

Study on the use of convolutional neural networks for the diagnosis of atrial fibrillation

M. Maione¹

(November 2020)

Abstract – In this study we will introduce atrial fibrillation, one of the cardiac arrhythmias, and see how it can be diagnosed using convolutional neural networks in combination with various methods and supervised learning models, including: gray-level co-occurrence matrix, short-time Fourier transform based spectrogram, support-vector machines, k-nearest neighbors, multi-layer perceptron, focal loss, multi-scale decomposition, time-frequency analysis.

Index terms – Biomedical monitoring, electrocardiogram, arrhythmia, atrial fibrillation, deep learning, convolutional neural network, classification.

1. Introduction

Cardiovascular diseases are a group of diseases affecting the heart and/or blood vessels. On a world level, and in particular in countries with a typically Western lifestyle, cardiovascular diseases are the cause of 17 million deaths per year.

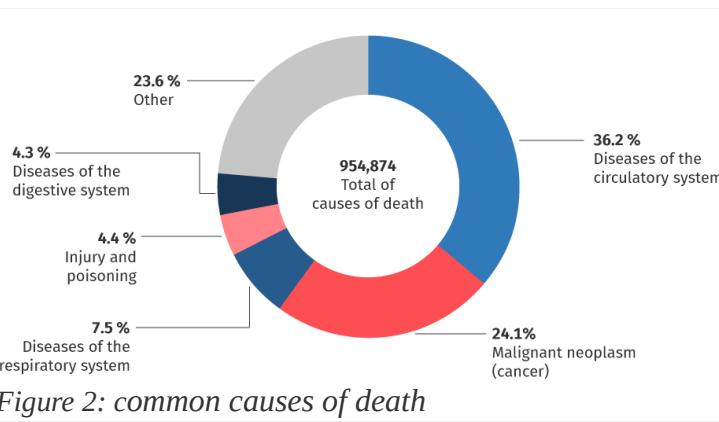


Figure 2: common causes of death

The narrowing, obstruction or excessive enlargement (aneurysm) of the blood vessels that can accompany this disease are in fact responsible for very widespread pathologies, such as coronary (angina pectoris and heart attack), cerebrovascular (stroke) and peripheral vascular diseases. In the family of cardiovascular pathologies, all congenital heart defects, rheumatic diseases involving myocardium, various forms of arrhythmia, pathologies affecting the heart valves and heart failure are also included. Cardiac pathologies are divided into: cardiovascular disease, coronary heart disease, heart "muscle" disease, heart valve disease, pericardial disease, heart conduction disease, vessel disease.

In this study we focus on heart conduction disease in particular on atrial fibrillation, one of the two arrhythmias^[b].

1.1. Electrocardiogram

The electrocardiogram (ECG) is the graphic reproduction of the electrical activity of the heart during its operation,

recorded at the level of the body surface. On the surface of the body, low intensity electric fields (1mV) are present and can be recorded, which in the individual at rest are mainly due to periodic depolarizations and repolarizations of the heart. To record an electrocardiogram it is necessary to have electrodes placed on the body surface, forming leads arranged in such a way as to be able to better analyze the variations of the heart's dipole vector. In order to record the potentials, 10 electrodes are placed on the body: 4 peripheral (wrists and ankles) and 6 precordial, so as to record 12 leads.

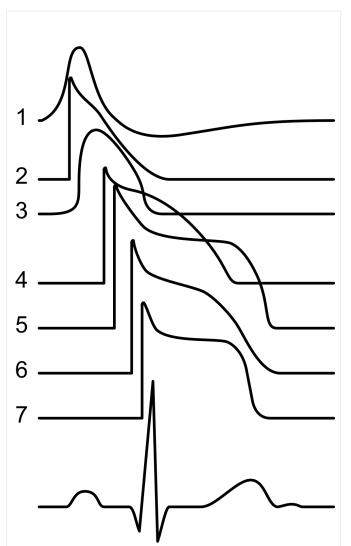


Figure 1: ECG related to the various action potentials of the heart

In general, the electrocardiogram (ECG) signals consist of six components that are designated as P, Q, R, S, T, and U:

- The P wave represents atrial depolarization [0.2 – 0.4mV];
- The QRS complex represents ventricular depolarization [1 – 2mV];
- The T wave represents ventricular repolarization [0.4 – 0.5mV];
- The U wave represents papillary muscle repolarization.

Usually an electrocardiograph has an amplification of a factor of 60db only in the useful band. For monitoring it uses frequencies between 0.05 – 50Hz, while for diagnostics purpose it goes up to 1kHz^[13].

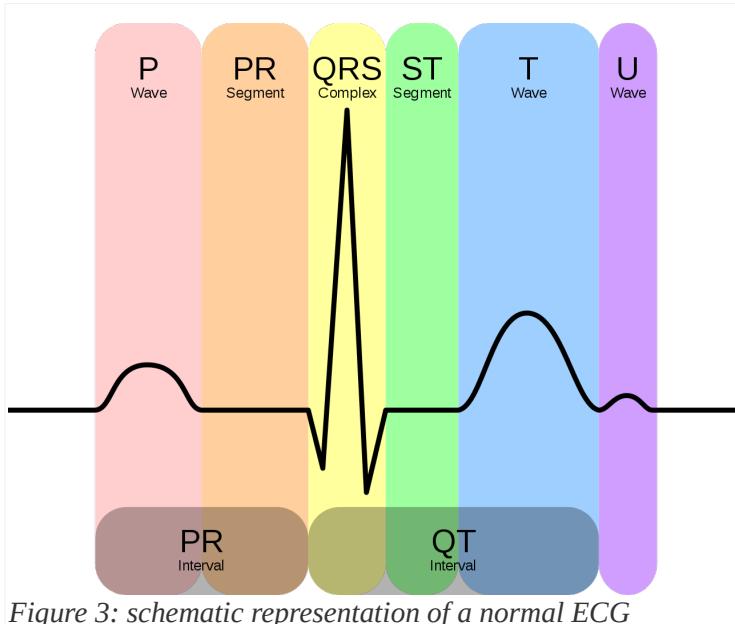


Figure 3: schematic representation of a normal ECG

1.2. Arrhythmia

Arrhythmia is a clinical condition in which the normal frequency or regularity of the heart rhythm is missing, or the physiological atrio-ventricular activation sequence is altered. Common symptoms are:

- Extrasystole: it is like a “void”, a missed beat;
- Tachycardia: the sensation is of an increase in beats, which can be regular but also irregular, fatigue, difficult breathing, dizziness;
- Bradycardia: fatigue, dizziness and possible loss of consciousness.

The alterations can have the following origins: at the level of the sinoatrial node, of supraventricular origin, at the level of the atrioventricular node, of ventricular origin. As for the supraventricular and ventricular origins, we have two types of fibrillation: atrial and ventricular^[c].

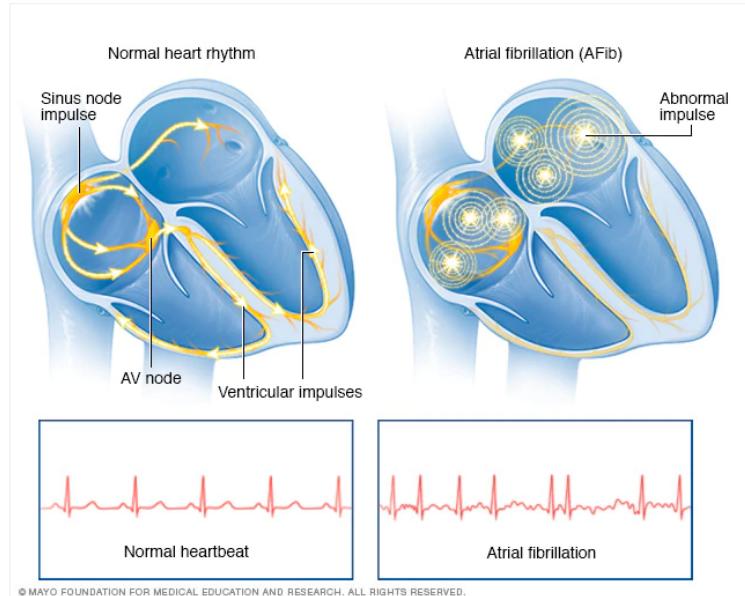
a) Ventricular fibrillation

Ventricular fibrillation (VF) is a very rapid, chaotic cardiac arrhythmia that causes uncoordinated contractions of the heart muscle. With the onset of this arrhythmia, blood circulation slows down considerably, up to cardio-circulatory arrest and subsequent respiratory arrest, until death if cardioversion of the rhythm is not intervened through defibrillation and cardiopulmonary resuscitation^[c].

b) Atrial fibrillation

Atrial fibrillation (AF) is a cardiac arrhythmia that originates in the atria. The electrical impulses that give rise to the contraction of the atria are activated in a totally chaotic and fragmentary way, giving rise to multiple wave fronts and disorganized and fragmentary contractions. The current clinical approach aims to treat symptoms by:

- rhythm control (i.e. recovery and maintenance of sinus rhythm with anti-arrhythmic drugs or catheter ablation);
- heart rate control with drugs that regulate the conduction of atrial stimuli to the ventricles associated with anti-thrombotic therapy.



© MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH. ALL RIGHTS RESERVED.

Figure 4: normal heartbeat & atrial fibrillation

From Figure 5 it can be observed that for AF patient there are tiny irregular fluctuations in the P-wave and QRS complex. Given the presence of these irregularities, for the purpose of arrhythmia screening various morphological features, including the peaks and widths corresponding to different ECG segments are typically used.

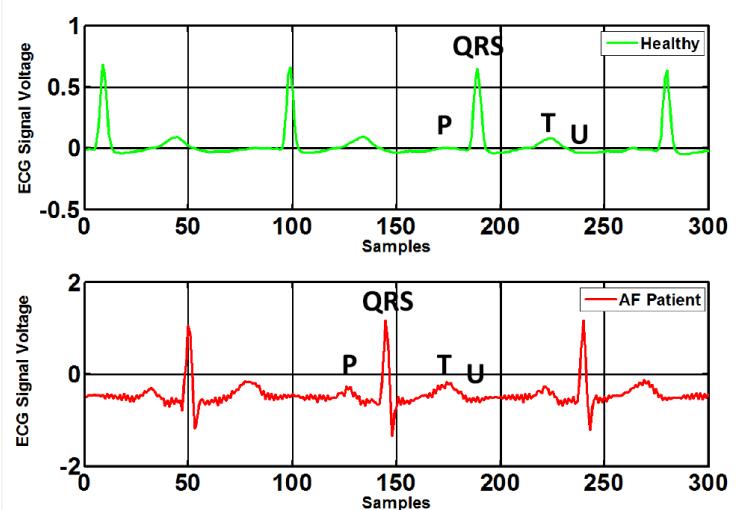


Figure 5: ECG signal of healthy person (top) and AF patient (bottom)

In Figure 6 are visible the F-waves with their saw-tooth appearance.

AF occurs in 2% of the population, and increases to 6% by the age of 65. The most serious complication of AF is thromboembolic stroke, which leads to permanent disability or even death. Since in many cases the symptoms are initially imperceptible to patients, diagnosing it as soon as possible becomes difficult, for this reason with the current technological progress, diagnostic tools can

be included in wearable devices, for example in smart watches or in affordable medical devices^[d].

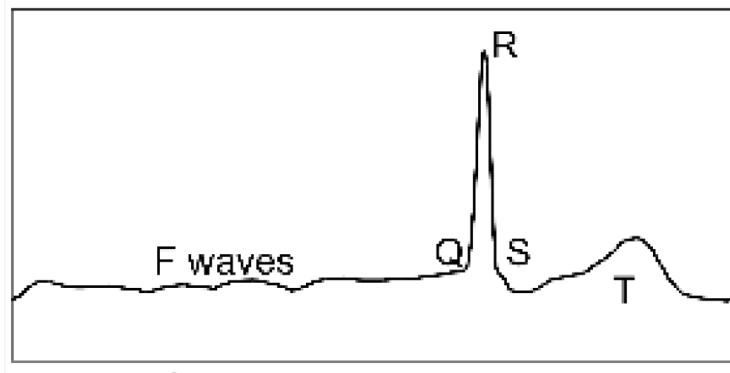


Figure 6: AF beat

2. Notation and relevant definitions

The introduction of the computational models used will now follow.

2.1. Artificial neural network

An artificial neural network (ANN) is an interconnected group of nodes, inspired by a simplification of neurons in a brain. As in Figure 7 each circular node represents an artificial neuron and a line represents a connection from the output of one artificial neuron to the input of another.

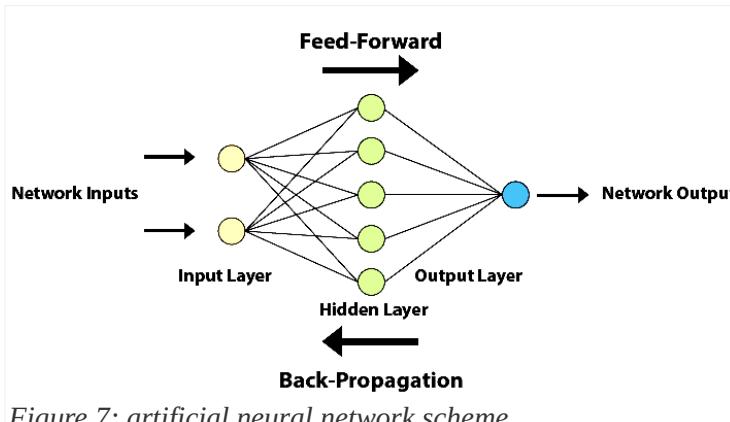


Figure 7: artificial neural network scheme

Typically neural networks are obtained through the combination of simple predictor of the form $g(x) = \sigma(w^T x)$. The function $\sigma: \mathbb{R} \rightarrow \mathbb{R}$ is known as activation function^[12].

To define the neural computation we must specify: the neural model, the model dimensions, the configuration procedure^[11].

a) Neural model

Neural model is composed by activation function and network topology.

Common activation function are: binary step, linear activation, sigmoid, tanH, ReLU, softmax, swish, etc...

The neural network topology represents the way in which neurons are connected to form a network. In other words, the neural network topology can be seen as the re-

lationship between the neurons by means of their connections. The topology of a neural network plays a fundamental role in its functionality and performance. Some famous networks are: feed-forward NN, regulatory feedback networks, radial basis function network, recurrent neural network, modular neural network, etc...

b) Configuration procedure

Configuration procedure is composed by: configuration algorithm, training set, validation set.

Configuration algorithm must be chosen based on the learning paradigm: supervised, unsupervised and reinforcement.

In supervised learning the data set contain both the data points X and the labels y . The learning task is to produce the desired output y for each input x . A cost function ℓ is used to estimate the correctness of the predicted label \hat{y} compared to the desired output y .

The simplest division of the data-set is using the Pareto principle, that divide in 80/20 the training and the test set.



Figure 8: 80/20 rule

When evaluating different settings for estimators, there is still a risk of over-fitting on the test set because the parameters can be tweaked until the estimator performs optimally. This way, knowledge about the test set can “leak” into the model and evaluation metrics no longer report on generalization performance. To solve this problem, yet another part of the data-set can be held out as a so-called “validation set”: training proceeds on the training set, after which evaluation is done on the validation set, and when the experiment seems to be successful, final evaluation can be done on the test set.

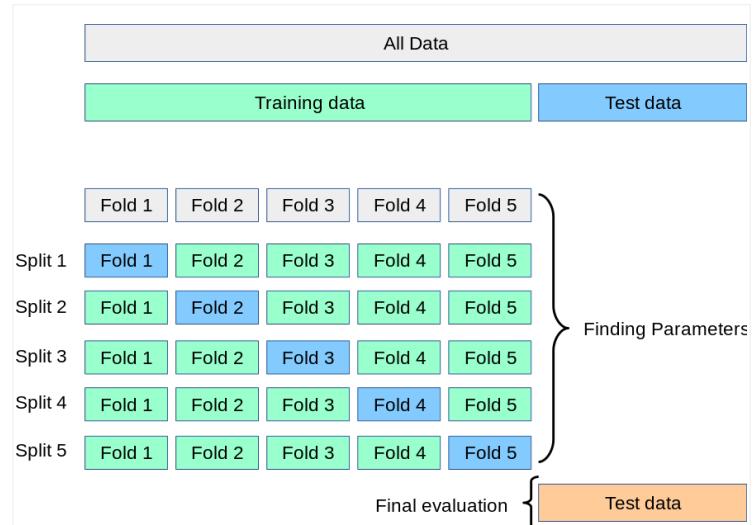


Figure 9: K-Fold CV

However, by partitioning the available data into three sets, we drastically reduce the number of samples which

can be used for learning the model, and the results can depend on a particular random choice for the pair of (train, validation) sets.

A solution to this problem is a procedure called cross-validation (CV). A test set should still be held out for final evaluation, but the validation set is no longer needed when doing CV. In the basic approach, called K-Fold CV (Figure 9), the training set is split into k smaller sets^[i].

