Double Assessment of Outflow Tract Ventricular Arrhythmias Using Deep Learning Techniques

Eric Berenguer, Anna Maria Llopart, Laura Montserrat

***Abstract*— *Objective:* The aim of this report is to comprehensively cover the clinical context, methodologies and processes that contributed to the creation of the two models: one used for the classification of the ventricle of origin (LVOT-RVOT) and another for two different specific sites of origin (SOO): Left Coronary Cusp (LCC) and Left Ventricular Summit (LVS) of outflow tract ventricular arrhythmias (OTVA). *Context:*. *Methods:* A large dataset from the Teknon Hospital containing multiple labeled ECGs from 190 different patients along with some of their clinical data was used to train and test the models from which the QRS complexes where extracted. The first model architecture used a three staged model with an initial Recurrent Neural Network (RNN) used to extract an embedding of the QRS which is then combined with clinical data from the patients and fed to a Binary Multi-Layer Perceptron (BMLP). For the SOO classification the model consisted of transforming into wavelets the QRS complexes of the selected leads to obtain a R^{2X2} graphical representation of the signals that is then used to train a Convolutional Neural Network (CNN) and correctly classify the different ECGs into LCC, LVS or others. *Results*: the results obtained are not perfect, among different iterations the RVOT-LVOT classification model obtained a peak accuracy over training data of 70.83%. For the SOO Classification the results were more satisfactory, achieving an accuracy of 87% for LCC vs LVS classification.**

INTRODUCTION

Disorders of the heart rate and rhythm abnormalities can occur in the absence of a structural heart disease. Known as idiopathic ventricular arrhythmias (IVAs), they include ventricular tachycardias (VTs) and premature ventricular contractions (PVCs), their mechanism is believed to be triggered activity (TA) [1]. The latter results from delayed afterdepolarizations (DADs) which are mediated by cyclic adenosine monophosphate (c-AMP) and intracellular calcium overload. It is when DAD reaches the cardiomyocyte threshold that another action potential is generated [2][3].

Said TA is commonly manifested through a focal source, in fact, most of IVAs originate from the ventricular outflow tract (OT) region. The right ventricle outflow tract (RVOT) is particularly more frequent (70-80%)[4], especially the septum and in less proportion the wall, than the LVOT, which primarily comprises the LVOT summit (LVS), the aortic coronary cusps (ACC) and the LVOT subvalvular region [5] (See Fig1).

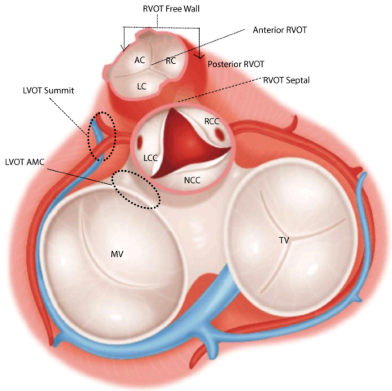

Figure 1: Sub-anatomic sites in LVOT and RVOT. LCC = left coronary cusp; RCC = right coronary cusp; NCC = non-coronary cusp; AMC = aortomitral continuity; AC = anterior cusp; LC = left cusp; RC = right cusp; MV = mitral valve; TV = tricuspid valve. Extracted from [6].

Although IVAs are not life-threatening and typically have a benign prognosis, this by no means implies that they should be overlooked. Quite the opposite, most IVAa are symptomatic and can potentially trigger tachycardia-induced cardiomyopathy, LV systolic dysfunction, or idiopathic ventricular fibrillation [1]. When one of these cases is met, the reliable therapy to treat them is through cardiac catheter ablation [5].

In order to minimize the risk of operative complications and possible re- interventions, reduce the procedure's duration and optimize its outcome, obtaining an accurate prediction of the SOO of OTVA before the surgery is crucial. This can be accomplished by analyzing the features of the electrocardiographic recordings [6], as their ECGs exhibit different characteristics according to their anatomical background.

To efficiently determine whether the ectopic foci originate from RVOT or LVOT, which is what the first section of this project seeks, to binarily classify OTVA's SOO into right or left ventricular OT with deep learning techniques, the bundle branch block pattern in lead V1 is an initial indicator to take into account. Concretely, right bundle branch block (RBBB) patterns are associated with origins in the LV, while LBBB patterns in the RV [5]. Another criterion is the R-wave transition, which refers to the first precordial lead exhibiting a dominant R wave. Because LVOT is posterior to the RVOT, LVOT VAa typically exhibit the precordial transition at V1 and V2. In contrast, in anterior RVOT SOO it can be found at V4 or later [7]. When it comes to lead V3, it is more difficult

to differentiate between both OTs [5].

In order to attempt to make the predictions take also into account clinical information, patient-specific features such as sex, age and hypertension, which other studies showed their relevance [8][9], were also introduced in the model along with the QRS complexes of the PVC rather than the whole ECG. Additionally, most researchers agree on precordial lead V2, followed by V3, in being the best leads for classification [10].

Needless to say, SOOs that are relatively close to each other, such as the LCC and LVS, tend to exhibit similar ECG patterns, which challenges their discrimination. Expanding on this insight, the second section of this paper is dedicated to classify the OTVAs from the LCC and LVS using wavelet transformation to take advantage of both frequential and temporal information of the ECG signals. Concretely, the segmented QRSs of the PVCs of the ECGs.

To do so, this study implements a 2D-CNN model using scalograms of QRS segments for classifying the proposed SOO, following a methodology derived from a study that classifies arrhythmias using continuous wavelet transformation (CWT) and deep learning[11]. CWT unfolds frequency spectrum of the temporal signal for specified time windows, providing an overcomplete representation of its characteristics. This is achieved through the convolution of the signal with a set of functions generated by the Morlet wavelet, commonly used in frequency-time analysis [12][13]. Subsequently, CWT coefficients are mapped into scalograms, which are 2d images containing the frequency components at each time interval through color-based visualization. Leveraging the powerful tools of deep learning in image classification, a 2D CNN is employed to identify the generated scalograms and classify them according to their corresponding SOO.

