; firstname.lastname@igd-r.fraunhofer.de
[2] custo med GmbH; firstname.lastname@customed.de

Abstract: In Holter monitoring, the precise detection of standard heartbeats and Ventricular Premature Contractions (PVCs) is paramount for accurate cardiac rhythm assessment. This study introduces a novel application of the 1D U-Net neural network architecture aimed at enhancing PVC detection in Holter recordings. Training data comprised the Icentia11k, INCART DB, and our custom dataset. The model's efficacy was subsequently validated against traditional Holter analysis methodologies across multiple databases, including AHA DB, MIT 11 DB, NST, and another custom dataset encompassing challenging real-world examples. The results underscored the 1D U-Net model's prowess in QRS complex detection, achieving near-perfect balanced accuracy scores across all databases. PVC detection exhibited variability, with balanced accuracy scores ranging from 0.909 to 0.986. Despite some databases, like the AHA DB, showcasing lower sensitivity metrics, their robust balanced accuracy accentuates the model's equitable performance in discerning both false positives and negatives. In conclusion, while the 1D U-Net architecture is a formidable tool for QRS detection, there's a clear avenue for further refinement in its PVC detection capability, given the inherent complexities and noise challenges in real-world PVC occurrences.

Keywords: Ventricular Premature Contractions (PVC) detection; 1D U-Net neural network; Holter monitoring

1. Introduction

Premature Ventricular Contractions (PVCs) are a prevalent cardiac arrhythmia frequently associated with an array of severe cardiovascular diseases. The swift, precise, and real-time identification of these events is pivotal for timely clinical interventions. With the increasing ubiquity of consumer-grade cardiac monitoring devices, there is an urgent demand for sophisticated detection algorithms that can operate effectively on these devices. Importantly, such an algorithm needs to handle the limitations of single-channel data, effectively mitigate noise interference, and function independently of cloud resources.

In this publication, we introduce a novel approach to PVC detection, conceptualizing the task as a one-dimensional (1D) segmentation problem. We employ a deep learning model rooted in the U-Net architecture [1]. Originally conceived for biomedical image segmentation, the U-Net model has demonstrated its adaptability and efficiency across a range of applications. In our research, we repurpose it to process 1D electrocardiogram (ECG) signals, underscoring its potential for robust, end-to-end PVC detection. Our objective was to devise an algorithm with the capability to consistently recognize the intra- and inter-individual variability that PVCs can exhibit.

This study's emphasis is on the model's ability to process noisy, single-channel ECG data, commonly obtained from consumer-grade devices. We evaluate our single-channel model on two and three-channel ECG data using a simple merging technique of beat lists. Processing a single ECG at a time may have multiple advantages over processing multiple leads at once. We see the following advantages:

1. **Simplicity**: Single-channel ECG processing is simpler and more straightforward, which can make the development and debugging of algorithms easier.

© 2023 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

2. **Data Availability**: In some situations, only single-lead ECG data might be available (e.g., Icentia11k DB). Many portable and wearable ECG devices only record a single lead, so algorithms designed for single-lead data can be more broadly applicable.
3. **Robustness to Noise**: Single-lead ECGs might be less susceptible to noise and artifacts that can affect multi-lead recordings. For instance, movement artifacts can affect different leads to different extents, potentially making multi-lead data more challenging to interpret. By analyzing each lead independently we may overcome this.

Although single-channel ECG data often contain a significant amount of noise, our model excels in detecting both normal and PVC beats with great precision, marking a substantial improvement over current techniques. A standout aspect of our approach is its operation solely on the existing hardware of physicians, thus rendering cloud support redundant. By doing so, our method addresses and solves latency problems linked with cloud-based processing while bolstering data privacy. Consequently, it facilitates prompt analysis of the ECG.

2. Related Work

The application of neural networks for automated detection of cardiac arrhythmias such as PVCs has been the subject of research for a number of years, with various models demonstrating improved outcomes over conventional signal processing techniques. These studies provide valuable insights for our current research.

One of the earlier significant works was presented by Kiranyaz, Ince, and Gabbouj [2], who developed a 1-Dimensional Convolutional Neural Network (1-D CNN) for patient-specific ECG classification, inclusive of PVC detection. Despite its impressive adaptability to unique ECG patterns of individual patients, the need for patient-specific training may hinder its scalability and applicability to real-time detection, especially on consumer-grade devices.

Following this, Acharya et al. [3] proposed a Deep Convolutional Neural Network (CNN) for automated detection of cardiac disorders, including PVCs. The model was trained using single-lead ECG signals, showing promising outcomes but handled only evaluated high frequency noise (gaussian).

More recently, Hannun et al. [4] developed a deep learning model, referred to as the Cardiologist-Level Arrhythmia Detector (CLAD), which employed a 34-layer CNN for multi-class detection of 14 types of cardiac arrhythmias from single-lead ECG records, including PVCs. While CLAD demonstrated remarkable performance, its high complexity could potentially limit its implementation on consumer-grade devices with constrained computational resources. Moreover, it was not specifically tailored for PVC detection.

While 1D ECG arrythmia detection approaches achieve good performance, 2D-image based ECG detecion methods also exists [5]. Either a image of a ECG recording or the analysis of a 2D spectrogram [6] is widely used to approach the detection of arrhythmia.

The authors of [7] present an approach for PVC classification that leverages a straightforward Support Vector Machine (SVM), achieving commendable sensitivity and specificity rates around 99%. Their methodology encompasses a range of preprocessing and feature extraction procedures. Nevertheless, there is a noteworthy detail in their experimental design that may influence the interpretation of these results. The team employed 10% of the PVC beats from the MIT dataset for training purposes, the very dataset they subsequently used for validation. Given the nature of the MIT dataset, which is characterized by substantial inter-subject variability contrasted by minimal intra-subject variability, this could potentially introduce the risk of overfitting and biased results. The issue is further compounded in real-world applications where ECG patterns showcase higher complexity and diversity. This raises concerns about the model's robustness and its ability to generalize in such scenarios. Interestingly, these reservations seem to gain some traction when we consider the performance drop by 2% seen on their experimental dataset, which comprises only 903 PVC beats.

We see further multiple issues in the current literature for PVC detection:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

1. **Real world settings are not considered**. The electrocardiogram (ECG) data inherently contain various types of noise, including baseline wandering, power line interference, muscle noise and other artifacts related to contact with the electrodes. These noise elements pose significant challenges to the extraction of robust features, consequently affecting the performance of PVC classification in real-world settings. Thus, an algorithm that performs well on a clean, noise-free dataset may not perform as well when deployed in a real-world setting where the noise level is higher or varies unpredictably.
2. **Testing datasets are not representative** : Gender differences in ECG are well documented in the literature. Men and women can have different heart rates, QRS complex durations, QT intervals, and T-wave morphologies, among other characteristics. These differences can affect the performance of PVC detection algorithms if they are not properly accounted for during algorithm development and testing [8,9].
3. **Training and testing datasets are not separated**. A notable limitation of many existing methods lies in their reliance on small or overlapped ECG datasets for training and testing. This practice raises questions about their efficiency and generalizability when applied to a large collection of ECG recordings, an issue that remains largely unaddressed in the literature [10].

In the light of these existing studies, our research aims to bridge the gap by delivering a robust, efficient, and highly accurate model for PVC detection, specifically tailored for noisy, single-channel ECG data from consumer-grade devices. The novelty of our approach lies in its unique application of the U-Net architecture to the problem of PVC detection as a 1D segmentation task. Moreover, our model operates independently of cloud resources, ensuring real-time detection with improved data privacy.