Instead unsupervised learning looks for previously undetected patterns in a data set with no pre-existing labels.

Reinforcement learning minimize long-term cost modifying the network's weights. At each point in time an action is performed and an observation is received with a cost. At this point the algorithm/agent decides whether to perform new actions to uncover their cost or to exploit prior learning. In this case NN is used as the learning component.

2.2. Feed-forward neural network

In a feed-forward neural network connections between units do not form loops and information only moves in one direction, forward, with respect to entry nodes, through hidden nodes (if any) to exit nodes. Feed-forward NN computes a function $f: \mathbb{R}^d \rightarrow \mathbb{R}^n$. A parameter $w_{ij} \in \mathbb{R}$ (called weight) is associated with every edge (i, j) . NNs are trained using algorithms that reduce the training error. Fixed a cost function ℓ , an example (x_t, y_t) , defined $\ell_t(W) = \ell(f_{G, W, \sigma}(x_t), y_t)$ and Z_t as the index of a random training example, then the standard training algorithm for NNs is stochastic gradient descent:

$$w_{i,j} \leftarrow w_{i,j} - \eta_t \cdot \frac{\partial \ell_{Z_t}(W)}{\partial w_{i,j}} : (i, j) \in E \quad (1)$$

This procedure is known as error back-propagation algorithm^[12].

2.3. Convolutional neural network

Convolutional neural networks (CNN) function like all feed-forward neural networks: an input layer, one or more hidden layers, which perform calculations using activation functions, and an output layer with the result. The difference is precisely the convolution in place of general matrix multiplication. The typical architecture of a CNN is formed by: convolutional layer, pooling layer, ReLu layer, fully connected layer and loss layer.

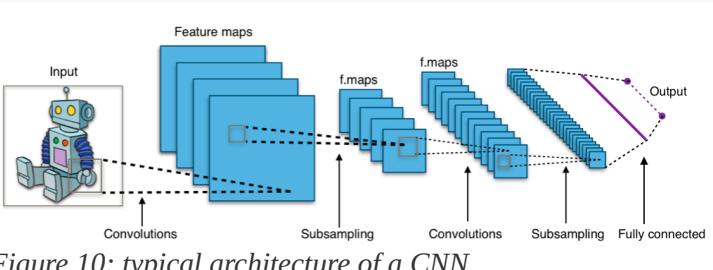


Figure 10: typical architecture of a CNN

a) Convolution

Discrete convolution is an operation between two functions f and h that produce a third function G which consists in integrating the product between the first and the second translated by a certain value. For complex-valued function f and h defined on the set \mathbb{Z} the discrete convolution G of f and h is given by^[h]:

$$G[m, n] = (f * h)[m, n] \quad (2)$$

$$(f * h)[m, n] \stackrel{\text{def}}{=} \sum_j \sum_k h[j, k] \cdot f[m-j, n-k] \quad (3)$$

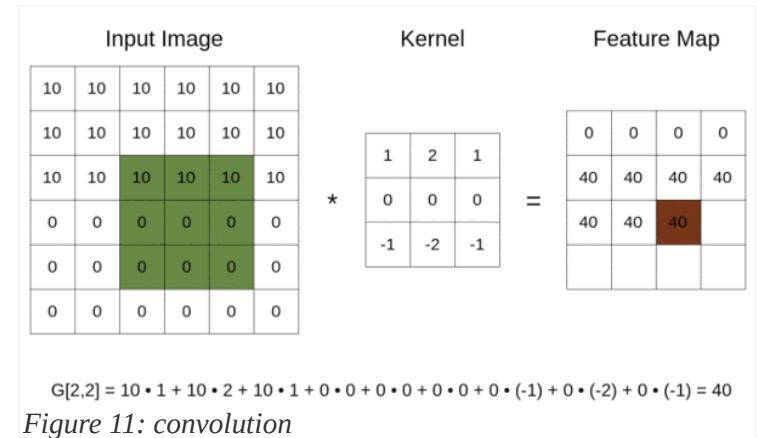


Figure 11: convolution

2.4. QRS recognition algorithms

QRS recognition algorithms are used to facilitate the annotation of the QRS complex in databases.

Various classes of QRS recognition algorithms have been proposed in the literature, which can be classified according to their complexity and performance. The Pan-Tompkins algorithm is one of the most used methods for real-time recognition of the QRS complex from the ECG signal^[13].

The performance of the method was tested on an annotated arrhythmia database and evaluated also in presence of noise. Pan and Tompkins reported that the 99.3% of QRS complexes was correctly detected^[e].

2.5. Assessment indicators

To evaluate the performance of the proposed classifiers, five statistical indicator are commonly used:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$\text{sensitivity} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{specificity} = \frac{TN}{TN + FP} \quad (6)$$

$$\text{positive predictive rate} = \frac{TP}{TP + FP} \quad (7)$$

$$F_1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (8)$$

3. CNN and AF detection

Below is the introduction of ten methodologies for AF detection using convolutional neural networks.

3.1. Time-frequency analysis and CNN (detect 12 heart rhythm)

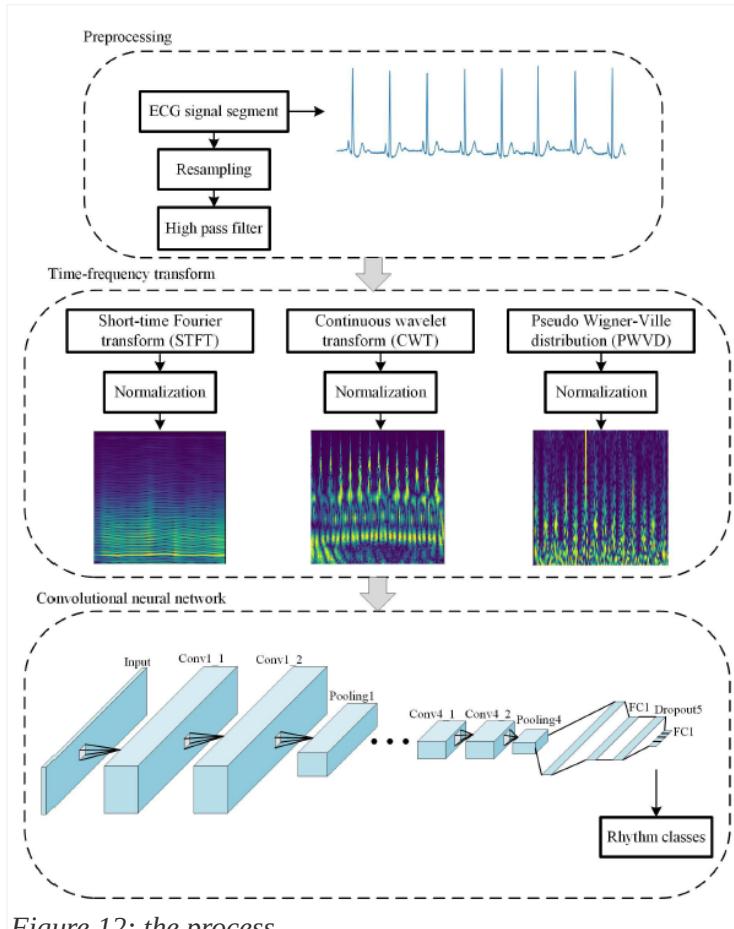


Figure 12: the process

Z. Wu, T. Lan, C. Yang, and Z. Nie^[1] proposed a method formed of three steps: preprocessing, time-frequency transform and the convolutional neural network. This method is able to recognize 12 different heart rhythms:

1. normal sinus rhythm (NSR);
2. paced rhythm (P);
3. atrial bigeminy (AB);
4. atrial fibrillation (AF);
5. atrial flutter (AFL);
6. ventricular flutter (VF);
7. first degree heart block (BI);
8. premature ventricular contractions (PVC);
9. sinus bradycardia (SBR);
10. ventricular tachycardia (VT);
11. supraventricular tachy-arrhythmia (SVTA);

12. noise and signal contaminated by noise.

The fact that this method has not specialized only on atrial fibrillation, and that its output consists in listing the percentage of the input belonging to these 12 classes, makes it in addition to having given excellent results (97%) one of the best.

a) Database

The authors used 6 different databases from PhysioNet² (that are labeled with the 12 different heart rhythms listed before):

1. MIT-BIH arrhythmia;
2. MIT-BIH malignant ventricular arrhythmia;
3. MIT-BIH atrial fibrillation;
4. Long-term AF;
5. MIT-BIH normal sinus rhythm;
6. MIT-BIH noise stress test.

b) Preprocessing

The ECG signal is formed of 12 leads and need to be preprocessed, so the signal was splitted in segments of 10s (using rhythms annotations included in the databases) and resampled to 125Hz. Then a second order high-pass filter with a cutoff frequency of 0.5Hz was used to remove the baseline of each signal.