To deepen the understanding of OTVA planning, this project presents a R-LVOT classification using both ECG data and patient-specific information as inputs for an innovative model which combines a Recurrent Neural Network (RNN) and a Binary Multi-Layer Perceptron (BMLP). Additionally, to tackle the issue of having similar ECG patterns of proximate SOOs, in particular LCC and LVS, the second objective of this paper is to classify LCC and LVS OTVAs employing wavelet transformations as inputs for a Convolutional Neural Network (CNN).

Methodologies

In this section the technical aspect of the project will be explained: from the data used in the classification models to the specifics of the different models themselves, as well as the evolution of them throughout the course of the project.

I QRS segmentation

Regarding the significance of PVC for arrhythmia detection, an automatic segmentation of the corresponding QRS is performed using two different methods. Considering that ECG signals within the datasets are arranged to present the PVC around second 2 of the recording, both algorithms were adapted to acquire the QRS of the targeted beat.

The first approach used the Pan-Tompkins algorithm, importing the module `from ecgdetectors import Detectors` to detect the R peaks from the ECG. Then the R wave more proximal to 2 s (or sample 2000) was selected along with 150 samples before and after it. Thus obtaining segmented QRS complexes of 300 samples.

To improve the precision of the segmentations, the ECGs were processed with a provided deep learning model which identifies different segments of the ECG. Once selected the PVC, padding was applied to obtain a uniform number of samples among all the collected signals, obtaining a final length of 338 samples for each signal.

II Data

The dataset used was the Teknon Dataset, available in the 'all_points_may_2024.pkl' file found in the subject's resource page.

The dataset contained data from 190 different patients undergoing a cardiac ablation procedure. Each entry had a variety of data from the patients: 'Sex', describing patient's sex; 'HTA', describing whether or not the patient had hypertension; 'Age', containing the age of the patient; 'PVC_transition', describing in which lead the PVC transition took place; 'SOO_chamber', containing in which chamber was identified the SOO of the arrhythmia; 'Height', containing the patent's height in Cm; 'Weight', containing the weight of the patient in Kg; 'BMI', with the patient's Body Mass Index; 'DM', containing weather or not the patient suffered diabetes; 'DLP', describing if the patient had dyslipidemia; 'Smoker', describing if the patient was a smoker or not; 'COPD', with information on if the patient suffered or not Chronic Obstructive Pulmonary Disease; 'Sleep_apnea', informing if the patient had sleep apneas; 'CLINICAL_SCORE', a coefficient calculated with varied clinical data from the patient; 'SOO', with the identified specific site of origin; 'OTorigin', describing if the patient's arrhythmia was located in the outflow tracts; and 'Structures', containing multiple ECGs from the patient taken at different moments of the intervention.

From this dataset, the information of the first 20 out of 108 patients with RVOT origin was destined for a testing dataset together with the data from the first 15 patients out of 81 with LVOT origin. The remaining patients' data was used for creating a training dataset.

Testing Dataset

For the testing dataset, for each one of the categories inside Structures, only one QRS was selected for testing. For the 20 patients with RVOT origin selected to use their information for the training dataset, a total of 207 different QRS are sampled, and from the 14 patients with LVOT origin, 159 samples were extracted. The two combined provide a slightly

biased dataset towards QRS from RVOT patients, comprising 56.56% of the dataset.

Training Dataset

For the training dataset, for each one of the categories inside Structures, at most the first 3 QRS were selected for training. For the 88 patients with RVOT origin selected to use their information for the training dataset, a total of 557 different QRS are sampled, and from the 67 patients with LVOT origin, 431 samples were extracted. The two combined provide a slightly biased dataset towards QRS from RVOT patients, comprising 56.37% of the dataset.

III Models

For each of the different two tasks: LVOT-RVOT classification (A), and SOO classification (B), a different model has been elaborated.

A. Right - Left Ventricular Outflow Tract Classification

For the task of classification between Left and Right ventricular outflow tract, a model composed of three stages: a Recurrent Neural Network, a vector enhancing and normalization stage, and a final Binary Multi-Layer Perceptron was proposed. (See Fig. 2)

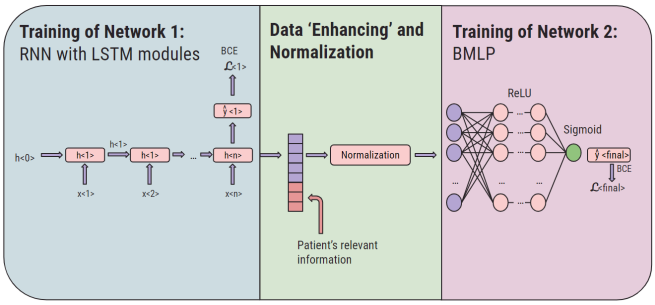

Figure 2: Schematic representation of the 3-staged model

A.1. Model stages

Each stage served a specific purpose for the correct classification of the arrhythmia origin:

Recurrent Neural Network

The first stage of the model consisted of a Recurrent Neural Network (RNN) architecture with Long-Short Term Memory Modules (LSTM). The aim of this stage was to obtain an embedding representation of the sequence introduced as input. For obtaining this representation, the RNN was trained to obtain a single prediction for the whole sequence and then a Binary Cross-Entropy loss function (BCE) was applied to backpropagate and update the weights, since a probabilistic approach to the classification was employed.

The core idea behind this lied in the fact that if the RNN network was capable of reducing the loss of the classification of the sequence, it meant that it was capable of constructing an embedding in its last hidden state $h<n>$ (see Fig. 1 in Appendix) for which it was significant enough to correctly classify the whole sequence. After the training, the classification provided by the RNN was ignored and the last hidden state representing the sequence was then passed to the next stage.

Data Enhancing and Normalization

The embeddings provided by the previous stage were then merged with relevant data of the patient by appending the data to the embedding vector, creating an enhanced vector. An important remark is that if data of a certain patient was not available, for instance it was missing the height or weight, and it was needed for the enhanced vector, it was filled by the average for that value across the dataset.