2.1. Beat Detection Performances

Table 1. Comparison for normal beat detection algorithms on MIT (VFib excluded) - adapted from [11].

Algorithm	Dataset	Se (%)	PPV (%)	F1
Pan and Tompkins [12]	MIT-BIH	99.76	99.56	99.66
Christov [13]	MIT-BIH	99.74	99.65	99.69
Chiarugi et al [14]	MIT-BIH	99.76	99.81	99.78
Chouakri et al [15]	MIT-BIH	98.68	97.24	97.95
Elgendi [16]	MIT-BIH	**99.78**	**99.87**	**99.82**
BeatLogic [11]	MIT-BIH	99.60	99.78	99.69
Liu et al. [17]	MIT-BIH	99.00	99.20	99.10
He et al. [18]	MIT-BIH	99.56	99.72	99.64
Martinez et al [19]	MIT-BIH VFib excluded	**99.80**	**99.86**	**99.83**
Arzeno et al [20]	MIT-BIH VFib excluded	99.68	99.63	99.65
Zidelmal et al [21]	MIT-BIH VFib excluded	99.64	99.82	99.73
BeatLogic [11]	MIT-BIH VFib excluded	99.60	99.90	99.75

2.2. PVC Detection Performances

Table 2. Comparison for PVC beat detection algorithms on MIT 11 records subset - adapted from [11].

Algorithm	Dataset	Se (%)	PPV (%)	F1
de Chazal et al [22]	MIT-BIH 11	77.5	90.6	83.5
Jiang and Kong [23]	MIT-BIH 11	94.3	95.8	95.0
Ince et al [24]	MIT-BIH 11	90.3	92.2	91.2
Kiranyaz et al [2]	MIT-BIH 11	95.9	96.2	96.0
Zhang et al [25]	MIT-BIH 11	97.6	97.6	97.6
BeatLogic [11]	MIT-BIH 11	**97.9**	**98.9**	**98.4**
Liu et al [17]	MIT-BIH (22 Records)	**91.6**	**95.6**	**93.6**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

3. Methods

In the methodology section of this paper, we delve into the specific techniques utilized for the detection and differentiation of normal heartbeats and Premature Ventricular Contractions (PVCs). Given the critical implications these heart rhythms have in clinical practice and patient care, the accuracy and efficiency of these detection algorithms are paramount. Our methodological approach incorporates a broad range of computational tools and techniques, each tailored to handle the specific characteristics of normal heartbeats and PVCs. These include, but are not limited to data preperation, signal preprocessing, machine learning algorithms, and evaluation metrics.

3.1. Overview

In our study, we employed a 1D U-Net architecture [1] for end-to-end detection of Premature Ventricular Contractions (PVCs). The U-Net consisted of four encoder layers, one bottleneck layer, and four decoder layers. The number of filters in the initial layer was set to 16 and was multiplied by 2 after each encoder layer. The U-Net convolution blocks used had a filter size of 9 and a dilation of 3. Instead of the conventional max-pooling operations in the encoding stages, we used strided convolutions, and each convolution was followed by a Rectified Linear Unit (ReLU) activation function.

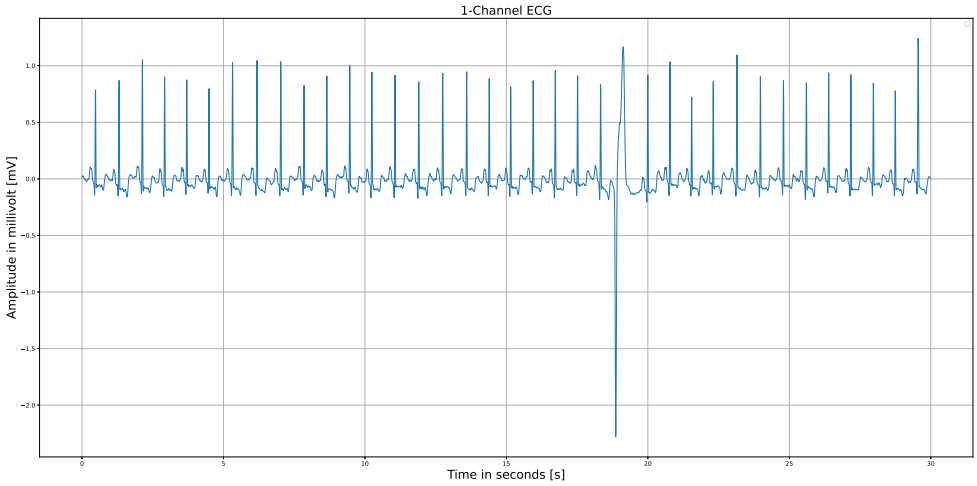
For the training of the U-Net, a composite loss function comprising the Dice loss [26] and the Focal loss [27] was used, with equal weights of 0.5 assigned to each loss. The Focal loss incorporated alpha coefficients of 1, 1, and 1.5. An initial learning rate of 0.001 was set, and the network was optimized using the Adam optimizer [28] with a betas value of 0.9 and 0.999, no weight decay.

Our dataset included the St. Petersburg 12-channel ECG dataset [29], the Icentia 11k dataset [30], and a custom dataset containing around 7500 10 to 30 second single channel ECG windows. We targeted the detection of normal and PVC heartbeats. We designed a segmentation mask where the class 1 was assigned to normal beats within a window of 200 ms around the R peak, while class 2 was assigned to PVCs detected 100ms before and 150ms after the R peak. The background was labeled as class 0.