The “MIT-BIH noise stress test” (it contains noise from baseline wander, muscle artifact and electrode motion artifact) was used to add noise to signals to make the network more reliable to noise. To the signal x was added the noise signal n multiplied by the gain a

$$y = x + a \cdot n \quad (9)$$

The gain a is calculated using the power of the raw signal and the power of the noise signal.

c) Time-frequency transform

Time-frequency analysis is one of the important methods to process non-stationary signals and provides information on time domain and frequency domain.

The authors use three different methods: short-time Fourier transform (STFT), continuous wavelet transformation (CWT) and pseudo Wigner-Ville distribution (PWVD).

The STFT of the sequence $x(t)$ is defined as:

$$STFT(t, w) = \int_{-\infty}^{+\infty} x(\tau) \cdot w(\tau - t) \cdot e^{-j\omega\tau} dt \quad (10)$$

where $w(t)$ is the window function of 2s and the step length is 0.08s.

The CWT is defined as:

² <http://physionet.org>

$$CWT(a, b) = \int_{-\infty}^{+\infty} \frac{x(t)}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) dt \quad (11)$$

where a is the scale factor and b is the time shift factor, $\psi(t)$ is the Morlet wavelet basis:

$$\psi(t) = \exp\left(-\frac{t^2}{2}\right) \cdot \cos(5t) \quad (12)$$

At last, the PWVD is defined as:

$$PWVD(t, w) = \int_{-\infty}^{+\infty} w(\tau) \cdot x(t + \frac{\tau}{2}) \cdot \bar{x}(t - \frac{\tau}{2}) \cdot e^{-jw\tau} d\tau \quad (13)$$

with \bar{x} the complex conjugate of x .

After the time-frequency transform (one of this) we can calculate the time-frequency distribution matrix as:

$$|Y(a, b)| = \sqrt{Y^2(a, b)} \quad (14)$$

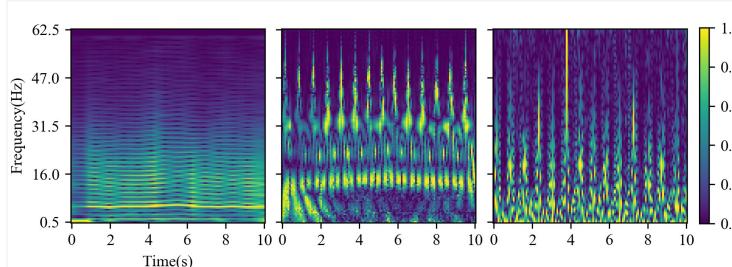


Figure 13: STFT, CWT & PWVD

Above the results of the three methods.

d) CNN

The network is formed by a convolutional block and a fully connected layer.

The convolutional block is formed by the concatenation of a block repeated 4 times. This block is formed by:

1. Convolution layer: 64 kernels (5×5 step 1), padding same, ReLU;
2. Convolution layer: 64 kernels (5×5 step 1), padding same, ReLU;
3. Max pool layer: (2×2);
4. Dropout layer: dropout rate 0.3.

The fully connected layer is formed by:

1. Flatten layer;
2. Fully Connected layer: 128 cells, ReLu;
3. Dropout layer: dropout rate 0.5;
4. Fully Connected layer: 12 outputs, Softmax.

The last layer outputs the probability for each of the 12 different heart rhythms.

e) Training method

The authors used a 5-fold cross validation with the cross-entropy as the loss function:

$$\ell = -\sum_{i=1}^n y_i \cdot \log(\hat{y}_i) \quad (15)$$

and the root mean square prop (RMSprop) method was used to update the weights of the network:

$$w_t = w_{t-1} - \alpha \cdot \frac{dw}{\sqrt{(s_{dw}) + \varepsilon}} \quad (16)$$

where α is the learning rate, s_{dw} is the accumulation of momentum and $\varepsilon > 0$.

f) Results

Four metrics were used to evaluate the three methods, the ranking was: STFT, CWT, PWVD.

Table 1: classification performance

Method	Accuracy	Sensitivity	Specificity	F1
STFT	96.65	96.47	99.68	96.27
CWT	95.26	94.71	99.55	94.80
PWVD	92.07	92.19	99.25	92.26

3.2. Lightweight CNN

D. Lai, X. Zhang, Y. Bu, Y. Su and C. Ma^[2] proposed a method formed of three steps: preprocessing, extraction of cardiac rhythms features and a lightweight CNN.

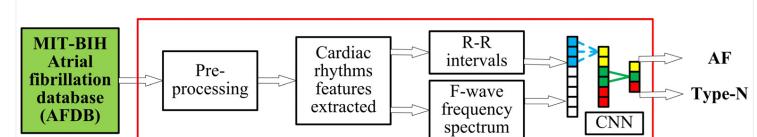


Figure 14: schematic diagram

a) Database

The database used was the MIT-BIH atrial fibrillation database from PhysioNet³ formed of 23 long-term ECG with a sample rate of 250Hz that are labeled with 4 different heart rhythms:

1. atrial fibrillation (AF);
2. atrial flutter (AFL);
3. atrial-ventricular junctional rhythm (AVJ);
4. other rhythms (N).

b) Preprocessing

In this step each recording is divided in segments of 10s, a median filter (0.5Hz) and a band-pass filter (100Hz) are applied, to reduce noise and increase precision and efficiency of the network. Early normalization was used on the data:

$$\hat{x} = \frac{x - \mu}{\sigma} \quad (17)$$

³ <http://physionet.org>

and batch normalization was used to accelerate the training during the back-propagation, using the following update equation:

$$y^{(k)} = \gamma^{(k)} \cdot \hat{x}^{(k)} + \beta^{(k)} \quad (18)$$

where β, γ are trained at each iteration on k iterations.

c) Extraction of cardiac rhythms features

A convolutional layer is used to obtain a feature maps c :

$$c = \sigma(b + \sum_n w \cdot x_{i+k}) \quad (19)$$

where σ is the activation function (ReLU), b is the bias of the activation map, n the size of the kernel, w is the weight and k is the stride. The result of the convolutional layer is fed into a pooling layer P , to reduce the dimension of the feature map c and the number of parameters, as follows:

$$P = \max_{t \in T} c_{i+s} \quad (20)$$

where t is the pooling window size and s the stride.

Then two different test where done (as shown in Figure 14), one using raw data and one using ECG signal analysis:

- R-wave detection;
- R-R interval calculation;
- F-wave transformation (in Figure 6).

The presence of numerous low-amplitude F-waves instead of P-wave can be found in AF, jointly with asymmetrical R-R intervals.

d) Lightweight CNN

A lightweight CNN was used as a binary classifier (details are shown in Figure 15), then back-propagation was used to reduce the loss.

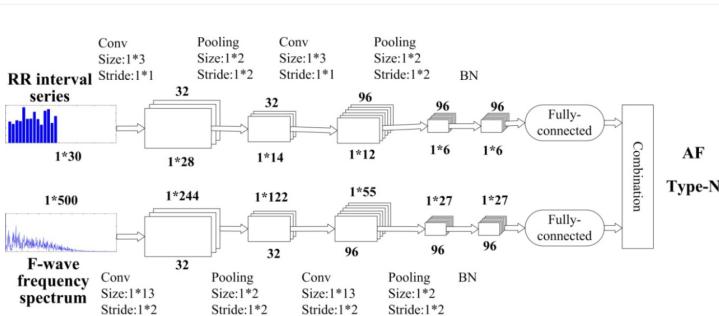


Figure 15: lightweight CNN

e) Results

A stratified 5-fold CV was used to tune both the model architecture and the hyper-parameters of the CNN and evaluate the model performance.