This stage was crucial to obtain a vector representative of both the QRS complex sequence and the patient clinical background. In various iterations of the model, different data from the patient was used according to different criteria, see *model iterations discussion* for more details. After the enhanced vectors were obtained, they were then normalized in each dimension independently, to equal all of the variances and differences in scale across the dimensions and passed to the next stage.

Binary Multi-Layer Perceptron

After obtaining the enhanced vector, a Binary Multi-Layer Perceptron (BMLP) was used to classify the vectors. Here a BCE loss was also used since the probabilistic approach is maintained. The output of the BMLP was then taken as the definitive output of the model.

A.2. Model iterations discussion

In the construction of the model, some iterations were produced in order to tune it and progressively build the whole model.

Models 0, 1 and 2

The initial models up to model 2 were designed to test the RNN stage and its capacity to extract embeddings from the inputs. No vector enhancing or BMLP were used. The insights obtained with these models will be discussed here instead of in the result section since they are not themselves a result for the classification but one step more in the process. At this stage the database used was a reduced version of the large dataset, containing only training data. This data was only used here for initial test purposes and will not directly influence coming models. The models 0, 1, and 2 all use a hidden size of 5. For the *Model 0* the whole 12 leads of the ECGs without any processing served as input for the RNN.

Upon inspecting the embeddings provided by the RNN in a reduced 2-dimensional Principal Components Analysis (PCA) representation (See Fig 2 in Appendix), no evident separability was found and thus the approach of using all the data unprocessed was quickly discarded.

For the *Model 1* the input was the sequence to the extracted QRS complex of the PVC at the V2 lead. In this model the reduced 2-dimensional PCA representation of the embeddings revealed a more separable plot (See Fig. 3 in Appendix).

For *Model 2* the input was a 2-dimensional vector of the QRS complexes of the V2 lead in the first dimension and of

the V3 lead in the second dimension. Here the points started to cluster in a more evident way, even in the 2-dimensional representation (See Fig.4 in Appendix).

Model 3.1

The third generation of models were already fully assembled using all of the stages described previously: RNN, vector enhancing and normalization, and BMLP. The data used from this model on is the one described in the data section prior to the model description. Since it was noted that when extracting the QRS complexes from the leads V2 and V3 the peaks did not always coincide, and thus it was decided to concatenate the data and produce a one-dimensional vector containing firstly the V2 QRS complex and the V3 QRS complex following it.

In this iteration the data used to enhance the embedding vector was chosen to be the Age, Height, and Weight of the patients, since it is data that quantifies a metric instead of being a classification. Some of the relevant hyperparameters can be observed at Figure 3. Where input data refers to the data passed at the first stage, *Input* in both RNN and BMLP refers to the vector dimension, *Embedding* in the RNN to the size of the vector embedding for the input sequence provided by the RNN, *Features* are the values chosen to be added to the embedding vector to create the unnormalized enhanced vector, notice that input size at BMLP is always the sum of the size of the embedding vector and the number of chosen features, and finally *Hidden s.* is an array containing the number of neurons in each of the hidden layers of the BMLP.

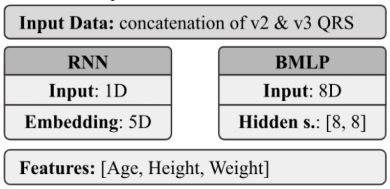

Figure 3: Relevant hyperparameters for the construction of model 3.1

Model 3.2

For this iteration of the model it was decided that even though the peaks were not synchronized, it could be useful to pass both leads as a 2D vector. The size of the embeddings and the BMLP were reduced to maintain a manageable number of parameters to train despite that the RNN now received a 2-dimensional input (See Fig. 4).

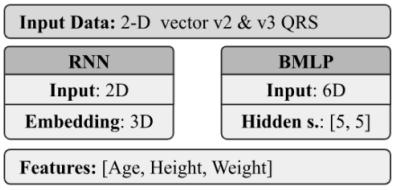

Figure 4: Relevant hyperparameters for the construction of model 3.2

Model 3.3

In the third iteration of the model, only the QRS complex of the V2 lead were passed, to check if the V2 lead alone would suffice. The hyperparameters were also tuned (See Fig. 5).

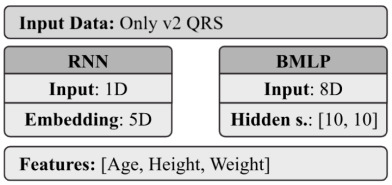

Figure 5: Relevant hyperparameters for the construction of model 3.3

Model 3.4

In this iteration the size of the embedding was enlarged considerably, to 10 dimensions, since the number of features selected had also been changed to a larger set containing also categorical indicators considered of importance for the LVOT-RVOT classification. Among other changes the perceptron now had 3 layers instead of 2 to handle the larger vectors (See Fig. 6). This iteration was significantly more parameter-heavy than prior ones.

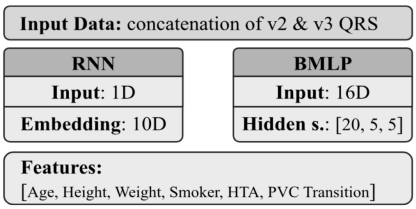

Figure 6: Relevant hyperparameters for the construction of model 3.4

Model 3.5

This last iteration also tried to capture information provided by the categorical patient features but it used a lower number of them than in the model 3.4 (3 instead of 6), reducing the number of trainable parameters of the network using the most relevant features by the reasons discussed at the introduction, for which also the hyperparameters had been slightly tuned (See Fig. 7).

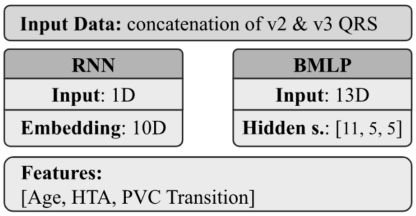

Figure 7: Relevant hyperparameters for the construction of model 3.5

B. Sites of Origin: Classification of LCC and LVS' VA

Concerning the proximity of both SOO, a deeper extraction of features was pursued to ensure accurate classification, as demonstrated in previous papers [11]. Therefore, this second part of the project proceeds with an independent pipeline described in the following sections. Particularly, two classifications were attempted, the first one was designed to discriminate between LCC, LVS and other SOO, whereas the second one consisted of a binary classification between LCC and LVS.