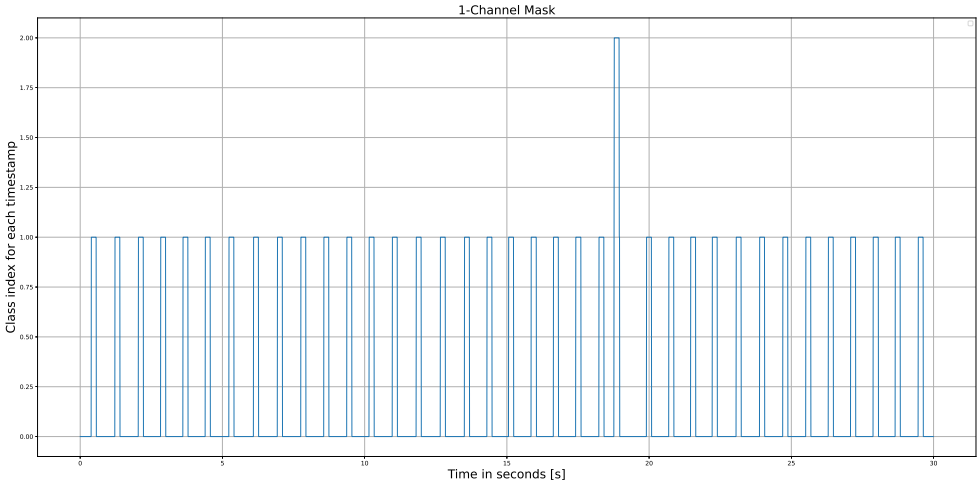
3.2. Data Preparation

The ECG signals were divided into discrete 30-second segments, without any overlap, and a corresponding segmentation mask was developed for each of these sections as previously described. If the ECG signals were less than 30 seconds in length, they were zero-padded to reach the desired 30-second duration. We made the decision not to normalize the amplitude of the ECG signals, opting instead to apply certain preprocessing steps. These included the utilization of a high-pass forward-backward Butterworth filter with a set cutoff frequency of 0.5 Hz as well as a power line filter. To further refine the signals, all were down-sampled to a frequency of 125 Hz using linear interpolation. These preprocessing steps were executed on the entirety of the signal prior to the division into the aforementioned 30-second segments. This procedural routine was consistently applied across all the datasets. The samples we selected only contained either a normal or a premature ventricular contraction (PVC) beat. We consciously omitted windows that contained beats which could not be accurately classified. Examples for 30 seconds ECG and their corresponding masks are depicted in Figure 1. The compilation of the training set resulted in approximately 4 million 30-second windows. In addressing the common issue of class imbalance inherent in medical datasets, an oversampling strategy was adopted specifically for the PVC class for training. The focal loss function, though employed as a countermeasure to class imbalance, does not eliminate the possibility of skewed learning towards the more predominant class. Consequently, to ensure an unbiased learning environment and enhance the PVC detection capabilities of our model, we endeavored to maintain a balanced representation of PVC and normal heartbeat classes within our training dataset.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1



(a) Example ECG from the MIT Arrhythmia dataset. A single PVC is located at around second 18.



(b) The corresponding mask containing background, normal beats and PVC encoded with 0, 1 and 2, respectively. Each beat is labeled with 100 ms before the r-peak and 100 after the r-peak. The PVC is labeled 100 ms before the r-peak and 150 ms after the r-peak to catch the whole property of a PVC beat.

Figure 1. .

3.3. Model Architecture

In our methodology, we employed a conventional 1D-UNet architecture for ECG, consisting of 4 encoder layers, a single bottleneck layer, and 4 decoder layers. We have chosen a one-dimensional architecture for processing ECG data. This allows analysis of single and multi-channel ECG data and is not limited to a particular hardware constellation. Although simultaneous inclusion of all channels would be desirable, a 1D architecture makes preprocessing and explainable AI more appropriate. In addition, a multichannel ECG may provide low-noise signals on one channel while the others are of poor quality. This interferes with simultaneous data analysis of the total data set. Therefore, a 3D architecture would be more error-prone or complex to design, while a 1D architecture is fast, efficient, and can easily handle and identify noisy or corrupt data. We consciously decided not to incorporate self-attention layers, preferring to retain standard dilated convolution layers. The rationale behind this

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

decision was to enable efficient real-time inference on consumer-grade devices, crucial for practical applicability. Moreover, we intentionally developed a localized solution, given that the transmission of sensitive health data to a cloud or server is generally not endorsed by the majority of cardiologists in Germany due to privacy concerns. U-Net is composed of two parts: the "contracting path" (or encoder) and the "expanding path" (or decoder). The bottleneck layer is the layer that connects these two paths. In a traditional U-Net architecture, the bottleneck consists of two convolutional layers followed by a ReLU activation function. This layer serves as a bridge between the encoder and the decoder, reducing the spatial dimensions of the input data and allowing the network to focus on the most important features. It provides an abstracted, high-level understanding of the input. The encoder layers capture the context in the ECG while the decoder layers enable precise localization using upsample layers and convolutions. Our convolutional blocks consists of standard convolutional layers with a kernel size of 9, a dilation of 3 and same padding. The second convolution layer in a block also has a stride of 2. We further use 1D Batch Normalization [31] and a ReLU activation function. Due to dilation we effectively utilize a receptive field of 25, which results in approx. 200 ms in the first layer. The output layer contains a 1x1 convolution which maps the 16 output channels from the previous convolutional layer to a N x C output matrix where N is the input length of the ECG and C is the number classes, in our case 3). All hyperparameters where identified using extensive grid search. The architecture are depicted in Figure 2.

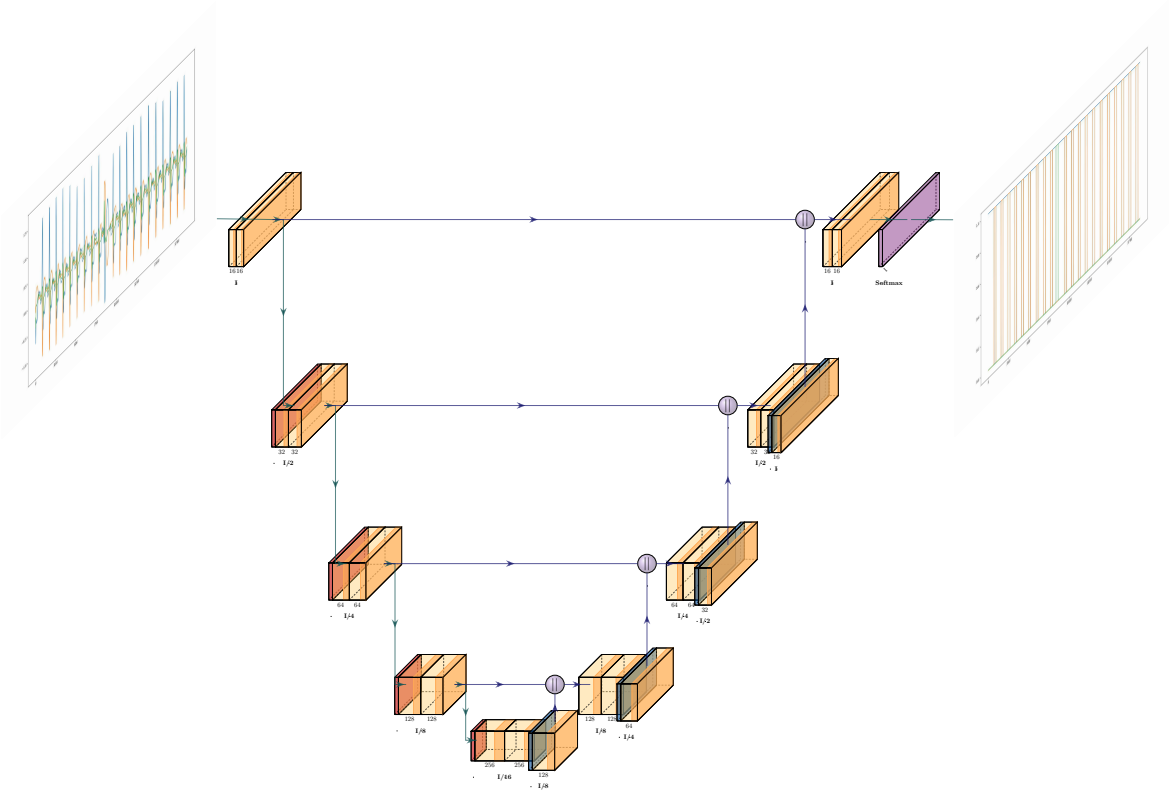


Figure 2. Our UNet 1D Architecture for ECG Beat Detection. Each block consists of a double convolution layer (1D Conv - Batch Norm - ReLU). For the decoder, we use simple linear up-sampling layer instead of a transposed convolution layer. The output of our model is a N x C matrix consisting of three masks (no beat, beat and PVC label for each timestamp).

3.4. Augmentation

During the training phase, a strategic augmentation of data was undertaken to promote a higher degree of model generalizability and resilience against variations in real-world scenarios. This augmentation process incorporated the random scaling of signal amplitudes, the infusion of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

random Gaussian, pink and brown noise, and the induction of minor baseline shifts. Notably, we refrained from employing other prevalent augmentation procedures, such as signal masking, temporal shifting, time compression and stretching, mixup [32], and cutmix [33].