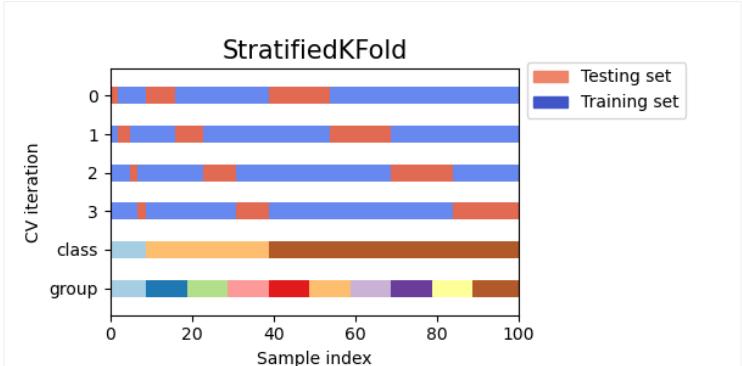


Figure 16: stratified K-Fold CV

Three metrics were used to evaluate the two methodologies:

1. R-R intervals + F-wave spectrum;
2. Raw data.

Table 2: classification performance

Low level features	Accuracy	Sensitivity	Specificity
RRI+FWS	97.5	97.8	97.2
Raw data	86.3	89.5	82.7

3.3. Multi-scale decomposition enhanced residual CNN

X. Cao, B. Yao and B. Chen^[3] proposed a signal decomposition via derived wavelet frames and two different CNN models: MSResNet and FDResNet.

a) Database

The database used from the PhysioNet⁴ Challenge 2017 is contributed by AliveCor (a manufacturer of single-channel ECG device) and is formed of 8528 single short ECG lead recordings, each of which is from individual customer of AliveCor, with a sample rate of 300Hz that are divided in 4 categories:

1. AF rhythm (A);
2. normal rhythm (N);
3. other rhythm (O);
4. noisy recordings (~).

b) Preprocessing

In this step each recording is divided in segments of 9s and derived wavelet frames (DWFs) is applied.

As you can see in Figure 17 implicit dual-tree complex wavelet packets (IWP) are constructed based on dyadic dual-tree CWP (DDCWP).

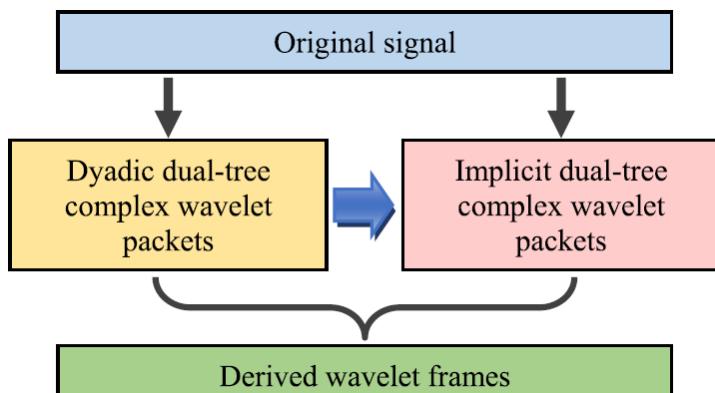


Figure 17: DWFs

To derive the IWP_s, assuming $x\{n\}$ is the ECG signal, then:

1. Perform a dual-tree wavelet packet decomposition, with k as the number of decomposition layers, j as the sequence number of the sub-signal. $x\{n\}$ is transformed into a set of sub-signals:

$$D_k = \{D_k^j(n) : j=1, 2, \dots, 2^k\} \quad (21)$$

2. Rearrange the content of D_k according to the central frequency as $R_k = \{R_k^j(n) : j=1, 2, \dots, 2^k\}$, let:

$$j = \sum_{m=0}^{k-1} 2^m \cdot n_m + 1 \quad (22)$$

the binary coding of the index j , and construct a new index as follows

$$j' = \sum_{m=0}^{k-1} 2^m \cdot n'_m + 1 \quad (23)$$

where the parameter n'_m is defined as

$$n'_m = \begin{cases} n_m, & m=k-1 \\ \text{mod}(n_m + n_{m+1}, 2) & m=0, 1, \dots, k-2 \end{cases} \quad (24)$$

3. Generate the implicit wavelet packet with the following equation:

$$iwp_k^j(n) = R_k^{2j}(n) + R_k^{2j+1}(n), \quad 1 \leq 2^{k-1} - 1 \quad (25)$$

In Figure 19 it can be seen that the center frequency of the derived wavelet packet is the band boundary of the traditional binary wavelet packet, thereby improving the ability of the algorithm to extract the information of the transition band. In Figure 18 the result of the DWFs on the AF rhythm signal.

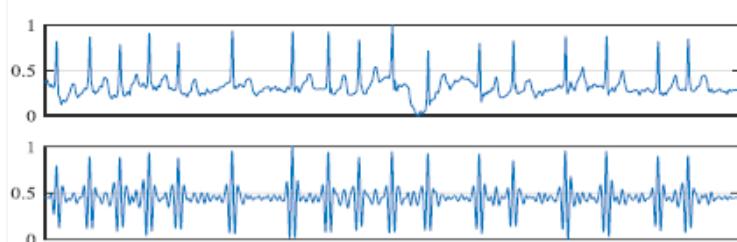


Figure 18: AF rhythm & the reconstructed sub-signal

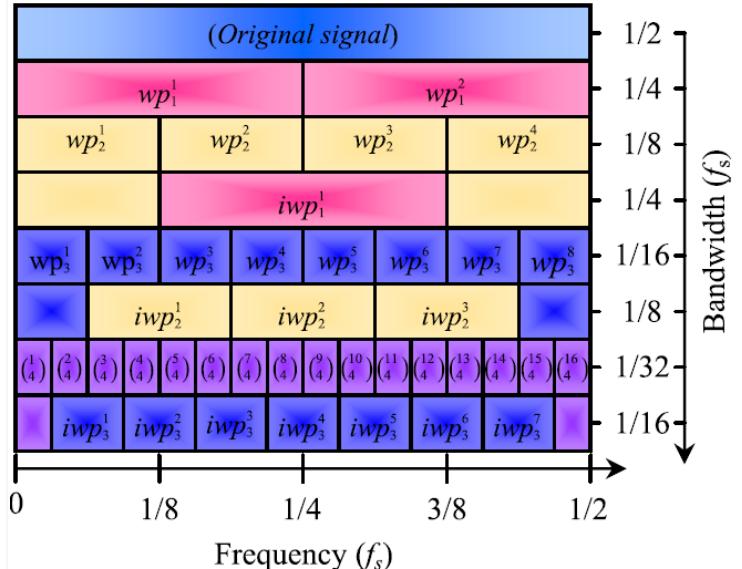


Figure 19: frequency-scale topology of the derived wavelet packet frame

c) MSResNet

Multi-scale decomposition enhanced fast down-sampling residual CNN consists of three parallel FDResNet, same structure but independently trained by reconstructed samples of different scales (as you can see in Figure 20), each of them has learned different features, and the independent classification capabilities are different. The predictions are connected into a small neural network that learn the end-to-end characteristics of the three, and higher recognition accuracy can be obtained.

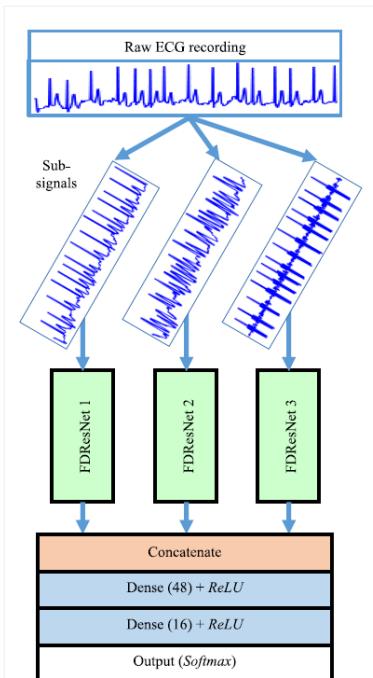


Figure 20: MSResNet

d) FDResNet

Fast down-sampling residual CNN is mainly composed of: a fast down-sampling module, a residual convolution module, and a classification module, as you can see in Figure 21 the full network structure is composed of 3 module:

1. Fast down-sampling convolutional module: is formed by $2 \times$ convolutional layer (ReLU) with random dropout layer and a batch-normalization layer to enhance the generalization of the model. This module effectively reduces the calculation of subsequent DN, reduces data redundancy and facilitates model learning;
2. $3 \times$ residual convolutional module: formed by convolutional layers in series and residual short circuit. The width of the 3 residual convolution

modules is gradually increased. All of them use a max-pooling layer to down-sample the feature vectors;

3. Classification module: consist of a convolutional layer (to reduce the dimension of the feature vectors), a flatten layer, a random dropout layer (to prevent overfitting), 2 full connection layers (ReLU) and a softmax classifier.

