Doctoral Dissertation
Doctoral Program in Metrology (34th Cycle)

Smart Health: Neural Models in Biomedical Applications

Annunziata Paviglianiti
* * * * * *

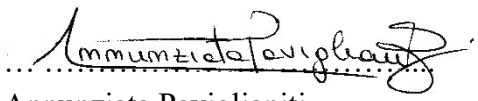
Supervisors
Prof. Eros Gian Alessandro Pasero Supervisor
Prof. Giansalvo Cirrincione, Prof. Alberto Vallan, Co-Supervisors

Doctoral Examination Committee:
Prof. Alessandro Vinciarelli, Referee, University of Glasgow
Prof. Salvatore Vitabile, Referee, Università degli Studi di Palermo
Prof. Sabrina Grassini, Politecnico di Torino
Prof. Claudio De Capua , Università degli Studi Mediterranea di Reggio Calabria
Prof. Francesco C. Morabito, Università degli Studi Mediterranea di Reggio Calabria

Politecnico di Torino
February 25, 2022

This thesis is licensed under a Creative Commons License, Attribution - Noncommercial - NoDerivative Works 4.0 International: see www.creativecommons.org. The text may be reproduced for non-commercial purposes, provided that credit is given to the original author.

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.



Annunziata Paviglianiti
Turin, February 25, 2022

Summary

Machine learning describes a class of algorithms that can combine inputs for prediction and classification purposes. Deep learning is a class of machine learning methods based on artificial neural networks with representation learning that uses multiple layers to progressively extract higher-level features from raw input. These algorithms, which represent subsets of the artificial intelligence (AI), have recently shown impressive results in a variety of domains, especially in medicine. Biomedical data are complex and often misunderstood. Hence, machine learning techniques can be particularly suited to solving these problems.
This thesis addresses some of these issues, by using both shallow and deep neural networks.

The first application deals with cardiovascular diseases, which represent the leading cause of deaths in the world. Arterial Blood Pressure (ABP) is a vital parameter that should be properly monitored for the purposes of prevention. The goal of this work is the continuous measurement of ABP through a non-intrusive approach. The approach is based on a neural network output-error and deep learning techniques to estimate ABP, starting from photoplethysmogram (PPG) and electrocardiogram (ECG) signals. The ABP was predicted first using PPG only and then using both PPG and ECG. The results show that the use of the ECG resulted in improved performance for each proposed configuration. The most performing configuration was obtained with a ResNet followed by three LSTM layers. It is also proven to be compliant with the ANSI/AAMI/ ISO 81060-2:2013 regulation for non-invasive ABP methods.

Another related problem is about the Short QT Syndrome (SQTS), an inherited cardiac ion channel disease linked with an increased risk of sudden cardiac death (SCD) in young and otherwise healthy individuals. Arrhythmic risk stratification is particularly challenging in asymptomatic subjects. AI-based electrocardiogram

(ECG) analysis has never been applied to SCD risk stratification in patients with cardiac channelopathies. The purpose of this study is twofold: on one hand, digitize the ECG paper-based in order to preserve the history and evolution of patients without the risk of the paper degrading (also facilitating the implementation of algorithms considering the vision of data as signals); on the other hand to analyse ECG features from SQTS patients with the aid of an AI system to evaluate its ability to discriminate between subjects with and without documented life-threatening arrhythmic events. The analysis of ECG features from SQTS patients with the aid of neural networks shows promising results in terms of discriminating between subjects with and without documented arrhythmic events (100 % negative predictive value). This could pave the way to a refined ECG-based risk stratification in this group of patients, potentially helping in saving the lives of young and otherwise healthy individuals.

This thesis also takes into considerations a completely different, but still very challenging medical problem: when nanomaterials are used in this field (bio-sensor, drug-delivery, etc.), their production through the electrospinning process requires careful inspection of the nano material, to ensure that no structural defects are created. The presence of anomalies prevents from the practical application of the electrospun nano-fibrous material in nanotechnology. A new classification system is proposed to distinguish homogeneous nanofibers (without anomalies) from non-homogeneous ones (with defects). Specifically, the image to be analysed are used as input for an unsupervised-supervised hybrid machine learning system. In the first stage, an automatic encoder (AE) is trained to generate code that represents the input image with a vector of relevant characteristics. Next, a multilayer perceptron (MLP) uses the extracted characteristics to classify images with non-homogeneous nanofibers (NH-NF) and with homogeneous nanofibers (H-NF). The resulting AE-MLP system has been shown to outperform other standard machine learning models, reporting a high rate of accuracy.

The last part of this work considers the current problem of Coronavirus disease, which is rapidly increasing, and contagions need to be kept under control to prevent their spread. Therefore, the development of algorithms for the diagnosis of COVID-19 is an open research area. Many studies have shown that Chest X-Ray images can be used for COVID-19 testing. In this work, a deep transfer learning technique is used to classify infected patients. Experimental results reveal that the proposed deep transfer learning-based COVID-19 classification model

provides efficient outcomes compared to other supervised learning models. In particular, the Vision Transformer achieved excellent performance with high test accuracy.

All the results obtained by the proposed neural models are better than the traditional approaches. This consideration justifies their use and paves the way to new possible applications in the medical domain.

Acknowledgment

I would like to acknowledge the professors who have made this PhD journey unrepeatable.
I would like to thank *Professor Eros Pasero*, my first boss, I wish everyone a boss like him: understanding, precise and stimulating. Entering at the Politecnico di Torino was not easy: new city, new job environment, new people. Thanks to Professor Pasero and my colleagues (who are now my dear friends), Politecnico has quickly become my home, and will always occupy part of my heart.
Thanks to *Professor Giansalvo Cirrincione* (Gians), my friend and confidant. I did not think that there were professors able to establish such a sincere relationship with their students. Gians is for me a model to be inspired by, a person with an immense baggage of knowledge who bases his life exclusively on love for science. He is a not superb mind always ready to help others. Thanks for the time spent reading my thesis, the nights writing papers and for the lunches offered, I will never be able to thank you enough.
Finally, thanks to *Professor Carlo Morabito*, my advisor in work and in life. Carlo was the first person to believe in my potential, the first person who introduced me to the world of Artificial Intelligence and the first person who supported my every decision. I hope I have made him proud of my (shared) achievement, it has always been one of the goals of my PhD. If today my work concerns what I love most is thanks to him. My acknowledgments will never be able to equate all the trust that he has responded in me, both as a person and as a professional.
Thanks to all those who rejoice at my successes as if they were their own.

To my beloved sister

For her emotional intelligence,
her shy courage and
her fragile strength.

Sommario

Introduction.....

1. Deep Learning Techniques for Arterial Blood Pressure Estimation 4

1.1 Cardiovascular Diseases (CVDs).....4

1.1.1 Cardiovascular System 5

1.1.2 Blood Pressure 7

1.1.3 Current Arterial Blood Pressure (ABP) Measurement Methodologies .10

1.2 Related Works.....12

1.3 Vital Parameters: Electrocardiogram and Photoplethysmogram 14

1.3.1 Electrocardiogram.....14

1.3.2 Photoplethysmogram 17

1.4 MIMIC Database Description.....19

1.4.1 Data Cleaning 19

1.4.2 Data Building.....21

1.5 Blood Pressure Estimation with Neural Networks Output-Error.....21

1.5.1 Neural Network Description.....22

1.5.2 Evaluation Metrics.....23

1.5.3 Results.....24

1.5.4 Future Works 26

1.6 A Comparison of Deep Learning Techniques for Arterial Blood Pressure 26

1.6.1 Methodology.....26

1.6.2 Convolutional Neural Networks 28

1.6.3 Recurrent Neural Network.....32

1.6.4 Results.....33

1.6.5 External Validation on Polito Dataset 37
1.6.6 Discussion 38
1.7 Conclusion 39
1.8 Acknowledgement 40
2. Development and Validation of a Shallow-Learning Model to Screen for Channelopathies from the Digitized Electrocardiogram with ECG-Dig GUI 41
2.1 ECG Digitalization: from ECG Paper to ECG Digital Signal 41
2.1.1 Related works 42
2.1.2 Paper Collection 42
2.1.3 ECG-dig Algorithm 44
2.1.4 Algorithm Validation Technique 48
2.1.5 Metrics 48
2.1.6 Results 49
2.1.7 ECG-dig Graphical Interface 52
2.2 Channelopathies: Short-QT 57
2.2.1 Short QT Definition 57
2.2.2 State of the Art 58
2.2.3 Shallow Learning for Event Prediction System 59
2.2.4 Study Populations 59
2.2.5 Dataset Description 60
2.2.6 AI Model Building and Internal Validation 62
2.2.7 Metrics 63
2.2.8 Results 64
2.2.9 Discussion 67
2.3 Conclusion 68
2.4 Acknowledgements 68
3. An Automatic Defect-detection System Based on Hybrid Unsupervised and Supervised Machine Learning for Electrospun Nanomaterials 70
3.1 Nanofibers in Industry 70
3.1.1 Biomedical Applications 71

3.2 Related Works 71
3.3 Electrospinning Process 73
3.3.1 Electrospinning Set-up 74
3.3.2 Dataset Construction 74
3.4 Filters Pre-processing 75
3.5 Hybrid Unsupervised-Supervised Machine Learning System 77
3.5.1 Networks Description 77
3.5.2 Metrics (ROC, Permutation Analysis) 80
3.6 Results 81
3.6.1 Permutation Analysis 83
3.7 Discussion 83
3.8 Conclusion 85
3.9 Acknowledgement 86
4. Classification of COVID-19 in Chest X-Ray Images Using Transfer Learning and Vision- Transformer (ViT) 87
4.1 Covid- 19 Pandemic 87
4.1.1 Covid-19 Clinical Stages 88
4.1.2 Vaccinations 89
4.1.3 Worldwide Covid-19 Pandemic Statistics 90
4.2 Related Works 91
4.3 Dataset Description: Chest X-Ray Images 92
4.4 Methodology 94
4.4.1 Transfer Learning 95
4.4.2 InceptionNet 95
4.4.3 XceptionNet 96
4.4.4 ResNet50 97
4.4.5 Vision Trasformer 97
4.5 Experimental Setup 98
4.6 Results 99
4.7 Conclusion 100

5. .....Conclusions 102
6. References.....105

List of Tables

Table 1.1 Hypertension classification suggested by the ESH based on blood pressure levels in adults aged 18 years or over. 10
Table 1.2. RMSE and MAE performances 25
Table 1.3. Errors (mmHg) for Direct SBP/DBP prediction task on SBP and DBP prediction for different setups with MIMIC databases 34
Table 1.4. Errors (mmHg) for Entire BP prediction task on SBP and DBP prediction for different setups with MIMIC databases 36
Table 1.5. LOO results on MIMIC Database with the best neural network (ResNet+LSTM). 37
Table 1.6. SBP and DBP prediction errors (mmHg) on Polito database using the best neural network (ResNet+LSTM) trained on MIMIC dataset (built using PPG and ECG lead I). 38
Table 1.7. Neural networks complexity order 39
Table 2.1. Dataset description.....44
Table 2.2. Pearson coefficient for each pathology.....49
Table 2.3. Variation of scale factor and fiducial points obtained by cropping 10 times the same ECG image (Normal sinus rhythm, 60 bpm). 51
Table 2.4 Comparison of scale factor and fiducial point for the same image in two formats 52
Table 2.5. Study population description 60
Table 2.6. Input dataset taxonomy.....65

Table 2.7 Classification performances: sensitivity, specificity, PPV, NPV, accuracy 67
Table 2.8. Classification performances: AUC 67
Table 3.1. Setup of the Experimental Electrospinning Parameters 75
Table 3.2. Classification Performance in Terms of Precision, Recall, F-score, and Accuracy of MLP with Different Hidden Layers (HL) and Hidden Units 82
Table 3.3. Classification Performance in Terms of Precision, Recall, F-score and Accuracy of the Proposed MLP (i.e.,), SVM and LDA Classifiers 82
Table 4.1 Covid-19 Classification Accuracy 99
Table 4.2 Classification indicators for each class 99

List of Figures

Figure 1.1 Heart anatomy 5
Figure 1.2 Circulatory/Vascular system 6
Figure 1.3 Cardiac cycle description 7
Figure 1.4 Arterial Blood Pressure (ABP) signal waveform: Diastolic Blood Pressure (DBP), Systolic Blood Pressure (SBP), Dicrotic Notch (DN) and PP (Pulse Pressure) 8
Figure 1.5 Invasive ABP monitoring 10
Figure 1.6 Auscultatory method 11
Figure 1.7 Electrocardiogram signal 14
Figure 1.8 Einthoven Triangle Representation, *bipolar leads* 16
Figure 1.9 Goldberger leads 17
Figure 1.10 Photoplethysmogram sensor 17
Figure 1.11 AC and DC components of blood volume variation 18
Figure 1.12 ECG, PPG and ABP signals 19
Figure 1.13 Data pre-processing pipeline 20
Figure 1.14 System identification procedure 22
Figure 1.15 The NNOE model structure 23
Figure 1.16 Waveform characteristic points 23
Figure 1.17 . The order index criterion evaluated with Lipschitz quotients method 25
Figure 1.18 NNOE - Target (blue solid line) and Output (red dashed line) signals with PPG (a) and ECG+PPG (b) as inputs. 26

Figure 1.19 Dataset distribution.....27
Figure 1.20 Receptive field of each neuron in different layers 28
Figure 1.21 Schematic convolutional neural network 28
Figure 1.22 Max pooling function example.....29
Figure 1.23 Skip connections.....30
Figure 1.24 Classic ReseNet Architecture 31
Figure 1.25 Dilated convolution layers.....31
Figure 1.26 LSTM memory block 32
Figure 1.27 ABP prediction on a validation set sample made with ResNet+LSTM trained with PPG dataset: original signal (light blue) vs network output (orange). 35
Figure 2.1. Data acquisition system.....43
Figure 2.2 Automatic algorithm pipeline.....45
Figure 2.3 ECG leads printed on graph paper 46
Figure 2.4 Squares in ECG graph paper 46
Figure 2.5 Algorithm validation scheme 48
Figure 2.6 Digital and scanner signal comparison for I, II, II, aVR, aVL, aVF leads 50
Figure 2.7 Digital and scanner signal comparison for precordial leads.....50
Figure 2.8 Reproducibility on normal person with 60 HB 52
Figure 2.9 Opening figure.....53
Figure 2.10 Figure selection 53
Figure 2.11 Opening ECG-image in ECG-dig tool 53
Figure 2.12 Lead I crop.....54
Figure 2.13 Displayed Lead I 54
Figure 2.14 Digitized signals 54
Figure 2.15 ECG folder 55
Figure 2.16 V5 lead digital signal.....55
Figure 2.17 Digital signal in .xls format.....56
Figure 2.18 Folder contents 56
Figure 2.19 Normal QT interval and Short QT interval 58
Figure 2.20 Shallow learning neural network scheme.....59
Figure 2.21 ECG signal fiducial points 61
Figure 2.22 Confusion matrix example: rows yield the real (actual) labels, columns the predicted ones, i.e., the network output.....64
Figure 2.23. Error over time for training, validation, and test subsets for the four input sets: QT (a), Twave (b), QT + Twave (c), All (d) 66
Figure 3.1 Electrospinning apparatus 74

Figure 3.2 Example of a SEM image sized 128 × 128 partitioned into four SEM nanopatches sized 64*64. In the reported example, all the sub-images belong to the homogeneous nanofiber (H-NF) class. 75
Figure 3.3 Architecture of the proposed hybrid unsupervised and supervised machine learning system. (a) Unsupervised processor composed of an AE [4096: 256: 4096]. The SEM nanopatch is reshaped into a single vector sized 1 × 4096 and used as input to the proposed AE that allows to extract the most relevant features (sized 1 × 256) from the input data. (b) Supervised processor composed of a MLP [256: 100: 80: 2]. The extracted features are the input to the proposed 2-hidden layers MLP for performing the 2-way classification task: NH-NF vs. H-NF. As an example, in the figure, a NH-NF SEM nanopatch inputs the hybrid unsupervised and supervised classification system. 80
Figure 3.4. ROC curves and corresponding AUC values of the proposed MLP, SVM and LDA classifiers for the NH-NF vs. H-NF classification. Note that the figure refers to the best MLP architecture 83
Figure 3.5. (a) Representation of the 256-dimensional features vector extracted by a SEM H-NF image (sized 64 × 64) via the proposed AE. (b) Examples of 10 reconstructed H-NF images. 84
Figure 3.6. (a) Representation of the 256-dimensional features vector extracted by a SEM NH-NF image (sized 64 × 64) via the proposed AE. (b) Examples of 10 reconstructed NH-NF images. 85
Figure 4.1. Covid-19 Disease Course 89
Figure 4.2 Population density affected by the pandemic worldwide....90
Figure 4.3 Coronavirus daily growth in the world....91
Figure 4.4. Normal lungs 93
Figure 4.5 Covid-19 pneumonia 93
Figure 4.6 Pulmonary opacity....94
Figure 4.7 Viral pneumonia 94
Figure 4.8 Covid Detection Methodology 94
Figure 4.9 Inception Module 95
Figure 4.10 InceptionV3 architecture 96
Figure 4.11 Xception Module....96
Figure 4.12 Xception architecture 97
Figure 4.13 ResNet50 architecture 97
Figure 4.14 Transformer Architecture 98
Figure 4.15 Vision Transformer Confusion Matrix....100

Introduction

Artificial intelligence (AI) is the ability of a computer system to simulate human cognitive functions, such as learning and problem solving. It uses mathematics and logic to simulate the reasoning adopted by people to learn from new information and make decisions. Artificial intelligence seeks to emulate planning, language understanding, object and sound recognition, learning and problem solving. Artificial Intelligence (AI) can be described as a system capable of simulating the functioning of the human brain and making decisions based on the analysis of the processed data. It thus becomes possible to make machines perform tasks such as speech and visual recognition, decision-making and predictive processes. Machine Learning (ML) is an application of artificial intelligence. It is the process constituted using mathematical models that allows a computer system to continue to learn and improve autonomously, based on experience (without direct instructions). Machine learning is essentially the ability of machines to receive a series of data and learn for themselves, changing algorithms as they receive more information about what they are processing. Machine learning automates the construction of the analytical model using statistical modelling methods and operational research to find hidden information in the data. Deep learning is one of the approaches to machine learning that took its cue from the structure of the brain, from the interconnection of various neurons (artificial neural networks). The artificial neural networks (ANNs) are based on direct acyclic graph (DAG) models consisting of a set of variables and conditional dependencies. Deep neural networks (DNN) perform very well complex data and they can be easily updated with new data using batch propagation. DNN architectures (i.e., number and structure of layers) can be adapted to many kinds of problems, and their hidden layers reduce the need for feature engineering.
The main advantages for the use of deep learning are:

- Automation of feature generation: DL algorithms can generate new functionality from a limited number located in the training dataset without further human intervention. This means that deep learning can perform complex tasks that often require extensive feature design.
- Working with unstructured data: DL algorithms are limited in their ability to analyse unstructured data, which means that this wealth of information

is often not exploited. In this sense, deep learning promises to have the greatest impact.

- Improved Self-Learning Skills: multiple layers in DNN allow models to become more efficient at learning complex functions and performing more intensive computational tasks in parallel. It can check the accuracy of his predictions / outputs and make any necessary changes. On the other hand, classical ML models require varying degrees of human intervention to determine the accuracy of the output.
- Advanced analysis: when DL algorithms are applied to data science, can offer better and more effective computing models. Its ability to learn without supervision drives the continuous improvement of accuracy and results. It also offers data scientists more reliable and concise analysis results.
- Scalability: DL is highly scalable due to its ability to process huge amounts of data (Big Data) and perform many calculations in a cost and time efficient manner. This improves efficiency by automatically scaling the number of nodes in use based on request traffic.

Artificial intelligence in healthcare is a general term used to describe the use of machine learning algorithms for presenting and understanding complex medical and health data. AI represents the automation of tasks to support the decision-making process of doctors, giving them more time for human interactions, a fundamental part in the medical field [1][2]. Artificial intelligence algorithms address automating arduous tasks and can sometimes outperform humans in tasks due to the experience made available by large amounts of clinical data that are collected by many clinical institutes to improve learning.

To generate an effective AI algorithm, computer systems are first fed with data that is typically structured, meaning that each data point has a recognizable label or annotation for the algorithm. After the algorithm has been exposed to enough data points and their labels, performance is analysed to ensure accuracy. This is done by entering test data for which programmers already know the answers, allowing them to evaluate the algorithms' ability to determine the correct answer. Based on the test results, the algorithm can be modified, fed with more data, or implemented to help make decisions for the person who wrote the algorithm. However, although some algorithms may compete with and sometimes outperform doctors in a variety of tasks, they have yet to be fully integrated into daily medical practice.

So far, algorithms in medicine have shown many potential benefits for both doctors and patients. However, adjusting these algorithms is a difficult task. The United States Food and Drug Administration (FDA) has approved some assist algorithms, but there are currently no universal guidelines for approval. FDA has strict acceptance criteria for clinical trials, which require extreme transparency on scientific methods. Many algorithms rely on very intricate and difficult to decode mathematics to get from the input data to the result. This lack of interpretability means the results must be accepted in itself. For this reason, these algorithms are

considered as black boxes. Failure to clarify the inner workings of an algorithm would impact the likelihood of the FDA approving an AI-based trial [3].
The first chapter concerns the use of deep learning algorithms to predict the arterial blood pressure signal. Given two physiological input parameters to the system, namely the electrocardiogram and the photoplethysmogram signals, the neural algorithm estimates the systolic and diastolic pressure values.
The second chapter is divided into two sub-projects: firstly, an automatic algorithm is implemented to transform the electrocardiogram signal on paper into a digital signal, secondly the features related to the ECG signal are used to estimate the probability of a cardiac event using shallow learning algorithms.
The third chapter, which mainly concerns the industrial sector, involves nanomaterials used for medical applications such as sensors, drug delivery and bio-compatible fabrics. In this case an advanced hybrid supervising, and unsupervised algorithm was applied to classify the quality of the material.
The fourth chapter relates to the implementation of very recent advanced neural networks, through the use of transfer learning, with the aim of diagnosing the coronavirus disease through X-Ray images.
The last chapter yields the conclusions and proposes new paths for future work.

Chapter 1

Deep Learning Techniques for Arterial Blood Pressure Estimation

1.1 Cardiovascular Diseases (CVDs)

Cardiovascular diseases (CVD) are disorders that involve the heart and the cardiovascular system. They are the first cause of death in the world, taking an estimated 17.9 million lives each year: 80% of CVD deaths are due to heart attacks and strokes while the remainder of these deaths occur prematurely in people under the age of 70 [1]. Diseases involving the heart and blood vessels are due to genetic and behavioural risk factors. On one hand, in recent years, cardiology research has focused attention on identifying and understanding the genetic basis of cardiovascular diseases. Because of the defects of a single gene, heart disorders involving all parts of the organ structure have been identified. A feature common to nearly all genetic cardiovascular diseases is the clinical heterogeneity observed in affected individuals within a household. Despite carrying the same genetic mutation, affected individuals can often exhibit marked clinical variability, ranging from symptom-free to premature death [2]. To this aspect some considerations are added concerning the gender of individuals suffering from cardiovascular diseases. Women develop disorders about 10-15 years later than men [3]. Today, the answer to this difference is focused on the hypothesis that endogenous estrogen is cardioprotective in women [4]. Therefore, the increase in coronary heart disease (CHD) rates after menopause and after ovariectomy confirms that endogenous estrogens can prevent CHD [5]. On the other hand, the behavioural causes that cause heart disease and stroke are an unhealthy diet, physical inactivity, tobacco use, and the harmful use of alcohol. The effects of behavioural risk factors can manifest in individuals as increased blood pressure (hypertension), increased blood sugar (diabetes mellitus),

increased blood lipids (dyslipidaemia), overweight and obesity. These risk factors strongly influence cardiovascular risk; consequently, the treatment and monitoring of these factors are the main steps in hospital care [6]. Drug therapy (such as aspirin, beta-blocker, diuretic and statins) is particularly effective because it can lead to a 75% reduction in myocardial infarction (heart attack).

1.1.1 Cardiovascular System

The heart can be considered one of the most important organs in human body, and even small anomalies in its functioning can lead to serious dysfunctions of different compartments of the body. It is located under the breastbone, in the centre of the chest and its main function is to transport oxygen and nutrients to all organs and tissues in the body. The heart is a muscle composed of involuntary striated muscle tissue but, unlike other muscles, it can generate the electrical signal that allows its movement independently. It is made up of several layers as shown in Figure 1.1 Heart anatomyFigure 1.1:

- pericardium, fluid-filled double-walled structure that provides mechanical protection;
- epicardium, membrane that completely covers the external surface of the heart, making it translucent and smooth;
- myocardium, striated intermediate muscle tissue;
- endocardium, the inner wall.

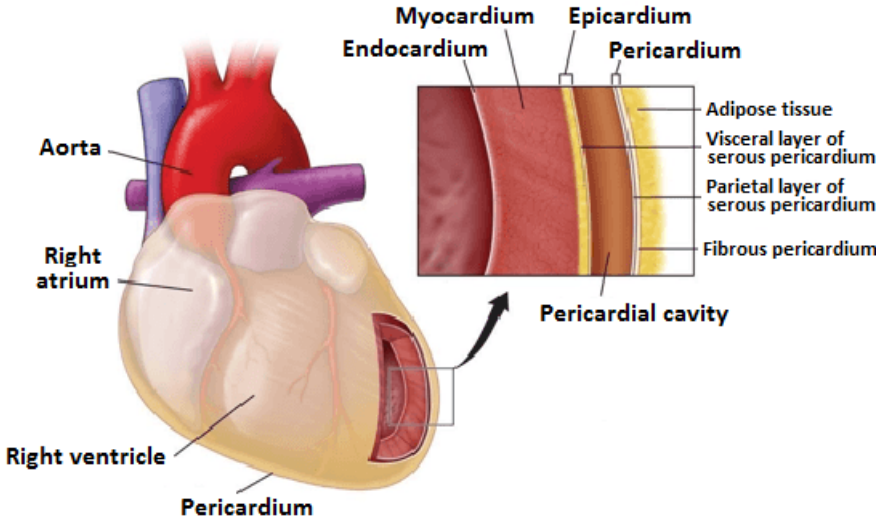


Figure 1.1 Heart anatomy

The heart is a hollow organ consisting of two atria and two ventricles. In addition, there are four valves: two atrioventricular valves (tricuspid and mitral valves) that allow the unidirectionality of blood between the atria and ventricles, and two semilunar valves (pulmonary valve and aortic valve) that control the flow between the heart and blood vessels [7].

The atria are cavities located in the upper part of the heart, they are not symmetrical but differ in position and size; they are separated from the atrial septum and communicate with the ventricles through two valves. The ventricles represent the main part of the organ and are located in the lower part; they are separated by the interventricular septum. In the right atrium, the venous blood comes rich in carbon dioxide, arriving from cells and tissues; through the tricuspid valve the blood passes from the right atrium to the right ventricle and is then pushed into the pulmonary artery. Inside the right atrium there is also the sinoatrial node which generates the impulses for the regulation of the heartbeat. The oxygen-rich blood from the four pulmonary veins arrives in the left atrium, passes to the left ventricle through the mitral valve and is finally pushed into the aorta, to be brought back to the cells and tissues. The left ventricle is characterized by thicker muscle walls than the right one.
The blood circulation can therefore be divided into systemic and pulmonary circulation (Figure 1.2):
• The systemic circulation (or large circulation) transports oxygenated and nutrient-rich blood to the body's cells, and then it to the right side of the heart once the exchanges have taken place.
• The pulmonary circulation (or small circle) carries the blood rich in carbon dioxide to the lungs and then returns it rich in oxygen to the left side of the heart, to be reintroduced into the systemic circulation [8].

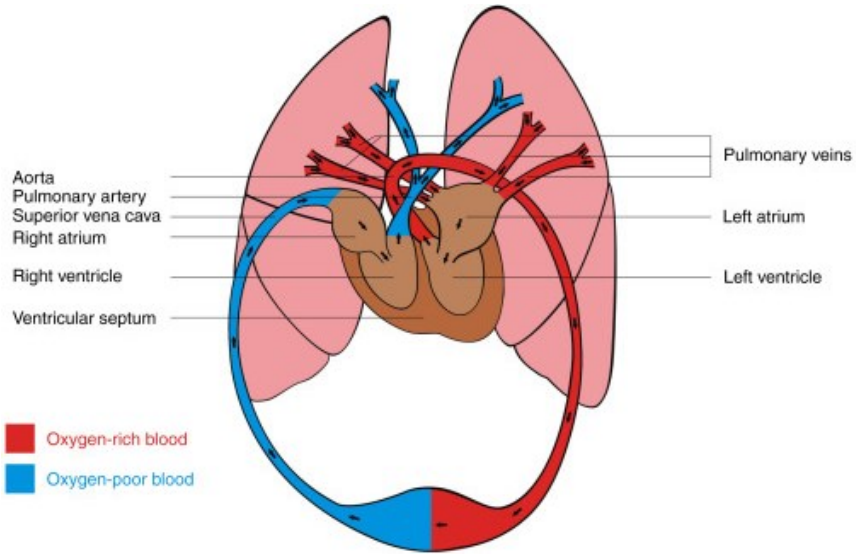


Figure 1.2 Circulatory/Vascular system

To allow the four chambers to contract in the correct way and with the right timing, a conduction system is needed that generates the impulses and makes them propagate appropriately among the cells of the myocardium. The action potential (excitatory signal) is generated within the sinoatrial node, located in the right atrium at the outlet of the superior vena cava. The heart rate can be modified, to adapt to the needs of the organism, through the intervention of the autonomic nervous system: the sympathetic stimulus (adrenergic) intervenes for example during physical exercise and allows to increase the speed of contraction and heart

rate, parasympathetic (or vagal) stimulation decreases the number of beats and intervenes in situations of rest or during sleep. The action potential propagates through the internodal tracts, causing the contraction of the atria, to reach the atrio-ventricular node.