To evaluate the effect of data augmentation, we apply a combination of the following transformations dynamically during training:

- Scaling the amplitude with a probability of $p = 0.75$ and scaling factor of $[0.6 \dots 1.4]$.
- Offset the amplitude with a probability of $p = 0.75$ and a offset value of $[-0.2 \dots 0.2]$.
- Adding gaussian, brown or pink noise with a probability of $p = 0.75$ and a offset value of $[-0.2 \dots 0.2]$.

3.5. Post-Processing

The model outputs are processed using the softmax function, an exponential function that generates three float values in range from 0.0 to 1.0, and their sum is 1.0. These values can be interpreted as the probability that the corresponding time point lies within no beat, a normal beat, or a PVC (Premature Ventricular Contraction). When the first, second, or third values of the triples are grouped together as a vector, three masks are created:

1. **The No Beat Mask**: This provides the probability for each time point that it does not lie within any beat.
2. T**he Normal Beat Mask**: This provides the probability for each time point that it lies within a normal beat.
3. **The PVC Mask**: This provides the probability for each time point that it lies within a PVC.

The PVC (Premature Ventricular Contraction) list is created by scanning the PVC mask until a threshold of 0.5 is reached or exceeded, marking the potential onset of a PVC. The termination of the potential PVC is detected when the threshold of 0.5 is consistently undershot for at least 2 samples. A PVC is acknowledged if the period between the onset and termination of the threshold exceedance lasts for a minimum of 5 samples (equivalent to 40 milliseconds). The midpoint between the onset and termination of the exceedance region is selected as the point of the beat. The list for normal beats is created in an analogous manner. Following this, the two lists are merged. Lastly, beats that lie within regions that clearly do not contain ECG, such as at the end of the evaluation when electrodes have been already removed, and where the neural network might have erroneously detected pseudo-beats, are canceled.

3.6. Training Data

We utilize three different datasets as training data. The first one is a custom 3-Channel ECG dataset collected from a wide variety of subjects. The second dataset is the Icentia11k dataset, consisting of single channel ECG data from 11,000 subjects, a large dataset for training a End-to-End PVC detection neural network. The third dataset is the 12-Channel St. Peterburgs INCART dataset.

3.6.1. Custo Med Training Dataset

We utilized a custo med flash 500/510 3-Channel Holter (see Figure 3), with a sampling frequency of 125 Hz and 5.6 microvolt/Bit resolution with 10 bit resolution.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1



Figure 3. Custo med flash 500/510 3-Channel Holter used for data acquisition of our custom dataset.

This data were obtained in an anonymized form from one of our clients. As such, we do not possess information regarding the age and sex of the individuals associated with the electrocardiograms. For the training data, approximately 1,000 ECGs were employed, albeit not all from unique patients. From these ECGs, we generated between 3 to approximately 30 snippets of varying lengths (ranging between 10 seconds to around 120 seconds) which resulted in 7500 ECG samples. These ECG snippets underwent careful evaluation, with corrections made to annotations as necessary to ensure accuracy. In addition, we utilized 36 different 24-hour ECG recordings taken from 36 unique patients for an extended, long-term monitoring evaluation. The design of this dataset was aimed at focusing on QRS and PVC classes specifically, thus only these were annotated (see Table 3).

Table 3. Count and description of ECG beat types - Custo med training dataset.

Symbol	Beat Description	Count
N	Normal	3,361,174
V	Premature ventricular contraction	163,592

3.6.2. Icentia11k

The Icentia11k dataset was primiraly desigend unsupervised representation learning for arrhythmia subtype discovery. The data originates from CardioSTAT, a single-lead heart monitoring device developed by Icentia [30]. This device records raw signals at a 250 Hz sampling rate with a 16-bit resolution in a modified lead 1 position. The dataset is compiled from the records of 11,000 patients predominantly from Ontario, Canada, who were monitored using the CardioSTAT device across various medical centers. The data underwent analysis by a team of 20 technologists from Icentia, using proprietary analysis tools. The initial beat detection was conducted automatically, following which a technologist manually examined the record, labeling beat types and rhythms by conducting a full disclosure analysis, wherein the entire recording is assessed. Each analysis underwent final approval by a senior technologist prior to inclusion in the dataset. The average age of the patient is 62.2±17.4 years. The Icentia11k data set consists of 2 Billion normal but only 17 Million PVC beats (see Table 4. We therefore sub-sampled the dataset that we only take 1% of segments with normal beats and 100% of the segments where atleast one PVC beat is present. Since PVC beats are rather rare in the dataset (less than 1%), we mostly settle with segments containing normal beats with little variation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

Table 4. Count and description of ECG beat types - Icentia 11k.

Symbol	Beat Description	Count
N	Normal	2,061,141,216
S	Premature or ectopic supraventricular beat	19,346,728
V	Premature ventricular contraction	17,203,041
Q	Undefined: Unclassifiable beat	676,364,002

3.6.3. St. Petersburg INCART 12-lead Arrhythmia Database

This database is composed of 75 annotated recordings, which have been derived from 32 Holter records. Each recording extends for a duration of 30 minutes and comprises 12 standard leads. These leads have been sampled at a rate of 257 Hz, with the gains fluctuating between 250 to 1100 analog-to-digital converter units per millivolt. The database's reference annotation files collectively contain more than 175,000 beat annotations.

The original records were gathered from patients who were undergoing tests for coronary artery disease. The patient demographic comprised of 17 males and 15 females, ranging between the ages of 18 and 80, with the mean age being 58. All patients in the cohort were devoid of pacemakers; however, most presented with ventricular ectopic beats. The records selected for inclusion in the database were primarily those of patients demonstrating ECG patterns indicative of ischemia, coronary artery disease, conduction abnormalities, and arrhythmias. Key observations from the selected records are as follows:

Table 5. Patient counts by diagnosis - St. Petersburg INCART.

Diagnosis	Patients
Acute MI	2
Transient ischemic attack (angina pectoris)	5
Prior MI	4
Coronary artery disease with hypertension	7
Sinus node dysfunction	1
Supraventricular ectopy	18
Atrial fibrillation or SVTA	3 (2 with paroxysmal AF)
WPW	2
AV block	1
Bundle branch block	3

Annotation of the data was initially performed by an automated algorithm, which was then manually rectified in accordance with standard PhysioBank beat annotation definitions. The algorithm typically positions beat annotations at the center of the QRS complex, as deduced from all 12 leads. Despite this, these locations have not undergone manual corrections, which might lead to occasional misalignments in the annotations.

3.7. Test data

The test data consists of publicly available and private data sets. No testing data is included in the training data, so we have a fairly good understanding of the generalizibility of our model. We utilize the standard MIT and MIT-11 subset dataset, the AHA dataset, the NST dataset to evaluate the performance of the model under noisy conditions and our own collected datasets, CST and CST Strips, where CST DB is a dataset consists of 18 records of 24h ECG samples and CST Strips contains 628 difficult real world ECG data samples (noisy, containing couplets and triplets PVC and other anomalies) from different subjects.