B.1. Continuous Wavelet Transform:

The QRS segments, which are one-dimensional temporal signals, were transformed into a 2D scalogram (see Figure 5 in Appendix) by computing its CWT. As demonstrated in Eq. (1), CWT is the convolution of a temporal signal ($f(t)$) and the

complex conjugate of a particular wavelet (ψ) that varies in function of a and b, representing scale and position parameters respectively. For this study, $f(t)$ represents the QRS segments, whereas ψ * denotes the Morlet2 wavelet, derived from the dot product of a Gaussian function and cosine and sine waves, representing the real and imaginary components, respectively. Therefore, the generated wavelets comprise a wave-like form with finite extremes tending to 0, allowing time windowing for computing the inner product with the temporal function. By scaling the wavelets as a function of parameter a, different frequencies can be captured (compressed and expanded oscillations lead to high-frequency and low-frequency component detection, respectively). Moreover, by varying parameter b, the wavelets get shifted along the signal to complete the convolution.

$$CWT(a,b) = \langle f, \psi_{a,b} \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \cdot \psi^* \left(\frac{t-b}{a}\right) dt \quad (1)$$

Both wavelet definition and CWT were performed in python using 'scipy.signal.morlet2' and 'scipy.signal.cwt' functions available in 'Scipy' library.

B.2. Convolutional Neural Network:

The scalograms (RGB images of size 227x227) were the input of a CNN consisting of four convolutional blocks. Each of these blocks contained a convolutional layer with ReLU activation function, max-pooling layer, followed by batch normalization and dropout to prevent overfitting (see Figure 6 in Appendix). Regarding the final output layer, for the multi-class classification, it consisted of a dense layer with 3 units and a *softmax* activation, whereas for the binary classification the number of classes was 1 and the activation function employed was *sigmoid*.

B.3. Attempts

To explore the extent of this approach, multiple attempts varying batch sizes, epochs and learning rates were performed in three different types of input data: unbalanced QRSs of PVC from leads V2 and V3, balanced data of the previous and balanced data from all the precordial leads V1-6 for both multiclass classification and binary classification of LVS and LCC. (see Tables 1-4). It is worth mentioning that the optimizer used was *Adam* as preliminary attempts showed its effectiveness among others. Additionally, several metrics were implemented to evaluate the performance of the CNN:

- Accuracy:

The accuracy assessment is a straightforward metric that reflects the number of correct predictions compared to the total number of predictions, as expressed in Eq. (2). Although not being representative of the model's performance, particularly in imbalanced datasets, it has been used to compare the different approaches performed initially and refine them to achieve enhanced and more robust models.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

- ROC:

Receiver Operating Characteristic (ROC) curve describes the proportion of correctly classified samples (true positives (TP)) on the Y axis with respect to the misclassified ones (false positive (FP) on the X axis. Therefore, a curve reaching the top left corner of the plot represents the ideal scenario of 100% correct predictions. In other words, the ROC curve is the plot representing the model's sensitivity versus its specificity [14]. Another quantitative assessment of this metric is the area under the curve (AUC), where higher values indicate better model performance. To implement this metric in a multi-class problem, One-vs-Rest multiclass ROC analysis was conducted, where the three classes are treated as binary classifiers to obtain the TP and FP. To mitigate biases due to imbalance data, micro-averaging is employed to assign equal weight to all instances, following the expressions of true positive rate (TPR) and false positive rate (FPR) of Eq. (3) and (4) [15] .

$$TPR = \frac{\Sigma_c TP_c}{\Sigma_c (TP_c + FN_c)} \quad (3)$$

$$FPR = \frac{\Sigma_c FP_c}{\Sigma_c (FP_c + TN_c)} \quad (4)$$

RESULTS

A. Right - Left Ventricular Outflow Tract VAs Classification

After training for 200 epochs, model 3.1 yielded an accuracy of 69.50% for training data and 68.17%. The prediction done by the model for testing data had a precision of 61.21% for LVOT, and 74.21% for RVOT; and a recall of 67.33% for LVOT and 68.78% for RVOT (see Fig. 8 A).

As for model 3.2, after training for 300 epochs, it yielded an accuracy of 61.95% on training data, no improvement over the past one and together with the issue of the lack of beat peak synchronization across extracted QRS complexes made this model not fit enough, and thus it was discarded.

Whereas model 3.3 yields, after training for 500 epochs, it obtained an accuracy of 67.29% for training data, not showing an improvement over model 3.1, and thus no further tests were needed to discard this iteration.

Regarding model 3.4, it achieved an accuracy of 69.18% for training data and 68.45% for testing data after training for 300 epochs. The prediction done by the model for testing data had a precision of 60.11% on LVOT and 77.84% on RVOT; and a recall of 75.33% for LVOT and 63.41% for RVOT.(see Fig. 8 B).

Finally, after training for 300 epochs, the model 3.5 yielded an accuracy of 69.81% for training data and 70.42% for testing data. The prediction done by the model for testing data had a precision of 62.03% on LVOT and 79.76% on RVOT; and a recall of 77.33% for LVOT and 65.36% for RVOT(see Fig. 8 C).

Also, after checking the weights for the input layers of the BMLP of the models 3.4 and 3.5 (See Fig. 7 in the appendix), no feature is more relevant than the others, making the selection of the features and the embedding equally important.