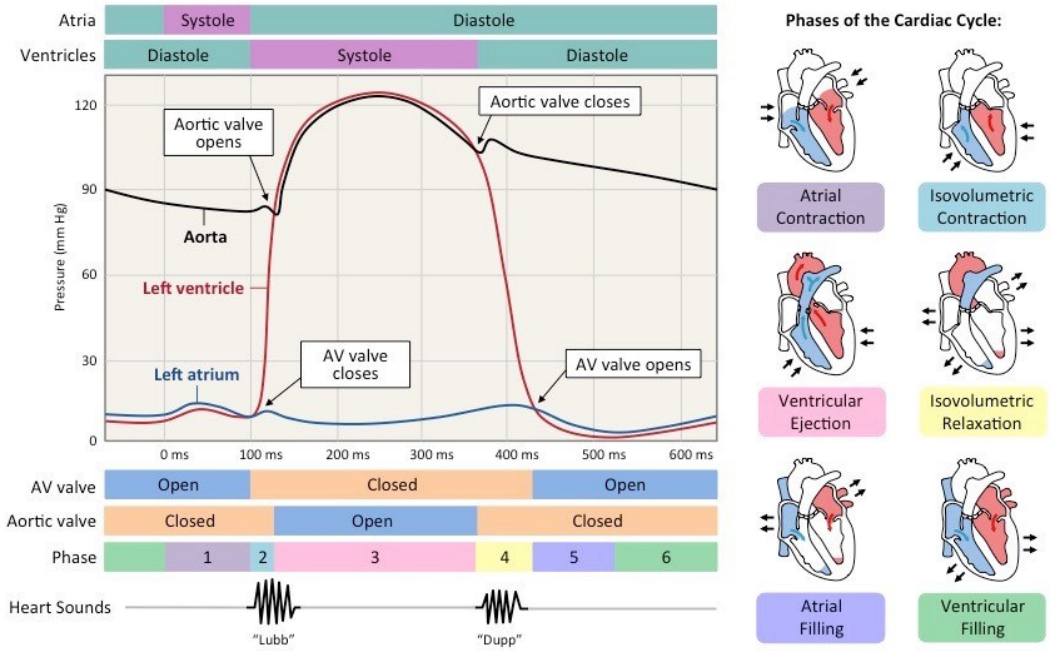


Figure 1.3 Cardiac cycle description

The cardiac cycle (Figure 1.3) is a set of coordinated movements divided into two main phases, which are repeated on average 70-80 times per minute in resting conditions: there is a contraction phase, called systole, and a relaxation phase, called diastole. The closing and opening of the valves that flow retrograde flow is controlled by the pressure that the blood exerts on these structures.

During diastole, blood flows from the superior and inferior vena cava into the right atrium. While the left atrium fills with oxygenated blood from the pulmonary veins, increasing the pressure. The opening of the mitral and tricuspid valves and subsequent filling of the ventricles occurs when the pressure in the atria exceeds that of the ventricles.

During systole, following the electrical impulse originating from the SA node, the atria contract, pushing blood into the ventricles. When the electrical impulse reaches the ventricles, these begin to contract (isometric contraction) but the force is not sufficient to open the aortic and pulmonary valves. Only when the pressure in the ventricles exceeds that of the aortic arch, the valves open allowing blood to circulate [9].

1.1.2 Blood Pressure

Arterial blood pressure ABP is one of the so-called vital signs and is accepted as an index of the circulatory condition.

Blood pressure can be measured in both arteries and veins, and, for this reason, it is called venous blood pressure (VBP) or arterial blood pressure (ABP). Since the veins do not receive the blood pushed by the heart, the venous pressure is always lower than the arterial pressure, in conditions of rest, and for this reason it is not considered as a descriptive parameter of the individual's state of health [10].
Blood pressure is one of the fundamental parameters for identifying several diseases of the circulatory system.
As shown in Figure 1.4, the waveform of the ABP signal can be considered a periodic signal in which peaks and valleys alternate, due to the different phases of the cardiac cycle, the greater the blood volume present in the aorta, the higher the blood pressure. Following the contraction of the ventricles, the systolic peak occurs and the blood is pushed into the aorta, the peak value represents the systolic pressure (also called maximum) which is normally around 120 mmHg.

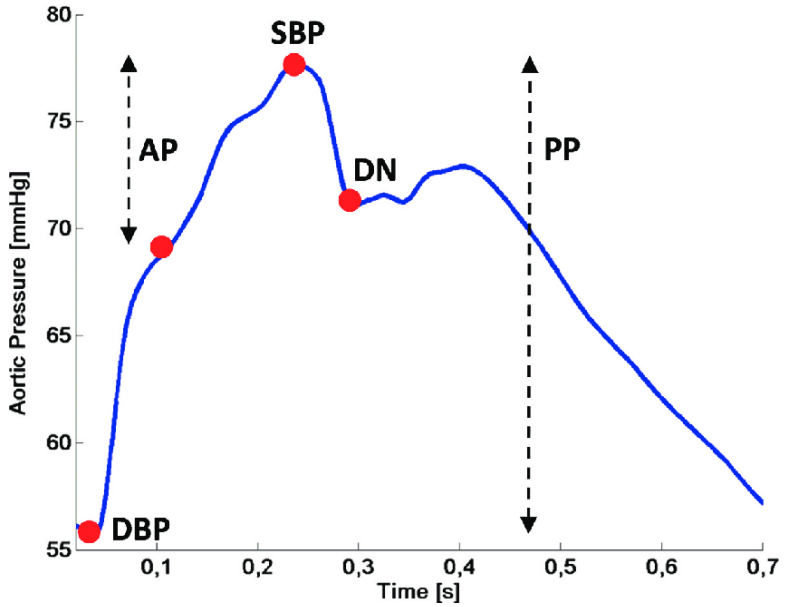


Figure 1.4 Arterial Blood Pressure (ABP) signal waveform: Diastolic Blood Pressure (DBP), Systolic Blood Pressure (SBP), Dicrotic Notch (DN) and PP (Pulse Pressure)

Then, there is a descending stroke since the ventricular ejection decreases and the blood moves towards the periphery, up to a depression (dicrotic notch) due to a small part of blood returning to the ventricle. Finally, it arrives at the lowest point which represents the diastolic (or minimum) pressure, usually around 80 mmHg [11] . The blood pressure values considered normal in the adult population are around 120/80 mmHg. If the blood pressure is lower or higher than the normal values, there incur diseases called hypotension and hypertension respectively.

Hypotension (Low Blood Pressure)

Hypotension is the condition in which the pressure is below normal values; a subject is considered hypotensive if his systolic pressure is less than or equal to 90 mmHg and his diastolic pressure is less than or equal to 60 mmHg. A sudden drop in blood pressure can cause a decrease in the blood supply to the brain and other organs with the possibility of fainting. In severe cases, hypotension may be due to

bleeding, heart failure, anaphylactic shock, severe infections, while in the less severe ones we may have orthostatic hypotension due to sudden changes in position. Younger people most commonly suffer from neuro-mediated hypotension, which may appear after long periods of standing. Hypotension usually occurs with the onset of dizziness, drowsiness, blurred vision, nausea, and confusion. It is important to treat hypotension to prevent organs from suffering due to a poor supply of blood and nutrients, and to avoid falls due to fainting [12].

Hypertension (High Blood Pressure)

Hypertension represents the condition in which blood pressure is higher than normal; it can cause damage related to cardiovascular diseases such as stroke or myocardial infarction. If both blood pressure values (systolic and diastolic) are higher than the norm we speak of systolic-diastolic hypertension, otherwise it is systolic or diastolic hypertension. Due to aging and increased stiffness of the vessels, the elderly suffer more often from isolated systolic hypertension, while diastolic hypertension is more common in young people [13]. 95% of hypertensive subjects have primary (or essential) arterial hypertension, which has no precise cause; it is thought to be due to imbalances in complex regulatory mechanisms, such as the autonomic nervous system, or to the presence of particular substances in the circulation.

The remaining 5% have secondary hypertension, caused by congenital diseases of the kidneys, heart, adrenals or vessels; this type of hypertension can also affect young people. One of the main problems of hypertension, also due to its high danger, is the fact that it often occurs asymptomatically, especially if it does not happen suddenly. In fact, the body gets used to high pressure values and does not give signs to the subject. In the case of the presence of symptoms these, being very common, are often associated with other causes; the main symptoms, in fact, are headache, tinnitus, nosebleed, altered vision and dizziness. It is therefore of fundamental importance to carry out periodic and, in some cases continuous, blood pressure checks to prevent this over time from creating serious damage to the circulatory system [14].

According to the consensus of the experts, the cardiovascular risk increases to the point of justifying a therapeutic intervention, including pharmacological, in the presence of blood pressure values equal to or greater than 140 mm Hg as regards the systolic pressure and/or equal to or greater than 90 mm Hg as regards diastolic blood pressure. Table 1.1 shows 2018 guidelines of the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH). Optimal systolic pressure is less than 120 mmHg and a diastolic pressure less than 80 mmHg, above 140 mmHg maximum and / or 90 mmHg minimum the subject is hypertensive, while it is isolated systolic hypertension when only the maximum is high (i.e. ≥140 mmHg) [15] .

Table 1.1 Hypertension classification suggested by the ESH based on blood pressure levels in adults aged 18 years or over.

	Systolic blood pressure (mmHg)	Diastolic blood pressure (mmHg)
Optimal	<120	<80
Normal	120-129	80-84
Normal-High	130-139	85-89
Grade 1 hypertension	140-159	90-99
Grade 2 hypertension	160-179	100-109
Grade 3 hypertension	≥ 180	≥ 110
Isolated systolic hypertension	≥ 140	≤ 90

1.1.3 Current Arterial Blood Pressure (ABP) Measurement Methodologies

Blood pressure in the arteries can be measured both directly (invasive) and indirectly (non-invasive). In the first case, a tube (catheter) is introduced into the artery and connected to a specialized measuring device (pressure transducer). This happens in particular circumstances, such as during surgery.
On the other hand, the common measurement of blood pressure is carried out indirectly, using special devices that can assess blood pressure from the outside, in a non-bloody way [16].

Invasive Method

The invasive monitoring is the gold standard for direct measurement of blood pressure (IBP) in patients in intensive care. It is an indicator to estimate the effect of drugs, blood gas values in patients supported by ventilation, to monitor patients at risk of a potential sudden situation or if, due to skin lesions, it is not possible to use other methods. The measurement takes place through the cannulation of an artery using an arterial cannula; everything is connected to a sensor and to a transduction system that outputs systolic (SBP) and diastolic (DBP) pressure on the monitor.

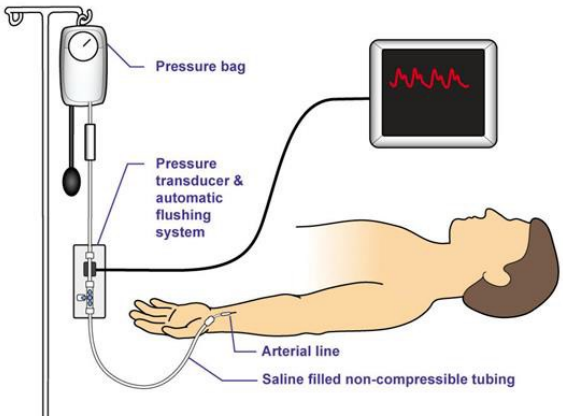


Figure 1.5 Invasive ABP monitoring

Non- invasive method

Among the non-invasive techniques the most used are the auscultatory method (generally identified as a sphygmomanometer), the oscillometric method and the no-load method.

Auscultatory method

In 1856 blood pressure was first recorded in humans using a U-shaped device, a manometric tube connected to a brass cannula connected directly to the artery.
However, there have been efforts since then to obtain a non-invasive BP measurement. The first instrument to accurately measure BP was the sphygmomanometer by Scipione Riva-Rocci in 1896 [17].
A sphygmomanometer consists of an inflatable cuff to collapse and then release the artery under the cuff in a controlled way, and is based on the auscultatory method. The auscultatory method consists of listening the Korotkoff sounds in the brachial artery (see Figure 1.6). When a cuff is inflated to a level above systolic pressure, the brachial artery is occluded. When the artery is fully compressed, there is no blood flow and no sounds are heard. Then the cuff is gradually deflated and blood flow is restored and Korotkoff sounds are first heard with a stethoscope under the cuff. When the sound is heard for the first time the pressure value is significant. Korotkoff's sounds will continue to be heard and the pressure drops further. However, when the cuff pressure reaches DBP, the sounds disappear. This method has always been the gold standard for clinical blood pressure measurement. It is performed by a qualified healthcare professional. Blood pressure values are obtained from an aneroid device.
There are many variables that affect the accuracy of this method, such as the size of the bracelet, and numerous studies have shown that doctors and healthcare professionals rarely follow up established guidelines for making correct manual blood pressure measurements [18]. When the cuff size is too small, the sphygmomanometer will emit more pressure, while when the cuff is too large it will produce lower values.

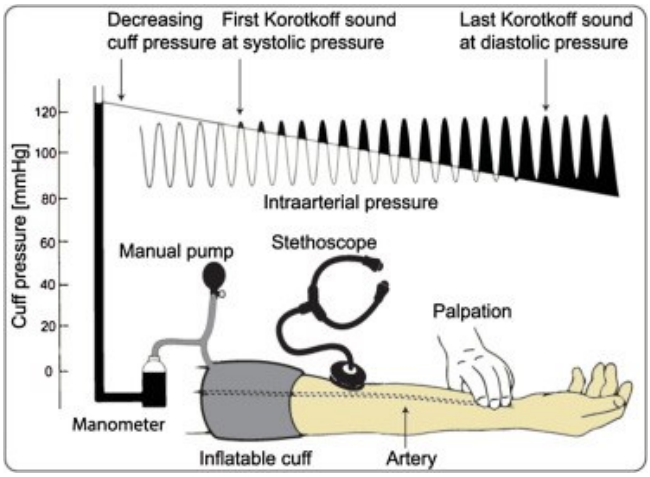


Figure 1.6 Auscultatory method

Oscillometric method

The oscillometric method was first introduced in 1876 and involves the observation of oscillations in the sphygmomanometer cuff pressure which are caused by the oscillations of blood flow.

This method uses a sphygmomanometer cuff, like the auscultatory method, but with an electronic pressure sensor (transducer) to observe cuff pressure oscillations[19]. It employs either deformable membranes that are measured using differential capacitance, or differential piezo resistance, and includes a microprocessor to automatically interpret the sensor results. The pressure sensor should be calibrated periodically to maintain accuracy. Initially the cuff is inflated to a pressure higher than SBP and then it is reduced to below DBP over a period of about 30 seconds. The sphygmomanometer records a constant pressure when the blood flow is blocked or unimpeded, while when the blood flow is present, but restricted the recorded pressure will vary periodically in synchrony with the cyclic expansion and contraction of the brachial artery.

This method makes it possible to create devices that can be used by the entire population, even non-experts, but has some limitations due to the accuracy of the algorithm and some factors that influence the measurement. In particular, the movements of the arm create unwanted vibrations, while the heart rhythm, for example in the case of arrhythmia, does not allow a correct measurement of the pressure.

Unloaded Method

The unloaded method was first developed by Penaz and works on the principle of the unloaded arterial wall and allows continuous BP measurements [20].

It is performed through a pulse oximeter, in general a photoplethysmography (PPG or PLETH). Indeed, it can measure finger blood volume changes using light. However, transforming volume changes into pressure is not easy because of the non-linearity of the elastic components in the finger (arterial walls and muscles).

The PPG output is used to drive a loop, which rapidly changes the cuff pressure to keep blood volume constant, so that the artery is held in a partially opened state. This method gives an estimate of the changes of systolic and diastolic pressure, although both may be underestimated when compared with brachial artery pressures.

1.2 Related Works

Two clinical gold standards exist to measure arterial blood pressure: the invasive and direct catheter system and the cuff-based methods. The former is very accurate but could cause pain and infection, the second is not invasive but the measure is subject to inaccuracy [21][22]. Continuous non-invasive arterial pressure (CNAP) measurement combines the advantages of the two approaches. For this reason, several non-invasive methods have been studied. One of the first approaches concerns the Pulse Wave Velocity (PWV) propagation estimates

blood pressure values by using the mathematical description by Moens and Korteweg [23]. In [24] it is shown the inverse proportionality between the blood pressure value and the PWV. However, in this case, the proposed mechanical-mathematical model uses patient physiological parameters that are difficult to detect, such as the artery diameter or the distance from heart to fingertip. Another useful parameter to detect the arterial blood pressure is the Pulse Transit Time (PTT), defined as the time the pulse wave takes to travel between two arterial sites within the same cardiac cycle and is another attribute for blood pressure estimation process [25]. The model proposed by [26] overcomes the problem of the availability of the patient's physiological parameters; however, the mathematical relation between PTT and BP is subject to approximations and non-linearity, which make the model not very general and robust. To overcome the non-linearity of the problem Neural Networks (NN) appear to be an ideal framework, for system identification. Neural networks are conceptually simple, easy to train and use and can approximate a target function in an excellent way; however, the biggest criticism is that the models produced are completely opaque (black box): it is therefore very difficult analyse the model and compute dynamic characteristics from the model. PPG has emerged as a potentially useful signal to measure arterial blood pressure [27]; indeed, many studies in the literature confirm a clear relationship between PPG and ABP. Since PPG and ECG can be easily integrated into wearable devices [28][29][30].

In [31], it is demonstrated neural networks can perform better than linear regression. They extracted a set of features from PPG recordings, taken out from MIMIC database. The achieved 3.80 ± 3.41 mmHg on SBP and 2.21 ± 2.09 mmHg on DBP on a very small dataset dataset (15000 heartbeats are analysed, which means roughly 4 hours of recordings). A complex recurrent neural network (RNN) has been employed in [32], which uses 22 ECG and PPG features extracted from MIMIC II [33]; the RMSE was 3.63 on SBP and 1.48 on DBP, which is an excellent result. Nowadays, thanks to advances in deep neural techniques, new approaches using raw vital signals are considered. In [34], ABP is measured by using only four features of the PPG signal in three different configurations: rest, exercise and recovery. This was one of the first studies based only on PPG and showed a good correlation between ABP and some PPG features: the results were good, but only a few healthy people were included in the cohort. Paper [35] had a huge impact on current research: the pre-processing approach is suitable for removing noise signals in MIMIC and the validation system that was adopted is the most robust applied to regression methods for estimating the ABP. However, the goal was limited to measure the ABP starting from PPG and, therefore, the ECG was not considered for better performance. This was one of the first approaches based on deep learning that analyses both temporal and spectral characteristics.

In this chapter, blood pressure has been estimated by using a typical regression approach with two configurations: the first adopts only the PPG signal as input, while the second employs both ECG and PPG in combination as input for the neural networks, whose output is the ABP.

1.3 Vital Parameters: Electrocardiogram and Photoplethysmogram

In recent years, to estimate blood pressure in a non-invasive way, a set of physiological signals taken from patients are used. Electrocardiogram (ECG) and photoplethysmogram signals are high correlated with arterial blood pressure (ABP).

1.3.1 Electrocardiogram

An electrocardiogram (ECG) is a graph which shows the electrical activity of the heart; it is obtained using electrodes placed on the skin. Rhythmic cardiac activity is based on repetitive depolarization and repolarization of the entire heart [36].
An ECG is an indirect indicator of heart muscle contraction, and is measured on the skin surface using electrodes. Through the analysis of the trace, it is possible to identify anomalies in the conduction of muscle fibers or other problems related to the heart. The examination at rest is carried out by applying electrodes in a variable number to the patient's skin, lying on a bed.

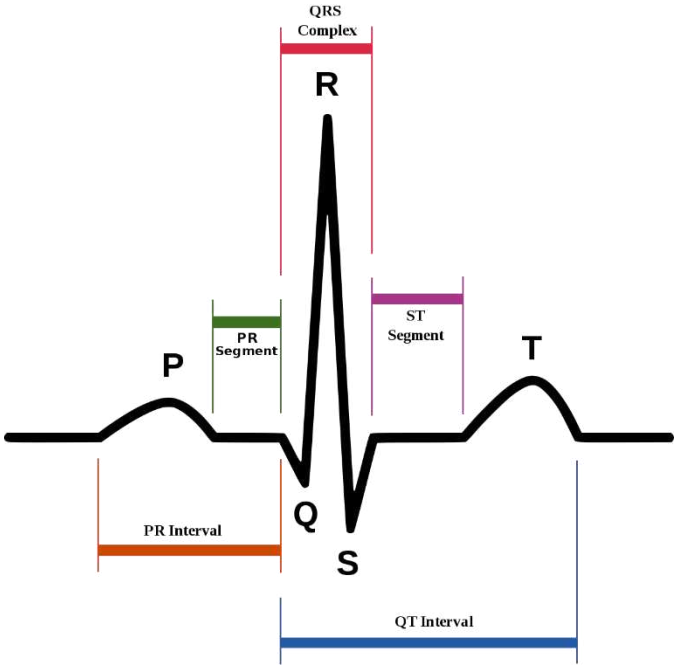


Figure 1.7 Electrocardiogram signal

Figure 1.7 shows the electrocardiogram signal, which usually has an amplitude that reaches up to 1-3 mV and is included in the 0.05-200 Hz band (it can also reach 1 kHz). The signal consists of intervals that represent the time between two cardiac events. Each segment corresponds to a function that the heart performs during the entire cardiac cycle [37]:

- The *P wave* begins when the sino-atrial (SA) node is activated and represents the depolarization of the atria.
- The *PR segment* is located on the baseline between the end of the P wave and the beginning of the QRS complex; it represents the transmission of the electrical impulse through the AV node to the Purkinje fibers.
  The *PR interval* represents the time interval between the beginning of the P wave and the beginning of the QRS complex; it contains all the electrical events ranging from the activation of the SA node to the depolarization of the ventricles.
- The *QRS complex* is composed of three different waves (Q, R and S) which, in some cases, may not all be present; it represents the depolarization of the ventricles. In conjunction with this event, the repolarization of the atria (wave TP) also occurs but it is not visible as the QRS complex prevails.
- The *ST segment* is located between the end of the QRS complex and the beginning of the T wave; it is an electrically neutral period for the heart, between the depolarization and repolarization of the ventricles. It is the time in which the heart remains contracted in order to be able to expel most of the oxygenated blood.
- The *T wave* represents the repolarization of the ventricles.
- The *QT interval* begins with the QRS complex and ends with the end of the T wave, it can have a variable duration depending on age, gender, and heart rate; it represents all events of ventricular systole.

Leads

The distance between two electrodes is called *lead*, which is associated with an electric vector that changes by following the electric field generated by the heart during the cardiac cycle.
The standard ECG is performed using 12 leads, which observe the same phenomenon from different points of view [38]. The Einthoven Triangle, shown in Figure 1.8, is the physiological principle on which the detection of the electrical activity of the heart is based. Einthoven's Triangle is based on the imaginary arrangement of an inverted equilateral triangle on the patient's chest, whose centre coincides with the heart. Each corner of the geometric figure is electrically coincident with a point of a specific limb which is assigned a name: VL (left arm) VR (right arm) and VF (left foot). The remaining limb, i.e., the right foot, is defined as neutral and does not participate in the formation of the triangle. Each of these points electrically looks at the heart from its own point of view, VL from the left, VR from the right and VF from below, but it is in the reciprocal vision of two points at a time that the cardiac bipole (a positive pole and a negative pole) manages to record the electrical events that unfold from the heart placed in the centre.

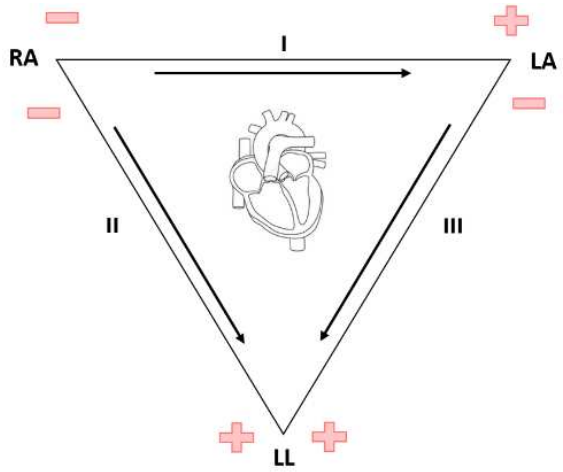


Figure 1.8 Einthoven Triangle Representation, *bipolar leads*

From the potential differences between two peripheral electrodes at a time, three *bipolar leads* are obtained:

- Lead I or D1: measured between the positive electrode on the left arm and the negative one on the right arm.
- Lead II or D2: measured between the positive electrode on the left leg and the negative one on the right arm.
- Lead III or D3: measured between the positive electrode on the left leg and the negative one on the left arm.

$$D1 = LA - RA \quad (1)$$
$$D2 = LL - RA \quad (2)$$
$$D3 = LL - LA \quad (3)$$

The sum of the voltages around any closed path around the triangle is equal to zero. Consequently, a virtual ground point can be derived from limb leads: the Wilson's central terminal, which is defined as the average of the vertices of the triangle.

$$Vw = \frac{1}{3}(LA + RA + LL) \quad (4)$$

The *unipolar leads* of the limbs (or Goldberger leads) allow to carry out an absolute measurement, but it is necessary to have a reference point (central terminal) with zero potential, Figure 1.9. This is obtained by joining the three electrodes at the ends (RA, LA, LL) through resistances, typically 5 kΩ. By joining each positive electrode to the indifferent electrode, the lead axes are obtained. The Goldberger leads (aVL, aVR and aVF) are also named increased as they have an amplitude increased by 50%, through a system incorporated in the electrocardiograph.

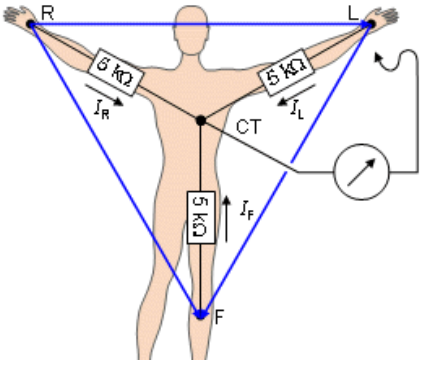


Figure 1.9 Goldberger leads

Unipolar precordial leads, placed on the patient's chest surround the heart and further increase the points of view of observation. In this way, it is possible to observe the electrical vector on a transverse plane, in a position close to the heart, with the possibility of identifying some anomalies not visible through the other leads. There are six leads (V1, V2, V3, V4, V5, V6) of the bipolar type, obtained through an indifferent electrode (zero potential). To position all the electrodes correctly, it is necessary to identify the landmarks between the intercostal spaces.

1.3.2 Photoplethysmogram

The photoplethysmogram is a method for measuring changes in blood volume in tissue microvascular. PPG, introduced in 1937 by Alrick Hertzman, is often obtained by using a pulse oximeter which illuminates the skin and measures changes in light absorption [39] . Pulse oximeter is one of the most popular wearable devices, it is mainly used to determine heart and respiratory rates, however, PPG signal is almost only displayed ICU applications [40]. As shown in Figure 1.10 the operation of the photoplethysmograph consists of a sensor consisting of an LED (Light Emitting Diode) applied to the subject's skin, which emits a beam of light at a certain intensity and wavelength. Part of the light, passing through the tissues, is absorbed and the remaining one reaches a photodetector, capable of translating the attenuation into a proportional electrical signal.

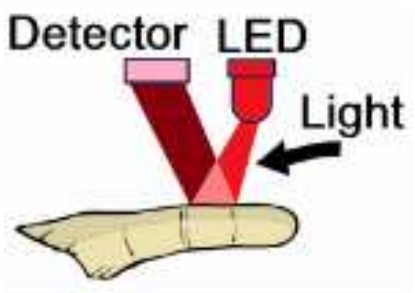


Figure 1.10 Photoplethysmogram sensor

The blood, by varying its volume inside the vessels with each heartbeat, produces a variable attenuation over time; on the contrary, the more superficial and less

vascularized tissues produce a practically constant attenuation. This results in an alternating component (AC) due to the variation of the blood inside the vessels, and a continuous component (DC) due to tissue absorption Figure 1.11[41].

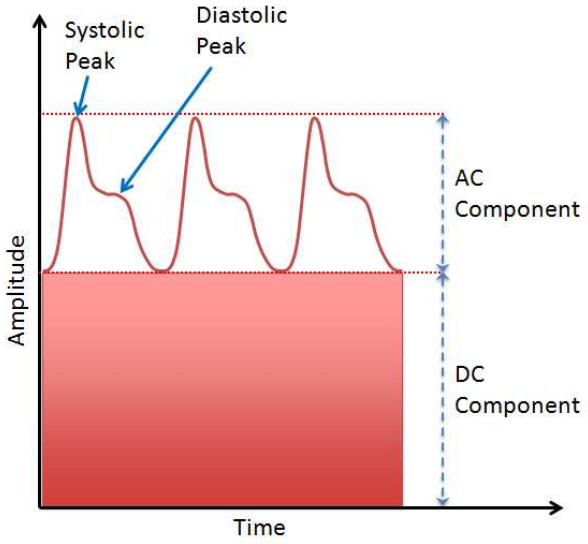


Figure 1.11 AC and DC components of blood volume variation

The DC component changes slowly with respiration, while the alternate component shows blood volume changes, which occur during the cardiac cycle. The fundamental frequency of the AC component depends on the HR and is added to the DC component. The interaction of light with biological tissue is quite complex and may involve scattering, absorption, and/or reflection, however, it can be fairly approximated by Beere Lambert's law:

$$A = \ln \frac{\mathrm{I_{in}}}{\mathrm{I_{out}}} = l \sum_{i=1}^{N} \varepsilon_i \; c_i \qquad (5)$$

where N represents the attenuating species in case of uniform attenuation. The absorbance A depends on the path length of the beam of light through the material sample l, the concentration c and the molar attenuation coefficient ε. The absorption clearly depends on the amount of light provided $\mathrm{I_{in}}$ and the amount of light, which is transmitted through the volume ($\mathrm{I_{out}}$, which refers to both reflected and transmitted light). The shape of the PPG waveform differs from subject to subject, and varies with the location and manner in which the pulse oximeter is attached. The PPG amplitude is the result of a complex interaction of several factors: stroke volume, vascular compliance, and tissue congestion effects. A large PPG pulse does not imply a high arterial pressure, absurdly, PPG amplitude can decrease during significant increases in blood pressure that are due to increased sympathetic tone, e.g. this phenomenon is usually seen in an incision on a subject under general anaesthesia.

1.4 MIMIC Database Description

The MIMIC Database (Multi-parameter Intelligent Monitoring for Intensive Care) is a collection of free access physiological signals recorded in the Intensive Care Units of medicine, surgery and cardiology at Boston's Beth Israel Hospital [42]. Recordings include continuous signals and periodic measurements ("numerics") taken from patient monitors. MIMIC Database is a representative clinical database because of the full range of pathophysiologies that result in sudden blood pressure changes [43]. This database consists of different physiological signals recorded from 121 ICU patients; however, only 72 patients were available. The data include signals and periodic measurements obtained from a bedside monitor as well as clinical data obtained from the patient's medical record. The recordings vary in length from 1 to 80 hours depending on patients. The data obtained from the bedside monitors are divided into files each containing 10 minutes of recorded signals, which can then be assembled without gaps to forma continuous recording [44]. The data were written in ten-minute segments in order to limit possible loss of data from power interruptions. The ECG, PPG and ABP signals (Figure 1.12) are sampled at 125 Hz with 12-bit precision and negligible jitter.