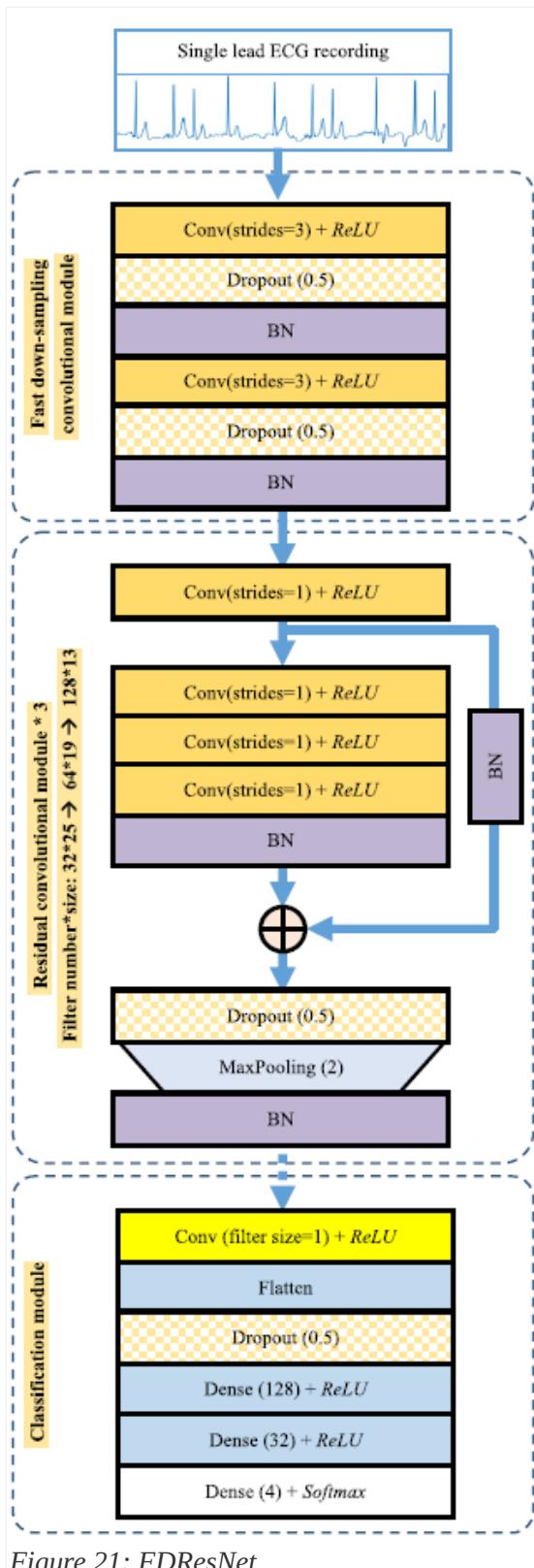


Figure 21: FDResNet

e) Results

6-Fold CV was used to train the network. FDResNet can learn effective classification features from time domain

ECG waveform.

Table 3: FDResNet classification performance

Sub-band	Normal	AF	F1
raw	0.8059	0.9816	0.8702
w p_3^1	0.7982	0.9757	0.8766
w p_4^1	0.8215	0.9751	0.8973

Based on different coupling strategies (concatenate layer in Figure 20) MSResNet gives the following results:

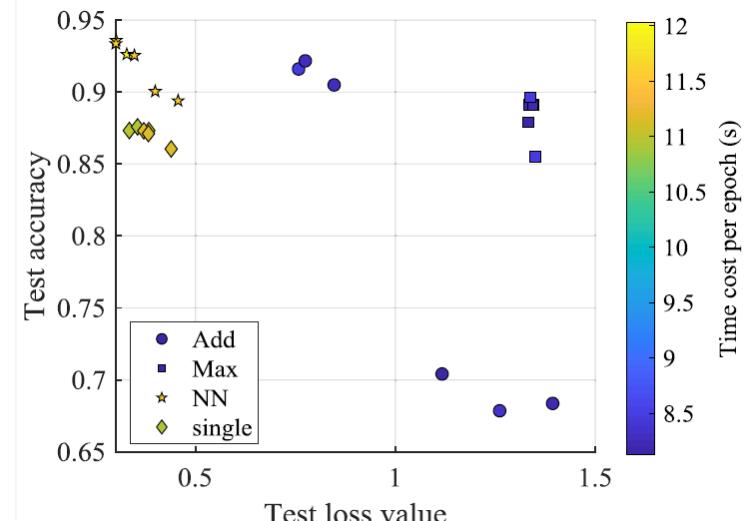


Figure 22: performance of different coupling methods

3.4. Dual heartbeat coupling based on CNN

X. Zhai and C. Tin^[4] propose to transform the beats into a dual beat coupling matrix (2D) as input of the CNN.

a) Database

The authors use the MIT-BIH arrhythmia database from PhysioNet⁵, 48 records of 30min two-channel ECG signals. These were filtered with a band-pass filter (0.1Hz – 100Hz) and digitized at 360Hz.

b) Preprocessing

The beat was segmented such that it was centered around the R peak (using a max interval of 20s) as in the left part of Figure 23. Because each segment can have different length, they were scaled into the same length M .

A series of three adjacent beats is taken into account. The first pair is $(\text{Beat}_{i-1}, \text{Beat}_i)$ denoted as $\text{DualBeat}_{i-1,i}$, the second pair is $(\text{Beat}_i, \text{Beat}_{i+1})$ denoted as $\text{DualBeat}_{i,i+1}$, as shown in the left part of Figure 23, then a coupling matrix (CM) with size $M \times M$ is computed as:

$$CM = [\text{DualBeat}_{i-1,i}[1], \dots, \text{DualBeat}_{i-1,i}[M]] \\ \times [\text{DualBeat}_{i,i+1}[1], \dots, \text{DualBeat}_{i,i+1}[M]]^T \quad (26)$$

the result is shown in the right part of Figure 23 (an example on a supraventricular ectopic beat is shown in Figure 24).

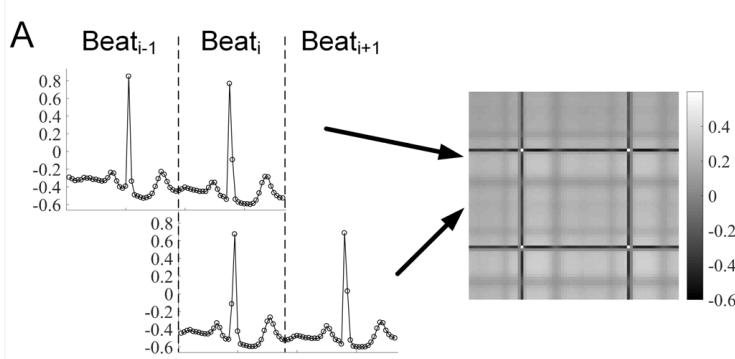


Figure 23: coupling matrix of beats originating from the sinus mode

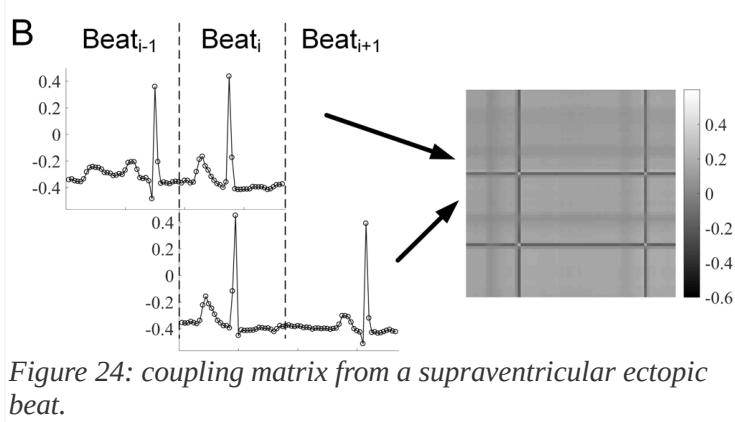


Figure 24: coupling matrix from a supraventricular ectopic beat.

c) CNN classifier

The network is formed by:

1. Input (73×73);
2. Convolutional l. with ReLU (50@66×66)
3. Maximum sub-sampling l. (50@33×33)
4. Convolutional l. with ReLU (100@24×24)
5. Average sub-sampling l. (100@8×8)
6. Convolutional l. with ReLU (150@4×4)
7. Fully connected layer with dropout (150);
8. Softmax loss layer (5).

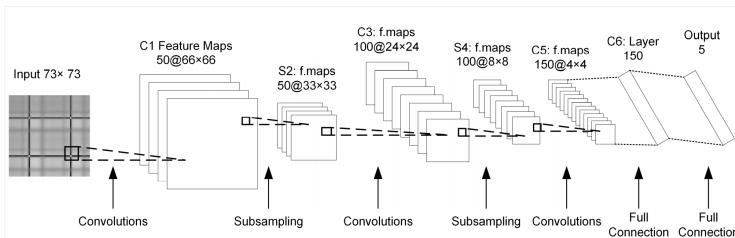


Figure 25: CNN scheme

d) Results

Results show that the method perform better on VEB.