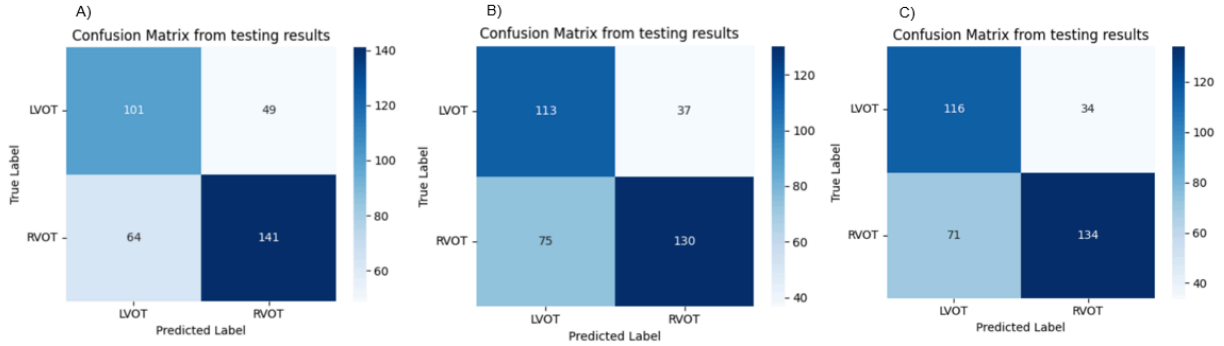

Figure 8: Confusion matrices of testing data for the cases: A) Model 3.1, B) Model 3.4, C) Model 3.5

B. *LCC and LVS' VAa Classification*

	Accuracy (Testing)	Accuracy (Validation)	AUC
Attempt 1	0.8214	0.7846	0.86
Attempt 2	0.8357	0.7230	0.83
Attempt 3	0.7714	0.7231	0.83
Attempt 4	0.7786	0.7077	0.85
Attempt 5	0.8714	0.8154	0.83

Table 1: Quantitative assessment of LCC and LVS classification

For the mentioned attempts of the multi-class segmentation, the resulting accuracies and AUC values for each class are collected in Tables 5-7 in the Appendix, as well as the ROC curves (Fig 8-11 in Appendix). All mentioned attempts, along with complementary and failed attempts are compiled in a supplementary Colab file (Final_wavelets.ipynb), which also includes the visualization of the metrics' curves.

Discussion

A. *Right - Left Ventricular Outflow Tract VAs Classification*

These results show that the last model iteration, the model 3.5, provides the best results, but still with an accuracy of 70%, leaving 30% of the cases identified incorrectly, too much for any sensible application. Therefore, further adaptations or a different approach are necessary to improve the performance. Alternatively, increasing the dataset size might help address this issue, given that the models may be too parameter heavy for the quantity of data and could work better with even more parameters.

B. *LCC and LVS' VAa Classification*

B.1. Multiclass classification LCC/LVS/others:

Regarding the multi-class classification (LCC, LVS and others) the approach which obtained best metrics' values were with unbalanced data from V2 and V3 (Table 5 in Appendix), particularly Attempts 1 (Accuracy test: 0.7421, Accuracy validation: 0.5705, ROC LCC: 0.93, ROC LVS: 0.96 and ROC others: 0.87) and 2 (Accuracy test: 0.6947, Accuracy validation: 0.7115, ROC LCC: 0.93, ROC LVS: 0.99 and ROC others: 0.90).

The second approach using balanced datasets (Table 6 in Appendix) failed to discriminate accurately, as well as using balanced data from the precordial leads (Table 7 in Appendix) but to a lesser extent. It must be mentioned that in preliminary attempts models trained with segmented QRS complexes from the 12 ECG signals were trained but they also did not succeed to classify them properly.

Nevertheless, there is consistency among all the attempts when it comes to the AUC values (Fig 8-10 in Appendix), in fact, they all obtain higher values for LVS prediction, meaning that the CNN is able to discriminate more easily VAs coming from LVS.

The better performance in V2 and V3 data indicates that the most valuable information might be captured in these two leads, since the placement of the electrodes are more proximal to the studied SOO.

B.2. Binary classification LVS/LCC

To purely discriminate the VAs originated from LVS and LCC gave significantly better results (see Table 1). This could be due to the fact that putting all the other sites of origin as one single class was difficulting the algorithm to classify correctly. Taking a closer look at the results, the best attempts were Attempt 1(Accuracy test: 0.8214, Accuracy validation: 0.7846, AUC: 0.86) and Attempt 5 (Accuracy test: 0.8714, Accuracy validation: 0.8154, AUC: 0.83).

Conclusion

In conclusion, this study implements multiple approaches based on deep learning models directly applied to biomedical challenges, specifically in arrhythmia detection. On the one hand, a complete approach for classifying LVOT/RVOT arrhythmias is followed, conducting several iterations for tuning the hyperparameters and model features. On the other hand, a signal processing analysis is opted for amplifying the features within the PVC QRS, using an scalogram-based CNN classification as an alternative approach.

References

[1] Liang, J. J., Shirai, Y., Lin, A., & Dixit, S. (2019). *Idiopathic outflow tract ventricular arrhythmia ablation: pearls and pitfalls. Arrhythmia & Electrophysiology Review, 8(2), 116.*

[2] Gaztañaga, L., Marchlinski, F. E., & Betensky, B. P. (2012). Mechanisms of cardiac arrhythmias. *Revista Española de Cardiología (English Edition)*, *65*(2), 174-185.

[3] Markowitz, S. M., Litvak, B. L., Ramirez de Arellano, E. A., Markisz, J. A., Stein, K. M., & Lerman, B. B. (1997). Adenosine-sensitive ventricular tachycardia: right ventricular abnormalities delineated by magnetic resonance imaging. *Circulation*, *96*(4), 1192-1200.

[4] Srivathsan K., Lester S. J., Appleton C. P., Scott L. R. P., Munger T. M. (2005). Ventricular tachycardia in the absence of structural heart disease. Indian Pacing Electrophysiol. J. 5, 106–121.

[5] Yamada, T. (2016). Idiopathic ventricular arrhythmias: relevance to the anatomy, diagnosis and treatment. *Journal of Cardiology*, *68*(6), 463-471.

[6] Zheng, J., Fu, G., Anderson, K., Chu, H., & Rakovski, C. (2020). A 12-Lead ECG database to identify origins of idiopathic ventricular arrhythmia containing 334 patients. *Scientific data*, *7*(1), 98.