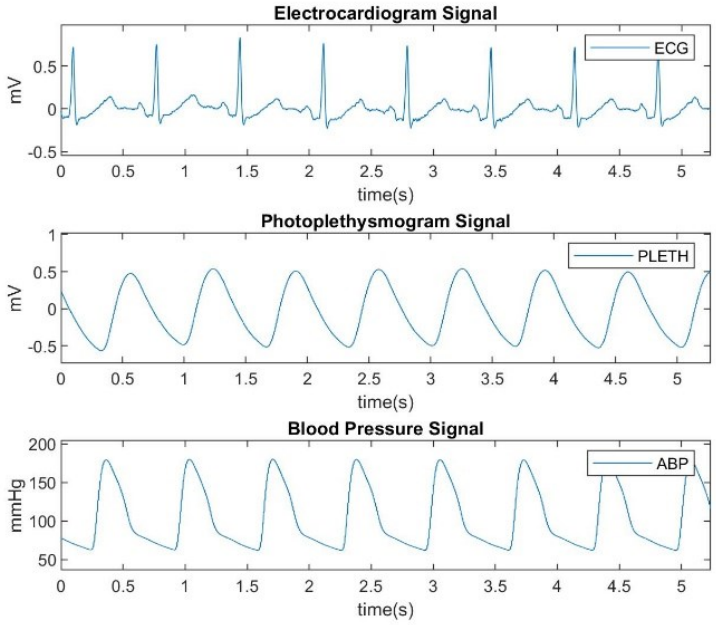


Figure 1.12 ECG, PPG and ABP signals

1.4.1 Data Cleaning

MIMIC was extracted via WFDB [42], a Python library supported by Physionet [39], and therefore any record without the required signals was discarded. Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP) values were extracted, representing the desired output.
The pre-processing pipeline is shown in Figure 1.13

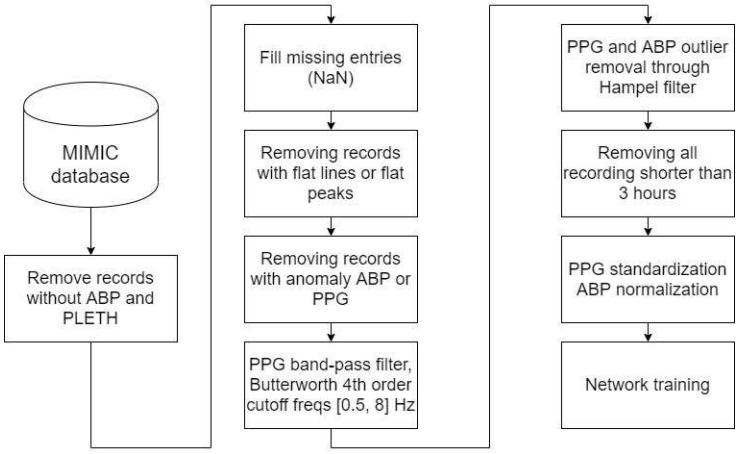


Figure 1.13 Data pre-processing pipeline

After deleting recordings without the requested signals (ECG, PPG, ABP), NaN values were managed replacing them with the first available value. The NaN values were simply replaced with the closest value because in some cases there were long stretches with missing values, so it was impossible to reconstruct the missing signal. Moreover, replacing the NaNs with the nearest value was not a problem because they were associated to flat lines, which are managed in the next step of the pipeline.

Then, it was necessary to exclude low quality recordings, i.e. those containing the so-called flat lines and flat peaks, which are recording errors mostly due to sensor problems, for example, a simple disconnection. *Flat lines* are long periods of time where the same value is always detected, while *flat peaks* are peaks with a flattened tip.

Subsequently, anomalies in ABP signal were managed: within the 10-minute recording the ABP signal should always be between a minimum of 15 and a maximum of 300 mmHg. In addition, a control on the pressure and plethysmography signal derivatives has been introduced, in particular the recordings that had first derivative always more than zero, or always less than zero, for more than 170 samples have been deleted. In practice, recordings in which the trend was increasing monotonous or decreasing monotonous for at least 1.36 seconds were eliminated. The remaining PPG recordings were then filtered through a 4th order band pass Butterworth filter with a bandwidth between 0.5 and 8 Hz, and then both PPG signals and ABP were filtered with a Hampel filter according to [45]. Those frequencies were chosen because anything below 0.5Hz is due to baseline wandering, while above 8Hz it is high frequency noise.

For computational reasons, the patients with less than 3.10 hours of recordings were discarded, while only the first 3.10 hours were taken from those with longer recordings. Out of 72 patients in the MIMIC dataset, only 61 had at least both ECG and PPG. Following the pre-processing pipeline, the dataset was created by extracting exactly 3.10 hours of the registration of each of the 50 patients remained. Also, in this case the distribution of systolic and diastolic BP was traced, and it is possible to notice that they are strongly biased towards

physiological values. Finally, the PPG signal was standardized and the ABP was normalized. Since the output is normalized, the predictions made by the networks must be denormalized using the minimum and maximum values calculated on the training set.

1.4.2 Data Building

Blood pressure, as mentioned, is closely related to photoplethysmography. However, ABP is a signal dependent on the cardiac cycle; for this reason, several prediction techniques use also the electrocardiogram signal. Two datasets have been created: the former uses the PPG signal to predict the arterial blood pressure, the latter uses both the PPG and the ECG (ECG could also help deep learning approaches). ECG lead V was used to obtain the largest possible data set, which represents the most frequent ECG lead recorded in the MIMIC database. An 8th order feedthrough and a type 1 Chebyshev filter were used to clean the ECG signals, with cut frequency of 2 and 59 Hz to avoid motion artifacts and alternating current artifacts. Out of 72 patients in the MIMIC dataset only 51 had at least both PPG, ECG lead V and PPG. After the pre-processing pipeline, the dataset was created by extracting exactly 3.10 hours of registration from 40 patients.

1.5 Blood Pressure Estimation with Neural Networks Output-Error

Arterial Blood Pressure (ABP) is an important physiological parameter that should be properly monitored for the purposes of prevention and detection of cardiovascular diseases, which represent one of the leading causes of death in the world. This section explains a blood pressure prediction technique based on system identification [46][47][48]. In order to predict the ABP values (both systolic and diastolic), electrocardiographic and photoplethysmographic signals are used as inputs of the networks. The Neural Networks Output-Error (NNOE) architectures are evaluated in terms of RMSE and absolute error. The neural networks output-error belong to the system identification algorithms and represents a synonym for mathematical modelling of dynamic systems using measurements of the input and output signals. Prediction-error methods are based on the observation that predictors which can be used to compare how well different LTI models can predict the desired output y(t). The main idea is to use some kind of measure of the distance between the predicted output and the desired output (target) *t* and to minimize this distance by adjusting some parameters in the model [49].

1.5.1 Neural Network Description

The purpose of Neural Network Output-Error (NNOE) is the identification of nonlinear dynamic systems in stochastic environment [50].
Figure 1.14 describes the procedure required to identify a dynamic system.

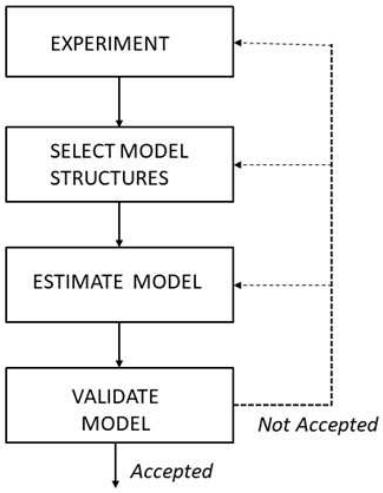


Figure 1.14 System identification procedure

The experimental phase is represented by the description of the dataset Z^N, which describes the entire system in its operating region with a proper choice of sampling frequency:

$$Z^N = \{[u(t), y(t)|_{t=1,\dots,N}]\} \tag{6}$$

$u(t)$ is the control signal, $y(t)$ represents the measured output signal and t specifies sampling instant number.
For the selection of the model structure, it is necessary to choose a set of regressors. The goal is to select a certain number of regressors based on the idea of a linear system identification and then determine the best possible network architecture. The selection of regressors as inputs of the neural network is carried out by Lipschitz method [15], as showed in Figure 1.15**Errore. L'origine riferimento non è stata trovata.**.
For NNOE, the shape of regression vector is given by:

$$\varphi(t) = [\hat{y}(t-1|\theta) \dots \hat{y}(t-n|\theta) \;\; u(t-d) \dots u(t-d-m)]^T \tag{7}$$

where θ is a vector containing the weights, n is the y-predicted lag, m is the input lag and d the delay to obtain the prediction (also called skip).

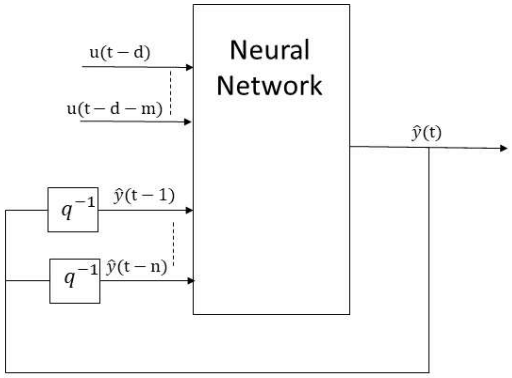


Figure 1.15 The NNOE model structure

The prediction vector is the following:

$$\hat{y}(t|\theta) = g(\varphi(t), \theta) \tag{8}$$

where g is the function realized by the neural network. The functions for estimating model are based on recurrent networks. The most common method of validation is to investigate the residuals (prediction errors) by cross-validation on a test set.

1.5.2 Evaluation Metrics

The output of the network is the ABP signal, estimated from an input which can be the PPG signal or both ECG and PPG signals (at the same time). The algorithm, implemented as supervised learning, associates, for each input, a real output (target) which is compared to the estimated one (output). The purpose of this first analysis is to estimate the waveform of the pressure signal and extract from it the minima and maxima points, i.e. the systolic and diastolic pressure values. An algorithm has been implemented for the identification of the characteristic points, i.e., the systolic peak, the diastolic peak, the dicrotic notch and the lowest point of the curve (foot), as shown in Figure 1.16.

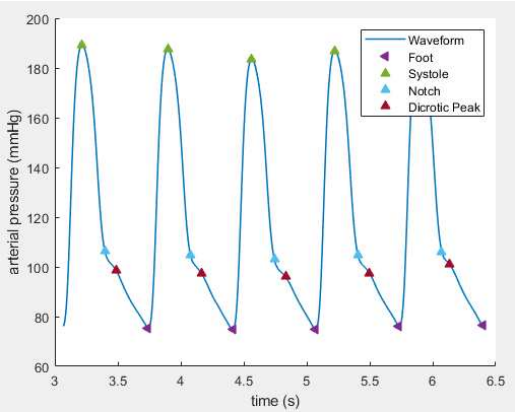


Figure 1.16 Waveform characteristic points

The characterizing parameters were extracted from the target signal and from the output signal and then applied different criteria for evaluating the results. The

points corresponding to the systolic pressure (systole) and diastolic (foot) pressure were taken into consideration. The parameters calculated to evaluate the errors made in the estimation of the peaks and valleys are the following:

- **RMSE (Root Mean Square Error)** measures the differences between values (sample or population values) predicted by a model or an estimator and the values observed. The RMSE represents the square root of the second sample moment of the differences between predicted values and observed values or the quadratic mean of these differences. These deviations are called *residuals* when the calculations are performed over the data sample that was used for estimation and are called errors (or prediction errors) when computed out-of-sample [51].
  RMSE is calculated as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2}{n}} \tag{9}$$

  where n is the number of measurements, y_i is the prediction and $\tilde{y}_i$ the desired true value (target).

- **MAE (Mean Absolute Error)** measures the errors between paired observations expressing the same phenomenon. Examples of y_i versus $\tilde{y}_i$ include comparisons of predicted versus observed, subsequent time versus initial time, and one technique of measurement versus an alternative technique of measurement [52]. MAE is calculated as:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}|y_i - \tilde{y}_i|}{n} \tag{10}$$

  where n is the number of measurements, y_i is the prediction and $\tilde{y}_i$ the desired true value (target).

1.5.3 Results

The NNOE networks presented in the previous section have been tested on systolic and diastolic blood pressure estimation task. ABP has been estimated with IBP as target and PPG as input and PPG+ECG as inputs, respectively. Then, their performances have been compared in terms of RMSE and mean absolute ABP error (in mmHg). At the beginning, the number of regressors of the network has been determined with the Lipschitz quotients method [53]. Figure 1.17 shows the Lipschitz graph, which can be used to reveal the order of the system. The indices for system orders from 1 to 10 are investigated; the system can be modelled by a sixth order model, since the slope of the curve is decreased for model orders greater than 5.

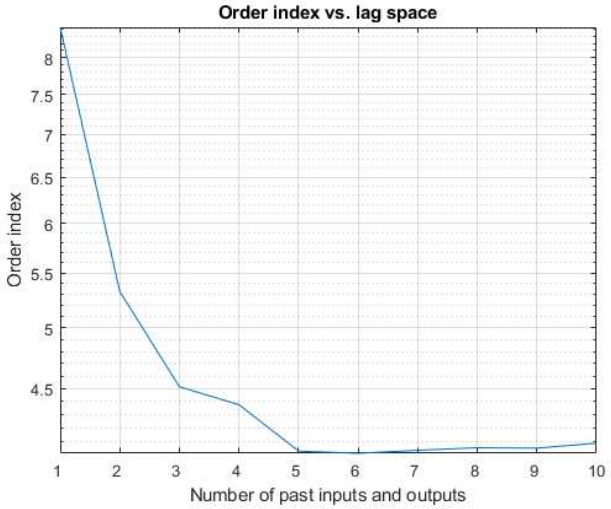


Figure 1.17 . The order index criterion evaluated with Lipschitz quotients method

In this analysis, the output-error architecture was implemented by a multi-layer perceptron (MLP), because of its capability to learn nonlinear relationship from a set of data. The hidden layer has 35 units and the activation function is the hyperbolic tangent. To train the network Levenberg-Marquardt method is used [54], which is the most effective method for feed-forward neural networks w.r.t. the training precision. It acts like a gradient-descent method when the parameters are far from their optimal value, and like the Gauss-Newton method when the parameters are close to their optimal value [55]. The error function is, of course, the sum of squared errors. The network is trained twice: first with PPG signal as input and then with both ECG and PPG signals as inputs. Before comparing the output of the network with the target, a moving mean filter (window length equal to 25 and 10, respectively) is applied to the output signal to remove noise artifacts. Figure 1.18 show the comparison between target (blue solid line) and output (red dashed line) signals with PPG and ECG+PPG as inputs, respectively. In both cases, the prediction is accurate. The model is evaluated in terms of RMSE; in particular, RMSE shows better performances for NNOE with ECG+PPG input than PPG. Table 1.2 summarizes the performances of the NNOE architecture.

Table 1.2. RMSE and MAE performances

	RMSE [mmHg]		MAE [mmHg]	
	Dias	Sys	Dias	Sys
PPG	1.43	1.73	1.06	1.54
PPG+ECG	**0.49**	**1.47**	**0.46**	**1.19**

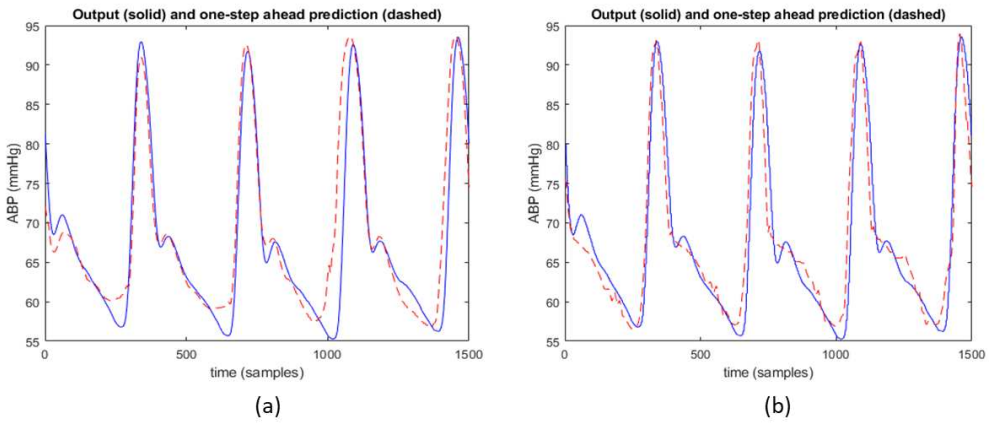


Figure 1.18 NNOE - Target (blue solid line) and Output (red dashed line) signals with PPG (a) and ECG+PPG (b) as inputs.

1.5.4 Future Works

The analysis compares the output-error neural networks (NNOE) for time series prediction in two configurations: PPG as input and PPG + ECG as inputs. NNOE is a network with a certain number of regressors as inputs, which try to identify nonlinear dynamic systems. The regression results are evaluated in term of RMSE and MAE. In particular, because of the regressors choice as input, NNOE approximates well the function that predicts the blood pressure value.
Future work will deal with improving the blood pressure estimation with NNOE method. In addition, an increase in the number of data is expected to ensure generalization.

1.6 A Comparison of Deep Learning Techniques for Arterial Blood Pressure

The purpose of this section is the continuous measurement of the arterial blood pressure (ABP) through the use of deep learning (DL) techniques with a cuffless and non-intrusive approach. Several deep neural networks (DNNs) are used to infer ABP, starting (also in this analysis) from photoplethysmogram and electrocardiogram signals [56]. The ABP was predicted first by exploiting only PPG and then by using both PPG and ECG. Convolutional neural networks (ResNet and WaveNet) and recurrent neural networks (LSTM) were compared and analysed for the regression task .

1.6.1 Methodology

In order to analyse the DNN architectures two different setups were implemented:
• *direct SBP/DBP prediction:* the network analyses 5 seconds of recording and then directly outputs a single value for SBP (peak) and another one for DBP (valley).
• *entire ABP signal prediction*: the network predicts the entire ABP in real time.

Predicting the entire signal could be better for hospital clinical applications, while, for commercial healthcare device implementation only systolic and diastolic values are predicted. DNNs for both setups were trained with both datasets and evaluated on a validation set, and then the best performing networks were cross validated using Leave-One-Out (LOO) since it is the most robust approach in terms of generalization performance [45]. DNNs were trained utilizing the Adam optimizer, the learning rate of 0.001, the Huber loss and the mini-batch training.
Because the Adam optimizer is an adaptive learning rate algorithm, it didn't require a lot of tuning and therefore the default learning rate was used, while Huber loss was chosen because it is a robust metric, unaffected by outliers, considering that the dataset did not have bell-shaped distribution as shown in Figure 1.19. So, Adam's learning rate using Adam optimizer helps to optimize the choice of the learning rate so that its variation does not particularly affect the algorithm's performance. The choice of architecture in order to find the best configuration through the validation set technique, in order to avoid overfitting of the network. While the choice of the number of training epochs has been defined by a stop criterion defining the accuracy you want.

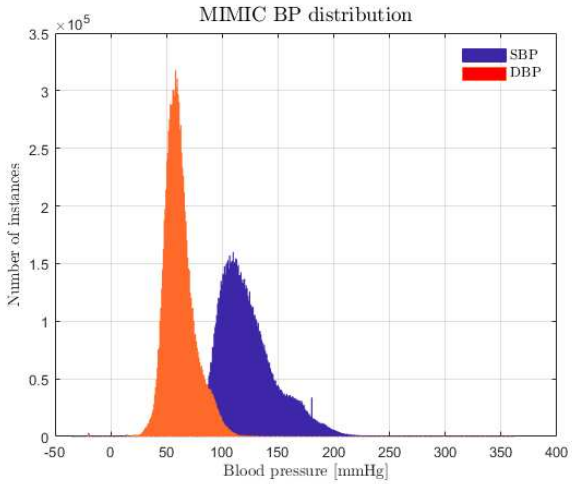


Figure 1.19 Dataset distribution

Data samples in *direct SBP / DBP prediction* are 5 seconds long, while in *entire BP prediction* are 2 seconds long. This difference is due to Long Short-Term Memory (LSTM); indeed, this cell has problems in managing too long sequences. LSTMs are used also in the first setup, however in this case it was possible to down-sample the input, through convolutional layers, since it was not necessary to output a value for every input.
In direct SBP / DBP prediction recordings were divided in 5 seconds chunks and hen on the samples was applied an algorithm that extracts SBP and DBP values. Since in 5 seconds usually there are between 4 and 6 cardiac cycles, it was taken the mean SBP and DBP among the cardiac cycles, as target values.

1.6.2 Convolutional Neural Networks

Convolutional neural networks (CNN) algorithms are inspired by the human visual cortex and are mainly known for their applications in image recognition. However, CNN are not limited to visual perception because they perform functions that are related to speech recognition, natural language processing and in general the analysis of time series.
The neurons in the visual cortex that they react only to visual stimuli localized in a limited region of the visual field (small receptive field). Neurons are organized on receptive field levels (Figure 1.20): because of higher-level neurons react to more complex patterns than they are combinations of lower level models they are led to the idea that the higher-level neurons rely on the outputs of neighbouring lower-level neurons [57].

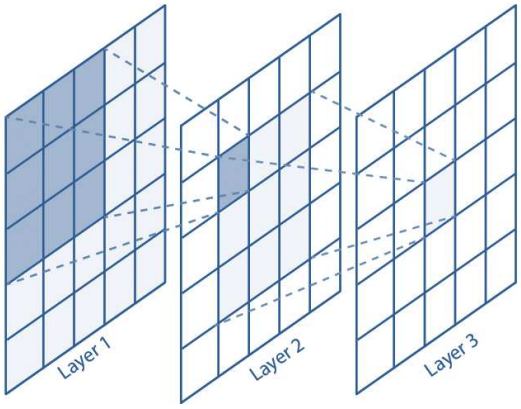


Figure 1.20 Receptive field of each neuron in different layers

The schematic architecture of a convolutional neural network is shown in Figure 1.21. The fundamental element of a CNN is the convolutional layer. Neurons in each convolutional layer are not connected to any single input point, but only to those in their receptive fields. The transition from one receptive field to the next is called step. This architecture allows the network to focus on small, low-level features in the first hidden level, then assemble them into larger higher-level items in the next hidden layer, and so on.

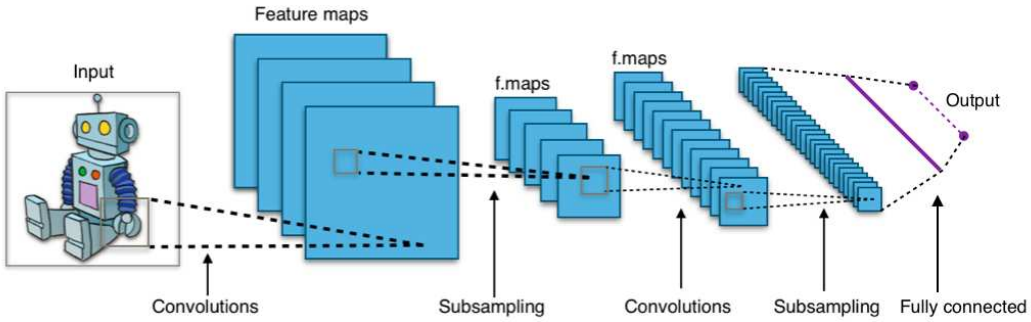


Figure 1.21 Schematic convolutional neural network

Convolutional levels are built on top of so-called convolutional kernels or filters, during the forward passage each filter is convolved on its inlet producing a 2-

dimensional activation map of that filter. This map is called a feature map. All neurons within a given feature map share the same weights and biases.
The receptive field of a neuron is the same as described above, and it extends over the feature maps of the previous levels. As a result, the network learns which filters are activated when it detects a specific type of feature in a spatial position in the input. Another important concept of CNNs is the pooling layer, which is a form of non-linear downsampling in order to reduce the computational load, the memory usage, and the number of parameters. Just like the convolutional layer, it has a receptive field, but it has no weights: it just aggregates the inputs using a dedicated function. The most common pooling function is the max pooling (Figure 1.22): in this case, the function partitions the input image into a set of non-overlapping rectangles and, for each such sub-region, outputs the maximum [58].

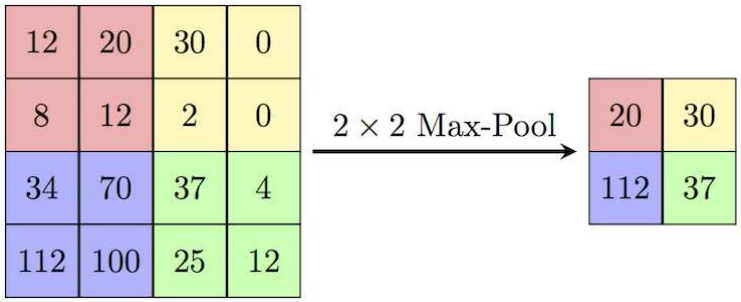


Figure 1.22 Max pooling function example

Fully Connected

Fully connected neural networks (FCNNs) are a type of artificial neural network where the architecture is such that all the neurons in one layer are connected to the neurons in the next layer. Such networks require many parameters and can be prone to overfitting. The major advantage of fully connected networks is that they are "structure agnostic", i.e., there are no special assumptions needed to be made about the input. While being structure agnostic makes fully connected networks very broadly applicable, such networks do tend to have weaker performance than special-purpose networks tuned to the structure of a problem space [59].

ResNet

The Residual Network (ResNet) is an architecture based on pyramid cells in the cerebral cortex, initially used for image classification. The effectiveness of neural networks is closely related to the depth of convolutional neural networks, however, ResNet are used to facilitate the training that is usually affected by the problem of evanescent/explosive gradients, degrading accuracy. Such degradation is not caused by overfitting and adding more layers to a suitably deep model leads to higher training error [60]. To avoid degradation, it introduced *skip connections* (Figure 1.23); the signal power in one layer is also added to the output of a layer located a little higher than the stack. This new technique made it possible to train

very deep networks like the original ResNet, a CNN made up of 152 layers. There are many variations of this network depending on about how deep it is. Usually, during the neural network training, the goal is to have a model target function *h (x),* which best approximates the network output starting from the input data. However, it is difficult to optimize the function when the network is very deep. For this reason, the input *x* is added to the network forcing output the network to learn the so-called residual map $f(x) = h(x) - x$.

When a regular neural network is initialized, its weights are close to zero, if there is a jump connection, the resultant network outputs only a copy of its inputs; in other words, it initially shapes identity function. If the target function is close enough to the identity function this will greatly accelerate the training [57] .

Thanks to skip connections, the signal can easily make its way through the entire network.

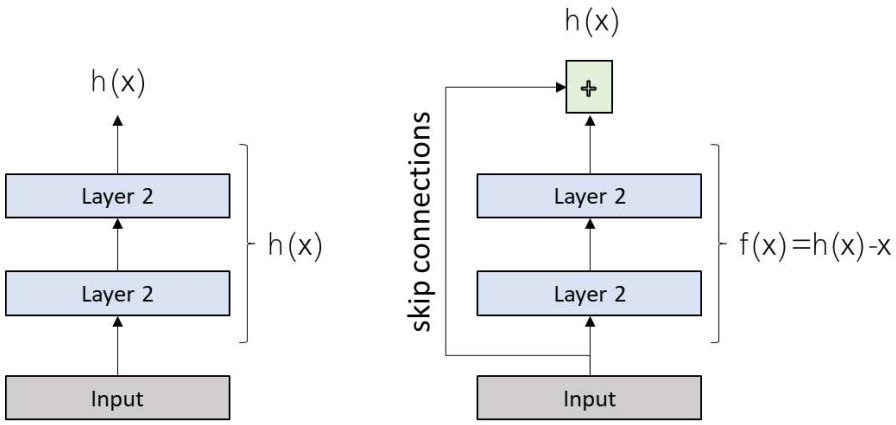


Figure 1.23 Skip connections

The classic ResNet architecture in illustrated in Figure 1.24. It starts with a convolutional layer and ends with a fully connected layer, if the final goal is a classification task the last layer is based on a softmax as an activation function, and in the middle there is only a very deep stack of simple *residual* units. Each residual unit is composed of two convolutional layers, with batch normalization and activation of ReLU, using 3×3 kernels and preserving the spatial dimensions.

There is no need to group levels into residual units because downsampling is performed directly from convolutional layers that have a stride of 2. Convolutional layers mainly have 3×3 filters and follow two simple design rules: for the same size of the output characteristics map, the layers have the same number of filters; if the size of the feature map is halved (using a convolutional level with step 2), the number of filters is doubled in order to preserve the time complexity per level. When the height and the width are halved, the inputs cannot be added directly to the outputs of the residual unit because they do not have the same shape. To solve this problem, inputs are passed through a 1×1 convolutional layer with stride equal to 2 and the right number of output capabilities maps [57].

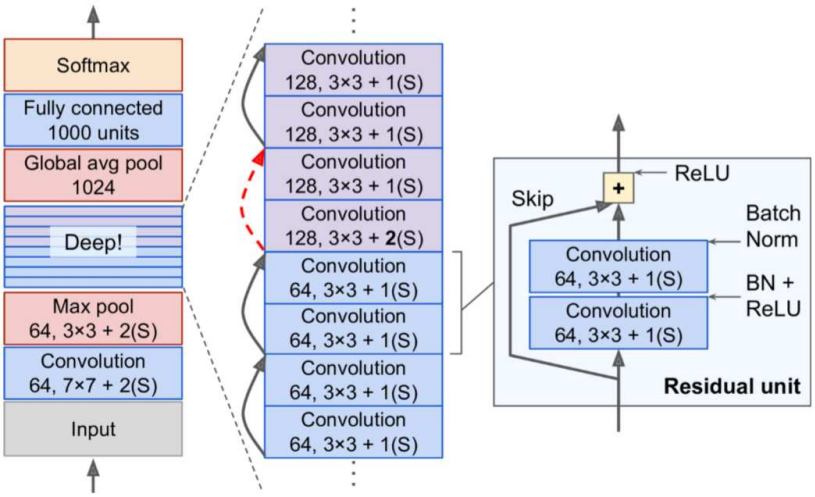


Figure 1.24 Classic ReseNet Architecture

WaveNet

The WaveNet architecture is a model originally designed to operate directly on the raw audio waveform. The classic model is composed of a stack of convolutional layers with no pooling layers and with causal padding. This padding allows the output to have the same temporal dimensionality as an entrance. Since this model does not require recurring connections, it is typically faster to train compared to RNN, especially when applied to very long sequences. However, one of the problems of causal convolutions is that they require many layers or large filters to augment the receptive field [61]. To solve this problem, WaveNet uses a dilation rate (Figure 1.25), which represents how far the inputs of each neuron are. A dilated convolution is a convolution in which the filter is applied over a larger area relative to its length by skipping the input values with a certain step. In this way, the lower layers they learn short-term patterns, while the upper layers learn long-term patterns. Thanks to the dilation rate of doubling, the network can process extremely large sequences in very efficient way.