3.7.1. AHA

The American Heart Association Database (AHA Database) is not available for download from PhysioNet or similar platforms but can be exclusively obtained from ECRI. Established during the late 1970s and early 1980s through a collaboration between the American Heart Association and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

the National Heart, Lung, and Blood Institute, the AHA Database serves as a valuable resource for evaluating ventricular arrhythmia detectors. Completed in 1985 without subsequent updates, it comprises 80 two-channel analog ambulatory ECG recordings. These were digitized at a 250 Hz per channel frequency, a 12-bit resolution over a 10 mV range, and classified into eight categories based on the severity of ventricular ectopy. These categories range from 'no ventricular ectopy' to 'ventricular flutter/fibrillation'. Each recording contains a thirty-minute segment annotated beat-by-beat, without differentiating supraventricular ectopic beats from normal sinus beats. Two versions of the database, short and long, are available, with the latter including 2.5 hours of unannotated ECG signals preceding each annotated segment.

In addition to the initial dataset, there exists a second set of 75 recordings (test set) assembled using identical criteria. Reserved for evaluations without the detectors being tuned for the test data, this test set was made available around 2003. The nomenclature of the records in the test set parallels the development set, with the first digit denoting the class and the second indicating the version (1 for the long version, 3 for the short) [29]. The AHA datasets contains a wide spectrum of beats, but we focus on the detection of normal and PVC beats. Our AHA test dataset therefore contains only N and V beats (see Table 6).

Table 6. Count and description of ECG beat types in AHA DB.

Symbol	Beat Description	Count
N	Normal	174,260
V	Premature ventricular contraction	16,296

3.7.2. NST

The dataset includes 15 recordings of half-hour duration each, comprising 12 Electrocardiogram (ECG) recordings and 3 recordings characteristic of noise typically found in ambulatory ECG recordings. The noise recordings are comprised of baseline wander, muscle artifact, and electrode motion artifact, each obtained from physically active volunteers using standard ECG apparatus. The ECG recordings are generated from two clean ECGs (118 and 119) from the MIT-BIH Arrhythmia Database. Noise from the electrode motion artifact record was then artificially incorporated into these clean ECGs to simulate real-world noisy conditions. This noise was added after the first 5 minutes of each ECG, alternating between two-minute segments of noise and clean intervals. The records have varied Signal-to-Noise ratios (SNRs) ranging from 24 dB to -6 dB during the noisy segments, to test algorithms under different noise conditions. As the original ECGs are noise-free, the accurate beat annotations are known, and these annotations serve as the reference, even in instances where the noise renders the ECGs visually unreadable [34]. Note that we utilize all noise levels of the dataset, unlike other works, e.g. [35].

3.7.3. MIT

The MIT-BIH Arrhythmia Database comprises 48 two-channel ambulatory ECG recordings, each of half-hour duration, from 47 subjects evaluated by the BIH Arrhythmia Laboratory between 1975 and 1979. These recordings were selected both randomly and purposefully from a larger set of 4000 records of 24-hour ECG recordings collected from a diverse patient population at Beth Israel Hospital. The digitization of these recordings was done at a rate of 360 samples per second per channel with an 11-bit resolution over a 10 mV range. Each record was annotated independently by two or more cardiologists to create computer-readable reference annotations for each beat. As of February 2005, all 48 complete records and reference annotation files are freely available.

The MIT testing dataset consists of 24.07 hours of data. We further utilized a subset of the MIT Dataset, the MIT 11-records subset which is also widely used to evaluate the performance. This dataset consists of around 5.52 hours of ECG data from 11 different patients. For our evaluation we only utilize 44 samples since 4 of them are not adequate for processing. The MIT Arrhythmia test dataset

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

contains a wide spectrum of beats, but we focus on the detection of normal and PVC beats. Our MIT test dataset therefore contains only N and V beats (see Table 7) [36].

Table 7. Count and description of ECG beat types - MIT Arrhythmia dataset.

Symbol	Beat Description	Count
N	Normal	100,718
V	Premature ventricular contraction	7,009

3.7.4. Custo Med Test Dataset

The CST Strips dataset consists of 627 records of 10 to 30 seconds ECG segments. The data is sampled from multiple patients and has a strong focus on noisy, hard to classify PVC beats, couplets, triplets or salves (4756 PVC beats). The number of normal and PVC beats may be found in Table 8.

Table 8. Count and description of ECG beat types - custo med test dataset.

Symbol	Beat Description	Count
N	Normal	39,133
V	Premature ventricular contraction	4,576

4. Results

The quantitative results on our test data are outlined in detail in Table 9. We utilized sensitivity (se), specificity (sp) and balanced accuracy (ba) to assess the performance of our model. We surpass the state of the art on the MIT 11 dataset and show stable performance across all data sets.

Table 9. Results of our model for normal and PVC beat detection.

DB	QRS_{Sp}	QRS_{Se}	PVC_{Sp}	PVC_{Se}	QRS_{BA}	PVC_{BA}
MIT DB	0.997	0.999	0.926	0.966	0.998	0.946
MIT 11 DB	0.999	0.999	0.976	0.991	0.999	0.986
AHA DB	0.992	0.997	0.972	0.857	0.995	0.915
NST DB	0.954	0.893	0.936	0.881	0.924	0.909
CST DB	0.983	0.999	0.950	0.973	0.991	0.962
CSTStrips DB	0.993	0.997	0.960	0.932	0.995	0.946

Where Sp denotes specificity, Se denotes sensitivity, and BA denotes balanced accuracy.

Sensitivity, also known as the true positive rate, is calculated as the proportion of actual QRS complexes or PVCs that the algorithm correctly identifies as such. A high sensitivity means the algorithm is good at catching these events, reducing the number of false negatives (instances where a QRS complex or PVC is present, but the algorithm fails to detect it). We calculated the TN based on if a certain beat is outside the range of 150 ms matching a corresponding RR peak.

Our dataset is strongly imbalanced towards the normal beat class (QRS). In the context of classification tasks, imbalanced datasets are those where the classes are not represented equally. For instance, in a binary classification task, you might have a dataset with 95% examples of class A and only 5% examples of class B. In such cases, a model might achieve high accuracy by simply predicting the majority class. For example, a model that always predicts class A would be 95% accurate, which could give a misleading picture of its effectiveness. In these scenarios, Balanced Accuracy comes to the rescue. It is the average of sensitivity (the true positive rate, or the proportion of actual positives correctly identified) and specificity (the true negative rate, or the proportion of actual negatives correctly identified), so it takes into account both false positives and false negatives. By doing this, it provides a more fair view of the model's performance across all classes, rather than favoring the majority class. The metrics are calculated as follows:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

$$Sp = \frac{TN}{TN + FP}; \quad Se = \frac{TP}{TP + FN}; \quad BA = \frac{Se + Sp}{2} \tag{1}$$

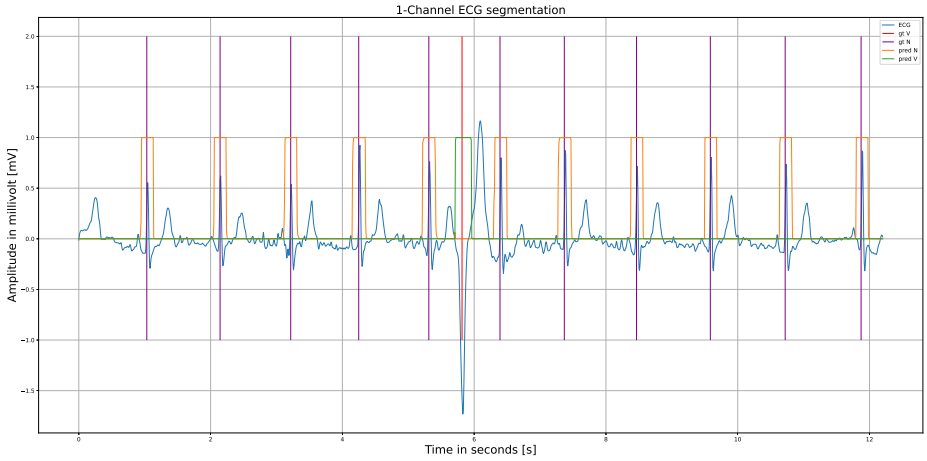
4.1. Evaluation Method

The objective of our evaluation is generate a list of PVC beat locations for each recording that align with the ground truth PVC annotations. For every reference PVC annotation, there should be a corresponding predicted PVC annotation within a 150ms interval centered around it. Note that reference PVC annotations present in the first or last 0.2 seconds of the recording are disregarded. Any detected PVC should fall within 150ms of its reference annotation.