[7] Ezzeddine, F. M., & Siontis, K. C. (2023). Localization of outflow ventricular arrhythmias from the electrocardiogram: educated guess, science, or both?. *Journal of Interventional Cardiac Electrophysiology*, *66*(8), 1775-1777.

[8] Penela D, Falasconi G, Carreño JM, Soto-Iglesias D, Fernández-Armenta J, Acosta J, et al. A hybrid clinical, electrocardiographic score to predict the origin of outflow tract ventricular arrhythmias. J Interv Card Electrophysiol. (2023) 66:1–12. doi: 10.1007/s10840-023-01507-per

[9] Korshunov, V., Penela, D., Linhart, M., Acosta, J., Martinez, M., Soto-Iglesias, D., ... & Berruezo, A. (2019). Prediction of premature ventricular complex origin in left vs. right ventricular outflow tract: a novel anatomical imaging approach. *Ep Europace*, *21*(1), 147-153.

[10] Doste, R., Lozano, M., Jimenez-Perez, G., Mont, L., Berruezo, A., Penela, D., ... & Sebastian, R. (2022). Training machine learning models with synthetic data improves the prediction of ventricular origin in outflow tract ventricular arrhythmias. *Frontiers in Physiology*, *13*, 909372.

[11] Mohonta, S. C., Motin, M. A., & Kumar, D. K. (2022). Electrocardiogram based arrhythmia classification using wavelet transform with deep learning model. *Sensing and Bio-Sensing Research*, *37*, 100502.

[12] *Thesis by Ahmet Esad TOP*. (s/f).

[13] Cohen, M. X. (2019). A better way to define and describe Morlet wavelets for time-frequency analysis. *NeuroImage*, *199*, 81–86. https://doi.org/10.1016/j.neuroimage.2019.05.048

[14] *Evaluating risk prediction with ROC curves*. Columbia University Mailman School of Public Health. (2023, March 13). https://www.publichealth.columbia.edu/research/population-health-methods/evaluating-risk-prediction-roc-curves

[15] *Multiclass receiver operating characteristic (ROC)*. scikit. (n.d.). https://scikit-learn.org/stable/auto_examples/model_selection/plot_roc.html

Appendix

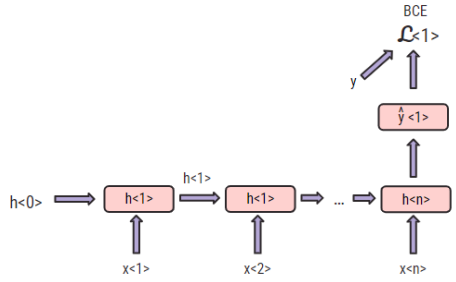

Figure 1: Simple graphical representation of the LSTM used

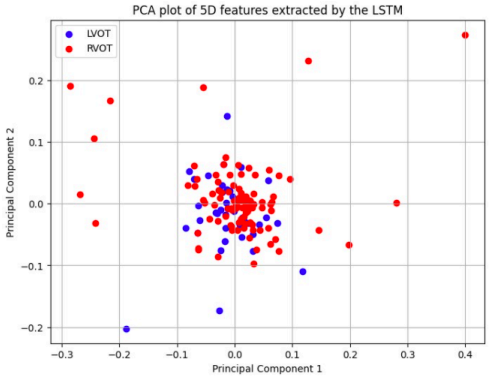

Figure 2: Model 0 embeddings reduced to 2 dimensions

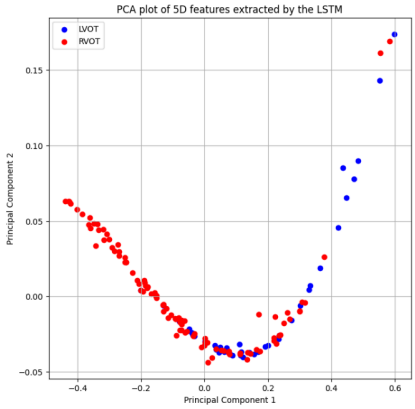

Figure 3: Model 1 embeddings reduced to 2 dimensions

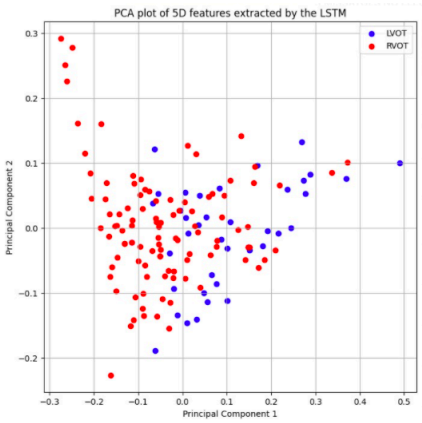

Figure 4: Model 2 embeddings reduced to 2 dimensions

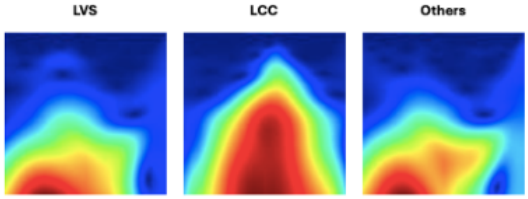

Figure 5: Scalograms from LVS, LCC and Others samples

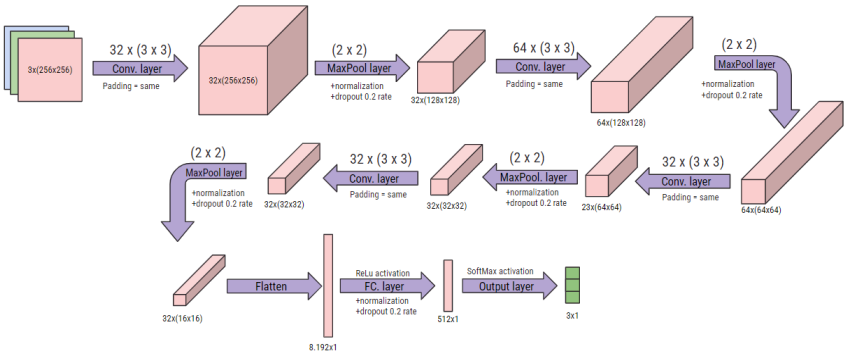

Figure 6: Graphical representation of the CNN employed

A)

