Department of BioMedical Engineering
Technion University

A feature extraction approach to the classification of 12-lead ECGs

Assaraf David, Intermediate report

19-05-2020

Contents

1. Pathologies

1.1 Atrial fibrillation

This pathology has been extensively studied by Armand. I have directly used Armand's features and got really high results, with a F_β higher than 0.90.

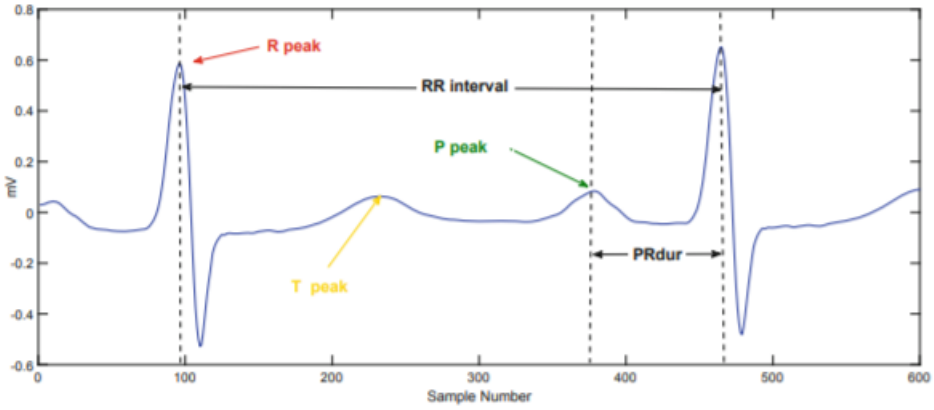
Feature	Definition
CosEn	Coefficient of Sample Entropy
AFEv, IrrEv, PACEv, OriginCount	Metrics derived from the Lorenz plot to assess irregularities in the RR intervals.
PoincSD1, PoincSD2	The standard deviation on the two principal axis of the ellipse on the Poincare plot.
minRR	The minimal RR interval in the segment.
medHR	The median heart rate in the segment.
AVNN	The mean RR interval over the segment.
SDNN	The standard deviation of the RR intervals over the segment.
SEM	Standard error of the mean.
PNN20, PNN50	The percentage of RR intervals shorter than 20 and 50 [ms], respectively.
RMSSD	The root mean square of the successive differences.
CV	Coefficient of variation.
PIP	Percentage of inflection points
IALS	Inverse average length of segments separated by inflection points.
PSS	Percentage of RR intervals that are in short segments.
PAS	Percentage of RR intervals that are in alternation segments of at least 4 intervals.

1.2 First degree Atrioventricular block

1.2.1 Pathology Presentation

First-degree atrioventricular block (AV block) is a disease of the electrical conduction system of the heart in which electrical impulses conduct from the cardiac atria to the ventricles through the atrioventricular node (AV node) more slowly than normal. First degree AV block not generally cause any symptoms, but may progress to more severe forms of heart block such as second- and third-degree atrioventricular block. It is diagnosed using an electrocardiogram, and is defined as a PR interval greater than 200 milliseconds [12], [9].

In [4], the features extracted focus especially on statistics related to these PR interval durations. They define the PR interval as the interval separating the P-peak from the R-peak.



.

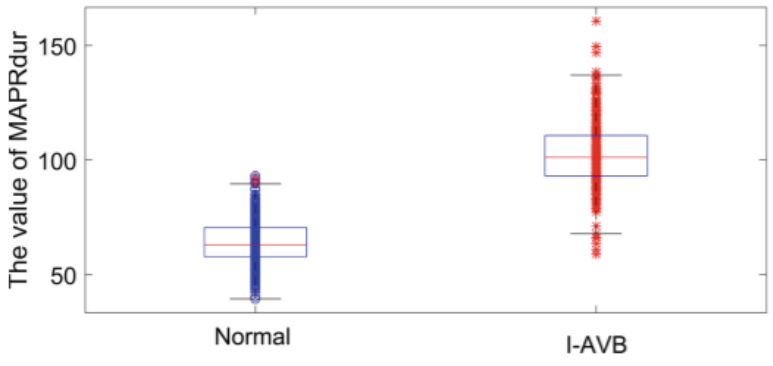
1.2.2 Features extracted

They extracted the 3 following features on lead 2, where $RAPR_{dur} = Mean([\frac{PR^i_{time}}{RR^{i-1}}])$

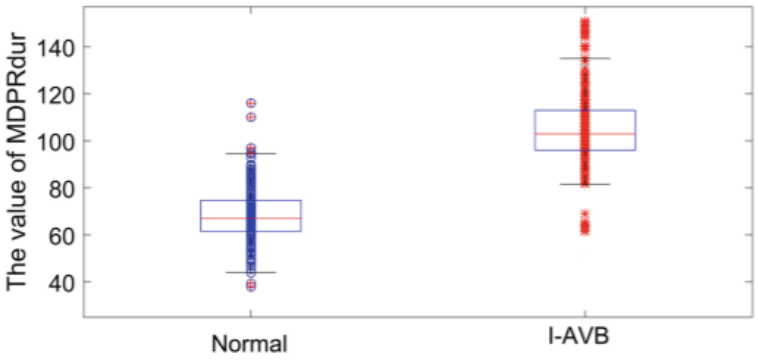
Features	Definition
$MAPR_{dur}$	Average PR interval duration
$MEDPR_{dur}$	Median PR interval duration
$RAPR_{dur}$	Renormalized PR durations

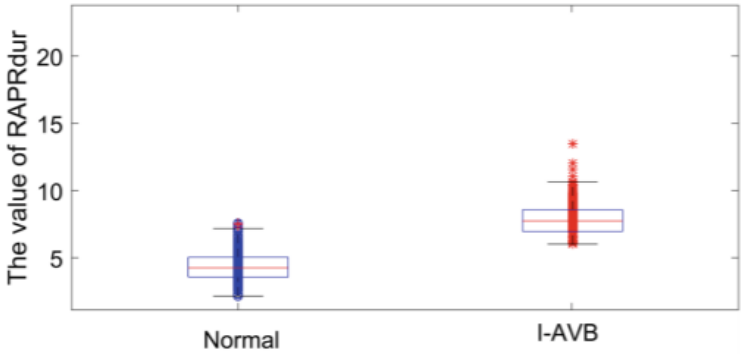
1.2.3 Results

The results they got on the CPSC2018 competition are the following:



.





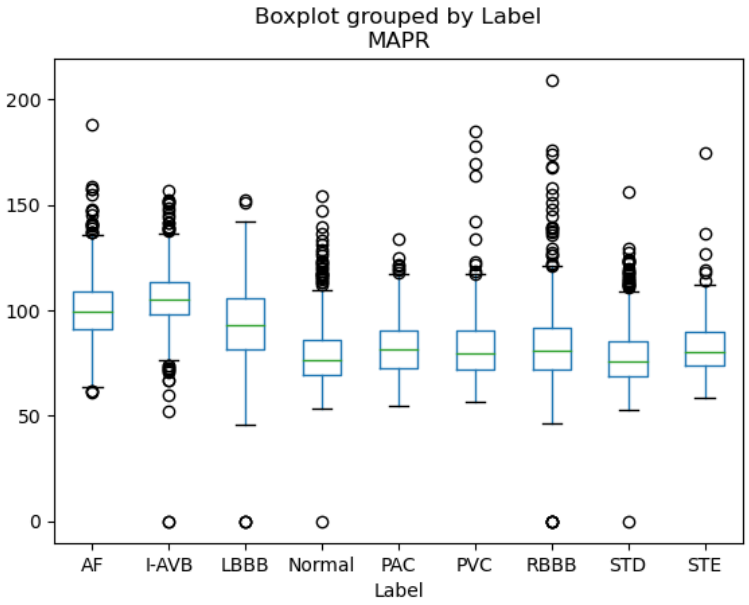
Therefore, those features allowed to differentiate well the **Normal** ECGs from the one with the **AVB** pathology. The classification score they had in separating Normal examples from AVB ones where really high when using SVM classifier:

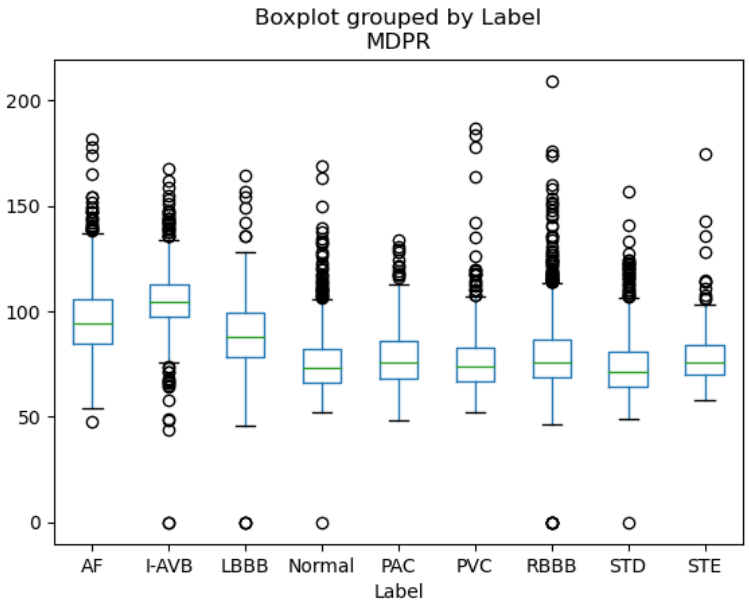
Model	ACC	SEN	SPE
$SVMclassifier$	98.5	98.7	98.3

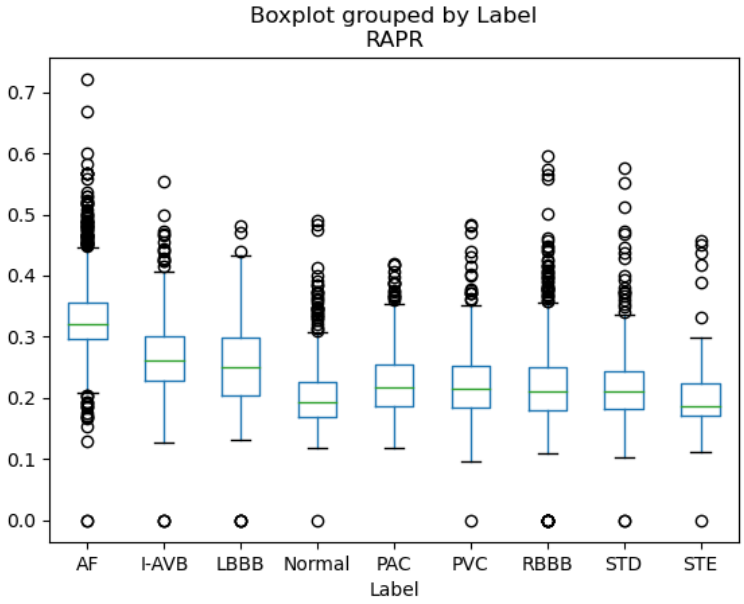
1.2.4 Limitations

Although this study has been led on the CPSC2018 competition dataset, they only performed the classification of Normal ECG vs AVB ones. Therefore, I will further study the distribution of these three features accross every pathology in order to see if these features are discriminative or not.

1.2.4.1 Features for all pathologies







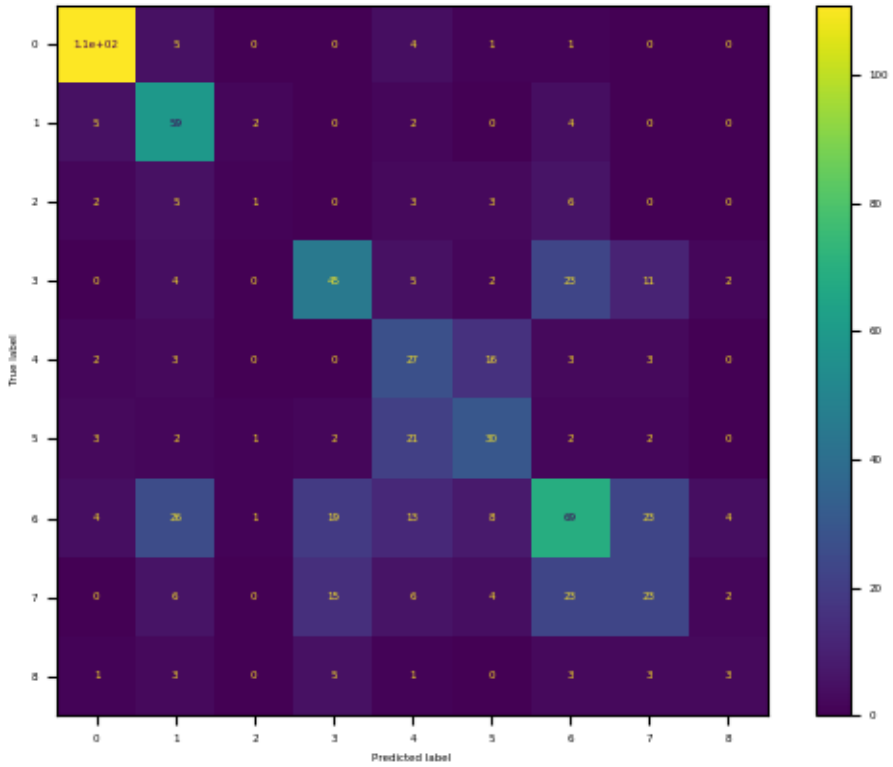
.

We can see that these features seem discriminative for the I-AVB pthology, expcept for the RAPR, where it seems that we have some troubles discriminating AF from I-AVB. I will further extract the features related to AF detection and see if we manage to discriminate AF and I-AVB from the rest of the pathologies while extracting solely these features (HRVB, MAPR, MDPR and RAPR).

The results we get are the followings (we only extract scores for I-AVB and AF):

$$F_\beta^{AF} = 0.92 \qquad F_\beta^{AVB} = 0.72 \tag{1.1}$$

I will further explore the confusion matrix to see why some AVB examples have been misclassified (AF or not?).



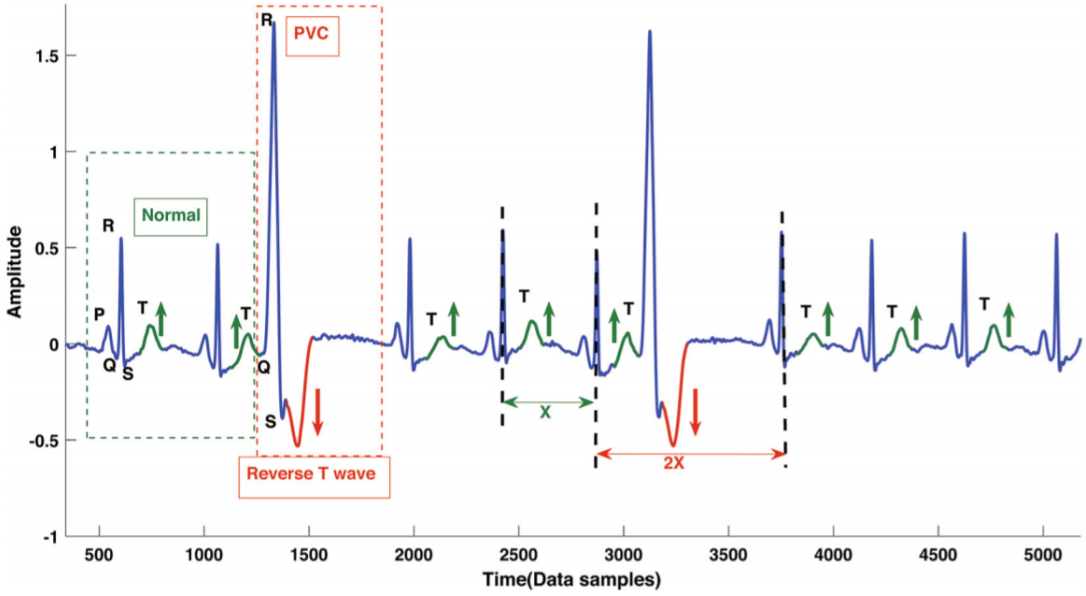
.

It is therefore clear that the problems in the AVB score do not come from conflicts with AF but with other pathologies (too much FN with RBBB). This is why I think that for now we should keep those features for AVB recognition.

1.3 Premature Ventricular Contraction

1.3.1 Pathology presentation

Premature ventricular contraction (PVC) is one of the most common arrhythmia diseases, which is caused by the ventricular activation in advance. Let us see what a PVC ecg looks like.