Table 4: classification performance

Beats	Accuracy	Sensitivity	Specificity	PPR ⁶
VEB ⁷	98.6	93.8	99.2	92.4
SVEB ⁸	97.5	76.8	98.7	74.0

3.5. CNN with SVM⁹

Z. Li, X. Feng, Z. Wu, C. Yang, B. Bai and Q. Yang^[5] proposed the following architecture (as in Figure 26): original data will be calculated through convolution layer by feed-forward channel, then the result will be fed to a max-pooling layer and stretched as a one-dimensional vector through a flatten layer. Finally, the vector is trained into a SVM classifier.

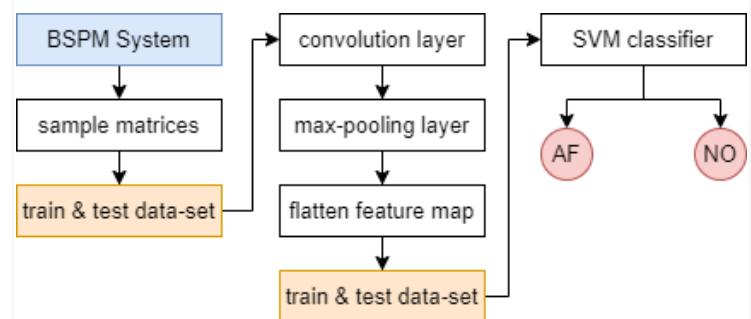


Figure 26: flow diagram

a) Database

The data-set were provided and supported by West China Hospital in collaboration with Sichuan University. It consist of ECG signals from 14 patients with AF, acquired via the preoperative body surface potential mapping (BSPM) and an integrated back-end acquisition system (NeuroScan). BSPM consist of 74 front electrode points and 54 on the back, with a sampling frequency of 1kHz. The method is compatible with the standard 12-lead electrode coordinates, and contains more spatiotemporal information.

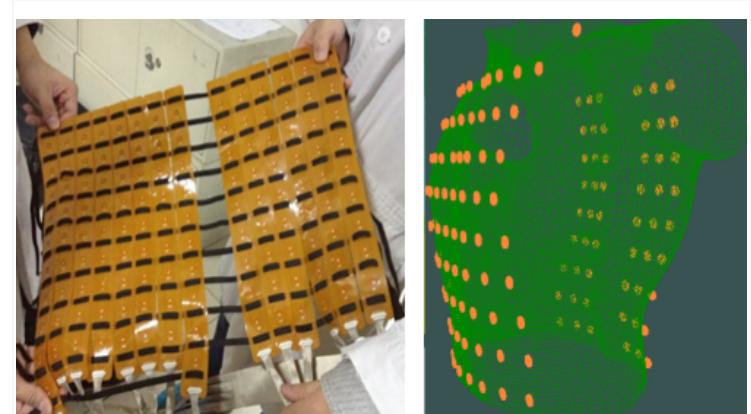


Figure 27: BSPM system

6 Positive predictive rate

7 Ventricular ectopic beats

8 Supraventricular ectopic beats

9 Support Vector Machine

b) Preprocessing

During the acquisition interference are removed via a 50Hz notch filter, then raw data were processed:

1. A zero-phase, third-order Butterworth filter (0.5–100Hz);
2. An absolute value filter;
3. A third-order Butterworth filter (20Hz cutoff).

Data are normalized (as in Eq. (17)), and then 160s are kept, from the center, of each acquisition. 10-fold cross validation is used.

c) CNN

Convolution layers use ReLU as activation function. The authors choose Adam^[f] (Adaptive Moment Estimation) as the algorithm for gradient descent with the following update rule:

$$w_{t+1} = w_t - \alpha \cdot \frac{\hat{m}_w}{\sqrt{\hat{v}_w + \epsilon}} \quad (27)$$

with gradient w , step-size α , momentum factor ϵ , and biased first moment estimate v_w , and second raw moment estimate m_w .

d) SVM

The decision function with kernel used in this project is:

$$g(x) = \text{sign}\left(\sum_{i=1}^N \hat{\alpha}_i \cdot y_i \cdot K(\tau, x) + \hat{b}\right) \quad (28)$$

The authors tried three different kernels:

1. Linear kernel

$$K(\tau, x) = (\tau \cdot x + 1) \quad (29)$$

2. Polynomial kernel, where q is the degree.

$$K(\tau, x) = (\tau \cdot x + 1)^q \quad (30)$$

3. Gaussian kernel, where δ is the width.

$$K(\tau, x) = \exp\left(-\frac{\|\tau - x\|^2}{2\delta^2}\right) \quad (31)$$

e) Results

As in Figure 28 the RBF (Radial basis function) kernel give the best result:

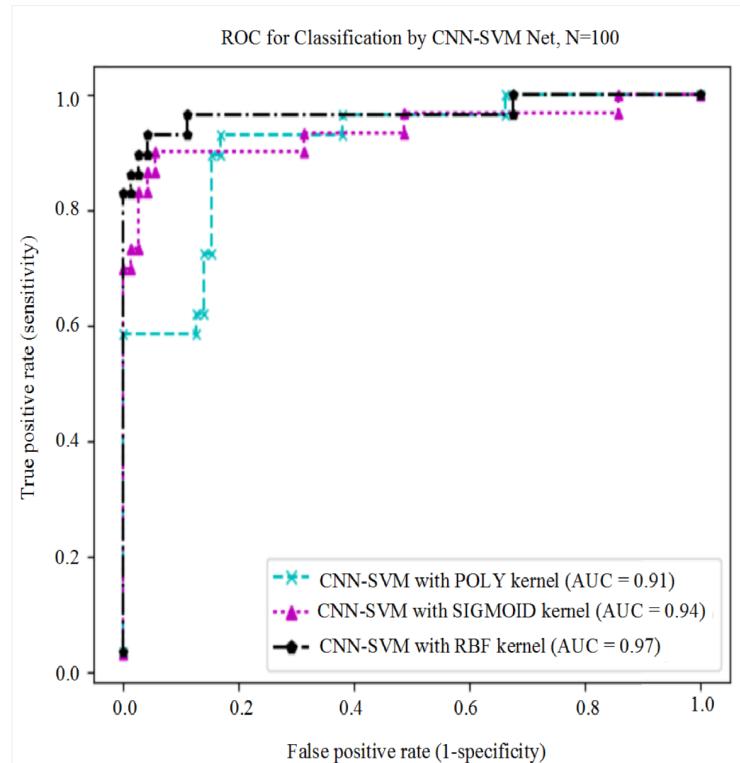


Figure 28: ROC curve for the 3 different kernels

3.6. Deep CNN

B. Pourbabae, M. Roshtkhari and K. Khorasani^[6] aa

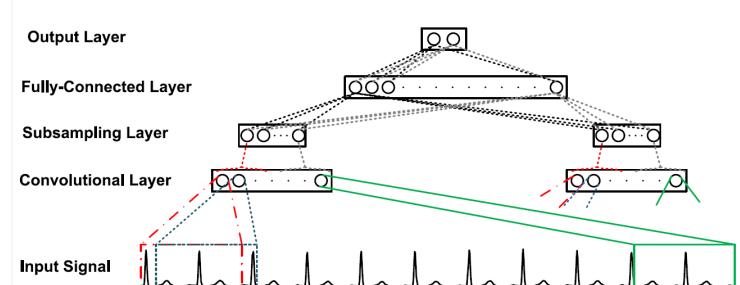


Figure 29: overall structure of the CNN

3.7. Dense CNN with focal loss and image generation

M. Al Rahhal, Y. Bazi, H. Almubarak, N. Alajlan and M. Al Zuair^[7] Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Accuracy	Sensitivity	Specificity
96.06	88.48	96.29