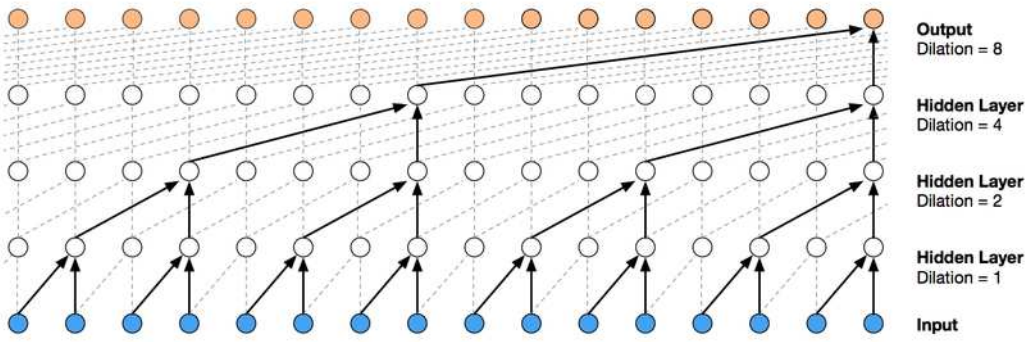


Figure 1.25 Dilated convolution layers

1.6.3 Recurrent Neural Network

A recurrent neural network (RNN) is an artificial neural network in which the connections among neurons make up a graph directed along a sequence. Thanks to this characteristic, it is possible to reconstruct a dynamic temporal behaviour of a phenomenon. Unlike feedforward neural networks, RNNs use internal memory to process input sequences. This makes them applicable to activities that extend over a certain period of time: handwriting recognition, speech recognition, automatic translation etc. The recurrent neural network aims to model a function that can provide output based on input. In most artificial neural networks, all inputs are independent of each other. But in RNN, all inputs are related to each other because it maintains all these relationships while the training.

LSTM

The Long Short Term Memory are a type of recurrent network that has the advantage of managing information in memory for a long period of time compared to RNN. Therefore, LSTMs are able to maintain long-term temporal dependencies, remembering less information (only the most important) for a long period of time. Indeed, LSTM networks remember past data in memory, resolving the vanishing gradient problem of RNN [62]. The LSTMs are based on special units called memory blocks (Figure 1.26).

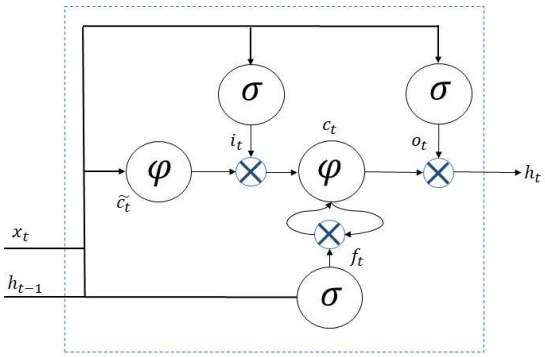


Figure 1.26 LSTM memory block

The working memory is represented by the hidden state h_t. The cell regulates its state using the gates: at first, there is a *forget gate* where some memories are dropped, then the memories are replaced with new ones selected by the input gate. The forget gate rule is given by:

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_t\right) \tag{11}$$

where W presents the weights vector, b the bias, σ the sigmoid function and t the current instant.

One copy of the new state is sent to the next iteration; the other one is passed through a *tanh* function and filtered by the *output gate*, as follows:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{12}$$

This is combined with the current inputs and the previous outputs to create the new output. The hidden state contains information on previous inputs and it is also used for predictions. Its output is given by:

$$h_t = o_t * \tanh(c_t) \tag{13}$$

The *input gate* recognizes important inputs and stores them into the long-term state, the forget gate deletes input that are no longer needed, and the output gate decides when to extract a specific input from the long-term state [63].
The current input and the previous output, also called short-term state, are fed to four different fully connected layers. The ones controlled by a sigmoid function are the layers that control the gates, their outputs range between 0 and 1 and are fed to element-wise multiplication operations; in this way, if they output is zero, they close the gate, while if the output is one, they open it. The forget gate controls which parts of the long-term state should be erased, the input gate controls which new memories should be added to the long-term state, and the output gate controls which parts of the long-term state should be read and output at this time step. The new memories are calculated in the layer controlled by the *tanh* function.

1.6.4 Results

In order to perform experiments and delve into the topic of blood pressure prediction, a study analysis was carried out with the neural networks analysed in the previous section and creating new hybrid networks that contain both recurrent and convolutional networks. Very simple convolutional neural networks and recurrent neural networks were initially trained, but the hybrid approach between CNN and RNN produced the best results.

Direct SBP/DBP Prediction

The first experiment was carried out by using a ResNet18 and testing different batch sizes. Smaller batches allowed a faster training and achieved better results, probably because they did not get stuck in some local minimum. Therefore, also for a regression task, mini-batch training is the best way to train a neural network. In particular, three settings were tested: in the first, 650 samples per batch were used maximum size permitted, before performing backpropagation; in the second, 128 samples, and in the third 32 samples were adopted. In every setup, the number of training steps is always the same: this is important because it represents the number of times the weights are updated; therefore, the networks are comparable only if their weights are updated the same number of times:

$$Training\ steps = epochs * \frac{Number\ of\ samples}{Batch\ size} \quad (14)$$

Classical feature selection has been automated by convolutional layers and skip connections; in addition, it is possible to stack layers creating a deep neural network, which can better analyse input data.

The next experiment for *Direct SBP/DBP prediction* employed a ResNet, like the previous one, followed by three LSTM layers, each one made up of 128 neurons. The first LSTM layer is bidirectional. Convolutional layers are activated when they are combined with recurrent layers: they extract features from a signal, and they can also downsample the input sequence using the right kernel size, stride and padding. The model can learn how to preserve the useful information dropping only the unimportant details and shortening the sequences; the convolutional layer may help the following recurrent layers to detect longer patterns. This network was the best performing to directly predict SBP/DBP values for both datasets. The results are shown in Table 1.3.

Table 1.3. Errors (mmHg) for Direct SBP/DBP prediction task on SBP and DBP prediction for different setups with MIMIC databases

Neural network	SBP	DBP	SBP	DBP
	MAE (mmHg)		RMSE (mmHg)	
ResNet (PPG)	9.556	4.217	13.572	6.012
ResNet (PPG+ECG)	4.667	2.445	6.227	3.042
ResNet +LSTM (PPG)	7.122	3.534	11.214	5.029
ResNet+LSTM (PPG+ECG)	**4.118**	**2.228**	**5.682**	**2.986**

Entire ABP Signal Prediction

Initially the entire ABP signal was tried to predict through a simple fully connected neural network. However, due to the simplicity of the model, good results were not achieved. Deeper models appear to converge faster, but still give high errors.

Another experiment for *Entire ABP signal prediction* task used a network which consists of three stacked LSTM layers, each with 128 cells. The first layer is bi-directional, while the output layer is a fully connected neuron with no triggering function. Bi-directional long-term memory (BLSTM) looks for contextual features both forward and backward, which is useful because the location of the feature the network wants to forget is not known. This approach is used to increase the amount of incoming information available to the network [32]. BLSTM is usually positioned as the top tier of the network because it has access to context on a much larger scale than the input sequence. BLSTMs greatly increase the computational cost; therefore, it is reasonable to use only one bidirectional layer. Each sample consists of 2 s of recording; as explained above: this value length was defined because LSTMs have trouble handling long sequences. It is difficult to remember long-term patterns if the sequence is too long; furthermore, this makes calculating the gradient over time too difficult. For

this reason, only the 2 seconds preceding the current instant t are taken as the input time window to predict a single output value at instant t.
The third experiment for the task used a simplified version of the WaveNet, composed of two blocks each one with four convolutional layers. The dilation rate is the double (from 1 to 8) in every convolutional layer inside a block. The output layer is a fully connected neuron without any activation function. Since the network is composed only by convolutional layers, it converges fast and, thanks to the doubling dilation rate, it can process extremely large sequences very efficiently. Afterwards, a second network was built stacking three LSTM layers (fourth experiment), each composed by 128 neurons, where the first layer was bidirectional. On top of this simplified WaveNet, convolutional layers extract features that are then analysed by the LSTM layers.
The last experiment was carried out using both LSTM and convolutional layers. In particular, here a ResNet is followed by three LSTM layers, being the first bi-directional. This network differs from the one presented in the *Direct Prediction* section of SBP / DBP because max-pooling levels are not used and convolutional levels have causal padding, like WaveNet. This is a crucial step: to predict the entire signal, it was necessary to output a sequence of the same length as the input sequence. This network performed best in predicting the entire signal. In this experiment, each ResNet is made up of four ResNet blocks. The convolutional layers have kernel size equal to 3 and strides equal to 2, while the number of filters increases in each block starting from 64 up to 512. Each layer consists of 128 cells. Each convolutional operation is here followed by a batch normalization, which centres and normalizes each input; then, resize and move the results using two new parameter vectors per layer: one for resizing, the other for moving. In other words, this procedure causes the model to learn the optimal scale and mean of each of the level inputs. As shown in Figure 1.27, this hybrid network was the best performing one to predict the entire BP signal for both datasets because the network was built specifically on this problem. It can be observed that the three peaks are well estimated (there is only a small error of 3 mmHg for the first peak).

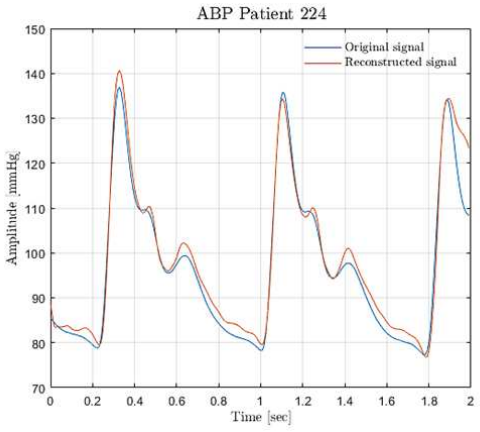


Figure 1.27 ABP prediction on a validation set sample made with ResNet+LSTM trained with PPG dataset: original signal (light blue) vs network output (orange).

A performance comparison among the analysed neural network architectures for *Entire BP prediction task* is summarized in Table 1.4.

Table 1.4. Errors (mmHg) for Entire BP prediction task on SBP and DBP prediction for different setups with MIMIC databases

Neural network	SBP	DBP	SBP	DBP
	MAE (mmHg)		RMSE (mmHg)	
Fully Connected (PPG)	36.559	10.602	45.013	13.417
Fully Connected (PPG+ECG)	29.753	12.759	39.330	15.198
LSTM (PPG)	12.118	5.018	17.875	6.890
LSTM (PPG+ECG)	7.603	3.688	11.846	5.320
WaveNet (PPG)	18.539	8.154	26.638	11.441
WaveNet(PPG+ECG)	14.501	7.224	22.922	10.477
WaveNet +LSTM (PPG)	14.353	6.311	21.323	9.150
WaveNet +LSTM (PPG+ECG)	8.812	3.471	12,967	4.864
ResNet +LSTM (PPG)	8.660	3.843	13.439	5.718
ResNet +LSTM (PPG+ECG)	**4.507**	**2.209**	**6.414**	**3.101**

Leave-One-Out (LOO) cross-validation

In order to understand the performance of the generalization, a Leave-One-Out (LOO) cross-validation was conducted on the best architecture, i.e., the hybrid network which consists of ResNet followed by LSTM, for both datasets: the one built using PPG only and the one built using both PPGs and ECG.

LOO is better than the k-fold cross validation [90] in case of very few data. Indeed, this is the best possible way to exploit the experimental dataset of this work. Specifically, the dataset, created using only PPG, had 50 patients, while the dataset, created using PPG and the ECG, had only 40 patients. During the training phase, it was important to have access to as much data as possible; like this, all available data were used; conversely, to compare performance, it was useful to have the same data set. The ECG signal generalization also improved when used at best ResNet + LSTM model, as shown in Table 1.5; however, errors were lower than the case in which nets were trained and tested on the same patients (different records). This phenomenon appears in many other types of research [64], and is generally called *personalization*. With individual calibration, PPG and ECG can be used to directly estimate SBP and DBP on new data obtained from the same individual. According to to the American National Standards Institute (ANSI) for the "Development of medical instrumentation" [65], in order to validate a new device, there should be an average difference of ± 5 mmHg between the standard and the newly developed device [66]. The mean square error for SBP is 5.682, while for DBP it is 2.986.

Table 1.5. LOO results on MIMIC Database with the best neural network (ResNet+LSTM).

Tested set	MAE	RMSE	MAE S	MAE D	RMSE S	RMSE D
	Direct SBP/DBP prediction					
PPG (50 pat)			23,5976	10,7459	27,6430	12,3444
PPG (40 pat)			24,2227	11,1056	28,2470	12,6419
ECG(40 pat)			20,3667	9,5484	23,0699	10,8475
	Entire BP prediction					
PPG (50 pat)	15,3419	19,1549	21,4666	10,6841	25,3825	12,3489
PPG (40 pat)	15,6788	19,5598	22,4095	10,8180	26,2460	12,4111
ECG(40 pat)	14,6093	18,0184	22,0995	10,1053	24,5865	11,5292

1.6.5 External Validation on Polito Dataset

To test the pressure prediction algorithm, a customized dataset (for external validation) was created at the Neuronica Lab of the Politecnico di Torino.
The dataset consists of nine healthy volunteers (5 males, 4 females, aged 22.84 ± 1.7 years) who were recruited to participate in the PPG, ECG and ABP signal acquisition experiment. Recordings were collected using a GE Healthcare B125 Patient Monitor, which is a certified clinical device, generally appreciated for its intuitiveness and reliability in a variety of acuities. The monitor offers proven NIBP technology, using GE patented smart cuff pressure control to improve measurement time, patient comfort and artifact rejection. It meets the requirements of both AAMI ISO81060-2 and IEC 80601-2-30. PPG, ECG and ABP were measured three times using the following recording protocol: first, PPG and ECG were recorded simultaneously; then, BP was measured using a sphygmomanometer. PPG and ECG recordings were 15 s long. The PPG was sampled at 300 Hz, while the ECG at 100 Hz; therefore, both signals were resampled both at 125Hz, respectively with the down-sampling (the samples were skipped) and the linear interpolation method. A sphygmomanometer was used because a CNAP system was not available, while invasive methods can only be performed by specific clinical staff.

Results on Polito Dataset

The best performing DNN (ResNet + LSTM) trained on PPG and ECG (Lead I) was used to predict SBP and DBP (point values) on Polito volunteers: MAE was 12.435 mmHg on SBP and 8.567 mmHg on DBP (see Table 1.6, which shows the results for the collection of data from Polito volunteers). The network was also trained using only the PPG. This configuration achieved MAE equal to 9.916 mmHg on SBP and 5.905 on DBP. In this case, the ECG improved the performance on the data extracted from the MIMIC database but did not affect generalization. Furthermore, it negatively affected the results on the Polito. The reason is probably due to the small training set; in fact, only 12 patients had ECG in MIMIC Database. Unexpected results on the Polito dataset could be due to different methods of pressure acquisition (invasive in the case of the MIMIC, non-invasive in the case of the Polito Dataset). In this case, for technical reasons, the

pressure was not acquired with a invasive method, but measured with a sphygmomanometer, which can introduce epistemic uncertainty.

Table 1.6. SBP and DBP prediction errors (mmHg) on Polito database using the best neural network (ResNet+LSTM) trained on MIMIC dataset (built using PPG and ECG lead I).

Tested set	MAE SBP	MAE DBP	RMSE SBP	RMSE DBP
	PPG			
Validation set	7,409	3,706	9,875	4,883
Leave-One-Out	15,706	7,251	17,792	8,171
Polito dataset	9,916	5,905	11,879	7,273
	PPG+ECG			
Validation set	4,546	2,515	5,766	2,982
Leave-One-Out	16,128	6,743	17,875	7,902
Polito dataset	12,435	8,567	14,082	10,211

1.6.6 Discussion

Blood pressure measurements are performed based on PPG only; however, the results are influenced by the inter-operator variability; to get around this problem and obtain a greater generalization, ECG signals should be considered. Different people show different ABP and PPG waves; however, the results also depend on the average pressure of the patients: the biggest mistakes were made on patients with the highest mean ABP. To obtain better results, it would be appropriate to use a larger dataset to have a Gaussian distribution of BP. Large datasets are of utmost importance in deep learning and the reason is clearly shown in Polito results: although the ECG importance was proven, it did not improve the performance, because the network was trained on a very small dataset.

Among the analysed neural networks, the hybrid networks ResNet+LSTM, which contain both recurrent and convolutional networks, have gained the best performances.

Table 1.7 shows the comparison of all the neural networks employed for arterial blood pressure detection in term of complexity. In particular, for convolutional neural networks, the complexity affects the length of the signal, the dimension of the input vector (1 for only PPG input and 2 for PPG and ECG as inputs), and the kernel size, while for recurrent neural network, the complexity affects only the length and the dimension of input vector. The ResNet + LSTM represents the best model in terms of performance, but, at the same time, the most expensive model in terms of computational complexity.

The goal for blood pressure estimation is twofold: on one hand to predict the entire pressure signal and on the other hand to estimate the point values of diastolic and systolic pressure. The objective is because it strictly depends on the application: the detection of the entire signal at a medical level is certainly more significant because it describes the morphology of the entire signal and allows to study the signal more thoroughly; instead, the detection of the values is useful for monitoring vital parameters in wearable devices, a simpler task as the purpose is to predict only two points and not an entire signal.

Table 1.7. Neural networks complexity order

Neural network	Complexity Order	Cost Estimation (FLOPs)	
		PPG	PPG+ECG
Fully Connected	$O(length \times (vector\ dimension)^2)$	~ 625	~ 2500
LSTM	$O(length \times (vector\ dimension)^2)$	~ 625	~ 2500
WaveNet	$O(length \times (vector\ dimension)^2 \times kernel\ size)$	~ 1850	~ 7500
WaveNet+LSTM	$O(length \times (vector\ dimension)^2 \times kernel\ size)$	~ 1850	~ 7500
ResNet+LSTM	$O(length \times (vector\ dimension)^2 \times kernel\ size)$	~ 4375	~ 17500

1.7 Conclusion

PPG-based techniques enable continuous and automated ABP measurements; they are also well tolerated by patients and are inexpensive and portable. These techniques are based on the direct detection of the blood volume in the arteries under the cuff. In this study, the ECG improved the performance of the PPG in each proposed configuration and allowed the network to generalize better: it is therefore important to collect ECG data, firstly with regressive approaches and secondly with deep learning approaches. Such systems are easy and non-invasive techniques for measuring blood pressure. The experiments were conducted on a subset of patients from the MIMIC database. In first section, the non-invasive blood pressure estimation technique is presented with a regressive neural network NNOE. To determine the best input between a single configuration (PPG) and a double channel (synchronized ECG and PPG) configuration, a comparative analysis is carried out. Because the goal is the ABP forecasting, the NNOE method is used based on MLP, employing only the previous predictions in the regression vector. The double channel configuration yields the best results w.r.t. mean absolute error, which results to be, on average, 2.42 mmHg and 3.17 mmHg for SBP and DBP, respectively; in this sense, this configuration is compliant with the legislation because the estimated values are within +/- 5 mmHg w.r.t. real invasive measurements. Despite the excellent results, regressive techniques have generalization problems; for this reason, in the second section an analysis was carried out by carrying out a comparison based on deep learning techniques.

In this more general case, thanks to the deep learning approach, the validation was in accordance with ANSI guidelines: the best performing network achieved an MAE of 4.118 mmHg on SBP and 2.228 mmHg on DBP. The selected network was also tested on a different custom dataset, created at Neuronica Labs (Politecnico di Torino), which performed better than the cross-validation of MIMIC LOO. This is likely since this dataset was smaller and, therefore, had a lower variance. Indeed, Polito's volunteers were all young and healthy subjects, while MIMIC is a particularly complicated dataset, because its patients have a huge variety of pathophysiologies that cause blood pressure changes.
The proposed neural algorithm can be incorporated into wearable portable devices to perform continuous health monitoring of blood pressure in order to prevent the onset of irreversible damage, such as cardiovascular disease and hypertension. Implemented in a device, this algorithm can prove to be a powerful tool for diagnosing aggressive covid-19 virus at an early stage.

1.8 Acknowledgement

This work has been partly supported by the Neuronica Lab and the PoliToBIOMed Lab - Biomedical Engineering Lab of the Politecnico di Torino.

Chapter 2

Development and Validation of a Shallow-Learning Model to Screen for Channelopathies from the Digitized Electrocardiogram with ECG-Dig GUI

2.1 ECG Digitalization: from ECG Paper to ECG Digital Signal

The electrocardiogram (ECG) signal describes the electrical activity of the heart, allowing to detect several health conditions, including cardiac system abnormalities and dysfunctions. Nowadays, most patient medical records are paper based. The importance of collecting digitized ECGs is twofold: firstly, all medical-engineering applications can be easily used if the ECGs are treated as signals; secondly, paper ECGs can deteriorate over time, therefore not guaranteeing a correct evaluation of the clinical condition of the patient. The goal of this paper is the realization of an automatic conversion algorithm from paper-based ECGs (images) to digitized ECGs signals. The algorithm has been tested on 17 scanned ECGs, also with pathologies. Quantitative analysis of the digitization method was carried out by evaluating the repeatability and reproducibility of the algorithm. The digitization accuracy was evaluated both on the entire signal and on ECG fiducial points (R-R peaks, QRS complex, QT interval, PQ interval, P wave amplitude, heart rate). Results demonstrate the quality of the algorithm with an average Pearson correlation coefficient of 0.94. Due to the promising experimental results, the algorithm could be embedded in a graphical interface, becoming a measurement and collection tool for cardiologists.

2.1.1 Related works

The digitization process is an essential process for the analysis and processing of signals. In recent decades the in-depth study of medical signals has been made possible thanks to its digital nature.
The fundamental advantages of digital signals are noted in terms of security, storage, and non-deterioration due to paper. Furthermore, saving the patient's history guarantees ease of knowing the patient's clinical evolution. Knowing the entire time series of each signal, it is possible to implement algorithms for the automatic detection of pathologies [67].
The ECG signal is an electrical signal that describes cardiac activity. The graph represents the trend of the heart potential over time. Nowadays, digital signals can be collected in the cloud and stored using the latest generation electrocardiographs. However, to know in depth the patient's medical history and build automatic algorithms it is essential to know the ECG signals of the past, which in most cases is paper based. Modern recording electrocardiographs cannot analyse preserved paper ECG records because it requires input in terms of digitized signal. For this reason, it is important to extract ECG signal from these preserved paper ECG records using digitization method. In [68], an entropy-based methodology is proposed. It is based on the bit plane slicing (EBPS) in which pre-processing is performed using dominant colour detection and local bit plane slicing. The adaptive bit plane selection based on maximum entropy is applied to the pre-processed image. Discontinuous ECG correction (DECGC) is then performed to produce a continuous ECG signal.
To increase the accuracy of the digitizing process, it is necessary to reduce noise with image-processing algorithms (often characterized by annotation include some characters) [69]. In [70], scanned images are enhanced by applying skew correction operation using the Hough transformation and noise removal is done using median filtration. Asymmetry and noise are common mistakes in the scans of the images and should be avoided to capture better results. Next, the grid is removed from the ECG images, using colour segmentation.
In this work, a Matlab-based tool, the *ECG-dig*, for digitizing paper-ECG data is presented. The conversion technique is validated by carrying out a similarity study based on the Pearson coefficient between the digital signal and the digitized one and evaluating the algorithm in terms of repeatability and reproducibility.

2.1.2 Paper Collection

The proposed data acquisition system is shown in Figure 2.1. Two medical instruments in cascade were used to build the database for our study: the ProSim 4 Vital Signs Simulator and the GE MAC 2000 electrocardiograph.

Figure 2.1. Data acquisition system

ProSim 4 patient simulator allows the simulation of several ECG functions, while the GE MAC 2000 is an electrocardiograph which, in addition to allowing the display of the simulated conditions through the monitor, guarantees the acquisition of signals both in digital format (.xlm file) and by printing the images on graph paper [71].

The simulated ECG signals had the following heart conditions:

• *Normal sinus rhythm*: the rhythm of a healthy heart. It means the electrical impulse from your sinus node is being properly transmitted.

• *Bradycardia*: the presence of a slow or irregular heartbeat, usually below 60 beats per minute.

• *Tachycardia*: the increase in the number of heart beats per minute (heart rate) under resting conditions.

• *Atrial Fibrillation*: the heart failure in which the heartbeat becomes irregular and, often, accelerated.

• *Atrial Flutter*: the heart failure when the electrical activity in the atria is coordinated. Therefore, the atria contract but at a much increased rate (250-350 beats per minute).

• *Muscle tremor artifact*: a type of movement artifact. It usually happens because the patient is trembling.

• *Premature Ventricular contractions* (PVCs): single ventricular impulses caused by abnormal automatism of the ventricular cells or to the presence of re-entry circuits in the ventricle.

• *Premature Atrial contractions* (PACs): a common cardiac dysrhythmia (similar to PVCs) characterized by premature heartbeats in the upper chambers of the heart, the atria.

• *Acute pericarditis*: the inflammation of the pericardium characterized by an accumulation of fluids in the pericardial space.

• *Supra Ventricular Tachycardia*: the high-rate heart rhythm originating above the ventricle.

• *Ventricular Fibrillation*: the hyperkinetic arrhythmia characterized by a high ventricular rate (100-150 beats per minute).

Table 2.1 summarizes the database artificially created with the ProSim4 simulator.

The images printed with the electrocardiograph GE MAC 2000 are scanned with the Kyocera TASKalfa 5053ci scanner, with a scanning speed of 220 ipm and a scan resolution of 600 dpi x 600 dpi, with 256 levels of grey per colour.
Finally, to make the image more suitable, the contrast and sharpness have been increased by 70%. The waveform speed of the ECG signal on the printout is set to 25 mm / s. The purpose of the digitization algorithm is to transform the signal printed on the graph paper into a digital signal that respects the measurements of mV (ordinate axis) and ms (abscissa axis). However, the conversion error due to the digitization process is combined with the error due to the electrocardiograph printing process of the signal on graph paper. In this last case, the error is not considered in the study because all cardiologist specialists evaluate the patient's health condition using the paper ECG trace.

Table 2.1. Dataset description

HR (bpm)	CONDITION	Amplitude (mV)
30	Bradycardia	1
45	Bradycardia	1
60	Normal sinus rhythm	0,5
60	Child normal sinus rhythm	1
60	Normal sinus rhythm	1
60	Normal sinus rhythm	0,5
80	Acute Pericarditis	0,2
100	Normal sinus rhythm	1
120	Sinus Tachycardia	1
76	Atrial Fibrillation	1
82	Atrial Flutter	1
60	Breath artifact	1
60	Muscle artifact	1
75	Premature Atrial contractions (PACs)	1
78	Premature Ventricular contractions (PVCs)	1
200	Supra Ventricular Tachycardia	1
152	Ventricular Fibrillation	1

2.1.3 ECG-dig Algorithm

In order to convert the image into digital signals, an algorithm was implemented by using MATLAB® platform. The algorithm is able to digitize the image, distinguishing the signals from the background and respecting time and voltage proportions of the ECG signals. It is based on a step-by-step automatic processing which involves the operations summarized in Figure 2.2.

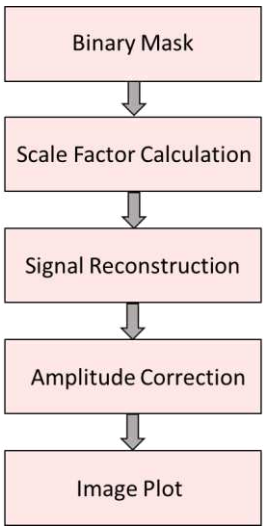


Figure 2.2 Automatic algorithm pipeline

Binary mask

The first block of the pipeline is the extraction of the binary masks, which has been inspirated by MathWorks Community [72]. It is based on the implementation of a colour threshold: some colour values have been chosen to extract the signal from the rest of the image. The result is a black and white image, where the signal will be afterwards smoothed, in order to have one black pixel in every column and a white background. The black pixels correspond to the signals.

Scale Factor (SF) calculation

The standard ECG leads are printed on a graph paper (Figure 2.3**Errore. L'origine riferimento non è stata trovata.**). When the image is scanned, the correspondence between pixels and millimeters is not always the same and it depends on some factors (e.g. the printer resolution and the available type of image). In order to find out how much a pixel is worth in each image, a specific function was created. To obtain the signals of all the ECG-leads, 12 crops (one for each lead) are made on the image by framing the image patch of the corresponding lead. Each crop is transformed in a grayscale image and the signal is extracted as in the previous paragraph (in this case, the signals become white).

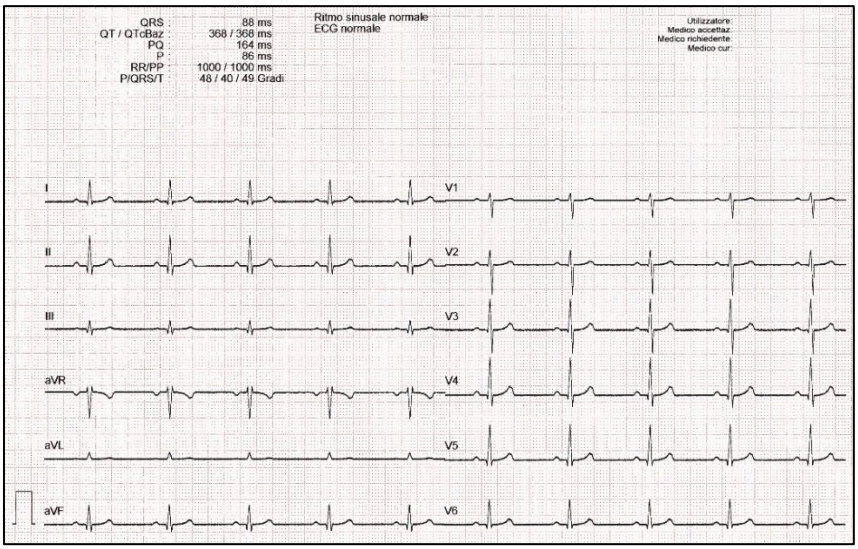


Figure 2.3 ECG leads printed on graph paper

Two thresholds have been chosen quite close to the greyscale extremes, in order to isolate the grid. Nevertheless, it is not certain that remaining black points and white background have a shade of grey exactly corresponding to the extremes. Therefore, the thresholds have been chosen not too much high (for white) or low (for black): in this way, black points and the white background and signal are excluded. In the proposed data set, the images have two grids, one less dense (with larger squares) and one denser. The first one is composed by dots, very close to each other, which form the perimeter of squares with a 5 mm side. In the second one, dots are further away and delimit squares with a 1 mm side as shown in Figure 2.4.