B)

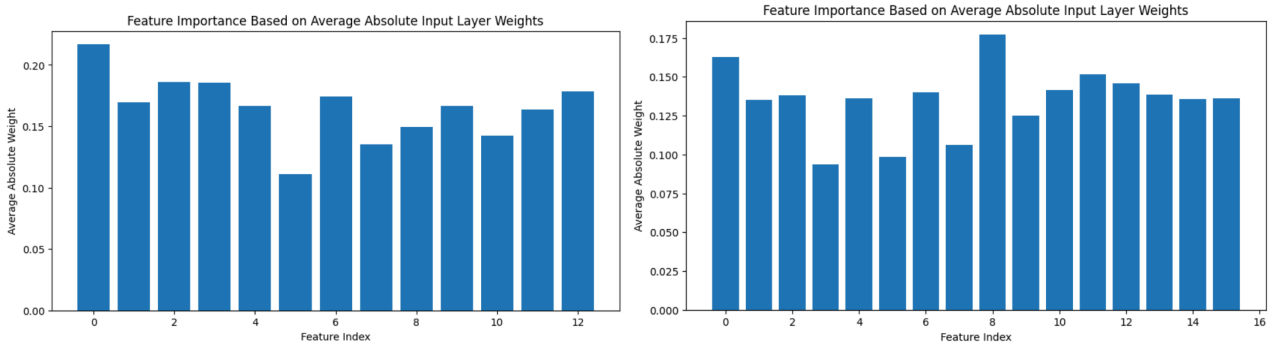

Figure 7: A) Input layer weights for model 3.5 B) Input layer weights for model 3.4

Tables 1-3 Classification SOO (LCC, LVS and others)

	Batch size	Epochs	Learning rate
Attempt 1	20	90	5e-4
Attempt 2	8	90	8e-4
Attempt 3	15	45	5e-4

Tabel 1: Unbalanced V2-V3 (762 images for training set, 190 for testing set and 156 for validation set)

	Batch size	Epochs	Learning rate
Attempt 1	20	70	5e-4
Attempt 2	15	80	8e-4

Table 2: Balanced V2-V3 (282 images for training set, 60 for testing set and 42 for validation set)

	Batch size	Epochs	Learning rate
Attempt 1	20	70	5e-4
Attempt 2	15	100	5e-4
Attempt 3	20	70	1e-3

Table 3: Balanced V1-V6 (846 images for training set, 210 for testing set and 93 for validation set)

Table 4 binary classification LCC vs LVS

	Batch size	Epochs	Learning rate
Attempt 1	20	60	1e-4
Attempt 2	10	40	5e-4
Attempt 3	20	50	1e-3
Attempt 4	20	80	5e-5
Attempt 5	20	80	7e-4

Table 4: Balanced V1-V6 (562 images for training set, 140 for testing set and 65 for validation set)

Tables 5-7: Quantitative assessment

	Accuracy (Testing)	Accuracy (Validation)	AUC LCC	AUC LVS	AUC others
Attempt 1	0.7421	0.5705	0.93	0.96	0.87
Attempt 2	0.6947	0.7115	0.93	0.99	0.90
Attempt 3	0.5053	0.4743	0.83	0.83	0.81

Table 5: Unbalanced V2-V3 (762 images for training set, 190 for testing set and 156 for validation set)

	Accuracy (Testing)	Accuracy (Validation)	AUC LCC	AUC LVS	AUC others
Attempt 1	0.4333	0.4523	0.43	0.87	0.60
Attempt 2	0.5667	0.4761	0.41	0.91	0.70

Table 6: Balanced V2-V3 (282 images for training set, 60 for testing set and 42 for validation set)

	Accuracy (Testing)	Accuracy (Validation)	ROC LCC	ROC LVS	ROC others
Attempt 1	0.6524	0.5376	0.70	0.92	0.73
Attempt 2	0.6143	0.5913	0.67	0.89	0.78
Attempt 3	0.5762	0.6021	0.68	0.83	0.69

Table 7: Balanced V1-V6 (846 images for training set, 210 for testing set and 93 for validation set)