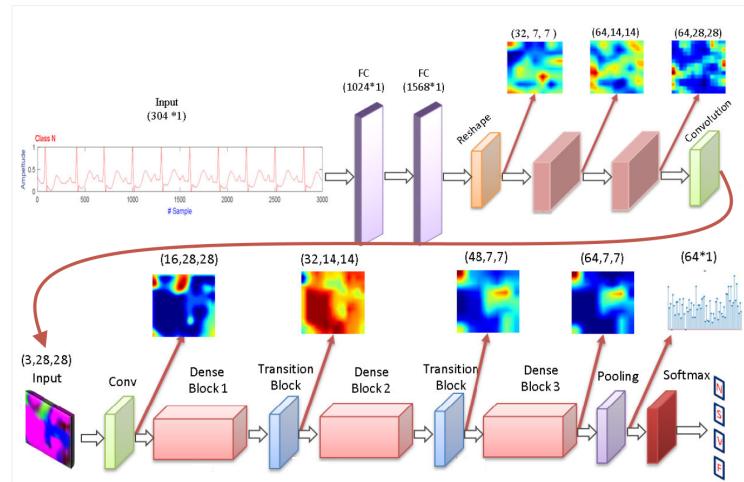


Figure 30: generator network & discriminative module

3.8. STFT-based spectrogram and CNN

J. Huang, B. Chen, B. Yao and W. He^[8] Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

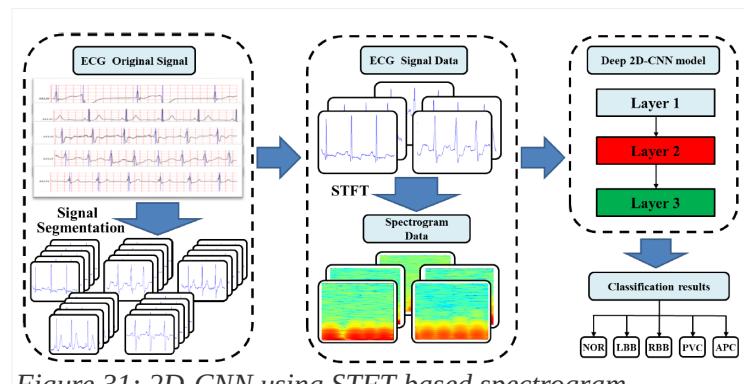


Figure 31: 2D-CNN using STFT-based spectrogram

3.9. Gray-level co-occurrence matrix enhanced CNN

W. Sun, N. Zeng and Y. He^[9] Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

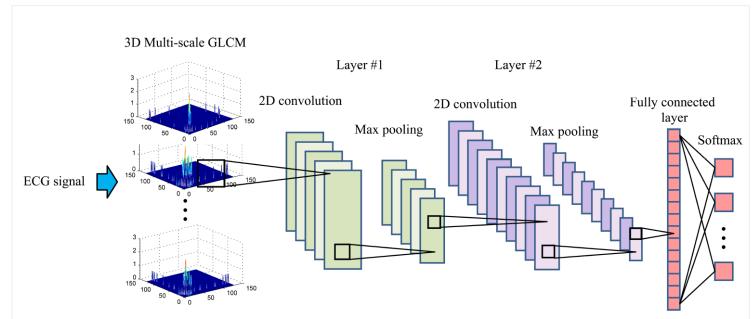


Figure 32: general architecture

3.10. Multiscaled fusion of deep CNN

X. Fan, Q. Yao, Y. Cai, F. Miao, F. Sun and Y. Li^[10]
Lorem ipsum dolor sit amet, consectetur adipiscing elit,
sed do eiusmod tempor incididunt ut labore et dolore
magna aliqua. Ut enim ad minim veniam, quis nostrud
exercitation ullamco laboris nisi ut aliquip ex ea com-
modo consequat. Duis aute irure dolor in reprehenderit in
voluptate velit esse cillum dolore eu fugiat nulla pariatur.
Excepteur sint occaecat cupidatat non proident, sunt in
culpa qui officia deserunt mollit anim id est laborum.

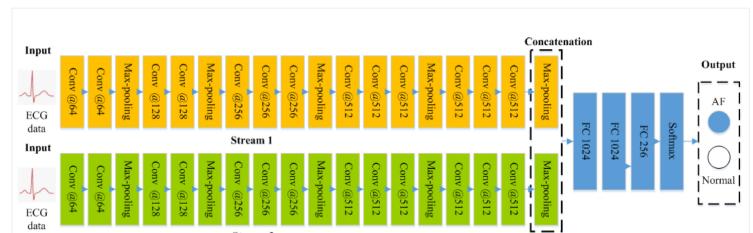


Figure 33: architecture of the MS-CNN

4. Results

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Consectetur adipiscing elit

Ut enim ad minim veniam

Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.

Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lore ipsum dolor sit amet, consectetur adipisic elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

5. Discussion

Lore ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lore ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lore ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lore ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

6. Conclusion

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

7. References

- 1) Ziqian Wu, Tianjie Lan, Cuiwei Yang, and Zhenning Nie, “A Novel Method to Detect Multiple Arrhythmias Based on Time-Frequency Analysis and Convolutional Neural Networks”, IEEE Access, vol. 7, pp. 170820-170830, December 2019
- 2) Dakun Lai, Xinshu Zhang, Yuxiang Bu, Ye Su, and Chang-sheng Ma, “An Automatic System for Real-Time Identifying Atrial Fibrillation by Using a Lightweight Convolutional Neural Network”, IEEE Access, vol. 7, pp. 130074-130084, December 2019
- 3) Xin-cheng Cao, Bin Yao, and Bin-qiang Chen, “Atrial Fibrillation Detection Using an Improved Multi-Scale Decomposition Enhanced Residual Convolutional Neural Network”, IEEE Access, vol. 7, pp. 89152-89161, December 2019
- 4) Xiaolong Zhai, and Chung Tin, “Automated ECG Classification Using Dual Heartbeat Coupling Based on Convolutional Neural Network”, IEEE Access, vol. 6, pp. 27465-27472, June 2018
- 5) Zhangjun Li, Xujian Feng, Ziqian Wu, Cuiwei Yang, Baodan Bai, and Qunqing Yang, “Classification of Atrial Fibrillation Recurrence Based on a Convolution Neural Network With SVM Architecture”, IEEE Access, vol. 7, pp. 77849-77856, December 2019
- 6) Bahareh Pourbabae, Mehrsan Javan Roshtkhari, and Khashayar Khorasani, “Deep Convolutional Neural Networks and Learning ECG Features for Screening Paroxysmal Atrial Fibrillation Patients”, IEEE Transactions On Systems, Man, And Cybernetics: Systems, vol. 48, n. 12, pp. 2095-2104, December 2018
- 7) Mohamad Mahmoud Al Rahhal, Yakoub Bazi, Haidar Almubarak, Naif Alajlan, and Mansour Al Zuair, “Dense Convolutional Networks With Focal Loss and Image Generation for Electrocardiogram Classification”, IEEE Access, vol. 7, pp. 182225-182237, December 2019
- 8) Jingshan Huang, Binqiang Chen, Bin Yao, and Wangpeng He, “ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network”, IEEE Access, vol. 7, pp. 92871-92880, July 2019
- 9) Weifang Sun, Nianyin Zeng, and Yuchao He, “Morphological Arrhythmia Automated Diagnosis Method Using Gray-Level Co-Occurrence Matrix Enhanced Convolutional Neural Network”, IEEE Access, vol. 7, pp. 67123-67129, June 2019
- 10) Xiaomao Fan, Qihang Yao, Yunpeng Cai, Fen Miao, Fangmin Sun, and Ye Li, “Multiscaled Fusion of Deep Convolutional Neural Networks for Screening Atrial Fibrillation From Single Lead Short ECG Recordings”, IEEE Journal Of Biomedical And Health Informatics, vol. 22, n. 6, pp. 1744-1753, November 2018
- 11) Vincenzo Piuri, “Intelligent Systems”, lecture notes, Università degli Studi di Milano, 2003
- 12) Nicolò Cesa-Bianchi, “Statistical Methods for Machine Learning”, lecture notes, Università degli Studi di Milano, 2019
- 13) Nicola Carbonaro, “Natural and Artificial Senses”, lecture notes, Università di Pisa, 2019

8. Web references

- a) Wikipedia contributors, “Electrocardiography”, Wikipedia, The Free Encyclopedia, 2020, <http://en.wikipedia.org>
- b) Wikipedia contributors, “Cardiovascular disease”, Wikipedia, The Free Encyclopedia, 2020, <http://en.wikipedia.org>
- c) Wikipedia contributors, “Arrhythmia”, Wikipedia, The Free Encyclopedia, 2020, <http://en.wikipedia.org>
- d) Wikipedia contributors, “Atrial fibrillation”, Wikipedia, The Free Encyclopedia, 2020, <http://en.wikipedia.org>

- e) Wikipedia contributors, “Pan–Tompkins algorithm”, Wikipedia, The Free Encyclopedia, 2020, <http://en.wikipedia.org>
- f) Wikipedia contributors, “Stochastic gradient descent”, Wikipedia, The Free Encyclopedia, 2020, <http://en.wikipedia.org>
- g) MIT Learning, “MIT Deep Learning 6.S191”, 2020, <http://introtodeeplearning.com>
- h) Towards Data Science, “Gentle Dive Into Math Behind Convolutional Neural Networks”, 2020, <http://towardsdatascience.com>
- i) Scikit-learn documentation, “Scikit-learn: machine learning in Python”, 2020, <http://scikit-learn.org>