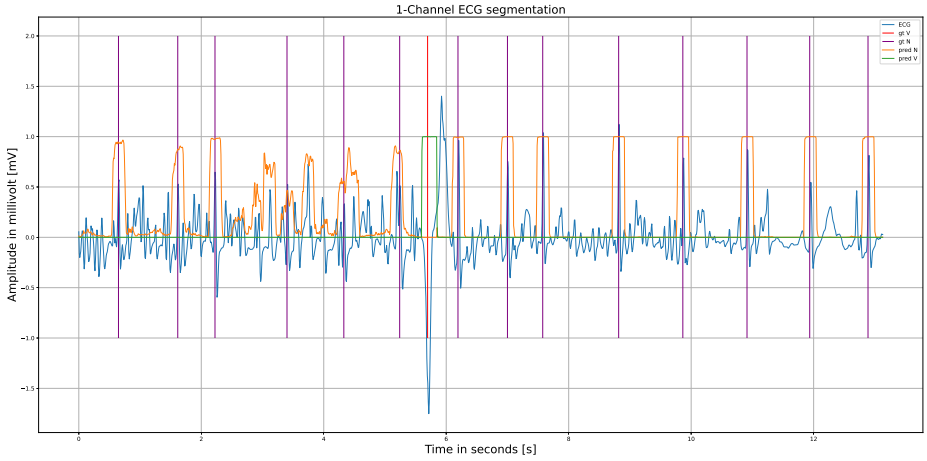
4.2. Model Output

The output of the system is a series of classifications for each timestamp in the input ECG data - as either a background, normal beat or PVC (see Figure 4. This end-to-end process allows for the automatic detection of PVCs from ECG data, reducing the need for expert manual review and potentially speeding up the diagnostic process. The model outputs for a given variable-length ECG a corresponding equal length segmentation mask. The minimum length of the ECG signal length is around 400 input timestamps which in our case equals to 400 / 125 = 3.2 seconds of input data due to the kernel size of 9, the convolutional stride of 2, same padding and a dilation of 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1



(a) Example ECG from our custom dataset segmented into Background, Normal beat and PVC beat. A single PVC is located at around second 6.



(b) Example noisy ECG from our custom dataset segmented into Background, Normal beat and PVC beat. A single PVC is located at around second 5. The Noral beat class mask (orange) is noisy, due to uncertainty.

Figure 4. Clean vs. Noisy ECG segmentation. The output masks show different behavior for clean and noisy ECG samples and require different post processing treatment.

4.3. Productive usage

We incorporated our model into the custo med software (see Figure 5) using the ONNX Runtime library for C++. The model size is around 13MB, consists of 836,387 trainable parameters and requires 63 GFLOPs (estimated using fvcore) for 1 hour of 1 Channel ECG data. A somewhat new CPU may run 60 or more GFlOPS which results in around 48 seconds for a 24h ECG. We tested the model on a i7-10750H where the model requires 60 seconds for a 3 Channel 24h ECG. For comparison, a standard 2D UNet for aerial image segmentation may require 16 GFlOPs for a single 256x256 pixel gray-scale image [37].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

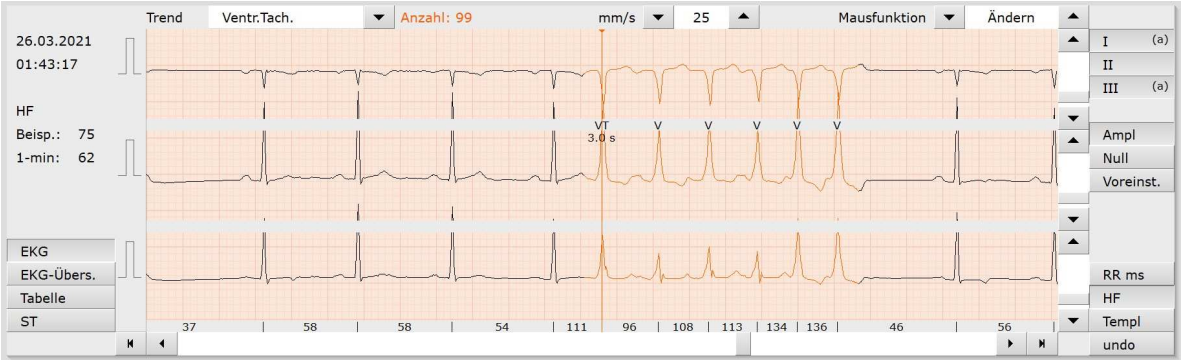


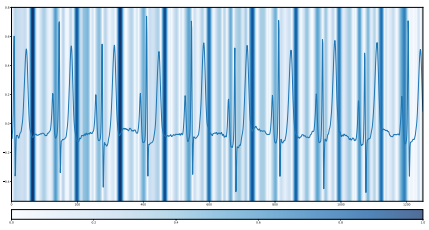
Figure 5. Screenshot from the custo diagnostics software. The model is enable to predict the normal beats and ventricular tachycardia (beats labeled as V).

5. Model Interpretability

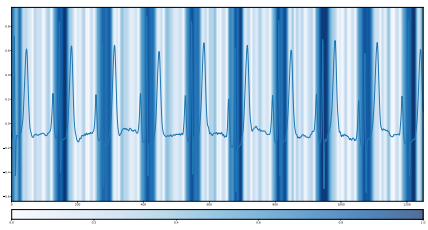
We utilize Layer-wise Gradient-weighted Class Activation Mapping (LayerGradCAM) [38] as a visualization technique for highlighting the important features in an input that contribute to the network's final decision. In the context of ECG data, LayerGradCAM can provide significant insights into which ECG segments are most influential for the model's segmentation decisions.

In the ECG domain, it's often crucial to understand not just the final output of a model, but also the 'why' behind that output. For instance, in the case of electrocardiogram (ECG) data, understanding which segments of the time series (i.e., specific heartbeats or signal patterns) are indicative of a certain disease can help clinicians make better-informed decisions.

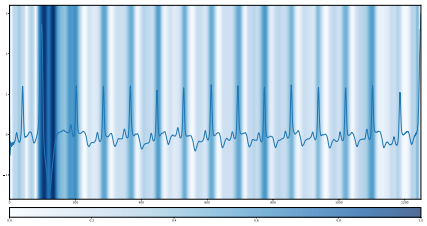
LayerGradCAM provides such an understanding by producing a heatmap of the input time series, with the intensity of the color (in our case blue) indicating the contribution of each time point towards the final decision (see Figure 6. This approach helps to uncover the model's internal decision-making process, thereby increasing the interpretability and transparency of the model.