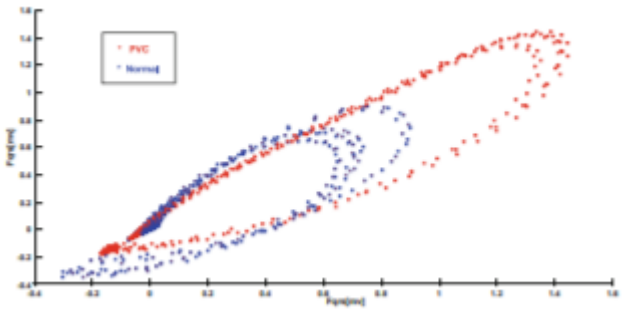


As we can see, PVC is characterized by several things, according to [3]:

- An advanced QRS-T complex (meaning that the RR intervals are not uniform and that we have compensatory pauses between peaks)
- High amplitude R-peaks
- A T-wave with opposite direction from the R-wave.
- A broad QRS complex

1.3.2 Features extracted

We follow the approach of [3]. In order to characterize the high amplitude R-peaks, they extract the **amplitude of** R_{peaks} and the **difference between amplitudes**. In order to quantify the broad QRS complex, they first focused on the **area of the QRS complex**, and then the **duration of the QRS complex**. Then a useful information comes when we plot the Poincare plot of PVC and normal examples.



We therefore see that using useful information from this ellipse will help us to differentiate PVC examples from normal ones: **major axis of the ellipse**. Last, in order to quantify those compensatory pauses, they focused on the **RR interval ratio**, in order to measure the deviation from the mean RR interval ratio.

Features	Definition
F_{max}	Maximum of the R_{peaks} amplitude
F_{mean}	Average of the R_{peaks} amplitude
F_{median}	Median of the R_{peaks} amplitude
F_{std}	Standard deviation of the R_{peaks} amplitude
DF_{max}	Maximum of the R_{peaks} amplitude variations
DF_{mean}	Average of the R_{peaks} amplitude variations
DF_{median}	Median of the R_{peaks} amplitude variations
DF_{std}	Standard deviation of the R_{peaks} amplitude variations
$Dqrs_{max}$	Maximum of the QRS complex duration
$Dqrs_{mean}$	Average of the QRS complex duration
$Dqrs_{median}$	Median of the QRS complex duration
$Dqrs_{std}$	Standard deviation of the QRS complex duration
$Sqrs_{max}$	Maximum of the QRS complex area
$Sqrs_{mean}$	Average of the QRS complex area
$Sqrs_{median}$	Median of the QRS complex area
$Sqrs_{std}$	Standard deviation of the QRS complex area
A_{max}	Maximum of the poincare plot major axis
A_{mean}	Average of the poincare plot major axis
A_{median}	Median of the poincare plot major axis
A_{std}	Standard deviation of the poincare plot major axis
IR_{max}	Maximum of the RR interval ratio
IR_{mean}	Average of the RR interval ratio
IR_{median}	Median of the RR interval ratio
IR_{std}	Standard deviation of the RR interval ratio

Therefore, we have 24 features for every of the 12 leads: 288 morphological features per example solely for PVC detection.

1.3.3 Results

The results they got using the CPSC2018 competition dataset were really high:

Model	ACC	FDR	Omission Ratio
$BPNN$	97.46	3.41	1.37

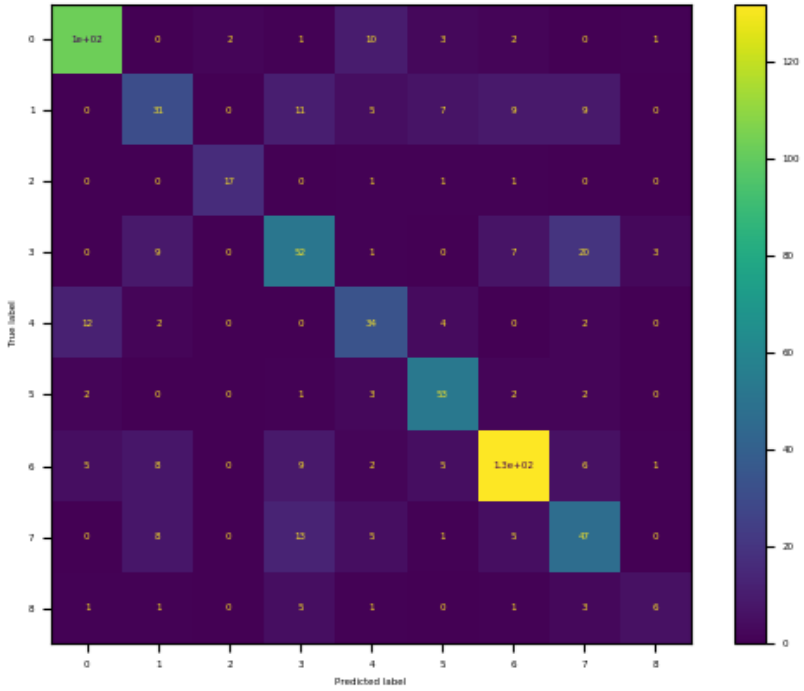
1.3.4 Limitations

1.3.4.1 Performance against other pathologies

Once again, these results have been obtained while performing Normal vs PVC classification. However, our classification task is more complex. The scores we get in classifying PVC with every other pathologies is:

Model	F_β^{PVC}	F_β^{RBBB}
$RFclassifier$	0.78	0.77

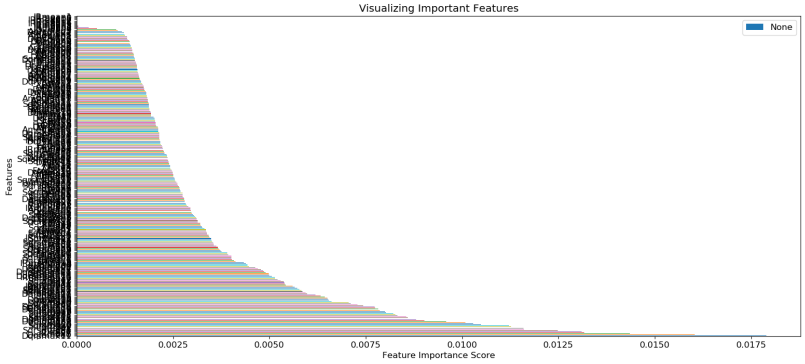
Therefore, the features we have been using for detecting PVC allow us to also somewhat detect RBBB.
The confusion matrix of this model on the test set is:



Therefore, we see that we do not have particular conflicts of PVC with another pathology.

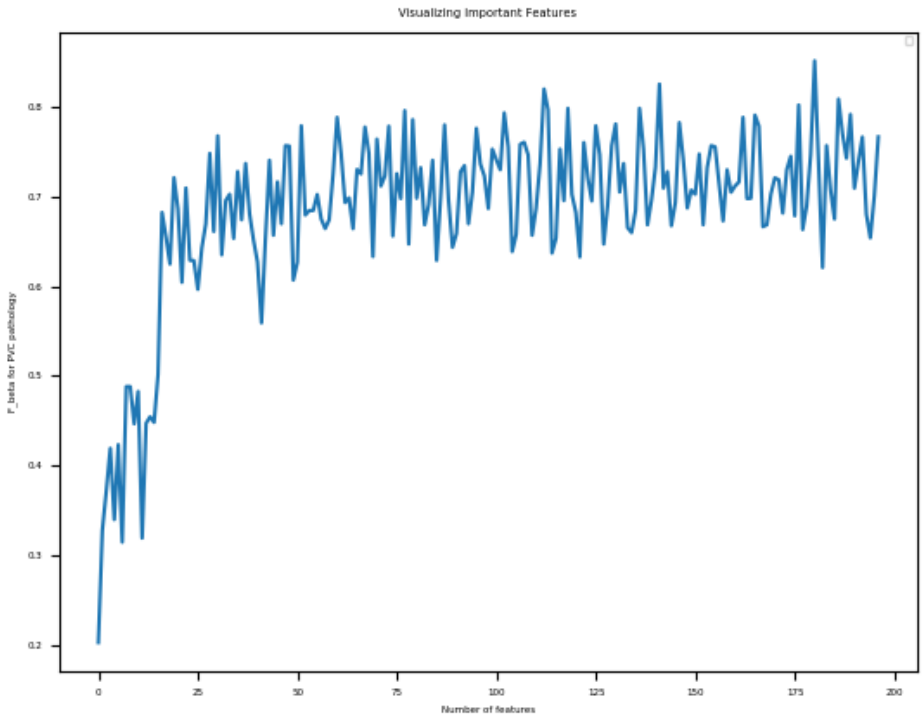
1.3.4.2 Performance when limiting the number of features

Moreover, the number of features is too large. We will need to perform feature selection while training a classifier for detecting PVC.



.