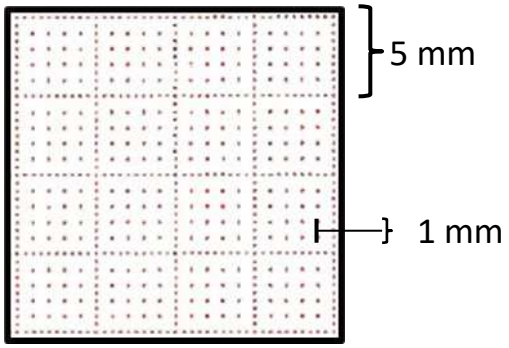


Figure 2.4 Squares in ECG graph paper

Thus, the algorithm joins the nearest to each other and deletes the furthest. In this way, we obtain a binary image with a grid, composed by horizontal and vertical lines, that form 5 mm side squares. The squares area (in pixels) is calculated as the mean of all the squares areas and, taking the square root, we have the inner side of the mean square. By doing the sum of the latter and the width of one line, the length in pixels of the square side is found. Knowing that it should be 5 mm, the scale factor SF is obtained as:

$$SF = \frac{5}{L} \tag{15}$$

where L is the measurement of the side in pixels and SF indicates how many millimetres a pixel corresponds to.
Since the SFs are not always the same (because of random inaccuracies during the binary grid creation), a mean value is calculated between the twelve scale factors. This one will be used in the next parts of the algorithm.

Signal Reconstruction
After obtaining the binary images, the algorithm takes only the pixels of the signal and uses their y-positions to know the amplitude value of each point. Then, the width is converted from pixels to millimetres, multiplying them by SF.
To align the isoelectric line on the $y = 0$ line, the mode of each lead is calculated and subtracted from the signal itself.

Amplitude Correction
The amplitude is often less than reality in the area where more black pixels are usually concentrated. This is especially the case with R peaks because the leads are generally tighter at that point.
The per-pixel reconstruction leads to a subsampling with respect to the signals produced by the electrocardiograph. The result is that the amplitude is sometimes less than reality. Also, this happens where more black pixels are usually concentrated, especially near the R peaks, because the leads are usually narrower here. Therefore, the algorithm automatically detects the positions of the R peaks and adjusts the amplitude value, adding 1mm to those points (when the peak is positive) or subtracting 1mm (when it is negative). This quantity was experimentally chosen by observing the differences between the reconstructed leads and the real signal and taking the average. Furthermore, this value agrees with the uncertainty due to the thickness with which the signal is printed ($\simeq$ 1mm).

Image Plot
Since 10 mm correspond to 1 mV, signals amplitude is converted from millimeters to voltage. About time scale length, each lead has a samples number equal to the number of pixels (voltage values). In order to create the time scale to be visualized, the samples are before converted in millimetres thanks to SF and then, in milliseconds, knowing that paper speed is 25 mm/s.
Each lead is plotted with a pink grid background which reproduces the graph paper: the x-axis is time (ms) and the y-axis is voltage (mV).
Lastly, the algorithm saves the images, voltage data of the 12 leads and the time samples.

2.1.4 Algorithm Validation Technique

The algorithm created to digitize and save the ECG signal in digital format of each patient must be validated. As the algorithm is intended to be a tool for clinical support it must be rigorously tested. To validate the algorithm, it must be evaluated in terms of *similarity* of the entire signals using the Pearson coefficient [73] and in terms of *repeatability* and *reproducibility* using the absolute error of the fiducial parameters of the ECG signal. Signal (digitized ECG and GE digital ECG) similarity refers to the provision of accuracy measures as an indicator of a qualitative assessment. The accuracy of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. The repeatability indicates the agreement between repeated tests performed with similar measurement conditions [74]. The reproducibility is defined as the agreement between two measurements done under different circumstances [75]. Figure 2.5 shows the algorithm validation scheme.

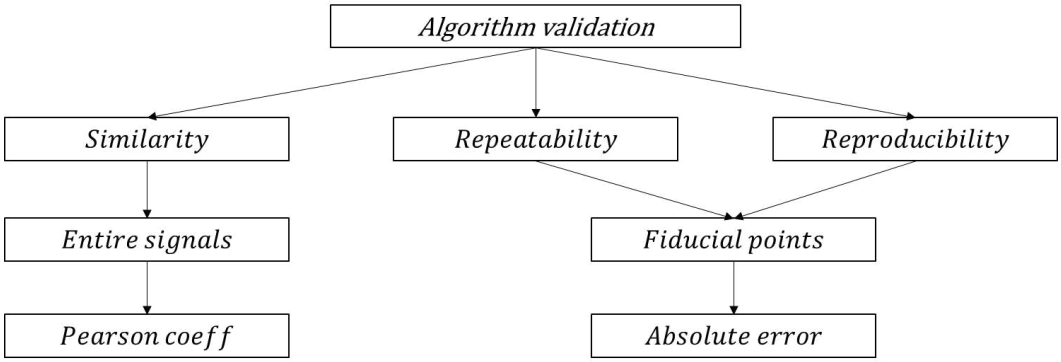


Figure 2.5 Algorithm validation scheme

2.1.5 Metrics

To assess the validity of the algorithm, both *Pearson correlation coefficient* (r) and the *absolute Error* (aE) between the entire sequence and ECG fiducial points of digital signal and the entire sequence and ECG fiducial points of digitized signal. The Pearson's correlation coefficient, used to test signals similarity, measures the statistical relationship between two continuous variables, using the covariance method [76] . It is defined as follow:

$$r = \frac{n * \sum_{i=1}^{n} y_i * \tilde{y}_i - \sum_{1=1}^{n} y_i \sum_{1=1}^{n} \tilde{y}_i}{\sqrt{[n * (\sum_{1=1}^{n} {y_i}^2)] - (\sum_{1=1}^{n} y_i)^2] * [n * (\sum_{1=1}^{n} {\tilde{y}_i}^2) - (\sum_{1=1}^{n} \tilde{y}_i)^2]}} \quad (16)$$

where y is the desired output (target), $\tilde{y}$ is the predicted values and n is the total number of data. It is ranged in [−1, 1]: $r = 1$ indicates perfect positive correlation between y and $\tilde{y}$; $r = -1$ perfect negative correlation; $r = 0$ no correlation.

Absolute Error is the amount of error in your measurements. It is the difference between the measured value and true value. It is regulated by the followed formula:

$$E = X_{\text{experimental}} - X_{\text{true}} \quad (17)$$

2.1.6 Results

Similarity

The signal similarity represents how close the result of the measurement of a quantity (scanned and digitized signal) is to the true value, in comparison with the reference samples (digital signal).

Table 2.2. Pearson coefficient for each pathology

Pathologies	Pearson coeff
30bpm	0.879820668516074
45bpm	0.944761283387962
60bpm__amplitude_0.5mV	0.925484754099535
60bpm__child	0.688046907949100
60bpm__sano	0.943397146029686
60bpm__ST_0.5mV	0.982079964352914
80bpm__ST_0.2mV	0.914497157149535
100bpm	0.914742016773946
120bpm	0.945884822424869
Atrial_Fibrillation_1c76bpm	0.924478232419820
Atrial_Flutter__82bpm	0.911780065556839
Breath__60bpm	0.968371447342140
Muscle_artifact__60bpm	0.908480703197358
Prem_Atrial_contractions75bpm	0.930018043693811
SupraVentricular_Tachycardia	0.923616360525127
Ventricular_Fibrillation152bpm	0.985244543581722

Figure 2.6 and Figure 2.7 show the similarity between the signals for the normal sinus 60 bpm case. In particular, Figure 2.6 shows the similarity between the scanned signal and the digital one for I, II, II, aVR, aVL and aVF leads. While Figure 2.7 shows the similarity between the scanned signal and the digital one for the precordial leads. As it is possible to notice the morphology is respected and faithfully adheres to the original signal. Table 2.2 illustrates the Pearson coefficient for each pathology. If the signals are highly correlated and superimposable, the Pearson coefficient is close to 1. In the best case (*normal sinus rhythm 60 bpm)* the Pearson coefficient is equal to 0.9821; in the worst case (*child 60 bpm*), the Pearson coefficient is equal to 0.6880. The reason is the morphology of the ECG of children, which is very different from that of adults.

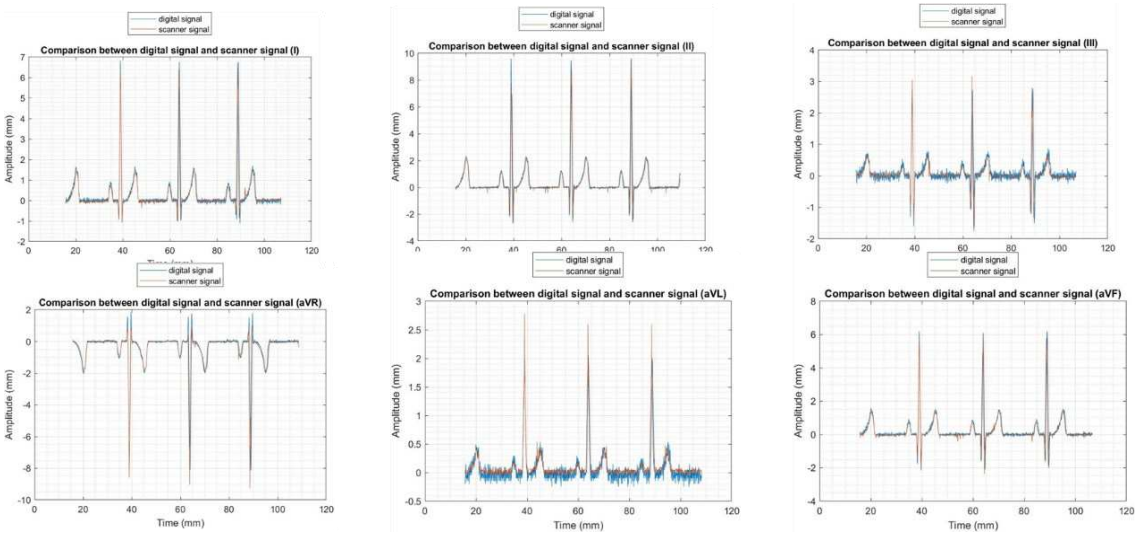


Figure 2.6 Digital and scanner signal comparison for I, II, II, aVR, aVL, aVF leads

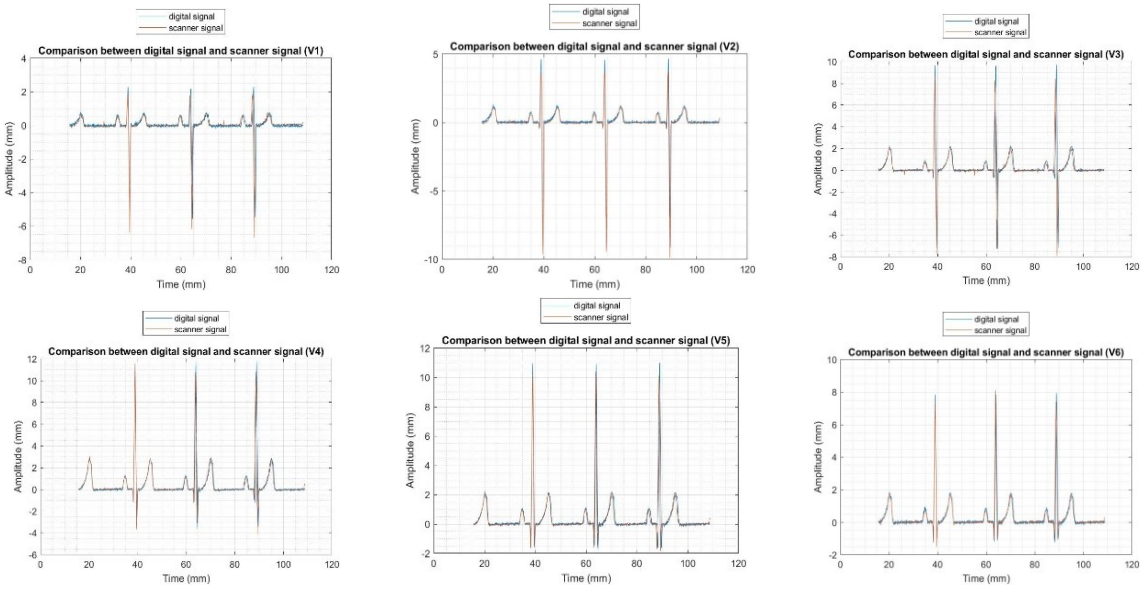


Figure 2.7 Digital and scanner signal comparison for precordial leads

Repeatability

In order to evaluate the repeatability, an image of the data set (normal sinus rhythm and 60 bpm) is chosen, and 12 leads crops are done for ten times on the same image. Since SF is slightly dependent on the made crop (causing a variation in the time length of the signal), each of the 10 times SF was collected and compared with the others. The variation of fiducial points was also analysed, by calculating mean and standard deviation of each point. Below, there is an example of how the mean heart rate is calculated:

$$HR_m = \frac{1}{10} \cdot \sum_{n=1}^{10} HR_m \quad (18)$$

where $n = 1,2, \ldots,10$ indicates the repetitions number and HR_m is the heart rate obtained with $m = 1,2\ldots, 12$ crop. The same was done for the other parameters.
Considering repeatability, *SF* is equal to (on average) 0.0430 ± 0.000524, with a minimum value of 0.042374 and a maximum value of 0.043802. Instead, the values of fiducial points, expressed with mean and standard deviations, are: 102.98 ms ± 7.95 ms (QRS complex); 369.65 ms ± 11.22 ms (QT distance); 176.38 ms ± 9.69 ms (PQ distance); 108.75 ms ± 1.34 ms (P-P points distance); 1012.18 ms ± 12.09 ms (R-R peaks); 59.30 ms ± 0.72 ms (heart rate). All the parameters are shown in Table 2.3Table 2.3. Last row has the values written on the original graph paper. During the 10 repeatability's tests, the maximum variation of SF was 0.01428.

Table 2.3. Variation of scale factor and fiducial points obtained by cropping 10 times the same ECG image (Normal sinus rhythm, 60 bpm).

SF	QRS complex (ms)	QT distance (ms)	PQ dist. (ms)	P-P dist. (ms)	R-R peaks (ms)	Heart Rate (bpm)
0.042829	96.51	360.33	183.31	108.50	1008 .62	59.49
0.043802	98.70	368.52	187.47	110.96	1031.53	58.17
0.042374	110.17	375.72	163.28	106.78	997.49	60.15
0.043552	113.82	386.75	167.24	109.75	1025.22	58.52
0.042838	97.10	361.55	182.21	108.52	1008.84	59.47
0.043690	113.59	387.39	168.35	110.10	1028.46	58.34
0.042555	95.89	358.03	182.14	107.81	1002.18	59.87
0.042388	110.77	376.41	162.77	106.82	999.78	60.13
0.042932	96.74	361.41	183.75	108.76	1011.05	59.34
0.042833	96.51	360.35	183.31	108.50	1008.65	59.49
-	88	368	164	86	1000	60

Reproducibility

In this case, the test was performed by using an image (the chosen patient is the same as repeatability) with two different formats: one is the JPEG produced by the electrocardiograph; the other is another JPEG with a different structure of graph paper (Figure 2.8). The resolution of the two formats is different.

Therefore, after the crops on the two images, SF and the fiducial points were extracted and compared.

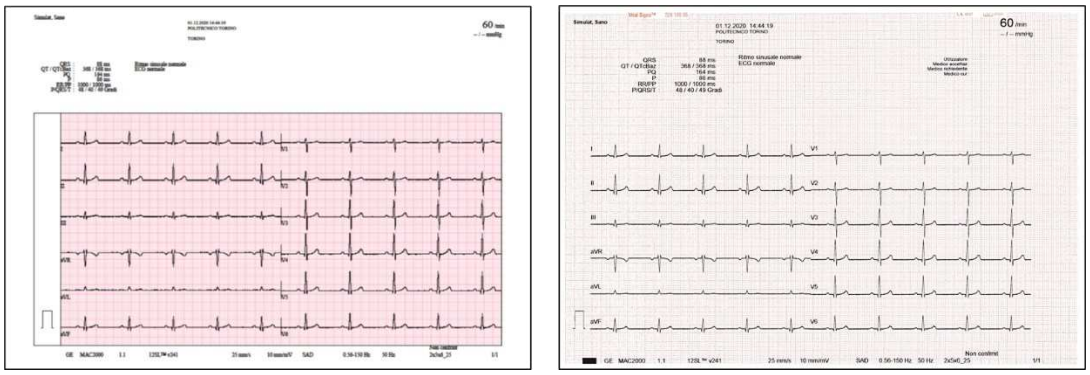


Figure 2.8 Reproducibility on normal person with 60 HB

Table 2.4Table 2.4 shows the comparison of scale factor and fiducial points obtained by cropping two versions of the same ECG image (normal sinus rhythm, 60 bpm). The bigger variation of SF (0.130231) is caused by the different images resolution. Besides, the maximum variation of the parameters is 22.45 ms (for R-R peaks), which corresponds to a difference of 0.56 mm on the graph paper. In this way, the heart rate calculation differs by 1.33 beats. The extent of each of these errors can be tolerate by cardiologists and considered acceptable according to our study. Indeed, it is in accordance with what we expected because of printing and reconstructions errors, i.e., an error < 1 mm.

Table 2.4 Comparison of scale factor and fiducial point for the same image in two formats

Parameters (mm)	JPEG image (1st version)	JPEG image (2nd version)	Absolute Error (ms)
SF	0.042285	0.172516	0.130231
QRS complex (ms)	110.51	101.21	9.3
QT distance (ms)	375.49	368.04	7.45
PQ distance (ms)	162.38	163.32	0.94
P-P distance (ms)	106.56	89.71	16.85
R-R peaks (ms)	995.40	1017.85	22.45
Heart Rate (bpm)	60.28	58.95	1.33

2.1.7 ECG-dig Graphical Interface

This section will present the graphical interface implemented as a support tool for the physician. The steps of the program are the following:

1) Open the ECG-dig software and save the patient data with the Fiscal Code which represent a unique key for the database construction (Figure 2.9).

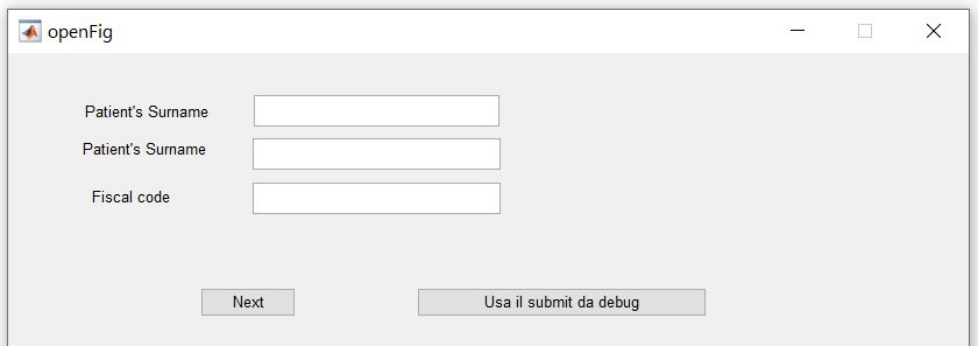


Figure 2.9 Opening figure

2) Select the scanned paper-based ECG (Figure 2.10)

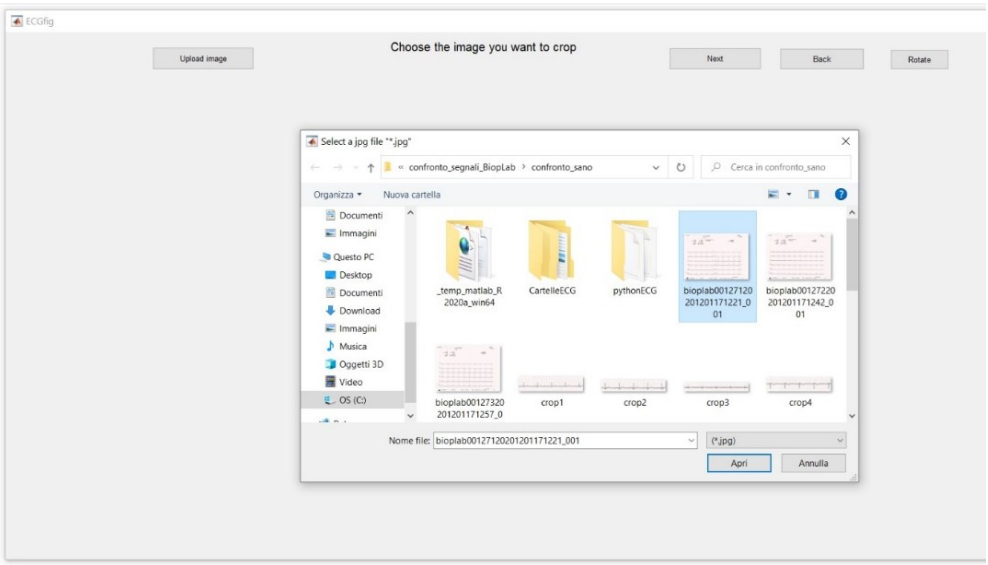


Figure 2.10 Figure selection

3) Open the scanned ECG image in the ECG-dig tool

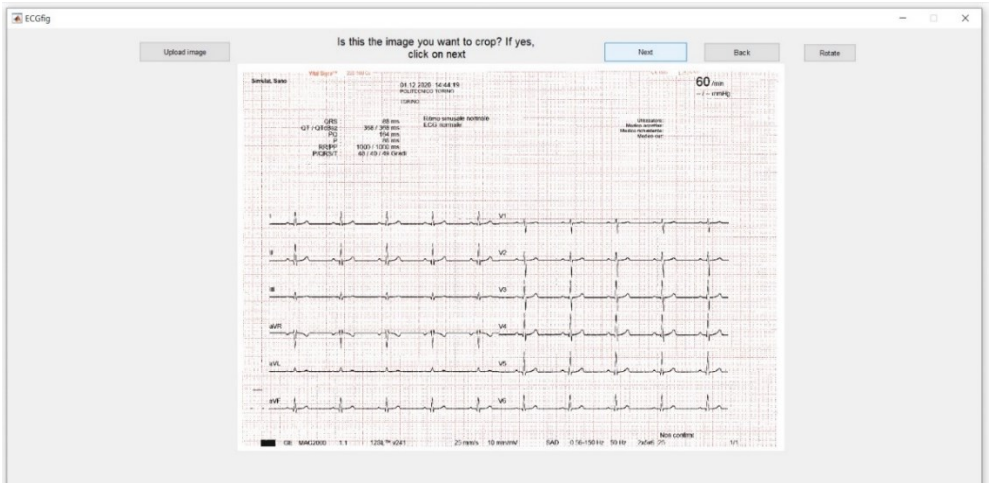


Figure 2.11 Opening ECG-image in ECG-dig tool

4) Make the crop for each lead (Figure 2.12).

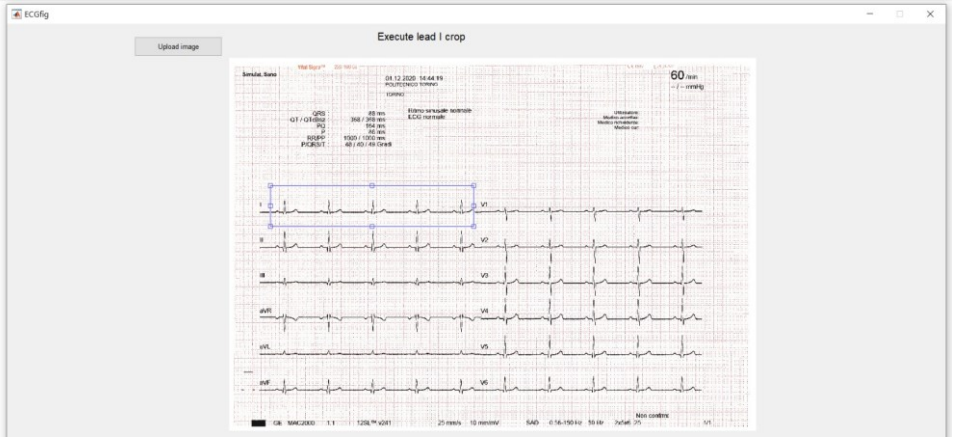


Figure 2.12 Lead I crop

5) Each crop is displayed individually, saved, and numbered (Figure 2.13).

Figure 2.13 Displayed Lead I

6) This represents the software output, the reconstructed and digitized signals are placed in a graph and displayed respecting the real proportions (Figure 2.14).

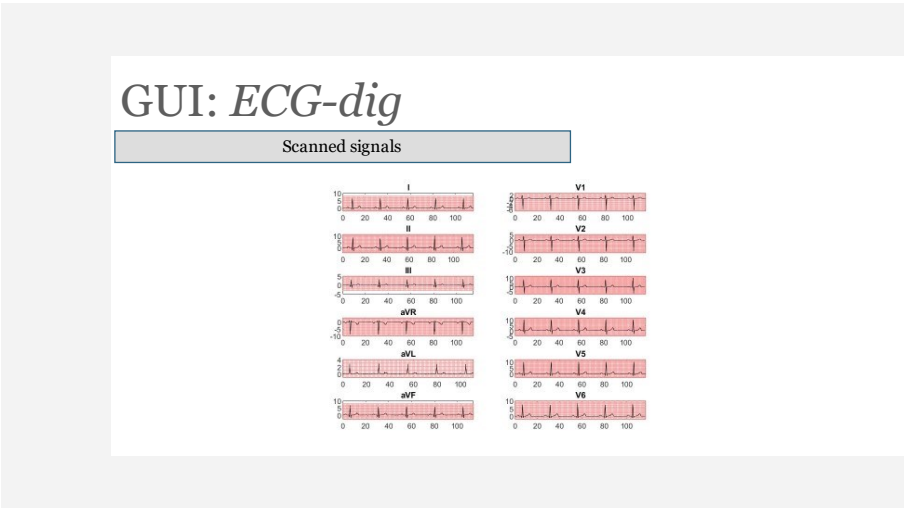


Figure 2.14 Digitized signals

7) For each patient, a folder with name and surname will be created in which both the images of each lead and an excel file that saves each lead in digital format will be saved (Figure 2.15).

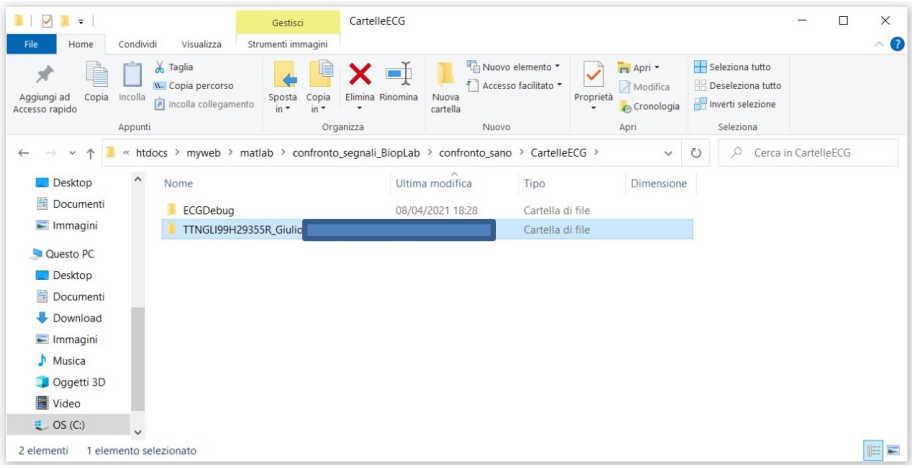


Figure 2.15 ECG folder

8) This is an example of the reconstructed V5 lead (Figure 2.16).

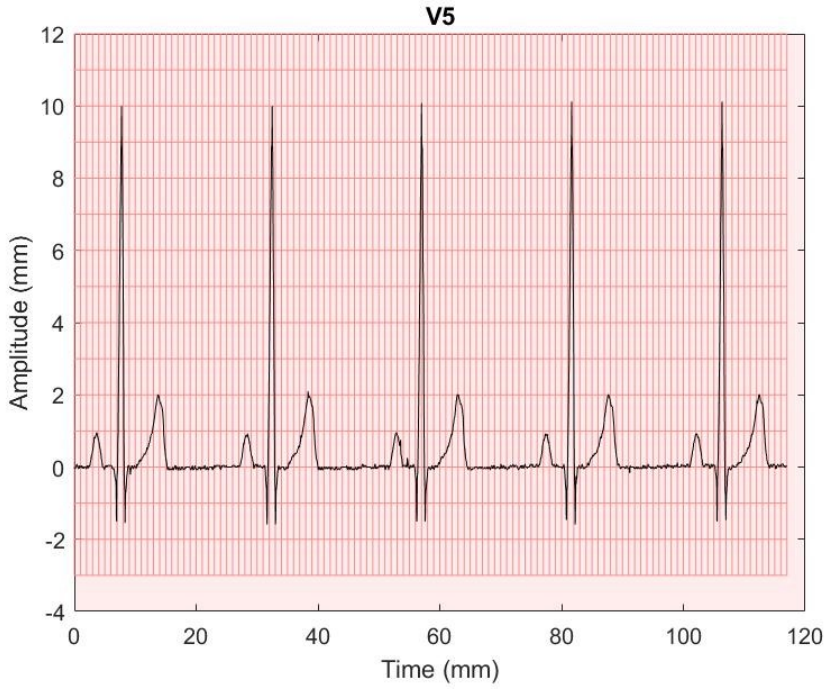


Figure 2.16 V5 lead digital signal

9) This is the .xls file (Figure 2.17).

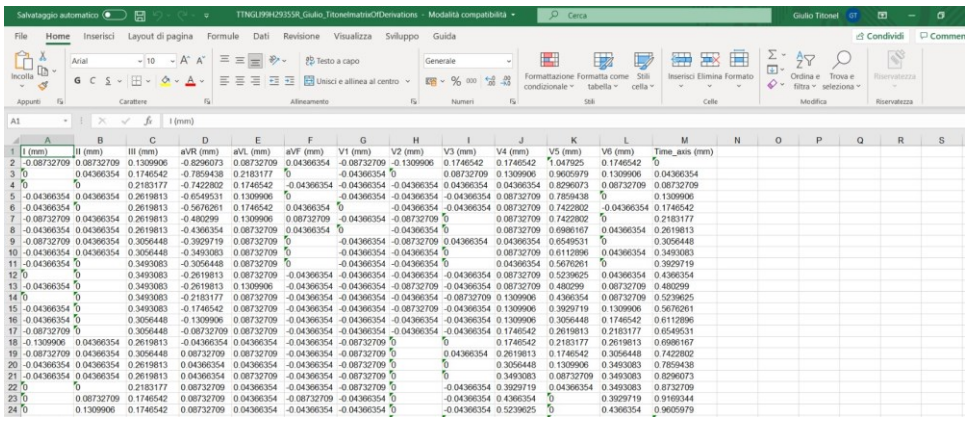


Figure 2.17 Digital signal in .xls format

10) The final contents of the folder are shown in Figure 2.18.