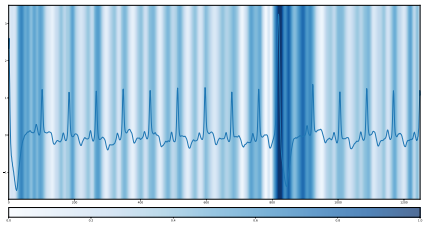
(a) Activations for the background class.



(b) Activations for the normal beat class.



(c) Activation for the PVC beat class.



(d) Activation for the PVC beat class.

Figure 6. Examples from the AHA test dataset using the LayerGradCAM attributions from the last encoder convolutional layer. Greater attribution to a certain class is depicted via darker shade of the color blue.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

6. Discussion

In the analysis of the different databases, several key observations emerge. The QRS detection specificity and sensitivity results are consistently high across all databases, with the specificity exceeding 0.95 and sensitivity exceeding 0.89 in all cases. This indicates a strong performance in the detection of QRS complexes.

Comparatively, the detection of PVCs exhibits some variability. While the specificity remains high, the sensitivity fluctuates between 0.857 (AHA DB) and 0.991 (MIT 11 DB). Such a disparity may reflect the inherent complexity and variability of PVC occurrences in ECG recordings, requiring more nuanced detection algorithms. Note, that the AHA Database has several false annotations and contains also a pacemaker and AFIB sections.

The balanced accuracy (BA) for QRS detection is exemplary across all databases, being closest to the ideal score of 1. This reflects an overall well-balanced performance between sensitivity and specificity, suggesting that the model doesn't lean excessively towards recall or precision. The lowest BA for QRS detection is seen in the NST DB (0.924), and the highest in the MIT 11 DB (0.999).

The PVC detection, on the other hand, showcases a broader range in balanced accuracy, from 0.909 (NST DB) to 0.986 (MIT 11 DB). It is noteworthy that despite the comparatively lower sensitivity in the AHA DB for PVC detection, its balanced accuracy still remains robust (0.915), demonstrating the model's relative equilibrium in handling false positives and negatives.

Overall, these results highlight the robustness of QRS detection across multiple databases and demonstrate a degree of variability in PVC detection, underscoring the need for additional algorithm refinement for this latter category. The lower performance metrics in the NST DB and AHA DB relative to the other databases might that the model does not perform optimally under strong noise conditions, suggesting an avenue for future investigation and algorithm optimization.

7. Limitations

While our study provides novel insights and adds value to the current literature, it is not without its limitations which offer avenues for future work.

1. **Pruning:** Our model did not incorporate pruning techniques during the training process. Pruning is a common strategy to reduce the complexity and size of deep learning models, improving computational efficiency and potentially reducing overfitting. Future studies might explore the impacts of various pruning techniques on model performance and efficiency.
2. **Absence of Attention Mechanism:** The model did not leverage any attention mechanism. Attention models have emerged as powerful tools in deep learning, enabling the model to focus on the most relevant parts of the input for a given task. Incorporating attention mechanisms could improve the model's performance, especially in tasks where certain parts of the input carry more informative content.
3. **Lack of Self-supervised Pretraining:** Our study did not exploit self-supervised pretraining using multiple data sets. This approach could potentially improve the robustness and generalizability of the model by exposing it to a wider range of data during pretraining.
4. **Limited Classification:** The scope of our model was confined to the detection of normal beats and Premature Ventricular Contractions (PVC). Although this focus has its own merits, the model's utility could be enhanced by expanding its classification capabilities to detect other types of cardiac events.
5. **Size of the Test Data sets:** Our test data sets were not particularly large. Larger test data sets would provide a more robust estimation of the model's performance and its ability to generalize to unseen data.
6. **Single Channel Model:** Our model was designed to work with single-channel ECG signals. While this design decision simplifies the model and its input requirements, it might limit the model's ability to detect cardiac events that are better characterized using multichannel ECG signals. Future research could investigate the benefits of a multi-channel approach.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

8. Conclusion

This study presented a comprehensive analysis of ECG signal processing for the detection of normal heartbeats and Premature Ventricular Contractions (PVCs). The employed methodologies combined state-of-the-art machine learning approaches, particularly convolutional neural networks, which showed promising results in extracting intricate patterns from the ECG signals, despite the inherent challenges of noise and non-stationarity. Several databases were employed to test and validate the proposed methods, including the MIT DB, MIT 11 DB, AHA DB, NST DB, CST DB, and CSTStrips DB. Each database presented unique challenges and nuances that offered valuable insights into the robustness of our methodologies. The results were presented with a focus on Sensitivity (Se), Specificity (Sp), and Balanced Accuracy (BA) - metrics that are critical for evaluating the performance of any diagnostic tool. In particular, the balanced accuracy proved to be of great importance as it balanced the trade-off between sensitivity and specificity, providing a more holistic view of the model's performance. We further analyzed ECG recordings from 36 different patients for a long-term monitoring evaluation. This unique dataset allowed us to simulate real-world scenarios and to evaluate the potential of our methodologies in clinical settings. For the successful deployment of the model in a production environment, special attention was paid to developing an end-to-end framework, which streamlines the process from raw ECG input to PVC detection. This framework would serve as a foundation for the seamless integration of the model into existing health monitoring systems. While the results obtained are encouraging, the complexity of ECG signal processing and the inherent noise and variability in ECG data always leave room for future improvements. As such, this research is part of an ongoing process of refining and expanding our methods to develop the most effective tools for PVC detection and, more broadly, for cardiovascular health monitoring. It is our hope that this research will contribute to improving the accuracy and reliability of ECG-based PVC detection, ultimately supporting the development of more effective, personalized treatments for patients with cardiovascular conditions.

Author Contributions: Conceptualization, D.K. and G.B.; methodology, D.K, G.B and P.R.; software, D.K and P.J.; validation, D.K and P.J.; formal analysis, D.K., P.J.; investigation, G.B.; resources, G.B. and P.R.; data curation, D.K, P.R. and P.J..; writing—original draft preparation, D.K. and P.R.; writing—review and editing, D.K. and P.R.; visualization, D.K.; supervision, G.B. and P.R.; project administration, G.B. and P.R.;

Funding: This research was funded by custo med GmbH.

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

References

1. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; Navab, N.; Hornegger, J.; Wells, W.M.; Frangi, A.F., Eds.; Springer International Publishing: Cham, 2015; pp. 234–241.
2. Kiranyaz, S.; Ince, T.; Gabbouj, M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. *IEEE Transactions on Biomedical Engineering* **2015**, *63*, 664–675.
3. Acharya, U.R.; Fujita, H.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Adam, M. Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. *Information Sciences* **2017**, *415*, 190–198.
4. Hannun, A.Y.; Rajpurkar, P.; Haghpanahi, M.; Tison, G.H.; Bourn, C.; Turakhia, M.P.; Ng, A.Y. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine* **2019**, *25*, 65–69.
5. Izci, E.; Ozdemir, M.A.; Degirmenci, M.; Akan, A. Cardiac arrhythmia detection from 2d ecg images by using deep learning technique. In Proceedings of the 2019 medical technologies congress (TIPTEKNO). IEEE, 2019, pp. 1–4.
6. Ullah, A.; Anwar, S.M.; Bilal, M.; Mehmood, R.M. Classification of arrhythmia by using deep learning with 2-D ECG spectral image representation. *Remote Sensing* **2020**, *12*, 1685.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