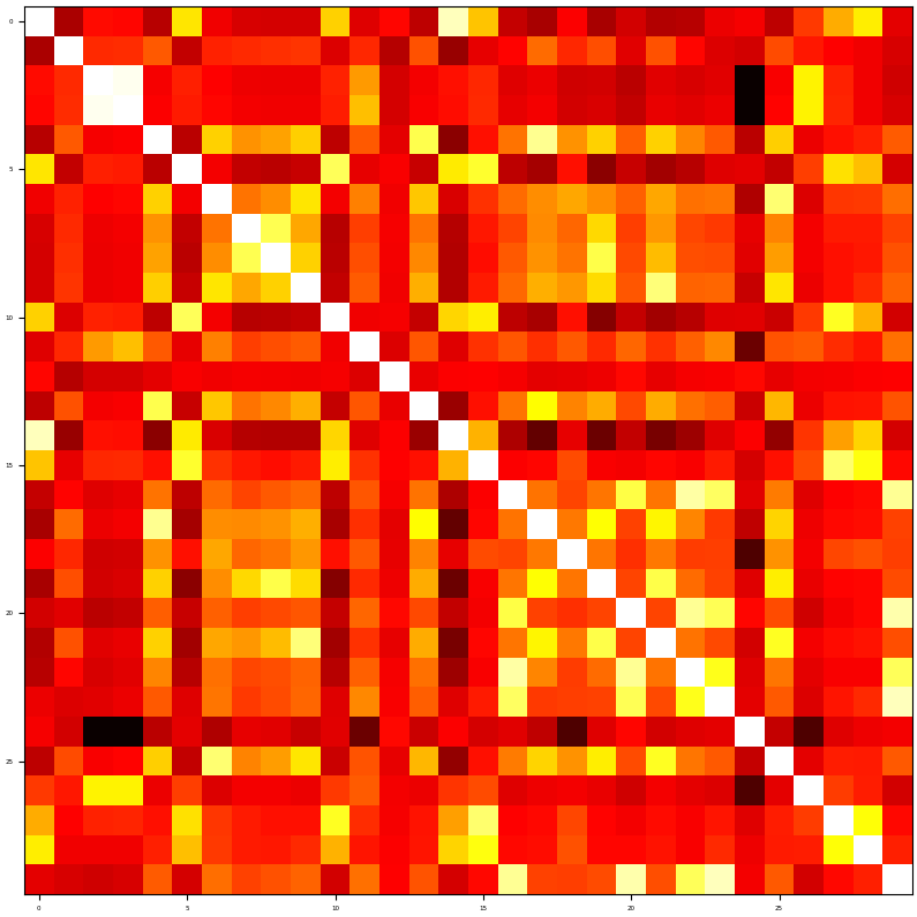
Our first experiment was to get the F_β^{PVC} score according to the number of features we have been using (using k features means that we have been using the k most important according to Random Forest feature importance). This will help us to select the first subset of features we will be using. The results we got are:



Therefore, we can confidently select the first 60 features from the important features selected by the Random Forest Classifier.

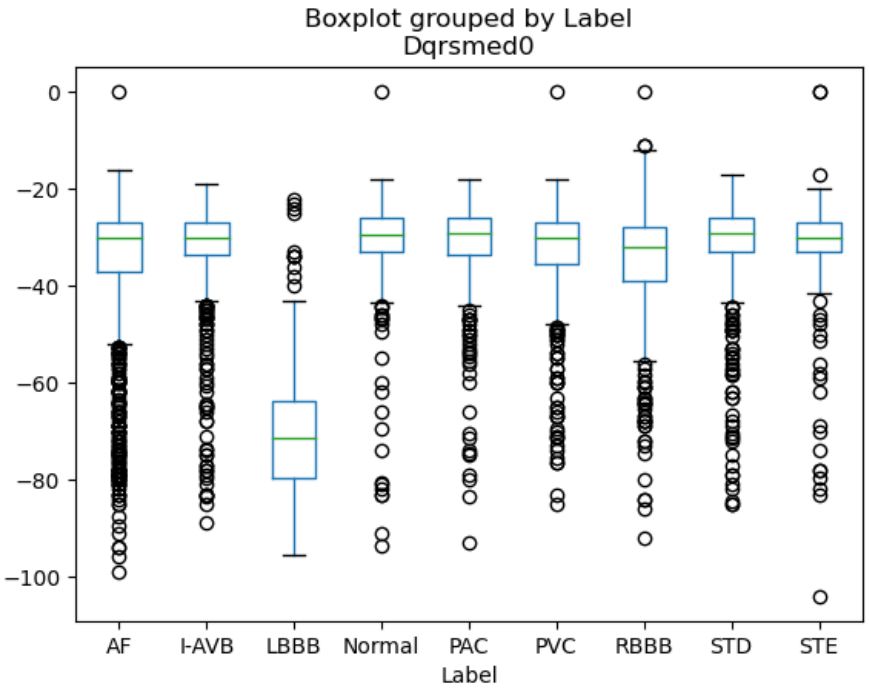
Number of features	F_β^{PVC}
40	0.70
60	0.73
80	0.77
290	0.74

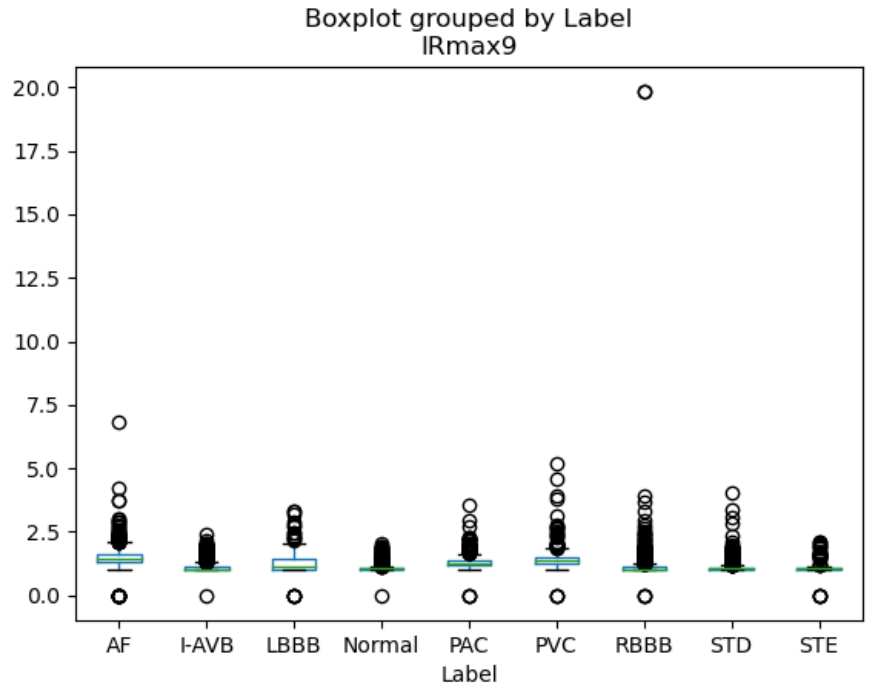
The second experiment was to remove from this first subset the features that were too highly correlated. The correlation matrix of the remaining features is:



We have grouped subset of features when their correlation was higher than 0.75. Therefore, our last set of features selected will be: 'Dqrsmed0', 'IRmax9', 'Age', 'Sqrsmean6', 'IRmax8', 'Dqrsmax11', 'IRmax7', 'Sqrsmax6', 'Dqrsmed11', 'Dqrsmed10', 'IRmax10', 'IRmax1', 'Sex', 'IRmax6', 'Sqrsmax1', 'IRstd9', 'Sqrsmed0', 'Dqrsstd10', 'IRmax5', 'Label'.

For instance, let us show the distribution of the two first features.