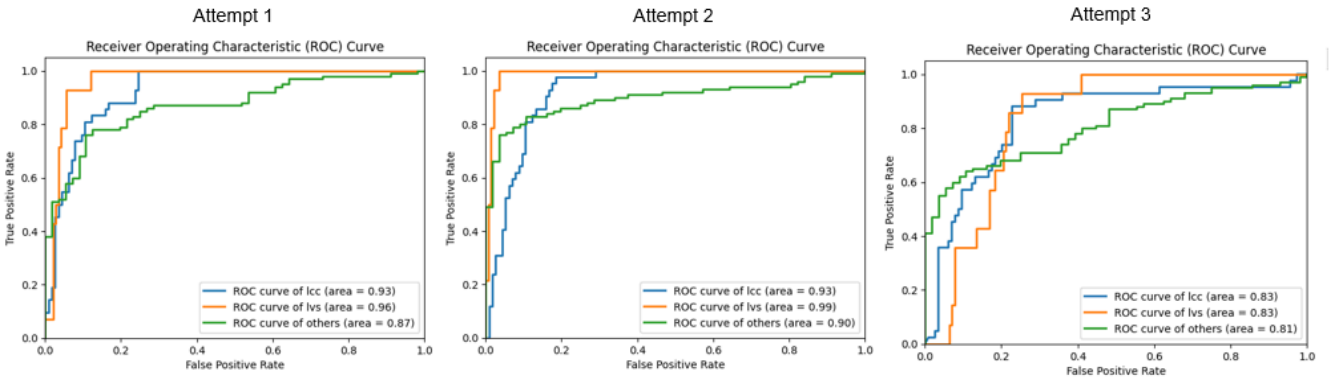

Figure 8: ROC curve of Attempts 1, 2 and 3 of Unbalanced V2-V3

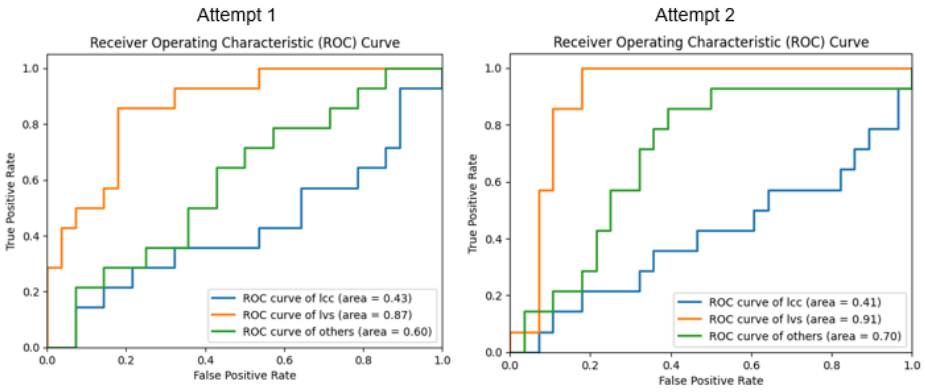

Figure 9: ROC curve of Attempts 1 and 2 of Balanced V2-V3

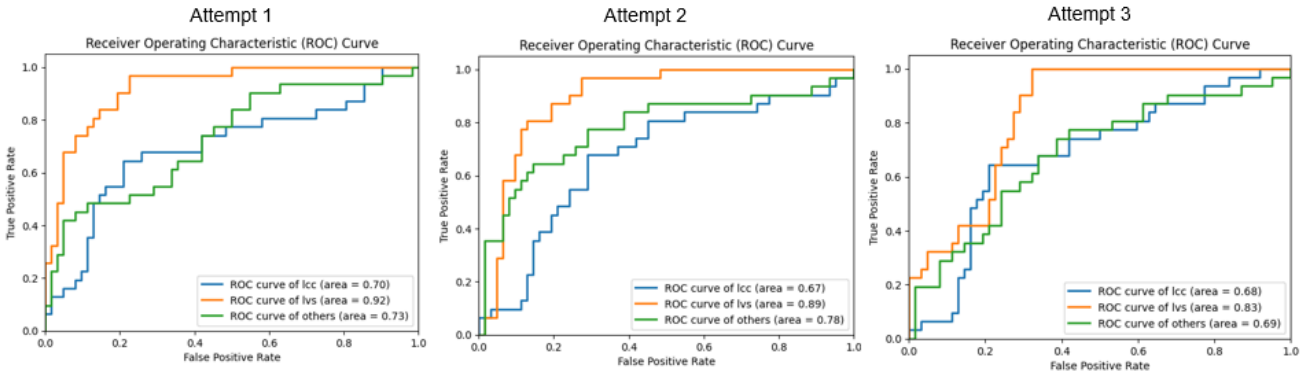

Figure 10: ROC curve of Attempts 1 and 2 of Balanced V1-V6

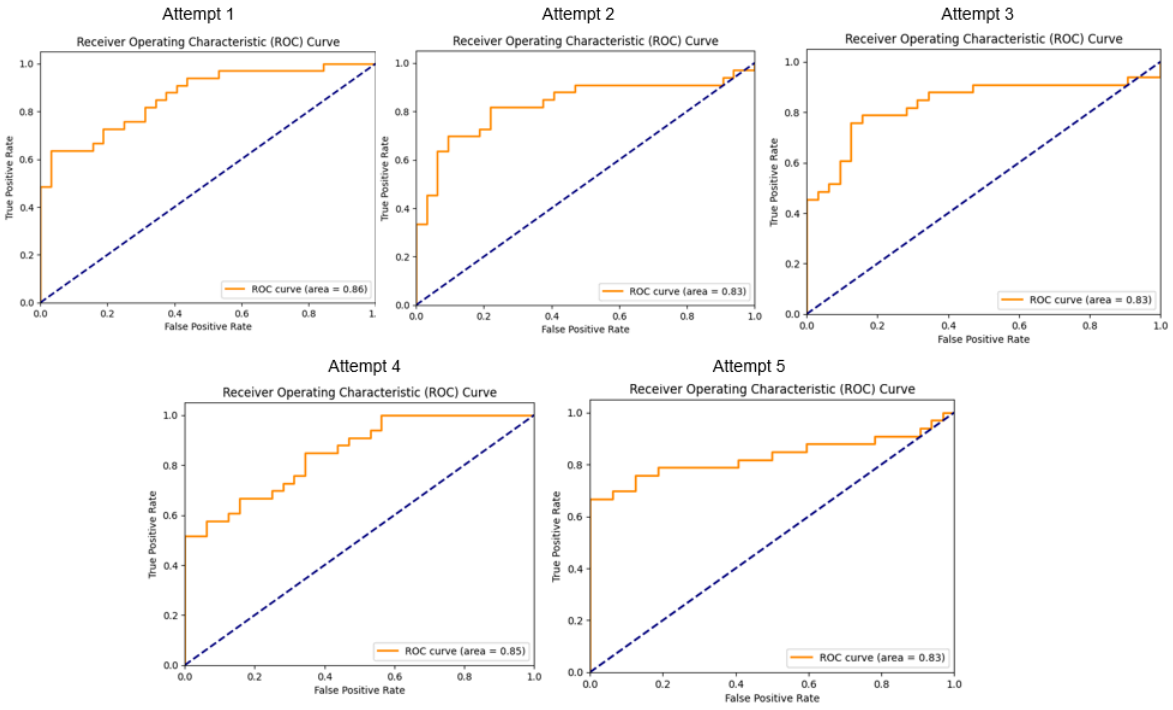

Figure 11: ROC curve of Attempts 1, 2, 3, 4, 5 Balanced V1-V6 (Binary classification)