7. Mazidi, M.H.; Eshghi, M.; Raoufy, M.R. Detection of premature ventricular contraction (PVC) using linear and nonlinear techniques: an experimental study. *Cluster Computing* **2020**, *23*, 759–774.
8. Rijnbeek, P.R.; van Herpen, G.; Bots, M.L.; Man, S.C.; Verweij, N.; Hofman, A.; Hillege, H.; Numans, M.E.; Swenne, C.A.; Witteman, J.C.; et al. Normal values of the electrocardiogram for ages 16-90 years. *Journal of electrocardiology* **2014**, *47*, 914–921.
9. Linde, C.; Bongiorni, M.G.; Birgersdotter-Green, U.; Curtis, A.B.; Deisenhofer, I.; Furokawa, T.; Gillis, A.M.; Haugaa, K.H.; Lip, G.Y.; Van Gelder, I.; et al. Sex differences in cardiac arrhythmia: a consensus document of the European Heart Rhythm Association, endorsed by the Heart Rhythm Society and Asia Pacific Heart Rhythm Society. *Ep Europace* **2018**, *20*, 1565–1565ao.
10. Zhou, F.y.; Jin, L.p.; Dong, J. Premature ventricular contraction detection combining deep neural networks and rules inference. *Artificial intelligence in medicine* **2017**, *79*, 42–51.
11. Teplitzky, B.A.; McRoberts, M.; Ghanbari, H. Deep learning for comprehensive ECG annotation. *Heart Rhythm* **2020**, *17*, 881–888. Digital Health Special Issue, https://doi.org/https://doi.org/10.1016/j.hrthm.2020.02.015.
12. Pan, J.; Tompkins, W.J. A real-time QRS detection algorithm. *IEEE transactions on biomedical engineering* **1985**, pp. 230–236.
13. Christov, I.I. Real time electrocardiogram QRS detection using combined adaptive threshold. *Biomedical engineering online* **2004**, *3*, 1–9.
14. Chiarugi, F.; Sakkalis, V.; Emmanouilidou, D.; Krontiris, T.; Varanini, M.; Tollis, I. Adaptive threshold QRS detector with best channel selection based on a noise rating system. In Proceedings of the 2007 Computers in Cardiology. IEEE, 2007, pp. 157–160.
15. Chouakri, S.; Bereksi-Reguig, F.; Taleb-Ahmed, A. QRS complex detection based on multi wavelet packet decomposition. *Applied Mathematics and Computation* **2011**, *217*, 9508–9525.
16. Elgendi, M. Fast QRS detection with an optimized knowledge-based method: Evaluation on 11 standard ECG databases. *PloS one* **2013**, *8*, e73557.
17. Liu, Y.; Li, Q.; He, R.; Wang, K.; Liu, J.; Yuan, Y.; Xia, Y.; Zhang, H. Generalizable beat-by-beat arrhythmia detection by using weakly supervised deep learning. *Frontiers in Physiology* **2022**, *13*, 850951.
18. He, R.; Liu, Y.; Wang, K.; Zhao, N.; Yuan, Y.; Li, Q.; Zhang, H. Automatic Detection of QRS Complexes Using Dual Channels Based on U-Net and Bidirectional Long Short-Term Memory. *IEEE Journal of Biomedical and Health Informatics* **2021**, *25*, 1052–1061. https://doi.org/10.1109/JBHI.2020.3018563.
19. Martínez, J.P.; Almeida, R.; Olmos, S.; Rocha, A.P.; Laguna, P. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Transactions on biomedical engineering* **2004**, *51*, 570–581.
20. Arzeno, N.M.; Deng, Z.D.; Poon, C.S. Analysis of First-Derivative Based QRS Detection Algorithms. *IEEE Transactions on Biomedical Engineering* **2008**, *55*, 478–484. https://doi.org/10.1109/TBME.2007.912658.
21. Zidelmal, Z.; Amirou, A.; Adnane, M.; Belouchrani, A. QRS detection based on wavelet coefficients. *Computer methods and programs in biomedicine* **2012**, *107*, 490–496.
22. De Chazal, P.; O'Dwyer, M.; Reilly, R.B. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE transactions on biomedical engineering* **2004**, *51*, 1196–1206.
23. Jiang, W.; Kong, S.G. Block-based neural networks for personalized ECG signal classification. *IEEE Transactions on Neural Networks* **2007**, *18*, 1750–1761.
24. Ince, T.; Kiranyaz, S.; Gabbouj, M. A generic and robust system for automated patient-specific classification of ECG signals. *IEEE Transactions on Biomedical Engineering* **2009**, *56*, 1415–1426.
25. Zhang, C.; Wang, G.; Zhao, J.; Gao, P.; Lin, J.; Yang, H. Patient-specific ECG classification based on recurrent neural networks and clustering technique. In Proceedings of the 2017 13th IASTED International Conference on Biomedical Engineering (BioMed). IEEE, 2017, pp. 63–67.
26. Sudre, C.H.; Li, W.; Vercauteren, T.; Ourselin, S.; Cardoso, M.J. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer International Publishing, 2017; pp. 240–248. https://doi.org/10.1007/978-3-319-67558-9_28.
27. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2999–3007. https://doi.org/10.1109/ICCV.2017.324.
28. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization, 2017, [arXiv:cs.LG/1412.6980].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 September 2023 doi:10.20944/preprints202309.1588.v1

29. Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet. *Circulation* **2000**, *101*, e215–e220, http://xxx.lanl.gov/abs/https://www.ahajournals.org/doi/pdf/10.1161/01.CIR.101.23.e215. https://doi.org/10.1161/01.CIR.101.23.e215.
30. Tan, S.; Androz, G.; Chamseddine, A.; Fecteau, P.; Courville, A.; Bengio, Y.; Cohen, J.P. Icentia11k: An unsupervised representation learning dataset for arrhythmia subtype discovery. *arXiv preprint arXiv:1910.09570* **2019**.
31. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International conference on machine learning. pmlr, 2015, pp. 448–456.
32. Zhang, H.; Cisse, M.; Dauphin, Y.N.; Lopez-Paz, D. mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412* **2017**.
33. Yun, S.; Han, D.; Oh, S.J.; Chun, S.; Choe, J.; Yoo, Y. Cutmix: Regularization strategy to train strong classifiers with localizable features. In Proceedings of the Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp. 6023–6032.
34. Moody, G.B.; Muldrow, W.; Mark, R.G. A noise stress test for arrhythmia detectors. *Computers in cardiology* **1984**, *11*, 381–384.
35. Vijayarangan, S.; R., V.; Murugesan, B.; S.P., P.; Joseph, J.; Sivaprakasam, M. RPnet: A Deep Learning approach for robust R Peak detection in noisy ECG. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020, pp. 345–348. https://doi.org/10.1109/EMBC44109.2020.9176084.
36. Moody, G.; Mark, R. The impact of the MIT-BIH Arrhythmia Database. *IEEE Engineering in Medicine and Biology Magazine* **2001**, *20*, 45–50. https://doi.org/10.1109/51.932724.
37. Wang, S.; Hou, X.; Zhao, X. Automatic Building Extraction From High-Resolution Aerial Imagery via Fully Convolutional Encoder-Decoder Network With Non-Local Block. *IEEE Access* **2020**, *8*, 7313–7322. https://doi.org/10.1109/ACCESS.2020.2964043.
38. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision* **2019**, *128*, 336–359. https://doi.org/10.1007/s11263-019-01228-7.

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.