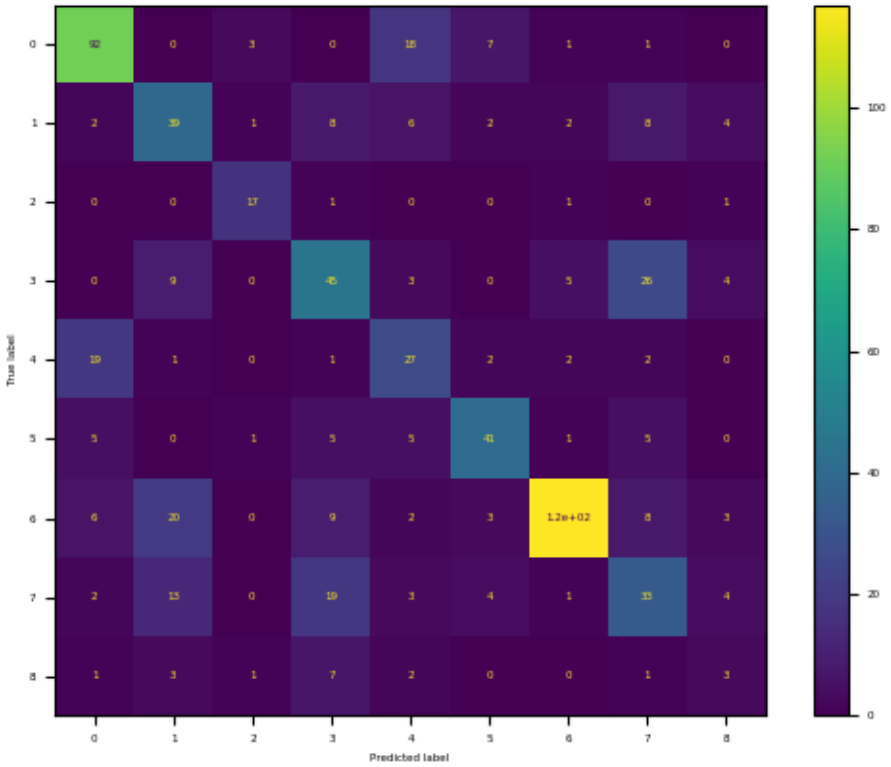




We need to clarify that those were the features selected when only trying to classify the PVC pathology. I think that the list of features selected will be different when inserting the features for all pathologies. Moreover, these two features selected do not seem to be really discriminative for the PVC pathology. We need to further considerate other considerations, like the fact that the T-wave is inverted or energy considerations ([6]).

The results we get with those features are: $F_\beta^{PVC} = 0.7$

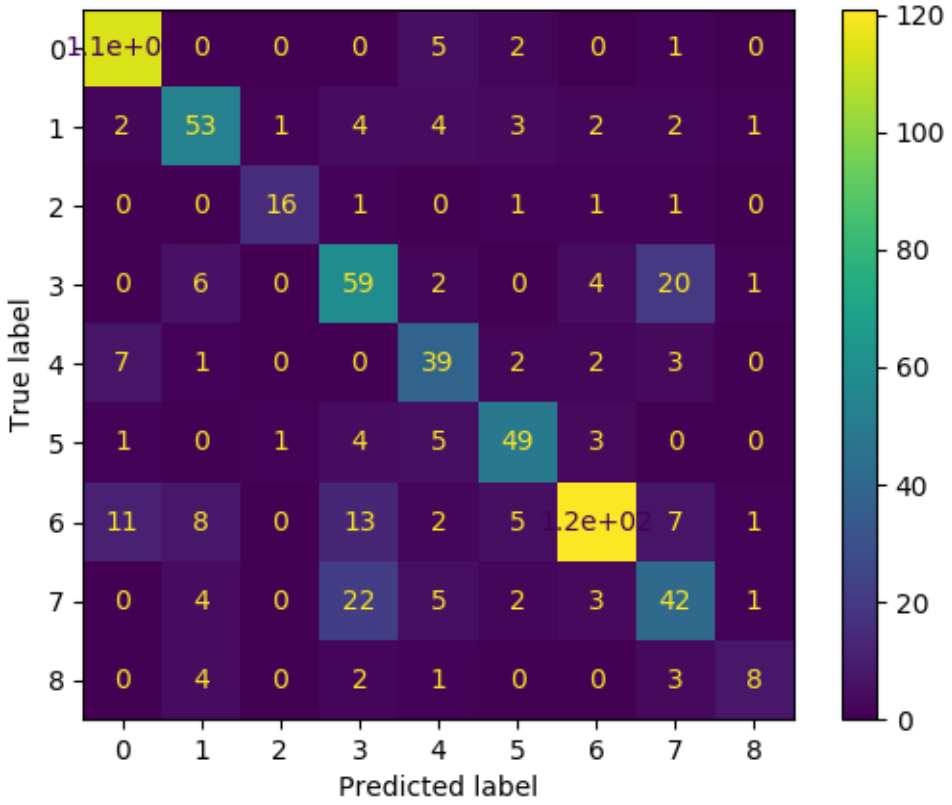
The confusion matrix of this model on the test set is :



We can see that we do have troubles separating the PVC pathology from any other one. I will further see if the features extracted for the PVC pathology are compatible with the AF and I-AVB pathologies, by extracting the features aforementionned for these 3 pathologies and see whether we can more efficiently separate them. I will extract all features for PVC detection and perform later feature selection. The results we got for this first global model are:

Metrics	**AF**	**I-AVB**	LBBB	Normal	PAC	**PVC**	RBBB	STD	STE
F_β	0.92	0.82	0.96	0.73	0.65	0.76	0.82	0.57	0.30
G_β	0.76	0.56	0.90	0.46	0.39	0.51	0.61	0.32	0.14

Note: I have ran several times the RandomizedSearchCV and the results are not really stable when looking at each classes. This is a result I got, not the best, not the worse.
Therefore, adding the other features contributes to having a better classification of PVC. For now, I will not perform feature selection. I will do that at the end, when we will have the global model. Let's take a look at the confusion matrix in order to see why there is such disrepancies between the F_β and the G_β (G_β penalizes more the FN than F_β).



As we can see, we still have troubles separating PVC from the other classes. I will definitely add other features for PVC detection, like the energy of a QRS complex, [6]

1.4 Premature Atrial Contraction

1.4.1 Pathology presentation

Premature atrial contractions (PACs), also known as atrial premature complexes (APC) or atrial premature beats (APB), are a common cardiac dysrhythmia characterized by premature heartbeats originating in the atria. While the sinoatrial node typically regulates the heartbeat during normal sinus rhythm, PACs occur when another region of the atria depolarizes before the sinoatrial node and thus triggers a premature heartbeat. PACs are often completely asymptomatic and may be noted only with Holter monitoring, but occasionally they can be perceived as a skipped beat or a jolt in the chest.
According to [12], PACs can be detected in ECGs thanks to precise morphological characteristics:

- Hidden ectopic P-wave, or a different P-wave morphology
- The PR interval can be longer
- The QRS complex can be narrower (usually combined with BBB pathology)
- Incomplete pause after a PAC beat

1.4.2 Features extracted

We follow the approach of [11], where they have extracted several features directly linked to the morphology of the ecg, and tried to characterize potential deviations in morphologies.

First, the pauses after a PAC beat have been engineered via the Interbeat RR-Interval Difference:

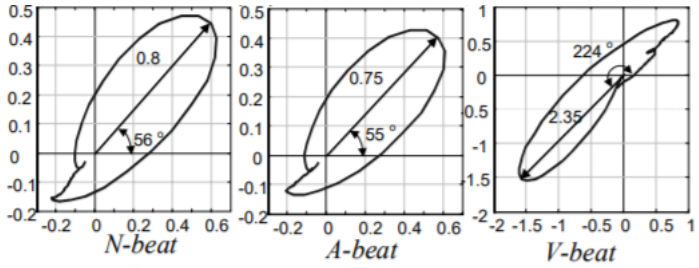
$$RRDiff_n = \frac{RR_n - RR_{n-1}}{(\sum_{i=n-7}^{n-2} RR_i)/5} \tag{1.2}$$

The difference in QRS global morphology have been engineered via the deviation from a reference QRS_{area} and QRS_{width}:

$$QRSWidthDiff_n = \frac{|QRSWidth_n - QRSWidth_{REF}|}{QRSWidth_{REF}} * 100 \tag{1.3}$$

$$QRSWidthArea_n = \frac{|QRSArea_n - QRSArea_{REF}|}{QRSArea_{REF}} * 100 \tag{1.4}$$

The last feature extracted comes from a representation of the two-leads ecg: the vectorcardiographic plane.



$$VECANGDiff_n = |VECAng_n - VECAng_{REF}| \tag{1.5}$$

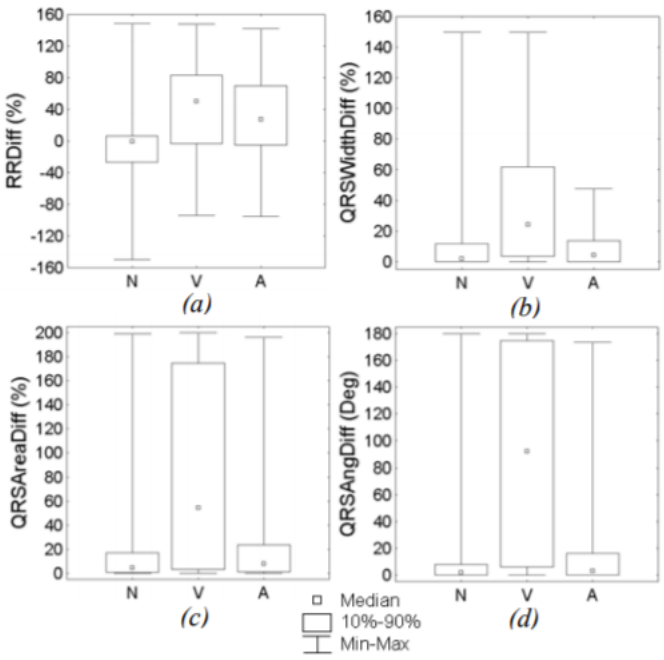
The reference value is always computed as the median of the 5 previous values of this quantity in the signal.
Note 1 : Here, the QRSWidth is defined as the difference between QRSoff and QRSon, whereas we defined it previously as the difference between the S_{point} and the Q_{point} for PVC detection: how can we choose?
Note 2 : Here, the QRSArea is defined as the sum of absolute values of the ecg between QRSoff and QRSon, whereas we defined it previously as the sum of values S_{point} and the Q_{point} for PVC detection: how can we choose?
Question: in the paper, how do we practically compute the VCG angle ?