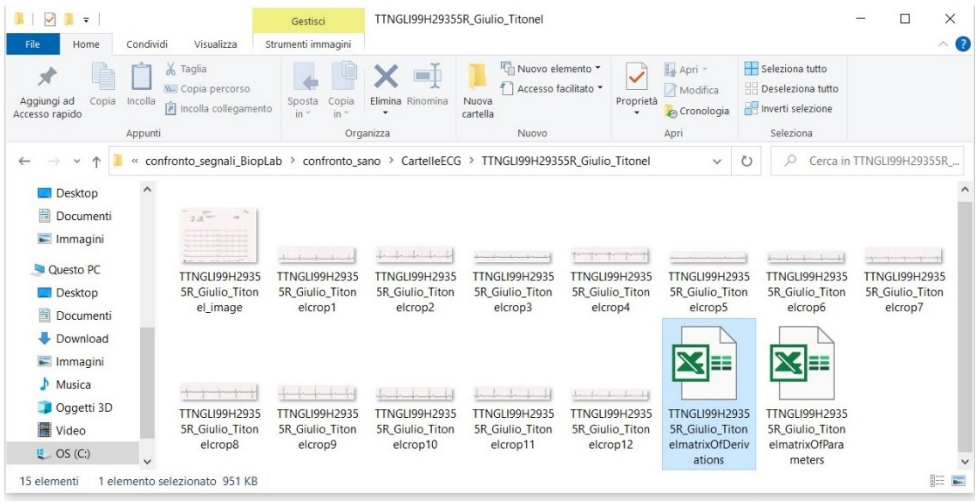


Figure 2.18 Folder contents

Table 2.5 summarizes the basic steps that characterize the software implemented for storing digital ECG traces.

Table 2.5 ECG-dig tool fundamental steps

Step 1	Saving patient data
Step 2	Scanned paper-based ECG selection
Step 3	ECG image in the ECG-dig tool display
Step 4	Crop for each lead
Step 5	Saving and Numbering each crop
Step 6	Software output representation
Step 7	Saving folder with name and surname
Step 8	Images with reconstructed signals
Step 9	Saving Excel files
Step 10	Showing final contents of the folder

2.2 Channelopathies: Short-QT

Short QT syndrome (SQTS) is an inherited arrhythmic disorder due to mutations in cardiac ion channels (cardiac channelopathies), characterized by the propensity to develop life-threatening ventricular arrhythmias in young and otherwise healthy individuals. An important difficulty related to the syndrome consists in correctly assessing the arrhythmic risk during the first contacts with patients. It is well known that, even among patients who share the same mutation that causes the disease, clinical manifestations can range from no symptoms to repeated cardiac arrest and sudden death (SD) [77]. So, the short QT syndrome constitutes a new clinical entity that is associated with a high incidence of sudden cardiac death, syncope, and/or atrial fibrillation even in young patients and new-borns [78]. This project focuses on the use of the electrocardiogram (ECG) signal, for the prediction of syncope, sudden death, and heart attack. Track analysis powered by artificial intelligence (AI), has recently gained tremendous momentum, showing promising results in the medical field. This research project aims to develop artificial neural network (ANN) algorithms for AI-enhanced ECG analysis in patients with cardiac channelopathies, to help refine risk stratification, particularly in asymptomatic patients. The analysis was carried out extracting features from ECG: RR interval, cardiac frequency, QT interval, corrected QT, QRS complex distance, J-Tpeak distance, J-Tend distance, Tpeak-Tend distance, predicted QT and QT/QTp.
This algorithm could be seen as a clinical support which will allow cardiologists to input ECGs and clinical data to obtain a patient-specific risk estimate based on AI-powered ECG analysis. Finally, to test the generalizability of our results, the models will be externally validated on independent datasets.

2.2.1 Short QT Definition

The risk of death and of presenting an event is highly correlated with the length of the QT segment of the ECG signal [79]. In 1993 Algra et al. observed that both prolonged and shortened corrected QT (QTc) intervals were associated with an increased risk of an event relative to intermediate QTc values [80]. An event was defined as the occurrence of SCD, major ventricular arrhythmias (ventricular tachycardia or ventricular fibrillation, requiring both resuscitation or defibrillation) and/or arrhythmogenic syncope. In 2003 the short QT syndrome (SQTS) was recognized as a new clinical entity related to the familiarity with autosomal dominant inheritance [78]. Although the upper limit of normal QT values is well defined, the lower limit has not yet been determined. Accordingly to European Society of Cardiology guidelines [81], diagnosis of SQTS can be established in the presence of a QTc $\leq$ 340 ms (or a QT/QTp ratio $\leq$ 88%), however should be considered in the presence of a QTc $\leq$ 360 ms and one or more of the following: confirmed pathogenic mutation, family history of SQTS, family history of SD at age 40 years; survival from a ventricular fibrillation in absence of

heart disease. Figure 2.19 shows the difference between normal QT interval and short QT interval.

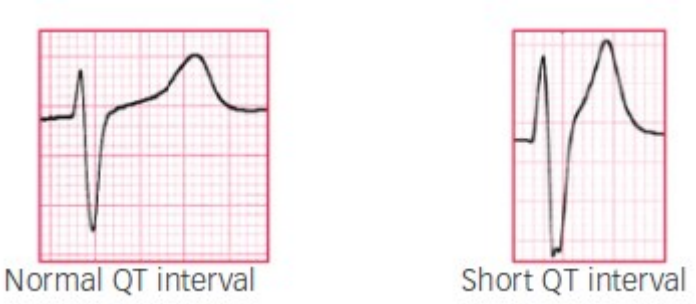


Figure 2.19 Normal QT interval and Short QT interval

2.2.2 State of the Art

Channelopathies are disorders caused by mutations in genes that code for cardiac ion channels and / or their regulatory proteins, which result in electrocardiogram (ECG) repolarization abnormalities that eventually lead to electrical instability and life-threatening ventricular arrhythmias in the absence of structural heart disease, especially in young and otherwise healthy people. Cardiac arrest is often the first clinical presentation in previously asymptomatic subjects with these syndromes. In SQTS, 30-40% of patients have a clinical presentation with sudden death (SD) or interrupted SD [82] [83]. The data mentioned above suggest the vital importance of predicting SD risk in these patients, as adequate assessment of arrhythmogenic potential can lead to improved treatment decisions that can potentially prevent premature death of young individuals. However, risk stratification in asymptomatic patients, who represent most of these subjects, is still far from satisfactory. Risk scores derived with traditional statistical techniques, such as logistic or Cox regression showed suboptimal results in their accuracy [84] [85]. The use of AI in medicine is relatively recent compared to other fields, but is rapidly gaining widespread interest due to the high expectations in terms of improving healthcare and reducing its costs [86]. In particular, the application of AI in ECG analysis and signal analysis has recently gained tremendous momentum.

However, despite its invaluable potential, AI has never been applied to risk stratification of sudden death in patients with Short-QT. This is essentially since Short QT syndrome is so rare that it is considered "familiar" and therefore not very well understood. The most precious resource for artificial intelligence systems to be efficient is the amount of data used to train AI systems: the greater the number of data, the greater the performance of the neural network used.

The algorithm implementation is intended to help clinicians stratify the risk of life-threatening arrhythmias in patients with cardiac channelopathies from a simple, readily available, and inexpensive tool such as the 12-lead ECG.

2.2.3 Shallow Learning for Event Prediction System

Shallow learning techniques are learning algorithms that learn the parameters of the statistical model directly from the characteristics of the examples in the training dataset. Most supervised machine learning algorithms belong to this category. The features extraction in Shallow Machine Learning is a manual process that requires domain knowledge of the data that we are learning from.
Deep Learning is a sub-class of Machine Learning algorithms whose peculiarity is a higher level of complexity. Shallow learning has a lower level of complexity because it works with a little amount of data. The goal of this research is a regression task: learning occurs using numerical labelled data to predict a quantity of an input.
Shallow neural networks are a type of neural network with a limited number of hidden layers. Understanding a superficial neural network gives us an idea of what exactly is going on inside a deep neural network. Figure 2.20 Shallow learning neural network scheme that shows a surface neural network with one hidden layer, one input layer and one output layer.

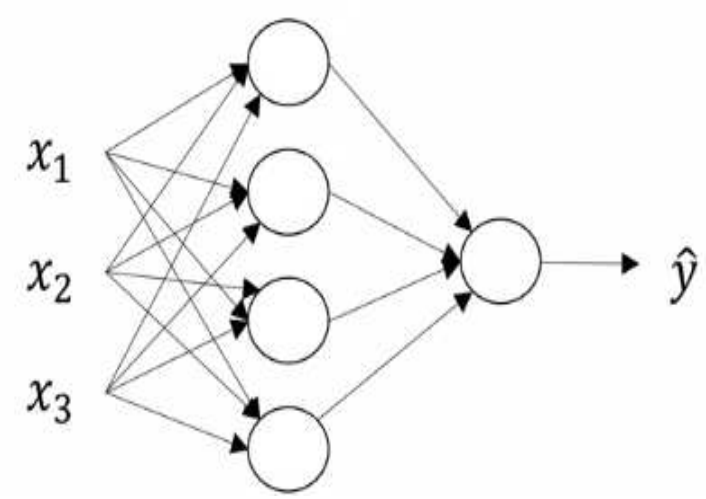


Figure 2.20 Shallow learning neural network scheme

2.2.4 Study Populations

The study group included a total of 155 subjects. 45 patients were symptomatic at presentation: 4 had died suddenly, 21 had an aborted SD, 3 had pre-syncope, 10 had syncope and 6 had palpitations (2 of whom had documented atrial fibrillation or flutter). In the following period, 23 patients, 9 of whom previously asymptomatic, reported the following symptoms: pre-syncope (n = 5), syncope (n = 5) and palpitations (n = 13). Overall, 39 patients developed an event, both at presentation and/or during follow-up. 94 patients presented with relevant familial history: 58 had familiarity for both SQTS and SCD, while the remaining showed familiarity only for SCD (n = 11) or SQTS (n = 25). Genetic test was performed in 82 patients, although pathogenetic mutations were found only in 43 patients, with most frequent genes affected being KNCH2 and KCNQ1. Overall, 38

patients developed a major arrhythmic event (i.e., SCD and/or arrhythmic syncope), both at presentation and/or during follow-up, while 117 did not. Table 2.6 summarizes the study population described.

Table 2.6. Study population description

Variables	N = 155
Family history, No. (%)	**94 (60,6)**
SCD **SQTS** **SCD and SQTS**	11 (7,1) 25 (16,1) 58 (37,4)
Symptoms at presentation, No. (%)	**44 (28,4)**
SD **aSD** **Pre-syncope** **Syncope** **Palpitations**	4 (2,6) 21 (13,5) 3 (1,9) 10 (6,5) 6 (3,9)
Symptoms during follow-up, No. (%)	**23 (14,8)**
Pre-syncope **Syncope** **Palpitations**	5 (3,2) 5 (3,2) 13 (8,4)
Event occurrence, No. (%)	**38 (24,5)**

2.2.5 Dataset Description

For each patient, data regarding both personal and family history were recorded. 12-lead ECGs with a paper speed of 25 and 50 mm/s and a gain of 10 mm/mV were collected. ECG parameters were measured independently by 3 expert cardiologists from lead V2 and from the lead with the highest T-wave amplitude (namely Vh; usually ranging from V2 to V5).
The features put into the network to predict the percentage of the occurrence of an event are the following:

- *RR interval (ms):* measured from the peak of an R-wave to the next one. It expresses the duration of a complete cardiac cycle.
- *HR (bpm):* derived from RR interval. It expresses the discharge frequency of the dominant pacemaker (usually sinus node):

$$HR = \frac{1}{RR(sec)} \tag{19}$$

- *QT interval (ms):* measured from first deflection of QRS complex to the end of T-wave (defined using tangential method). QT interval expresses

global duration of ventricular electrical activity, although used almost exclusively to evaluate ventricular repolarization.

- *QTc interval (ms):* calculated with Bazett's formula [87]. QT correction formulas are necessary, as QT duration varies considerably with HR:

$$QTc = \frac{QT}{\sqrt{RR}} \tag{20}$$

- *QRS (ms):* measured from first deflection of QRS complex to the beginning of the ST segment (Jpoint). It expresses the duration of ventricular depolarization.
- *Jpoint-Tend (ms):* measured from J point (junction between the termination of the QRS complex and the beginning of the ST segment) to end of T-wave (defined using tangential method).
- *Jpoint-Tpeak (ms):* measured from J point to the peak of the T-wave, representing early repolarization.
- *Tpeak-Tend (ms):* measured from the peak of the T-wave to its end (defined using tangential method). This interval as emerged as a correlate of global dispersion of repolarization. J-Te, J-Tp and Tp-Te values have been also corrected with Bazett's formula.
- *QTp (ms):* predicted QT value using Rautaharju et al. formula [88]:

$$QTp = \frac{QT(120 + HR)}{180} \tag{21}$$

- *QT/QTp* (%): ratio between measured and predicted QT interval.
- *T-wave amplitude (mV):* measured from isoelectric line to T-wave peak.

Figure 2.21Figure 2.21 shows the characteristic points for the definition of the features

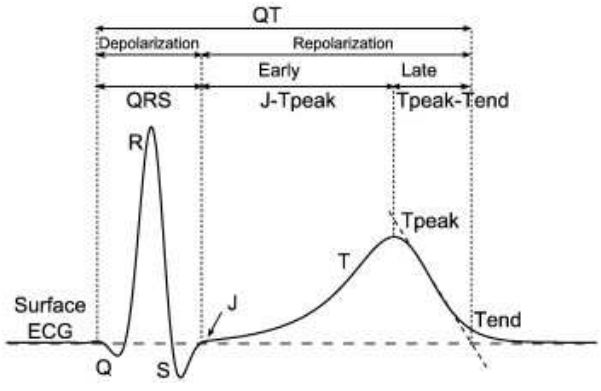


Figure 2.21 ECG signal fiducial points

2.2.6 AI Model Building and Internal Validation

Due to the dataset size and the use of human-engineered features, it was chosen to use a shallow learning model based on neural networks.

Neural networks are modelled on the human brain functions, designed to recognize the relationship between the input and the labelled output (target) [89]. This AI model is based on a standard multi-layer perceptron (MLP) configuration. The main layers of this system are the input layer, the hidden layer and the output layer. The *input layer* works as an entry point to the neural network. It consists of passive nodes, which do not modify the input, but only transmit the information to each neuron of the subsequent layer (also known as fully connected). The *hidden layer* has an arbitrary number of neurons, which depends on the complexity of the problem at hand. Each hidden node combines the information received from each unit of the input layer to achieve a complex representation of the phenomenon under investigation. At this purpose, a non-linear activation function is employed, such as the hyperbolic tangent sigmoid. At the end, the *output layer* yields the input data classification by means of the softmax function [90]. A feed-forward fully connected neural network with one hidden layer has been designed using MATLAB® R2020b [91] and trained on a single CPU of LENOVO Y50-70 workstation with 16 GB RAM. To compute the best performance several hidden layer sizes have been tested using a *trial-and-error* approach. The best performing architecture has 10 and 1 neurons in the hidden and output layers, respectively, while the input layer size depends on the experiment set up.

Hidden units were equipped with hyperbolic tangent sigmoid transfer function, while the output layer used softmax to yield classification. The network training was performed using the scaled conjugate gradient (SCG) technique to minimize the cross-entropy error function. In order to guarantee the quality of input data, avoid bias and reduce noise the input dataset has been processed during the network training.

Many machine learning algorithms attempt to find trends in the data by comparing features of data points. For this reason, data have been statistically normalized (Z-score) so that the network was able to intrinsically determine each input feature importance for classification.

Z-score normalization is a strategy of normalizing data that avoids this outlier issue[92]. The formula for Z-score normalization is below:

$$Z - score = \frac{value - \mu}{\sigma} \tag{22}$$

Here, μ is the mean value of the feature and σ is the standard deviation of the feature. If a value is exactly equal to the mean of all the values of the feature, it will be normalized to 0. If it is below the mean, it will be a negative number, and if it is above the mean it will be a positive number. The size of those negative and positive numbers is determined by the standard deviation of the original feature. If the unnormalized data had a large standard deviation, the normalized values will

be closer to 0. Without this step, it would have been possible that some features masked some others, preventing the network to understand the real contribution of each input attribute to SQTS.
Both the input and target have been randomly divided into three sets as follows: 70% for training; 15% to validate that the network is generalizing and to stop training before overfitting; and the remaining 15% to independently test the network classification performance. To ensure that the distribution of the input data (equally distributed amount of event and non-event cases) was preserved across the three sets, random splitting was performed separately for the event and non-event subsets. Indeed, due to a strong target classes imbalance (117 without event and 38 with event), it was necessary to fictitiously increase the amount of data with event. At this purpose, there exists several possible data augmentation techniques, such as interpolation or data replication. Since they were medical records, it was chosen to use the latter approach to avoid introducing fake data in the cohort, which could have misled result interpretation. Therefore, the event data were replicated three times.

2.2.7 Metrics

The classification performance was estimated analysing the confusion matrices and the associated AUC.
The former measures the number of times the network correctly classifies the input; in this sense, it yields, an estimate of how much a single class, i.e., a medical condition (event/non-event), was understood by the neural model. Therefore, to better analyse the network performance, also the True Positive rate (sensitivity) and the True Negative rate (specificity), the false positive ratio (probability of false alarm) and the false negative ratio (probability of no alarm) were computed.

- *True Positive (TP)* is the number of correct predictions that an example is positive which means positive class correctly identified as positive.
- *False Negative (FN)* is the number of incorrect predictions that an example is negative which means positive class incorrectly identified as negative.
- *False positive (FP)* is the number of incorrect predictions that an example is positive which means negative class incorrectly identified as positive.
- *True Negative (TN)* is the number of correct predictions that an example is negative which means negative class correctly identified as negative.

The advanced classification metrics based on confusion matrix are mathematically expressed as follow:

- *Sensitivity* is also referred as True Positive Rate or Recall. It is a measure of positive examples labelled as positive by classifier. It should be higher. For instance, proportion of emails which are spam among all spam emails. In medicine, highly sensitive tests are generally used for screening purposes, due to their ability to rule out the disease/event occurrence.

- *Specificity* is also known as True Negative Rate. It is a measure of negative examples labelled as negative by classifier. There should be high specificity. For instance, proportion of emails which are non-spam among all non-spam emails. In medicine, highly specific tests are typically used for confirmation purposes, due to their ability to rule in the disease/event occurrence
- Positive predictive value (PPV), also known as Precision: the ratio between the total number of correctly classified positive examples and the total number of predicted positive examples. It yields the correctness achieved in positive prediction.
- Negative predictive value (NPV): the ratio between the total number of correctly classified negative examples and the total number of predicted negative examples. It yields the correctness achieved in negative prediction.
- Accuracy is the proportion of the total number of predictions that are correct.

Figure 2.22 shows the key to understanding a confusion matrix.

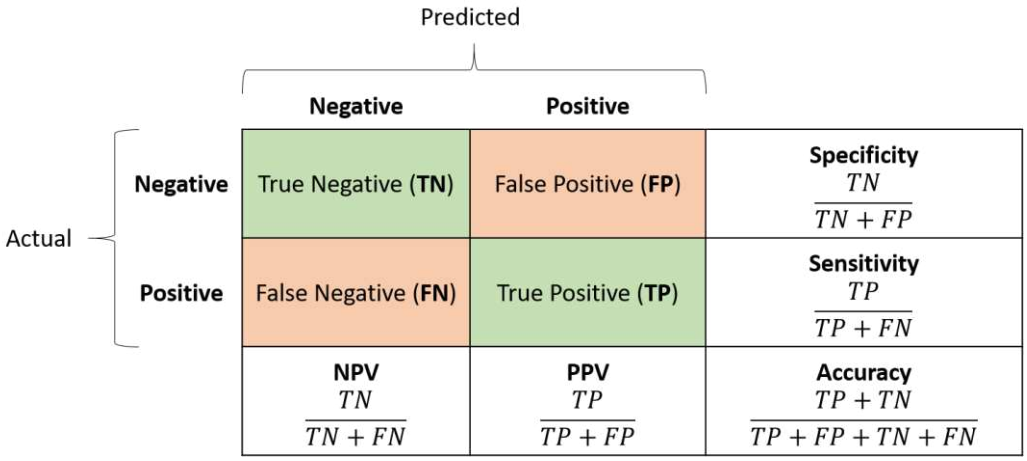


Figure 2.22 Confusion matrix example: rows yield the real (actual) labels, columns the predicted ones, i.e., the network output.

Despite accuracy provides a single global measure of the classification quality, it is just an average value of the network performances. On the contrary, the area under the ROC curve (AUC) yields a more precise measure (the higher the best) of the predictive accuracy because it represents the probability that a randomly chosen positive sample is ranked higher than a corresponding negative one.

2.2.8 Results

The classification ability of the proposed neural system has been tested on different input configurations, i.e. different input features, in order to study which features were the most relevant to correctly discriminate among subject who will have an event from those who will not. In this sense, it was investigated the importance of the QT interval and the T wave in distinguishing the two classes

(i.e., non-event/event). Therefore, the experiments could be grouped into four categories as per:

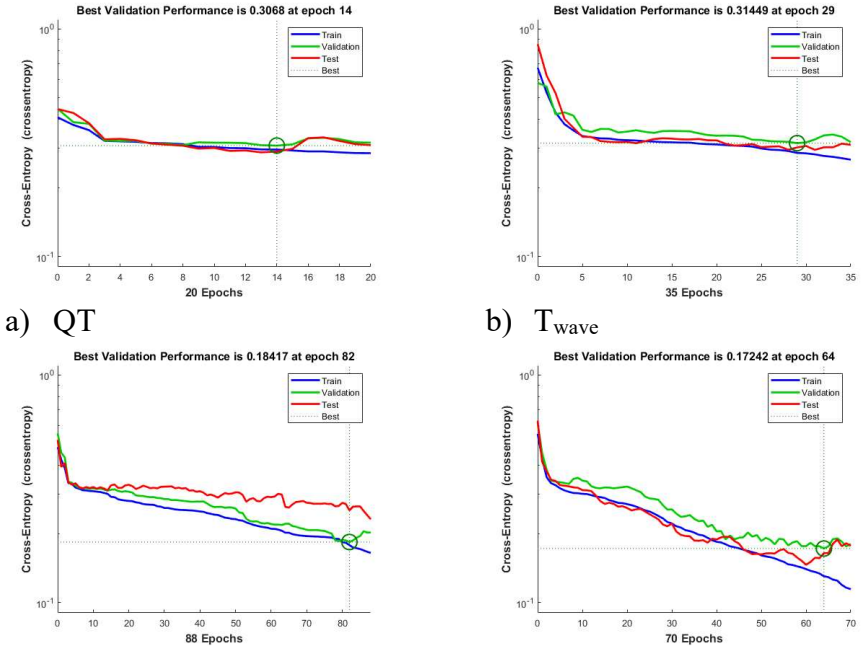
- QT: only the QT related features were considered,
- Twave: only the T wave features were considered,
- QT + Twave: both the QT related and T wave features were considered,
- All: all the input features were considered.

Table 2.7 shows the input dataset taxonomy.

Table 2.7. Input dataset taxonomy

	Datasets			
Feature	**QT**	**T_{wave}**	**QT + T_{wave}**	**All**
RR (ms)	✓		✓	✓
HR (bpm)	✓			✓
QT (ms)	✓		✓	✓
QTc (ms)	✓		✓	✓
QTp (ms)	✓			✓
T_{amp} (mV)	✓	✓	✓	✓
QRS (ms)			✓	✓
J-T_p (ms)		✓	✓	✓
T_p-T_e (ms)		✓	✓	✓
J-T_e (ms)		✓	✓	✓
cJ-T_p (ms)				✓
cT_p-T_e (ms)				✓
cJ-T_e (ms)				✓
QT/QTp				✓

Figure 2.23 shows the error over time (epochs) for the training, validation, and test subsets. The QT (Figure 2.23.a) and Twave (Figure 2.23.b) behaviour is quite similar: the best performance is around 0.31 and the error is coherent in the three subsets; in this sense the networks seems to have learnt the phenomenon at hand. On the contrary, despite the best performances reach similar values in Figure 2.23.c (~ 0.18) and Figure 2.23.d (~ 0.17), the network behaviours are quite different, as clearly shown by the three curves. The network trained on the QT + Twave dataset learn properly the training set (blue and green curves are superimposed), while slightly worsen the test set (red curve stays above the best performance intersection point). Figure 2.23.d exhibits an opposite behaviour: the network validation performance is much closer to the test one rather than training, which keeps the lowest error values of all the experiments.



a) QT b) T_{wave}

c) QT + T_{wave} d) All

Figure 2.23. Error over time for training, validation, and test subsets for the four input sets: QT (a), Twave (b), QT + Twave (c), All (d)

In order to deepen the network classification performances, a more specific analysis was made by means of sensitivity, specificity, PPV, NPV, and accuracy. The results can be summarized as follows (Table 2.8):

- Sensitivity: it is identical in the *QT* and T_{wave} cases, while increases up to 80.0 % and 100.0 % for $QT + T_{wave}$ and *All* datasets, respectively. In the latter case, it means the network always predicts an event when it occurs.
- Specificity: except for the *All* dataset, this metric is always lower in training than in test and never above 90 %. In this sense, the negative predictive power should be further investigated.
- PPV: the first two input configurations do not allow the network to properly predict an event, while, when combining the QT related features with the T wave shapes, it raises above 80 %.
- NPV: in case of the fourth and most complete configuration (All), a negative output from the network has up to 100% probability of being correct; in other words, it means that the network does not predict an event when it does not occur.
- Accuracy: w.r.t. precision, the first two datasets reach good but not satisfactory performances, while the other two exhibit a much higher classification quality even on the test subset (82.9 % and 88.6 %).

Table 2.8 Classification performances: sensitivity, specificity, PPV, NPV, accuracy

	QT		***T_{wave}***		***QT + T_{wave}***		***All***	
	Training	Test	Training	Test	Training	Test	Training	Test
Sensitivity	73.8 %	60.0 %	73.4 %	63.2 %	92.7 %	80.0 %	94.9 %	**100.0 %**
Specificity	69.1 %	85.0 %	78.0 %	81.3 %	81.0 %	**85.0 %**	88.0 %	76.5 %
PPV	70.2 %	75.0 %	76.3 %	80.0 %	83.5 %	80.0 %	88.1 %	81.8 %
NPV	72.7 %	73.9 %	75.3 %	65.0 %	91.4 %	85.0 %	94.8 %	**100.0 %**
Accuracy	71.4 %	74.3 %	75.8 %	71.4 %	87.0 %	82.9 %	91.3 %	88.6 %

Table 2.9 reports the AUC values for the four datasets. As before, the QT and Twave datasets have similar performances but not enough to properly discriminate between the two classes. The QT + Twave behaves quite well, reaching an impressive 0.87 in test, while the dataset made by all features further improves this value.

Table 2.9. Classification performances: AUC

	QT		***T_{wave}***		***QT + T_{wave}***		***All***	
	Training	Test	Training	Test	Training	Test	Training	Test
AUC	0.72	0.70	0.76	0.73	0.92	0.87	0.96	0.90

2.2.9 Discussion

In this study, the objective is the estimation, given the patient's ECG, of the probability of presenting an event. To this end, considering the small number of data made available (since short-QT disease is present in very few individuals in the world), a shallow learning system was trained.
The neural network was developed to perform event and non-event classification.
To compare the performance of the different network configurations, the study was divided into four macro areas. In the first case the performances are related to a neural network with input features exclusively related to the duration of the QT and of all the measurements concerning this measurement (RR, HR, QRS, QTc, QTp, QTc / QTp). In the second case, instead, the features related to the T wave (J-Tpeak, J-Tend, Tpeak-Tend) were exclusively analysed, the network is the consequence of some observations by cardiologists according to which patients with short-term QT they would present during the disease a little more accentuated than the height of the T wave with consequent shortening of the J-Tend interval. The third configuration considers both the QT related and T wave features. The fourth configuration, which achieved the best performance, all the input features were considered.

2.3 Conclusion

The first part of this project is based on the implementation of a Matlab-based graphic interface for the digitization of ECG traces. The need for digitization arises for two fundamental reasons: the digital archiving of the traces to avoid deterioration and the use of digital signals for research purposes based on neural networks. The GUI developed in this research project, ECG-dig, can be widely adopted as an aid to medical diagnosis. The GUI therefore represents a single digital container, always updated, which represents the patient's medical record. The medical support tool is designed to maximize the impact of the proposed research and make its dissemination as pervasive as possible.

The second part of the project concerns the realization of a predictive neural system, based on a shallow learning approach, to estimate the probability of the occurrence of a cardiac event. The implantable cardioverter defibrillator (ICD) still represents the mainstay of treatment for SQTS patients who have survived SD or who have documented spontaneous sustained ventricular tachycardia (VT), despite the significant risk of device-related complications. For this reason, in this work, a non-invasive system was created to monitor SQTS cases. The results suggest that AI-based ECG analysis could help refine risk stratification in SQTS patients, supporting clinical decision making.

The main part of the research is centered on the detection of the QT-short pathology given the fiducial points of the ECG by means of a shallow learning algorithm. With the second part, research is enriched by implementing a platform for digitizing the ECG signal from the printed ECG image.

Future studies should focus on the automatic calculation of the characteristics of digital ECG recordings. This will ensure greater reproducibility than manually extracting relevant ECG features. In parallel, the supervised deep learning model that evaluates the entire row digital ECG signal should also be explored, as it could potentially show greater accuracy than a features extraction approach.

2.4 Acknowledgements

This work has been supported by the Molinette Hospital and by Professors and Cardiologists Fiorenzo Gaita and Carla Giustetto. A special thanks also to the PhD student Andrea Saglietto and Pierre Maynet.

Chapter 3

An Automatic Defect-detection System Based on Hybrid Unsupervised and Supervised Machine Learning for Electrospun Nanomaterials

3.1 Nanofibers in Industry

The production of nanomaterials by the the electrospinning process requires a detailed inspection of related scanning electron microscope (SEM) images of the electrospun nanofiber, to ensure that no structural abnormalities they are produced. The presence of defects prevents the success of applications of electro-spun nano-fibrous material in nanotechnology. Hence, automatic tracking and quality the control of nanomaterials is an objective challenge in the context of Industry 4.0. In the last decade, nanostructured materials have gained interest in both scientific research and industrial contexts, for their versatility. The combination of nanotechnology and information and communication technologies (ICT) represent the frontier of the fifth industrial revolution: in fact, the small size of products can facilitate the automation of tasks, previously limited by physical restrictions [93].