Features	Definition
$RRDiff_{max}$	Maximum of the Interbeat RR Interval Difference
$RRDiff_{mean}$	Average of the Interbeat RR Interval Difference
$RRDiff_{median}$	Median of the Interbeat RR Interval Difference
$RRDiff_{std}$	Standard deviation of the Interbeat RR Interval Difference
$QRSWidth_{max}$	Maximum of the QRS duration, defined as the length between QRSon and QRSoff points.
$QRSWidth_{mean}$	Average of the QRS duration, defined as the length between QRSon and QRSoff points.
$QRSWidth_{median}$	Median of the QRS duration, defined as the length between QRSon and QRSoff points.
$QRSWidth_{std}$	Standard deviation of the QRS duration, defined as the length between QRSon and QRSoff points.
$QRSArea_{max}$	Maximum of the QRS area, defined as the ecg sum of absolute values between QRSon and QRSoff points.
$QRSArea_{mean}$	Average of the QRS area, defined as the ecg sum of absolute values between QRSon and QRSoff points.
$QRSArea_{median}$	Median of the QRS area, defined as the ecg sum of absolute values between QRSon and QRSoff points.
$QRSArea_{std}$	Standard deviation of the QRS area, defined as the ecg sum of absolute values between QRSon and QRSoff points.

Therefore, we have 12*2 = 24 morphological features for the PAC pathology: we will also need to perform feature selection (we used only leads II and V1 like they did in the paper [11]).

1.4.3 Results

The features extracted in this article and the results they got allowed to separate well PVC beats and PAC beats from normal beats.

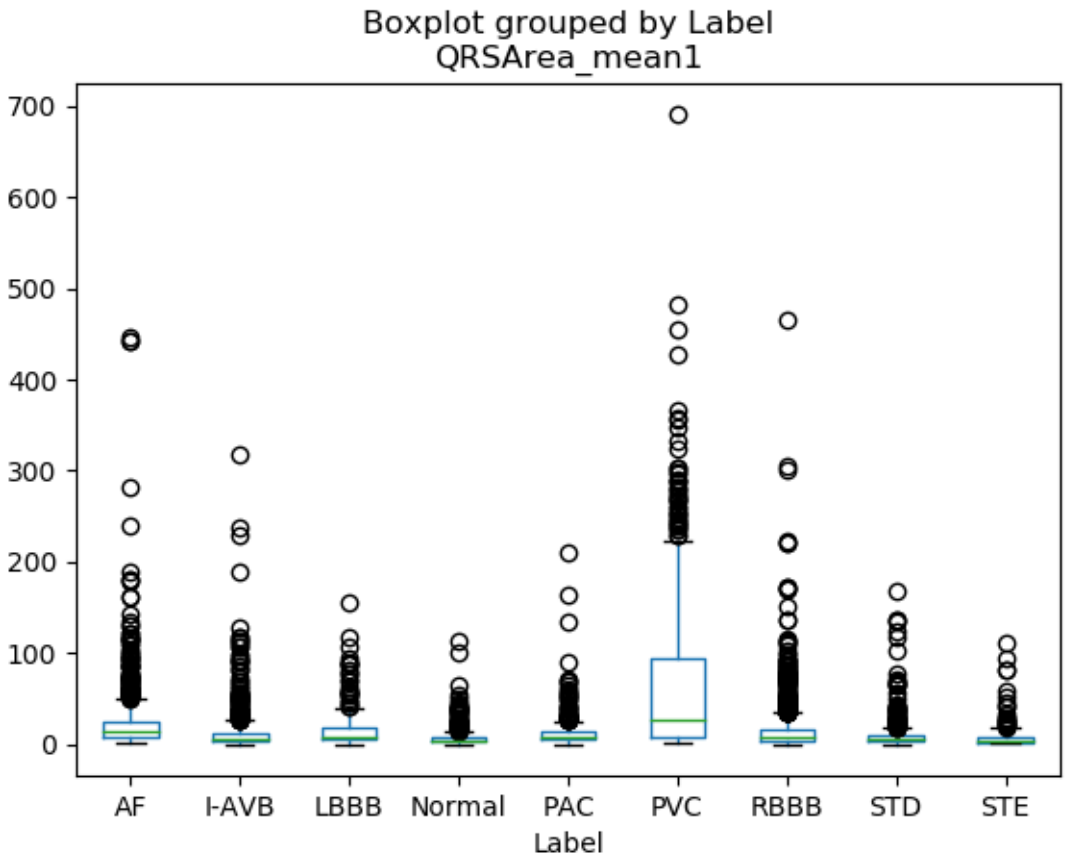


Model	Sen	Spe
$Deterministicclassification$	92.2	96

1.4.4 Limitations

1.4.4.1 Pathology classification

Once again, our task is more complicated since we need to separate the PAC class from 9 other classes. Let us see if the features selected allow to separate efficiently PAC from other classes.

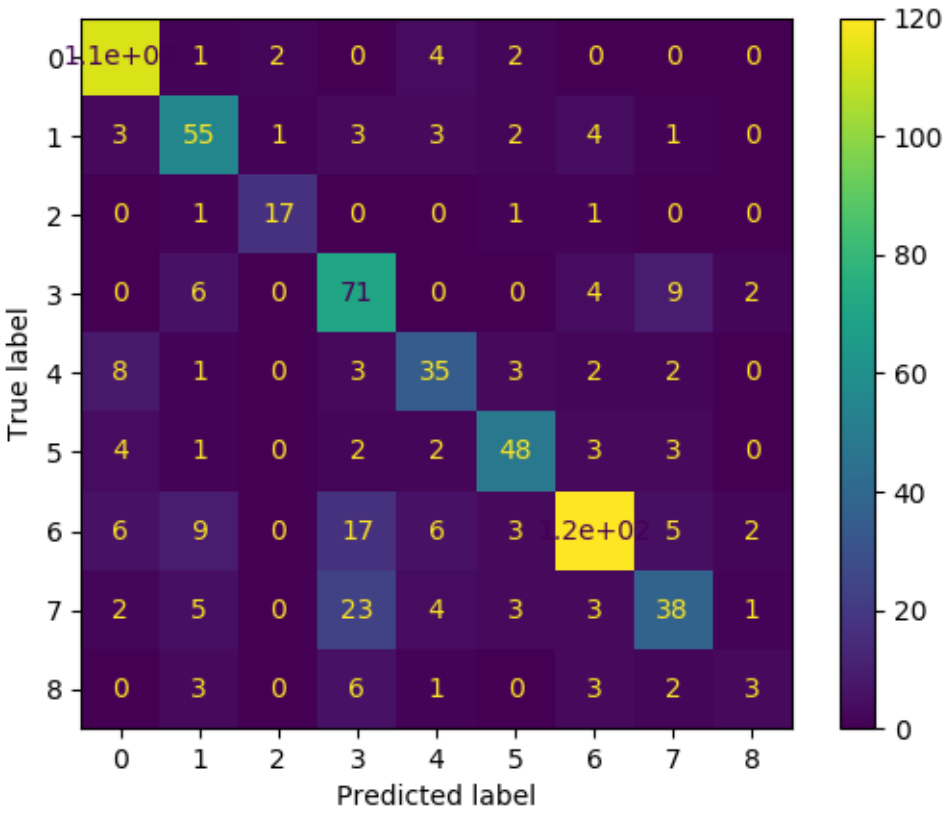


The vast majority of features extracted presents differentiated value for **PVC** only. However, according to experiences performed when studying the PVC pathology, we differentiated PVC from PAC without many problems. Therefore, these features are not really relevant for PAC detection and we need to add other features in order to detect efficiently PAC, like the VECGAngle or other features we will find in the litterature. The score we get for PAC classification against all others is :

Model	F_β^{PAC}	F_β^{PVC}
$RFclassifier$	0.6	0.5

We will see if these features coherently behave with the ones we have extracted until now, by extracting the features for AF, I-AVB, PAC, PVC and see the results we get.

Metrics	**AF**	**I-AVB**	LBBB	Normal	**PAC**	**PVC**	RBBB	STD	STE
F_β	0.92	0.76	0.85	0.71	0.63	0.77	0.74	0.45	0.20
G_β	0.74	0.48	0.65	0.41	0.36	0.52	0.51	0.22	0.08

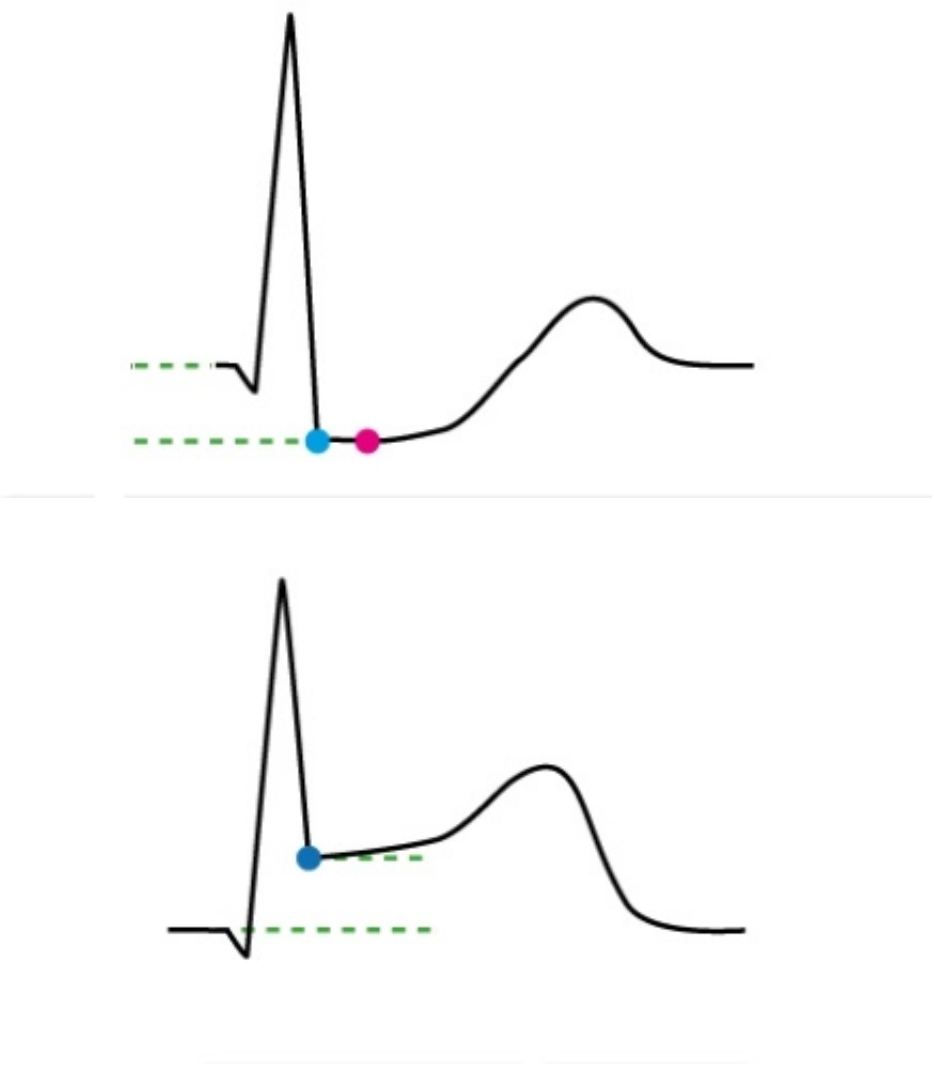


As we can see, the performances of our model where we have included PAC features are slightly worse than our model without PAC features. This is surely dued to the fact that this new model has too much features and is overfitting our training set: we need to perform **feature reduction**.

1.5 ST Depression and Elevation

1.5.1 Pathology presentation

ST Depression and Elevation are not pathologies by themselves, they are markers of other conditions. This is why there is such poor litterature on these pathologies, they have been asked only for the CPSC2018 and PhysioNet 2020. From what we can see from last year's competition-winning entry [2], the STE pathology was the hardest to classify, first because of the poor number of examples (202, [10]), and also because of its hard recognition pattern (depend on the quality of several fiducial points on the ECG).



Therefore, the terms "elevation" and "depression" are defined relatively to the amplitude of the QRSon point. Since STD and STE are not pathologies, there are no research paper displaying a feature engineering approach in order to detect those patologies. Therefore, I searched what STE and STD were commonly the consequence of: Myocardial Infarctions. Thus, detection of MI will consist in deploying techniques to measure ST deviation. Therefore, I will follow the approach of [5] and see how well they perform on our dataset.

1.5.2 Features extracted

In this paper, the features are extracted from the 12 leads. Th ST segment is defined as starting from the QRSoff point to the Ton point. The ST deviation is usually measured 60 or 80 ms after the QRSoff point. Therefore, the features extracted are, for every beat of the 12-leads ECG:

- The amplitude of the ecg at the QRSon point
- Statistics of the amplitude of the ST segment
- The amplitude of the ST segment 60ms after the QRSoff point
- The amplitude of the ST segment 80ms after the QRSoff point

Features	Definition
Amp_{max}	Maximum of the amplitude of the ST segment
Amp_{mean}	Average of the amplitude of the ST segment
Amp_{median}	Median of the amplitude of the ST segment
Amp_{std}	Standard deviation of the amplitude of the ST segment
$QRSon_{max}$	Maximum of the ecg amplitude on the QRSon point
$QRSon_{mean}$	Average of the ecg amplitude on the QRSon point
$QRSon_{median}$	Median of the ecg amplitude on the QRSon point
$QRSon_{std}$	Standard deviation of the ecg amplitude on the QRSon point
$QRSonPoint1_{max}$	Maximum of the ecg amplitude 60ms after the QRSoff point
$QRSonPoint1_{mean}$	Average of the ecg amplitude 60ms after the QRSoff point
$QRSonPoint1_{median}$	Median of the ecg amplitude 60ms after the QRSoff point
$QRSonPoint1_{std}$	Standard deviation of the ecg amplitude 60ms after the QRSoff point
$QRSonPoint2_{max}$	Maximum of the ecg amplitude 80ms after the QRSoff point
$QRSonPoint2_{mean}$	Average of the ecg amplitude 80ms after the QRSoff point
$QRSonPoint2_{median}$	Median of the ecg amplitude 80ms after the QRSoff point
$QRSonPoint2_{std}$	Standard deviation of the ecg amplitude 80ms after the QRSoff point

Therefore, we have 16*12 = 192 features extracted for this pathology.

1.5.3 Results

In this article, the database used is the PTB ECG database. The results obtained are:

Model	Sen	Spe
KNN	99	99

1.5.4 Limitations

1.5.4.1 Pathology classification

Let's see if the features extracted behave coherently with the other pathologies we wish to classify.

Metrics	AF	I-AVB	LBBB	Normal	PAC	PVC	RBBB	**STD**	**STE**
F_β	0.83	0.46	0.95	0.74	0.31	0.80	0.82	0.72	0.52
G_β	0.60	0.22	0.86	0.45	0.14	0.59	0.62	0.45	0.3

2. Work to be done on this approach

2.1 Sharpening of PVC and PAC detection

2.1.1 PVC

We will need to find new features for detecting PVC more accurately: we will take into consideration the fact that when a PVC beat is occuring, the T wave is inverted, and also try to implement energy considerations.

2.1.2 PAC

For now, I wish to determine how to implement the VECG angle. I think that it will help to discriminate PAC from other pathologies. Maybe I should also consider on the fact that the wavedet algorithm is a state of the art algorithm. Therefore, the proportion of P waves detected should also be an indicator of the presence of a different P wave morphology, or a hidden ectopic P-wave.

2.2 LBBB and RBBB classification

For now, as we can see, we already have relatively high scores for LBBB and RBBB pathologies. However, the absence of specific features forthese pathologies can conduct to the misclassification of other pathologies as RBBB and LBBB, leading to FN labels, being a big drawback of our current classification model. I will therefore need to implement features for classifying these pathologies.