Nanomaterials working at molecular level, allow generating wide structures endowed with different properties, useful to improve quality of life in different areas [94]. In biomedical engineering, the huge progress of nanomaterials research suggests that they could yield interesting alternative solutions to many healthcare problems. In tissue engineering application, nanofibers are used for the reproduction of tissue architecture at the nanoscale, thus giving an impulse to wearable applications. Nanofiber materials act as excellent structures for

adhesion, proliferation and cells scaffolding differentiation for musculoskeletal, skin, vascular and tissue engineering and as potential vectors for the controlled delivery of proteins and DNA [95]. Among many nanofibers synthesis techniques, electrospinning appears to be the most promising technology able to meet these industrial objectives. Electrospun fibers can indeed be applied to study drug delivery, encapsulating the therapeutic agent in the fibers and maintaining the integrity and bioactivity of molecules due to slight processing parameters. Indeed, as the drug release depends on the degradation of the polymer fibers, it can be adequately controlled. In bioengineering, the nanofibers could allow to include substrate-based optical antenna systems for improved bio-sensing applications [96].

3.1.1 Biomedical Applications

The use of electrospun nanofibers materials is widely spread in different fields due to their significant characteristics: the high surface-to-mass (or volume) ratio and the porous structure with excellent pore-interconnectivity. These properties make electrospun nanofibers employable in advanced applications [97][98]. One of the most important applications of nanomaterials concerns the biomedical sector. Most of the studies focused on their safety and their bio-compatibility with human tissues. Nanofibers are used for tissue engineering for reconstructing damaged tissues or organs. Notably, cells are generally seeded on biomaterial scaffolds and microenvironments in order to enhance tissue development; biomedical nanomaterials play a key role in this medical application because they may better support tissue regeneration [99]. Due to the high specific surface area and the high variability of the process variables, electrospun nanofibers are also used in drug delivery. Electrospun nanofibers prepared with an accurate analysis of polymers are used to deliver antibiotic and anticancer agents, DNA, RNA, proteins and growth factors. During drug delivery, the electrospun nanofibers enclose the medicinal agent to maintain the integrity and bioactivity of the drug molecules and reduce the side effects of the drugs through a localized inoculation [100].

3.2 Related Works

The automatic detection of any possible structural anomaly plays a key role in industrial manufacturing of nanomaterials. Indeed, the automatic recognition of the materials quality entails an acceleration in the production chain for the use in the industrial sector. In this context, ML-based systems have been emerging for a more efficient anomaly detection. Conventional algorithms identify the surface of a specific defect by extracting suitable features and more specifically analysing texture, skeleton, edge and spectrum of the image.

In [101], spatial correlation functions (conveniently defined between the bands of a sensor) are used to recognize the colour structure. A linear model for surface spectral reflectance is used to show that changes in illumination and geometry

correspond to a linear transformation of both correlation functions and their coordinates. In [102] a low computational method for classification of gray scale and rotation invariant texture based on local binary models and nonparametric discrimination, is presented. The technique is based on recognizing the local binary models, defined as uniform, that represent the main properties of the local image texture. Other automatic defects detection algorithms are based on the selection of a suitable threshold. In this regards, the Otsu method is typically employed to perform thresholding by exploiting bimodal distributions [103].
Therefore, the method fails when the image histogram is unimodal. At this purpose, to improve the Otsu method, the weighted object variance method (WOV) is proposed by [104], capable of detecting surface defects, by means of the defect occurrence cumulative probability weighted on the variance between classes. In [105] a new double-visual geometry group16 (DVGG16) is first developed to the automatic classification and localization of surface defects. Next, gradient-weighted class activation mapping (Grad-CAM) is applied to the original images. The achieved heat maps are processed through a threshold segmentation method in order to automatically detect anomalies in the input image. Recently, advanced ML techniques (i.e., deep learning (DL) [106],[107]) are used to reduce production times and simultaneously increase the quality of nanomaterials. It is worth noting that DL has been successfully employed in several applications, such as cyber security [108], neuroscience [109], sentiment analysis [110], remote sensing [111], image decomposition [112], and fault detection systems [113]. With regards to the automatic nanomaterial anomalies detection in SEM images, some works have been proposed in the recent literature. Boracchi *et al.* [114] addresses the issue of automatic detection of anomalies in SEM images, allowing an intelligent system to control independently the validity of the data acquired by a sparse-based representation. Carrera *et al.* [115] proposed a detection algorithm based on a dictionary of normal patches, subsequently used to detect defects in a patch-wise mode. Napoletano *et al.* [116] presented a region-based method for detection and anomaly localization in SEM images.
The degree of anomaly is assessed by means of a CNN, considering a dictionary generated from anomalies-free sub-images belonging to the train set. The automatic detection of defects by using DL models has been addressed also by [117] and [118]. In particular, a CNN has been proposed to automatically classify SEM images of H-NF and NH-NF, interpreted as two different categories. As a difference with most of the previous papers, in this work both samples with and without anomalies are analysed. This approach appears more significant as the images are typically generated with different sets of configuration parameters, which implies a variety of possible ranges of presentation for the nanofibers also in absence of anomalies [117] [119]. However, being it a fully data-driven approach, it is data hungry, requiring the collection of lot of examples through suitably designed laboratory experiments. In contrast, in this work, a novel automatic classification system based on *hybrid unsupervised and supervised* ML able to discriminate H-NF/ NH-HF SEM sub-images (i.e., nanopatches) was created, by avoiding the use of the redundant full SEM representation. It is worth

noting that the use of sub-images allows to improve the detection of possible defects and reduces the computational complexity and cost of the network. Further, the originality of the proposed approach lies also in extracting the most relevant features via *unsupervised* learning, hence, without using the class information, that is not known in advance during realtime use. In addition, the cardinality of the available dataset is augmented by generating extra-latent vectors: this is carried out by corrupting available data with white Gaussian noise. This procedure enabled a rough emulation of new electrospinning experiments, eliminating the requirement for a costly laboratory test. Experimental results reported encouraging performance achieving accuracy rate up to 92.5%.

3.3 Electrospinning Process

Electrospinning is the most successful process for nanofibers fabrication as it is characterized by the ability to improve the product's performance allowing specific modifications for each type of application [120]. The nanofiber fabrication method requires an instrumental apparatus (Figure 3.1) comprised of a high-voltage supply, an extruder and a grounded metallic collector screen where the fibers are collected. A polymeric solution is initially contained into a dosing syringe, regulated by the volumetric pump, which allows controlling the flow-rate. A high-voltage (HV) is applied between the needle of the syringe (anode) and the collector (cathode), which are electrostatically charged to a different electric potential. By increasing the applied voltage, the surface charge of the polymeric solution increases while the radius of the polymeric solution drop decreases, until a critical voltage value. At this moment, the drop takes the form of a cone, referred to as Taylor cone [121]. Due to the electric field, a jet is generated from the cone to the collector; meanwhile, the solvent evaporates and is deposited on the collector in the form of nanofibers. Viscosity, electrical conductivity and surface tension of the polymer solution affect the diameter and the morphology of the generated fibers [122]. Specifically, increasing the viscosity also increases the diameter of the fibers, because the solution opposes more resistance to the elongation by the electric field, and consequently the jet stabilizes and makes a shorter path. The increase of the electrical conductivity of the solution causes a greater repulsion of charge jet, and a higher ironing of the fibers, which decrease in diameter. Hence, in order to produce the nanofibers, the applied electrical charge must exceed the surface tension of the solution.

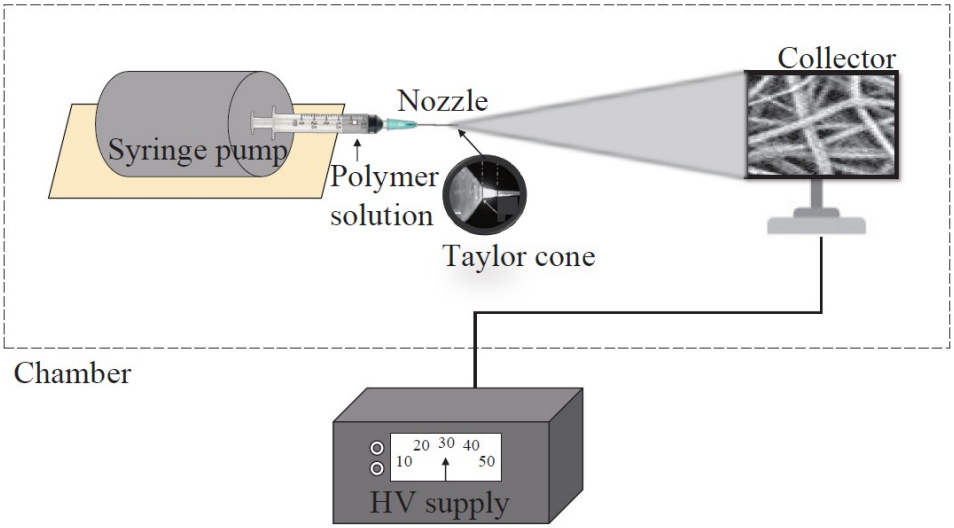


Figure 3.1 Electrospinning apparatus

3.3.1 Electrospinning Set-up

The main parameters used in electrospinning to control the morphology of nanofibers are: concentration (p_1), applied voltage (p_2), flow rate (p_3), and tip-to-collector distance (TCD, p_4). In the laboratory experiments here carried out, a CH-01 Electrospinner 2.0 (Linari Engineering s.r.l.) was used with a 20 mL glass syringe, equipped with a stainless steel needle of 40 mm length and 0.8 mm thick. The solution was instead composed by polyvinylacetate (PVAc) as polymer and ethanol (EtOH) as solvent. In order to obtain a homogeneous polymer solution, it was placed in a test tube and then in a magnetic stirrer (a tool used to mix solvent and solute, by rotating a magnetic latch). To analyze the materials produced by electrospinning, the scanning electron microscope (SEM) Phenom Pro-X was used. It is an electro-optical instrument based on the emission of an electron beam on material surface. After the material production, the Fibermetric SEM images analyzer was used to evaluate the average diameter, the distribution of the nanofibers and the presence of anomalies (i.e., structural defects).

3.3.2 Dataset Construction

Sixteen laboratory experiments were carried out at different working conditions, as reported in Table 3.1. It is to be noted that the experiments were developed by varying the aforementioned parameters (p_1, p_2, p_3, p_4) in the well-defined working range: p_1 [10; 25] %wt; p_2 [10; 17.5] kV; p_3 [100; 300] µL/min; p_4 [10; 15] cm. The e-th nanofibrous material (with e = 1, 2, ..., 16), under- went to the SEM analyzer and 10 relevant and representative areas were selected by an expert operator. Hence, a total of 16 × 10 = 160 SEM images (sized 128 × 128) were collected [34]. Each SEM image was then partitioned into four patches (hereinafter referred to as *nanopatches*) of the same size 64 ×64 (as shown in Figure 3.2). Each SEM patch was manually classified by the nanomaterials expert in two different classes: H-NF and NH-NF images. It is worth mentioning that

homogeneous nanomaterial fabrication is typically observed with high values of voltages and concentrations; while non- homogenous nanomaterial fabrication is affected by the presence of anomalies, such us beads or films, that can occur when the polymeric solution is made up with low values of concentrations or when the tip-to-collector distance is too high.

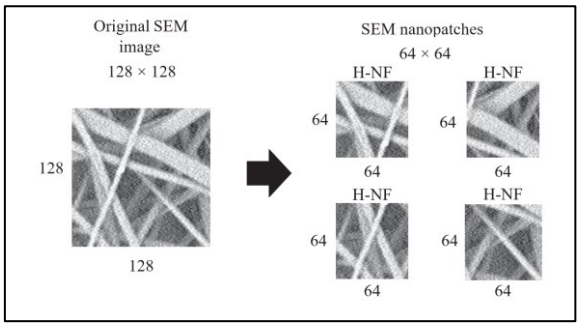


Figure 3.2 Example of a SEM image sized 128 × 128 partitioned into four SEM nanopatches sized 64*64. In the reported example, all the sub-images belong to the homogeneous nanofiber (H-NF) class.

Table 3.1. Setup of the Experimental Electrospinning Parameters

#	*p*1 (%wt) *Concentration*	*p*2 (kV) *Voltage*	*p*3 (µL/min) *Flow-rate*	*p*4 (cm) *TCD*
1	10	15	100	10
2	15	10	100	10
3	15	13.5	100	10
4	15	15	100	10
5	15	15	200	10
6	15	15	300	10
7	15	15	100	12.5
8	15	15	100	13.5
9	15	15	100	15
10	20	10	100	10
11	20	11.5	100	10
12	20	13.5	100	10
13	20	15	100	10
14	20	16	100	10
15	20	17.5	100	10
16	25	15	100	10

3.4 Filters Pre-processing

In order to make the classification task easier each NHNF/HNF SEM image I(x; y), has been pre-processed by reducing the number of grey-scale levels but, simultaneously, maintaining the texture of the individual image as much as possible . For this reason, edge detection techniques are excellent candidates as they segment images based on information on the edges providing information on the object contours using some edge- detection operators finding discontinuity in the grey levels, colour, texture, etc. the edge pixel (x; y) are pixel in which the intensity of brightness, f(x; y), of the image changes abruptly and the edges (or

segments of edge) are sets of connected pixels . By means of Sobel technique [, edge detection is achieved by of a differential operator consisting of two convolution matrices $3 * 3$ with integer values, $Gx = [0\ 1\ 2; -1\ 0\ 1; -2\ -1\ 0]$ and $Gy = [-2\ -1\ 0; -1\ 0\ 1; 0\ 1\ 2]$, which convoluted with the image I calculate an approximate value of $\nabla f(x, y) = [fx(x, y), fy(x, y)] = [Gx\ I, Gy\ I]$ identifying the direction of greater variation of $f(x, y)$, $\theta = \tan{-1}(fy(x, y)/fx(x, y))$ together with its speed in the same direction identified by its magnitude $|f(x, y)| = \sqrt{fx(x,y)2 + fy(x,y)2}$. According to Marr and Hildred, instead, edge detection can be implemented using the filter $\nabla^2 G$, with:

$$G(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{23}$$

which represents the LoG filter (Laplacian of the Gaussian).

However, to reduce the computational complexity of LoG, usually a convolution matrix *5*5*, such as [0 0 -1 0 0; 0 -1 -2 -1 0; -1 -2 16 -2 -1; 0 -1 -2 -1 0; 0 0 -1 0 0], is used that approximates $\nabla^2 G$. Fuzzy edge detection is an alternative approach to edge detection which considers the image to be fuzzy because, often, in most of the images the edge are not clearly defined, so that detection can becomes very difficult. In this paper, a modified Chaira and Ray approach [123] exploiting the fuzzy divergence between the image window and each of a set of 16 convolution matrices (3 * 3, whose elements belong to the set {0:3; 0:8} to ensure good edge detection) which represent the edge profile of different types is presented. Specifically, after normalizing the image I, the center of each convolution matrix is place on each pixel (*x; y*) of *I*. Then, fuzzy divergence measure, *Div(x; y),* between each of the elements of the image window and the template is calculated and the minimum value is selected. This procedure is repeated for all of 16 convolution matrices selecting the maximum value among the 16 divergence values obtained. Then, we obtain a divergence matrix on which a threshold technique must be applied. For this purpose, in this paper a new entropic 2D fuzzy thresholding method based on minimization of fuzzy entropy is proposed. In particular, for each threshold *T*, set a square matrix *W* of size *r* centred on (*x; y*) and considered another window W0 of the same dimensions centred on another pixel (*x0; y0*), their distance is first calculated by the fuzzy divergence. The average value of all the fuzzy divergences obtained by moving (*x0; y0*) in all possible positions is then calculated. Moreover, we calculate the further average value obtained by moving (*x; y*) in all possible position. $Mean_r$ is the latter average value obtained. The procedure was repeated for square windows of size $r + 1$, obtaining $Mean_{r+1}$. Then, Fuzzy Entropy depending on *T*, *FE(T)* can be computed as $FE(T) = ln(Mean_r = Mean_{r+1})$ so that the optimum threshold, $T - optimum$ can be computed by means of $Toptimum = arg;\ min_T\ |FE(T)|$.

Obviously, if necessary, a pre-treatment such as contrast enhancement could be implemented to improve the image quality globally [124][125][126].

3.5 Hybrid Unsupervised-Supervised Machine Learning System

The proposed ML-based architecture is a series of the below detailed network topologies, i.e., an autoencoder and a multilayer perceptron.

3.5.1 Networks Description

Autoencoder (AE)

Autoencoders (AEs) are neural networks trained with unsupervised learning technique that are commonly used for the tasks of representation learning and dimensionality reduction [127][128]. The most typical topology includes an encoding and a decoding stage. AEs commonly exploit backpropagation learning algorithm with a suitable cost function with the objective of making the output as similar as possible to the input while building an internal latent representation of reduced size. AEs thus project the input image into a lower-dimensional hidden layer (called latent- space representation) and then try to reconstruct the output from this reduced representation. After the compression phase, the number of neurons of the hidden layer should be smaller w.r.t. the input layer and the output layer. In the encoding stage, the network is forced to learn the hidden features behind the input data. In the decoding stage, the AE reconstructs the input layer data at the output layer with optimal accuracy [129]. In this way, the internal representation extracts the most significant aspects (i.e., features) of the image presented at the input by exploiting its redundancy. AE works in two steps: an encoding stage represented by the function $y = f(x)$ and a decoding one that generates the reconstructed original vector/image $z = g(y)$. In short, AEs can be described by the function:

$$g(f(x)) = z \tag{24}$$

where z is as close as possible to the original input x. The encoder contains the input layer and the hidden layer, where input data is mapped to obtain a deterministic latent-space representation y.

$$y = \sigma(W^T x + b) \tag{25}$$

where σ is typically a sigmoid or other nonlinear functions; x the input image, W represents the encoder's weight matrix and b is an offset vector. The decoder consists of the hidden layer and the output layer. In this case, the latent space representation is inversely mapped to obtain

$$z = \widehat{W}^T x + \hat{b} \tag{26}$$

Where $\widehat{W}$ is the reconstruction decoder's weights $\hat{b}$ matrix and is the reconstruction offset vector. Finally, in order to reproduce the outputs more and more similar to the inputs, the error function $J(x, z)$ is minimized.

$$J(x,z) = \frac{1}{2}(x-z)^2 \tag{27}$$

The ideal AE should be sensitive enough to the input to build an accurate reconstruction, while, at the same time, insensitive enough to it in order to avoid the model may simply overfit the training data. This trade-off is achieved by taking advantage of the redundancies of the input [130].

Multilayer Perceptron (MLP)

The second stage of the proposed network is a well-known multilayer perceptron (MLP). MLP is the commonest feedforward neural network that consists of an input layer, an output layer and of one or more hidden layers. If the MLP is used for classification, the successive layers are trained to build a complex decision boundary between the classes. It belongs to the supervised learning networks that exploit the class label information to minimize a loss function through standard gradient-based backpropagation technique. Each neuron in a MLP computes a weighted sum of all its inputs that is passed through a non-linear activation function to determine its output. In a classification problem, the output yields the probability that the input vector belongs to a specific class.

Hybrid Model

Figure 3.3 shows the architecture of the proposed hybrid unsupervised and supervised machine learning system for SEM images produced by electrospinning procedure. Specifically, the proposed system includes two main modules: the *unsupervised* processor (Figure 3.3(a)), i.e., an AE that performs the features extraction operation, and the *supervised* processor (Figure 3.3 (b)), i.e., a MLP that performs the classification task: NH-NF vs. H-NF. The AE extracts a reduced representation of the input, i.e., a feature vector. Figure 3.3(a) illustrates the architecture of the proposed AE-based *unsupervised* processor employed for features extraction. It includes an AE [4096: 256: 4096]. Notably, given the *n*th NHNF/ H-NF SEM sub-image (i.e., *nanopatch*) sized 64 × 64, it is flattened into a vector 1 × 4096. Next, the AE compresses the input representation (x, sized 1 × 4096) into a latent-space (y , sized 1 × 256) subsequently used to decode the same input space ($z \approx x$, sized 1 × 4096). In this work, the AE [4096:256: 4096] is trained in unsupervised learning mode for 10^3 epochs on a workstation Intel (R) Core (TM) i7-8700K CPU @ 3.7 GHz with 64 GB RAM and 1 NVIDIA GeForce RTX 2080 Ti GPU installed (training time ≈120 s). The hyperbolic tangent is employed for the encoder and the linear function for the decoder module. Actually, the hidden layer size (1 × 256) of the AE was set after several experimental tests, by estimating the minimum reconstruction error. In particular, the minimum mean squared error was of 0.4416. Hence, overall, a features matrix of 640 × 256 (i.e., number of SEM patches by the number of features) was

extracted (respectively, 320 belonging to NH-NF and 320 to H-NF). However, due to the limited size of such datasets with respect to the number of free parameters of the network, the accuracy in the training and test phases were quite different; hence, a simple data augmentation technique was used to enlarge the database size. Specifically, all the features data vectors were corrupted by a white Gaussian noise at a SNR = 10 dB, and the generated vectors were included in the dataset. A grand total of 1280 × 256 instances were taken into account (640 belonging to NHNF and 640 belonging to H-NF). Figure 3.3(b) shows the proposed MLP. Specifically, the features vector (sized 1 × 256) previously extracted from the unsupervised processor is used as input to a MLP with 2 hidden layers of 100 and 80 hidden. units, respectively. Note that the hyperbolic tangent is used as activation function for each hidden neuron. The network ends with a softmax output layer employed to perform the 2-way classification task: NH-NF vs. H-NF. The architecture, here referred to a MLP with 100 neurons in the first hidden layer and 80 in the second one, was trained over 103 epochs on the aforementioned workstation (i.e., Intel (R) Core (TM) i7- 8700K CPU @ 3.7 GHz with 64 GB RAM and 1 NVIDIA GeForce RTX 2080 Ti GPU installed). Training time was on average of about 40 minutes using the leave-one-out (LOO) technique over the whole dataset.

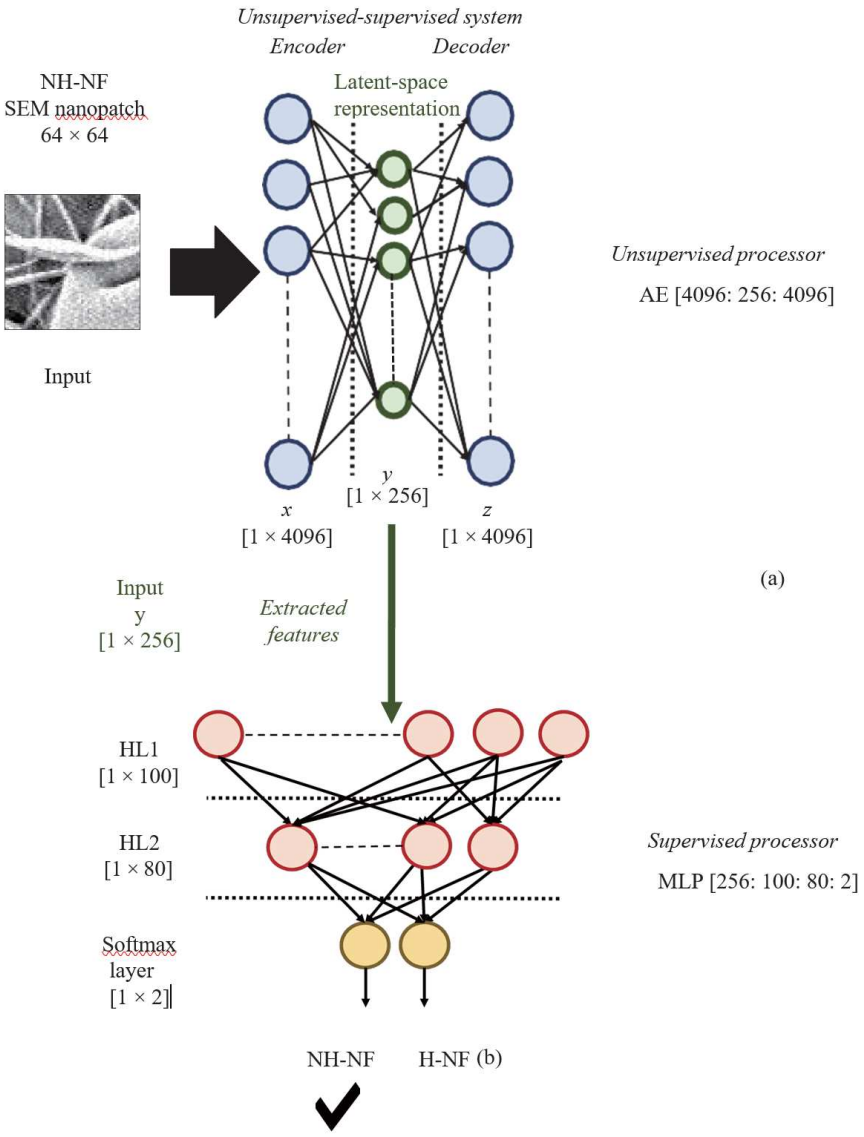


Figure 3.3 Architecture of the proposed hybrid unsupervised and supervised machine learning system. (a) Unsupervised processor composed of an AE [4096: 256: 4096]. The SEM nanopatch is reshaped into a single vector sized 1 × 4096 and used as input to the proposed AE that allows to extract the most relevant features (sized 1 × 256) from the input data. (b) Supervised processor composed of a MLP [256: 100: 80: 2]. The extracted features are the input to the proposed 2-hidden layers MLP for performing the 2-way classification task: NH-NF vs. H-NF. As an example, in the figure, a NH-NF SEM nanopatch inputs the hybrid unsupervised and supervised classification system.

3.5.2 Metrics (ROC, Permutation Analysis)

The performance of the proposed hybrid unsupervised and supervised ML system were assessed using a set of standard metrics, i.e., Precision, Recall, F-score, and Accuracy, defined as follows:

$$Precision\ (Pr)\ = \frac{TP}{TP + FP} \tag{28}$$

$$Recall\ (Rc)\ = \frac{TP}{TP + FN} \tag{29}$$

$$F - score = 2 \times \frac{Pr \times Rc}{Pr + Rc} \tag{30}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{31}$$

where TP, FP, TN, FN are the acronyms of true positive, false positive, true negative, false negative, respectively. In this study, TP denotes SEM images with defects correctly identified as NH-NF; FP denotes SEM images of homogeneous nanofibers misclassified as NH-NF; TN is the number of SEM images of homogeneous nanofibers correctly identified as H-NF; FN is the number of SEM images of nanofibers with defects misclassified as H-NF. The augmented features dataset of 1280 instances (640 belonging to H-NF and 640 belonging to NH-NF) was used as input to our proposed MLP. It is worth mentioning that the LOO technique was applied to validate the efficiency and generalization ability of the developed model. Specifically, LOO consists in partitioning repeatedly the dataset into train set, composed of all instances excluded the *i*th, and test set composed of the *i*th left-out observation. Here, the LOO procedure was applied to the whole dataset. Hence, N= 1280 networks were trained on N–1 data-points and tested on the held-out case.

3.6 Results

Note that the best MLP architecture was determined using a trial-and-error approach, namely, estimating the performance of different numbers of hidden neurons and hidden layers.

Table 3.2 reports comparative classification performance in terms of Precision, Recall, F-score, and Accuracy. First, the 256-dimensional input representation was used as input to MLP classifiers with 1-hidden layer of different size.

Specifically, 40, 60, 80, 100, 120, 140 hidden units were tested. Experimental results show that the 1-hidden layer MLP with 100 neurons (denoted as MLP_{100}) achieved the highest F-score and Accuracy: 92.04% and 91.80%, respectively.

The metrics used in the experiments are chosen to consider classification performance. In particular, the F-score is the metric that best expresses the goodness of a classifier compared to the images it analyses. This research, unlike the others, does not really have a medical background, in fact it was conducted with the aim of evaluating the goodness of nanomaterials that can be used for medical purposes of biocompatibility, drug delivery and other medical application.

Table 3.2. Classification Performance in Terms of Precision, Recall, F-score, and Accuracy of MLP with Different Hidden Layers (HL) and Hidden Units

Model	HL1	HL2	HL3	*Precision*	*Recall*	*F-score*	*Accuracy*
MLP_{40}	40	-	-	94.06%	89.58%	91.76%	91.56%
MLP_{60}	60	-	-	94.53%	88.19%	91.25%	90.94%
MLP_{80}	80	-	-	94.68%	88.85%	91.67%	91.40%
MLP_{100}	100	-	-	94.84%	89.40%	92.04%	91.80%
MLP_{120}	120	-	-	94.37%	89.61%	91.93%	91.71%
MLP_{140}	140	-	-	92.34%	86.03%	89.07%	88.67%
$MLP_{100,40}$	100	40	-	93.59%	89.81%	91.66%	91.48%
$MLP_{100,60}$	100	60	-	92.03%	88.70%	90.34%	90.16%
$MLP_{100,80}$	**100**	**80**	**-**	**95%**	**90.48%**	**92.68%**	**92.50%**
$MLP_{100,80,60}$	100	80	60	93.44%	88.46%	90.88%	90.63%
$MLP_{100,80,40}$	100	80	40	91.25%	88.75%	89.98%	89.84%
$MLP_{100,80,20}$	100	80	20	90.31%	85.76%	87.98%	87.66%

Next, additional layers were used in order to find out possible better configurations. In particular, MLP classifiers with 2-hidden layers were tested, that is: $MLP_{100,40}$, $MLP_{100,60}$, and $MLP_{100,80}$.

As can be seen, among these architectures, $MLP_{100,80}$ reported the highest F-score and Accuracy: 92.68% and 92.50%, respectively. Finally, MLP classifiers with 3-hidden layers were tested: $MLP_{100,80,20}$, $MLP_{100,80,40}$, and $MLP_{100,80,60}$. Here, the highest scores were achieved by $MLP_{100,80,60}$with F-score of 90.88% and Accuracy of 90.63%. Hence, comparative results show that the 2-hidden layer $MLP_{100,80,60}$ achieved the best classification performance in terms of Precision (95%), Recall (90.48%), F-score(92.68%), and Accuracy (92.50%).

The proposed $MLP_{100,80}$ was also compared with other standard ML techniques. Notably, support vector machine with linear kernel (SVM, [131]) and linear discriminant analysis (LDA, [132]) were developed to perform the 2-way discrimination task (NH-NF vs. H-NF). For fair comparison, LOO procedure was applied to the whole dataset.

Table 3.3. Classification Performance in Terms of Precision, Recall, F-score and Accuracy of the Proposed MLP (i.e.,), SVM and LDA Classifiers

Model	*Precision*	*Recall*	*F-score*	*Accuracy*
$MLP_{100,80}$	**95%**	**90.48%**	**92.68%**	**92.50%**
SVM	64.22%	67.60%	65.87%	66.72%
LDA	62.66%	65.52%	64.06%	64.84%

Table 3.3 reports the performance of each classifier evaluated on the test sets: $MLP_{100,80}$, SVM, and LDA. Specifically, SVM classifier achieved F-score of 65.87% and Accuracy of 66.72%; whereas LDA classifier achieved F-score of 64.06% and Accuracy of 64.84%. As can be observed from Table 3.3, the proposed $MLP_{100,80}$, outperformed all of the other models. In support of this outcome, the receiver operating characteristic (ROC) and the corresponding area under the curve (AUC) measure of the developed MLP, SVM, LDA based classifiers were evaluated. As can be seen in Figure 3.4, MLP achieved the highest AUC score of 0.90.