2.2.1 LBBB

I have found interesting research papers for LBBB feature extraction [8], [1], both having the same feature extraction approach: **QRS duration measurement**, **QRS Morphology Determination** and **QRS Notching and Slurring Detection**.
Question : How can we implement a robust QRS morphology detector?

2.2.2 RBBB

To this day, I have not found any interesting research papers for RBBB detection.

2.3 Sharpening of AVB classification

As we saw, our scores for I-AVB classification were not spectacular ($F_\beta = 0.72$). This is why I need to find more features for AVB detection, since we only used 3 features of only a single lead for this pathology classification.

2.3.1 Sharpening of STD/STE classification

I will add additional features in order to try to directly specify the ST segment elevation or depression.

2.4 Work on features

An obvious issue of our model is that we currently have too many features in order to spot the 9 pathologies. I have not done this work yet, but I am pretty sure that our model is overfitting our training set (work to be done within the next few days). I see several ways of performing feature reduction.

2.4.1 Feature unification

We can avoid having several definitions of morphological characteristics. For instance, the QRSWidth is sometimes defined as the length between Q point and S point whereas it is also defined as the length between QRS_{on} and QRS_{off} point.

2.4.2 Lead Selection

For now, I have just reproduced the experiments made in the research papers I have found. When the leads on which they were working were unspecified, I extracted the features on every lead.
This is why I think that I should focus my efforts on specific leads for specificpathologies. However, **I have not found in the litterature** such pieces of information.

2.4.3 Feature selection

2.4.3.1 Computational approach

Once we would have done all this work, we will be able to perform feature selection. I will use the **MRMR approach** since it is a unified way to select the best performing and independent subset of features.

2.4.3.2 Medical approach

Ask to a cardiologist which features are more useful in practice or not, and see if we are missing some valuable ground truth.

2.4.4 Additional Features

I will also merge our current features set with the following features, that have all been suggested by Joachim A. Behar.

Features	Definition
bSQI	Signal quality of the overall recording (9)
CosEn	Coefficient of sample entropy (11)
AFE	AFEvidence (12)
OrC	Number of points in the bin containing the Origin (12)
IrE	Irregularity Evidence (12)
PACe	PAC Evidence (12)
$\min_{rr}$	Minimum RR interval
$\max_{rr}$	Maximal RR interval
median_{rr}	Median RR interval
$\text{nb}_{outliers}$	RR-interval outliers. An outlier was defined as a sample exceeding 20% of a window average of size 12 beats.
medR	Median R-peak amplitude (mV)
stdR	Standard deviation of the R-peak amplitude (mV)
medQT	Median distance from Q_{on} to T_{off}.
medQT_b	Median QT interval corrected using the Bazett's formula
medQT_{fre}	Median QT interval corrected using the Frederica's formula
medQT_{fra}	Median QT interval corrected using the Framingham formula
medQT_{hod}	Median QT interval corrected using the Hodge formula
medQS	Median QRS interval length
stdQS	Standard deviation of the QRS intervals
medP	Median P-wave length defined as the distance from P_{on} to P_{off}
stdP	Standard deviation of the P-wave length
medPR	Median PR interval defined as the distance from P_{on} to Q_{on}
stdPR	Standard deviation of the PR interval
medPamp	Median P-wave amplitude defined as the amplitude of the P-wave computed from P_{off} to the peak of the P-wave.
medPRseg	Median PR segment defined as the distance from P_{off} to Q_{on}
medT	Median T-wave length defined as the distance from T_{on} to T_{off}
stdT	Standard deviation of the T-wave length
medTamp	Median T amplitude computed as the amplitude in mV between the T_{off} to the peak of the T-wave.
stdTamp	Standard deviation of the T-wave amplitude
medST	Median segment defined as the distance between QRS_{off} and T_{on}
medSTvar1/2	Amplitude of the ST segment defined as the median amplitude between the PR segment and the ST segment. The amplitude of the PR segment is computed as the median amplitude between P_{off} to QRS_{on}. The amplitude of the ST segment is computed as the median amplitude between the QRS_{off} and T_{on} or alternatively QRS_{off}+60 ms to the T_{on}.
AVNN	Average NN interval duration (ms) (13)
SDNN	Standard deviation of NN interval duration (ms) (13)
RMSSD	Root-mean-squared difference between adjacent NN intervals (ms) (13)
pNN50	Percent of NN interval differences greater than 50 milliseconds (%) (13)
SEM	Standard error of the mean NN interval (ms) (13; 14)
PIP	Percentage of inflection points (%) (13; 14)
IALS	Inverse average length of segments (13; 14)
PSS	Percentage of NN intervals that are in short segments (13; 14)
PAS	Percentage of NN intervals that are in alternation segments of at least 4 intervals (%) (13; 14)
ratio	Ratio of the power spectral frequency in the band 5-9 Hz normalised by the total power frequency computed on the PQRST cancelled signal
$\max_{freq}$	Peak frequency in the band 4-45 Hz from the power spectrum computed on the PQRST cancelled signal

2.5 Statistical feature extraction approach

A major drawback of our approach is the loss of information while extracting statistics of a specific feature. For instance, since I can not use the R-peaks amplitudes of an ecg lead, I will use as features the median, maximum, standard deviation and mean of those amplitudes. But doing so, I lose potential valuable information.

2.5.1 Beat per Beat approach

Some research articles perform a beat per beat classification [7]. This poses some computational complications, but this is an approach I wish to try, I still don't know what will be the modalities of this approach

2.5.2 Representative QRS complex

Some research articles select a representative QRS complex [8] and extract all the relevant features on this specific QRS complex. However, they do not specify how they select this QRS complex.

2.6 Signal processing

For now, we have only found sub-optimal preprocessing techniques. The complication is that every research paper uses a different signal processing technique in order to extract the interesting features they have been focusing on. Therefore, there is two ways of approaching this problem: define a specific signal preprocessing technique for every pathology, or define a unified preprocessing signal in order to extract all the interesting features. There are several techniques proposed:

- Wide Bandpass filter with additionnal Notch filter in order to suppress the powerline interferences
- Wavelet denoising
- A bandpass filter where we determine the cutoff frequencies as a hyperparameter of our global model score
- Drift and baseline wandering suppression

Also, we need to perform zero-padding at the beginning of every lead because the wavedet algorithm systematically skips the two first QRS complexes.

2.7 Imbalanced classification

I need to evaluate if the performances of our model on underrepresented classes (STE: 200/6800) are dued to the fact that we do not have enough examples. There will be various methods in order to overturn this complication: resampling (SMOTE), weighted models.

2.8 Challenge metrics

We need to get a better understing of the FN of our classification. We lose too much power in the G_β.

3. What needs to be explored

3.1 Multi-Label classification

For now, we only perform single label classification. This may explain the number of FN we have on our predictions. At a later stage, I will try to implement multi-label classification for our models. We have already worked on this subject during the unofficial phase but we only got relatively poor results.

3.2 Transfer Learning

I will use the AF models trained by Armand.

Bibliography

[1] Juan Pablo Martinez Alba Martin-Yebra. Automatic diagnosis of strict left bundle branch block using a wavelet-based approach. 2019.

[2] Lina Zhao Xiangyu Zhang Feifei Lieu, Chengyu Liu. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. 2018.

[3] Luning Mao Jieying Wei Jiangling Song Hao Chen, Jiaqi Bai and Rui Zhang. Automatic identification of premature ventricular contraction using ecgs. 2019.

[4] Jiaqi Bai Jieying Wei Qiang Li Luning Mao, Hao Chen and Rui Zhang. Automated detection of first-degree atrioventricular block using ecgs. 2019.

[5] Fayyaz A. Afsar Muhammad Arif, IjazA. Malagore. Detection and localization of myocardial infarction using k-nearest neighbor classifier. 2010.

[6] Awadhesh Pachauri and Manabendra Bhuyan. Wavelet and energy based approach for pvc detection. 2009.

[7] Richard B. Reilly Philip de Chazal, Maria O'Dwyer. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. 2004.

[8] Ivo Viscor Josef Halamek Filip Plesinger Magdalena Matejkova Radovan Smisek, Pavel Jurak. Automatic detection of strict left bundle branch block. 2019.

[9] Hein J.J. Wellens Sinjin Lee. Paroxysmal atrioventricular block. 2009.

[10] Chih-Han Huang Tsai-Min Chen. Detection and classification of cardiac arrythmias by a challenge-best deep learning neural network model. 2020.

[11] Ivaylo I. Christov Vessela T. Krasteva, Irena I. Jekova. Automatic detection of premature atrial contractions in the electrocardiogram. 2006.

[12] Frank G. Yanowitz. Introduction to ecg interpretation. 2018.