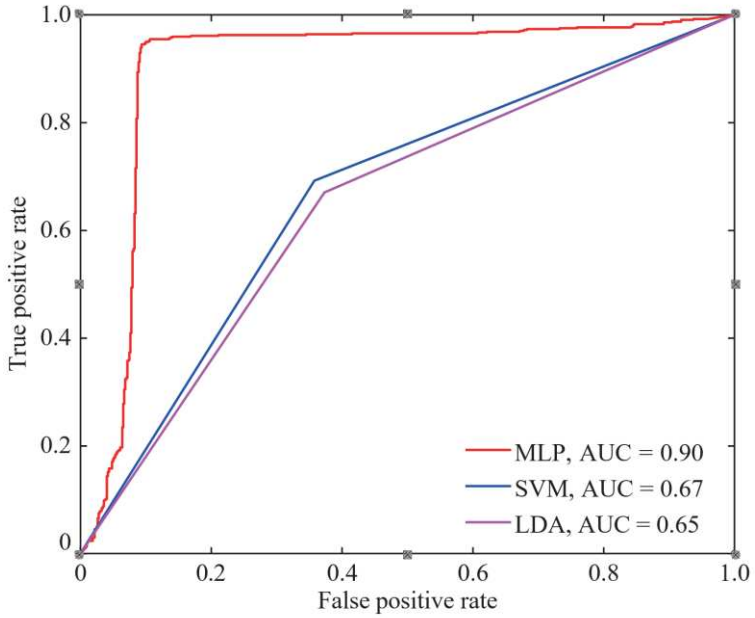


Figure 3.4. ROC curves and corresponding AUC values of the proposed MLP, SVM and LDA classifiers for the NH-NF vs. H-NF classification. Note that the figure refers to the best MLP architecture

3.6.1 Permutation Analysis

In order to assess the dependency of the proposed classifier on the available dataset the standard permutation-based p value statistical test is carried out [133]. This test estimates the p-value under a certain null hypothesis that is: features and class labels are independent. Specifically, the labels are repeatedly permuted and for each iteration the statistical metric of interest (here, the accuracy A_i , with i = 1, 2,..., N_{perm}) is computed. Finally, p-value is empirically calculated as the total number of all A_i equal or greater than the performance estimated with the original dataset (i.e., accuracy A_0), divided by the number of permutations (N_{perm}). p-value smaller than a threshold (typically, α= 0.05) results in rejecting the null hypothesis and consequently concluding that the classifier is statistically significant. It is worth noting that, ideally, all of the possible labels permutations should be taken into account in order to evaluate the exact p-value. As this is computationally expensive, in this study, N_{perm}= 100 were performed [133]. Experimental results reported that p-value = 0.00/100 = 0.00 < 0.05. Hence, the null hypothesis is rejected. In conclusion, the proposed classifier is statistically significant.

3.7 Discussion

An innovative hybrid unsupervised and supervised ML system is proposed aiming to automatically reject defective electrospun nanofibers by processing the related SEM images. The dataset here used for training the classification system is composed of 160 SEM images of PVAc nanofibrous materials [134]. However, in order to reduce the complexity of the classification task, the available full images,

generated by the microscope, were divided into four sub-patches. The cardinality of the resulting dataset is now 640: 320 images belonging to NH-NF and 320 to H-NF classes, respectively. Each SEM image sized 64 × 64 was reshaped into a single vector (sized 1 × 4096) and used as input to the first module of our proposed hybrid ML system, i.e., the AE. The developed AE [4096: 256: 4096] was trained off-line using unsupervised learning and was employed to automatically extract the most relevant features from the input representation. As an example, the presence of beads in the original image reflects in segment of the feature vector with consecutive high values, whilst the presence of a regular texture gives rise to a quasi-periodic representation with low and high values.

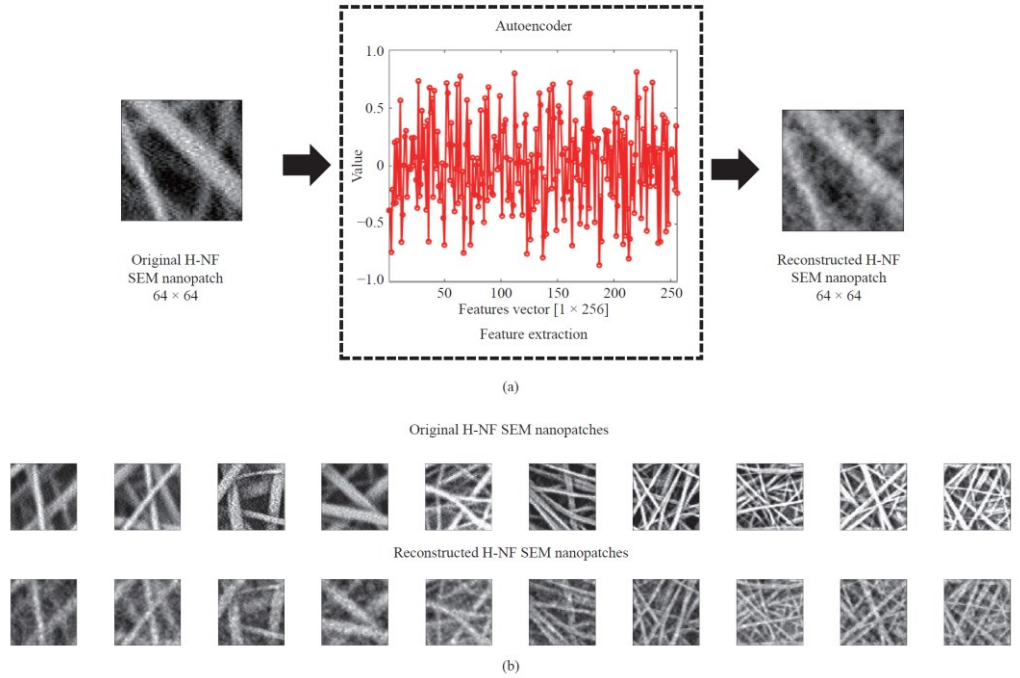


Figure 3.5. (a) Representation of the 256-dimensional features vector extracted by a SEM H-NF image (sized 64 × 64) via the proposed AE. (b) Examples of 10 reconstructed H-NF images.

Next, the compressed 256-dimensional features vector was used as input to the second (supervised) module of the hybrid ML system, i.e., the MLP. The proposed 2-hidden layer MLP performed the binary discrimination task: NH-NF vs. H-NF. The decomposition of the original SEM image in sub-patches simplifies the processing and allows to zoom in small sections of the image that can include some defects. This contrasts with the standard processing of the full image information. Furthermore, the originality of the proposed methodology lies in coding the information of the SEM subregions (i.e., their texture) into a compressed features' vector achieved by the AE processor, by using only unlabeled data Such unsupervised data compression allowed to facilitate the supervised training of the classification processor (i.e., MLP).

It is worth noting that the average reconstruction error of the AE [4096: 256: 4096] was very small, namely, of only 0.4416; thus, the loss of information in the compression stage is rather acceptable as being finalized to reveal the presence of

defects, not to regenerate the original image. As an example, Figure 3.5(a) and Figure 3.6(a) report the representation of the 256-dimensional features vector extracted by a H-NF and a NHNF images with the proposed AE. The figures also show the decoded images of 10 H-NF (Figure 3.6(b)), and 10 NH-NF (Figure 3.5(b)). As can be seen, the original NH-NF/H-NF image and the corresponding reconstructed NH-NF/H-NF image (produced by the AE) are visually similar. Note that the size of AE was empirically defined after several trial-and-error simulations. Furthermore, in order to find out the best MLP architecture, different numbers of hidden layers and hidden units were also tested (Table 3.2).

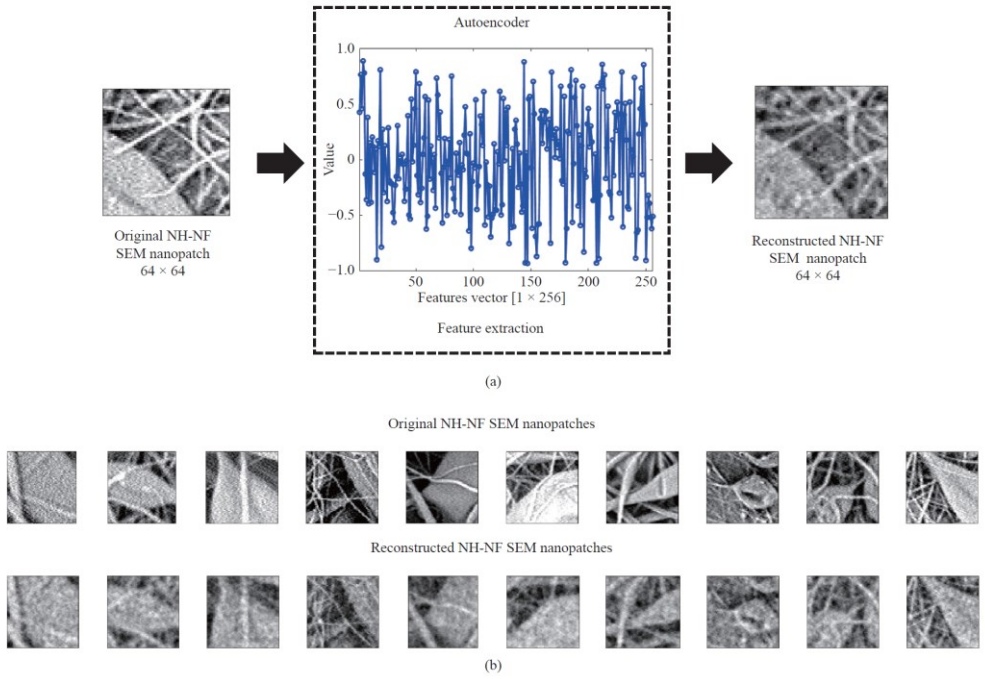


Figure 3.6. (a) Representation of the 256-dimensional features vector extracted by a SEM NH-NF image (sized 64 × 64) via the proposed AE. (b) Examples of 10 reconstructed NH-NF images.

Experimental results show that our proposed hybrid unsupervised and supervised ML system, that is, the combination of AE and MLP architectures, reported the highest performance when compared with other ML-based classifiers (i.e., SVM, LDA). Specifically, the $MLP_{100,80}$ proposed achieved F-score and Accuracy rate up to 92.68% and 92.50%, respectively. Furthermore, a permutation test was carried out to assess the statistical significance of the estimated classification accuracy.

3.8 Conclusion

In this work, a novel automatic classification system for homogenous (anomaly-free) and non-homogenous (with defects) nanofibers is proposed. The inspection procedure aims at avoiding direct processing of the redundant full SEM image.

Specifically, the image to be analysed is first partitioned into sub-images (nano-patches) that are then used as input to a hybrid unsupervised and supervised machine learning system. In the first step, an autoencoder (AE) is trained with unsupervised learning to generate a code representing the input image with a vector of relevant features. Next, a multilayer perceptron (MLP), trained with supervised learning, uses the extracted features to classify non-homogenous nanofiber (NH-NF) and homogenous nanofiber (H-NF) patches. The resulting novel AE-MLP system is shown to outperform other standard machine learning models and other recent state-of-the-art techniques, reporting accuracy rate up to 92.5%. In addition, the proposed approach leads to model complexity reduction with respect to other deep learning strategies such as convolutional neural networks (CNN).

To evaluate the quality of the material, an evaluation of this type will be made: if at least one quadrant is not homogeneous, the entire SEM image will be classified as non-homogeneous and therefore automatically discarded. Furthermore, the proposed methodology based on unsupervised AE can form the basis of a generative model (e.g., [135], [136]) which will allow to increase the database by designing a reduced number of new expensive laboratory experiments.

3.9 Acknowledgement

I would like to thank Professor Morabito and the researchers of the NeuroLAB of the University Mediterranea of Reggio Calabria (Italy) for the support and the closer collaboration. Indeed, thanks the researchers at the Materials for Environmental and Energy Sustainability Laboratory at the University Mediterranea of Reggio Calabria (Italy) for use of the electrospinning system to generate the SEM image database adopted in this work.

Chapter 4

Classification of COVID-19 in Chest X-Ray Images Using Transfer Learning and Vision-Transformer (ViT)

4.1 Covid- 19 Pandemic

Coronavirus disease (COVID-19) was confirmed as a pandemic disease on February 11, 2020. The pandemic has already caused thousands of victims and infected several million people around the world. The aim of this work is to provide a Covid-19 infection screening tool. Currently, the most widely used clinical tool for detecting the presence of infection is the reverse transcription polymerase chain reaction (RT-PCR), which is expensive, less sensitive and requires the resource of specialized medical personnel. The use of X-ray images represents one of the latest challenges for the rapid diagnosis of the Covid-19 infection. This work involves the use of advanced artificial intelligence techniques for diagnosis using algorithms for classification purposes. The goal is to provide an automatic infection detection method while maximizing detection accuracy. A public database was used which includes images of COVID-19 patients, patients with viral pneumonia, patients with pulmonary opacity, and healthy patients [137][138]. The methodology on which the study is based is transfer learning or pre-trained networks to alleviate the complexity of calculation. In particular, three different types of convolutional neural networks, namely InceptionNet and XceptionNet, ResNet50 and the Vision Transformer are implemented. The high accuracy of this computer-assisted diagnostic tool can significantly improve the speed and accuracy of COVID-19 diagnosis.

On December 31, 2019, Chinese health authorities reported an outbreak of pneumonia cases of unknown aetiology in the city of Wuhan (Hubei Province,

China). Many of the initial cases reported exposure to Wuhan's South China Seafood City market. For this reason, the involvement of live animals in the chain of transmission was hypothesized. Over the last year, research has been carried out to validate this hypothesis, without finding a specific link with the Wuhan fish market. More recently, it has been speculated that the virus had a natural reservoir in bats, but that these animals are likely to be in Wuhan a year ago. On January 9, 2020, the China CDC (the Center for Disease Control and Prevention of China) identified a new coronavirus (tentatively named 2019-nCoV) as the etiological cause of these diseases. Chinese health authorities have also confirmed the inter-human transmission of the virus. On 11 February, the World Health Organization (WHO) announced that the disease transmitted since 2019-nCoV has been called COVID-19 (Corona Virus Disease). The Coronavirus Study Group (CSG) of the International Committee on Taxonomy of Viruses has officially classified with the name of SARS-CoV-2 the virus provisionally named by the international health authorities 2019-nCoV and responsible of cases of COVID-19 (Corona Virus Disease). The CSG - responsible for defining the official classification of viruses and the taxonomy of the Corona viridae family, after evaluating the novelty of the human pathogen and on the basis of phylogeny, taxonomy and established practice, has formally associated this virus with the coronavirus it causes severe acute respiratory syndrome (SARS-CoVs, Severe acute respiratory syndrome coronaviruses) classifying it as Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2)[139]. After assessing the severity levels and global spread of the SARS-CoV-2 infection, WHO declared that the COVID-19 epidemic can be considered a pandemic.

4.1.1 Covid-19 Clinical Stages

As shown in Figure 4.1, during Covid-19-induced disease, three distinct stages have been demonstrated [140]:

- The first phase is that in which the virus, through the respiratory tract, enters the organism and replicates inside the cells. The symptoms of this phase are the same as those related to flu syndromes: malaise, widespread arthralgia, fever, dry cough. When the disease stops at this stage, spontaneously or thanks to drugs, the prognosis is excellent, and the course is benign.
- The second phase is that of interstitial pneumonia, which unlike classic lobar pneumonia, affects the two lungs very extensively, both through the direct effects of the virus and the inflammatory response of our body. In this phase, very important respiratory symptoms may appear, associated with a reduction in oxygen saturation and hospitalization is often necessary. The prognosis of this phase is variable and depends, in addition to the treatment, on the type of patient affected: evidently those with pre-

existing cardiac or pulmonary pathologies, the older ones and those with chronic pathologies of any kind are more at risk.

- Finally, the disease evolves into a third phase, characterized by a worsening clinical picture, caused by a hyper-inflammatory response that determines, among other things, a picture of both arterial and venous vasculopathy, with a state of hypercoagulability, small vessel thrombosis, evolution towards even extremely serious and potentially permanent pulmonary lesions (fibrosis) and extra-pulmonary involvement. In this case the prognosis can be very bad, even in less elderly patients and without associated pathologies.

From a cardiovascular point of view, inflammation can highlight a pre-existing picture or cause the appearance of a new disease involving the heart: myocarditis, heart attack, heart failure and arrhythmias. Furthermore, some drugs used to treat the infection, such as azithromycin and hydroxychloroquine, have rare but possible effects on the heart, being able to cause arrhythmias, even lethal through the lengthening of the QTc interval.
A very frequent consequence was pulmonary thrombo-embolism, which manifests itself through a sudden worsening of the clinical picture with oxygen desaturation, confirmed by CT pulmonary angiogram. Neurological complications have also been described, including ischemic and even haemorrhagic stroke.

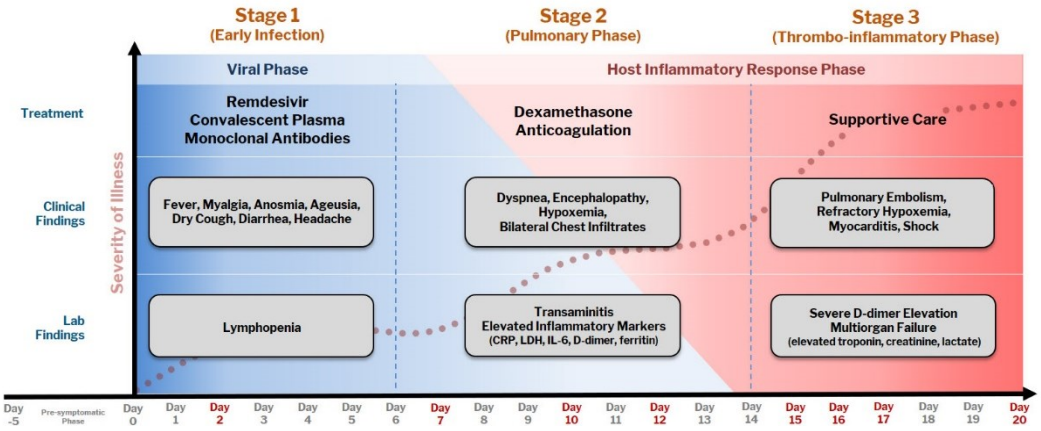


Figure 4.1. Covid-19 Disease Course

4.1.2 Vaccinations

A variety of COVID-19 vaccines are currently in use around the world. Actually. in Italy, the following vaccines are administered:

- Pfizer-BioNTech (BNT162b2), approved by the United States Food and Drug Administration (FDA).
- Modern (mRNA-1273)
- Janssen (Johnson & Johnson; formerly JNJ-78436735, also known as Ad26.COV2.S)

Vaccines target the virus's characteristic spike protein which is essential for the virus to attack host cells through various methods. The Pfizer-BioNTech and Moderna vaccines do not contain the viral antigen, but instead provide a small synthetic fragment of mRNA that codes for the targeted antigen (the spike protein). After being taken up by the cells of the immune system, the mRNA vaccine degrades after instructing the cell to produce the viral antigen. The antigen is then released and triggers the desired immune response to prevent a serious infection following subsequent exposure to the virus. The Janssen vaccine uses an adenoviral vector platform that contains a piece of DNA, or genetic material, which is used to produce the SARS-CoV-2 virus signature spike protein which then triggers the desired immune response.
Pfizer-BioNTech and Moderna vaccines are contraindicated in people with a known history of severe allergic reactions (e.g., anaphylaxis) to a previous dose of the vaccines or any component of the vaccines (including polyethylene glycol). Janssen vaccine is contraindicated in individuals with a history of severe allergic reactions to some of its components (including polysorbate 80).

4.1.3 Worldwide Covid-19 Pandemic Statistics

This section presents the coronavirus statistics in the world, updated to October 2021, in particular the situation of those infected with Covid-19, the situation of the dead, the data of the healed and the currently positive (active) coronavirus cases [141]. The Figure 4.2 represents the coronavirus map with the affected states. The size of the diameter of the circle explains the contagions by territorial extension and therefore the population density affected by the pandemic.

Figure 4.2 Population density affected by the pandemic worldwide

The graph in Figure 4.3 shows the daily growth of the coronavirus from January 2020 (first alarms on the circulation of the virus) until today. As can be seen from the graph, after the peak of April 2021, there is an attenuation, considering the start of vaccinations. Currently, there are about 246,459,876 infected in the world, about 4,999,695 dead and about 223,302,678 recovered.

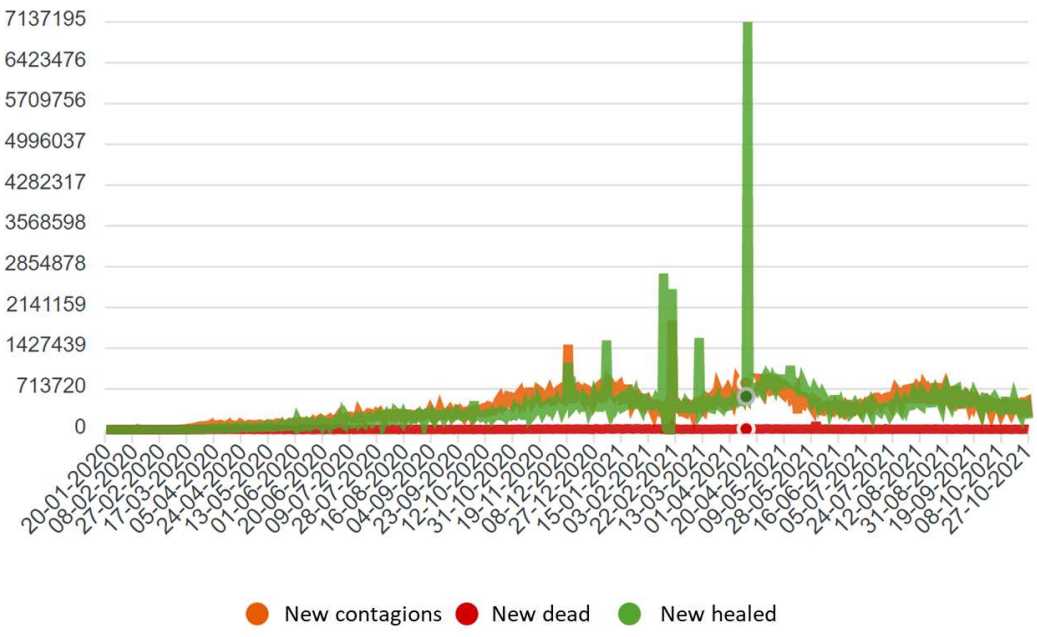


Figure 4.3 Coronavirus daily growth in the world

4.2 Related Works

After the World Health Organization (WHO) declared the rapid spread of the aggressive COVID-19 virus, the world of scientific research went to great lengths to propose a solution for the early diagnosis of the virus [142]. Indeed, the rapid detection of COVID-19 can help control the spread of the disease.

Nowadays, the most used and most reliable method of diagnosing infection is the Reverse Transcription-Polymerase Chain Reaction (RTPCR). A sample is taken by nose / mouth and pharyngeal swab and analysed by real-time molecular methods through the amplification of the viral genes most expressed during the infection. This analysis can only be carried out in highly specialized laboratories, identified by the health authorities and requires on average from 2 to 6 hours to return a result. Another category of tests that have a lower sensitivity and specificity than the previous molecular tests, are the antigen swabbing. This type of test is based on the search for viral proteins (antigens) in respiratory samples. The sampling methods are the same as for molecular tests (nasal and throat swab) but the response time is shorter (about 15 minutes). Finally, the serological tests highlight the presence of antibodies against the virus and tests reveal that there has been exposure to the virus; but only in a few cases can they detect that an infection is in progress. In the current state of scientific development, serological

tests cannot replace molecular tests based on the identification of viral RNA [143].
In recent times, the attention for the diagnosis of infection is focusing on imaging tests. Chest X-ray (CXR) and computed tomography (CT) are the most popular imaging techniques for diagnosing COVID-19 disease. The historical conception of diagnostic imaging systems has been fully explored through several approaches ranging from automation engineering to deep learning[144].
Convolutional neural network (CNN) is one of the most popular and effective approaches in the diagnosis of COVID-19 from digitised images. Several reviews have been carried out to highlight recent contributions to COVID-19 detection [145]. In [146] pre-trained CNN models were used for feature extraction using SVM classifiers with various kernel functions. Then, several pre-trained CNN models were further trained using chest X-ray images for COVID-19 detection. The accuracy of the classification was used to evaluate the performance of the proposed methods. The pre-trained deep CNN models used in the study were ResNet18, ResNet50, ResNet101, VGG16, and VGG19. Since testing the study, the deep characteristics model (ResNet50) and SVM with linear kernel function produced an accuracy score of 94.7%, which was the highest of all results. Test results for fine-tuning the ResNet50 model and end-to-end training of the developed CNN model were 92.6% and 91.6% respectively. Since the number of COVID-19 X-Ray samples is limited, transfer learning appears as the reference method for classifying disease data to develop accurate automated diagnosis models. In this context, networks are able to acquire knowledge from pre-trained networks on large-scale image datasets or alternative data-rich sources. In [147] the classification algorithm based on transfer learning acquired results with an accuracy of 97.66% and an F1 score of 97.61%.
The studies suggested that transfer learning can extract significant features related to the COVID-19 disease diagnosis. In this study, the classification and diagnosis of Covid-19 was by means of transfer learning, first with the classic convolutional neural networks and subsequently through the Vision Transformer.

4.3 Dataset Description: Chest X-Ray Images

The clinical database used for the study was created with chest X-ray images for COVID-19 positive cases along with normal and viral pneumonia images. The database consists of 3616 COVID-19 positive cases, 10,192 normal images, 6012 pulmonary opacity (non-COVID lung infection), and 1345 viral pneumonia[137][138]. The lungs are the two organs responsible for supplying oxygen to the body and for the elimination of carbon dioxide from the blood, or the gaseous exchanges between air and blood (a process known as haematosis). Located in the thoracic cavity, they are surrounded by a serous membrane, the pleura, which is essential for the performance of their functions. The lungs are separated by a space between the spine and the sternum, the mediastinum, which includes the heart, oesophagus, trachea, bronchi, thymus and great vessels. each of the two lungs has at the upper end, an apex that extends upwards to the base of the

neck and, at the lower end, rests on the diaphragmatic muscle. Their main blood is to receive the load of carbon dioxide and waste products from the peripheral circulation and to clean it up: once cleansed the blood is then sent to the heart, from where it is sent to organs and tissues. Figure 4.4 shows an X-Ray image of healthy lungs (without any lung ailments).

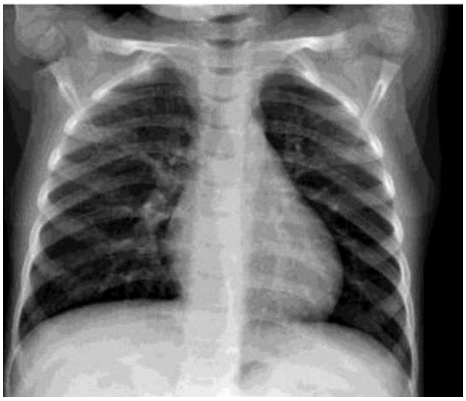


Figure 4.4. Normal lungs

In general, in pneumonia, the lungs fill with fluid and become inflamed, causing difficulty breathing. For some cases, breathing problems can become severe and require hospitalization with oxygen and ventilator treatments. Pneumonia caused by COVID-19 tends to take hold in both lungs. The air sacs in the lungs fill with fluid (Figure 4.5), limiting their ability to absorb oxygen and causing shortness of breath, cough, and other symptoms. Even after the disease has passed, lung lesions can cause breathing difficulties that could take months to improve. The white portions on the are scars inside the lungs. In those areas the lung is not functioning.

Figure 4.5 Covid-19 pneumonia

Pulmonary opacity is represented by spots that appear on the lungs and usually do not exceed 3 cm in diameter (as shown in Figure 4.6). In most cases they are benign, meaning they are not cancerous. A pulmonary nodule is usually seen by means of chest X-ray or computed tomography (CT). They may appear as single nodules or there may be several. A cancerous lung lump is usually larger than 3 cm and can be irregular in shape.

Figure 4.6 Pulmonary opacity

Viral pneumonia is defined as a pathological entity in which there is the viral cause of abnormal oxygen and carbon dioxide gas exchanges in the alveoli, secondary to virus-mediated inflammation and / or immune response [148]. As shown in Figure 4.7, in the Xray image, areas of the chest are visible in which the region or regions are highlighted pneumonia (lighter, whitish spots).

Figure 4.7 Viral pneumonia

4.4 Methodology

Figure 4.8 shows the proposed methodology. The new methodology of the transformer, in particular the Vision Trasformer (ViT), is compared to the classic convolutional neural networks, in particular InceptionNet.

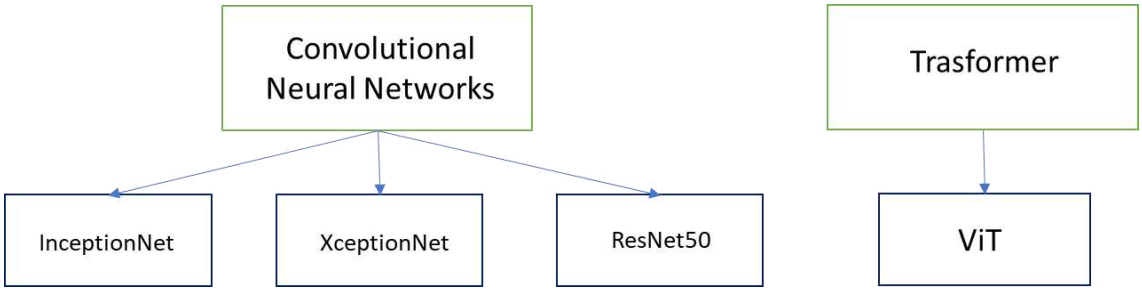


Figure 4.8 Covid Detection Methodology

4.4.1 Transfer Learning

In recent years, deep learning has achieved great success for its wide range of applications, as it is possible to extract models from pre-trained data to predict or classify future outputs [149] [150].
Machine learning leverages training datasets and test data with the same data distribution. When there is a difference in data distribution between the training data and the test data, the performances do not show promise. [151]. Given the large amount of data, obtaining training data that matches the functionality space and expected data distribution characteristics of the test data can be difficult and costly. Thus, using transfer learning it is possible to implement a high-performance learner for a target domain trained by a related source domain.
Through non-technical experiences, it is possible to learn from the real world to understand because learning can be transferred.
So, the need for transfer learning occurs when there is a limited supply of target training data. There are many machine learning applications in which transfer learning has been successfully applied to including text sentiment classification [152], image classification [153], human activity classification [154], software defect classification [155], and multi-language text classification [156].

4.4.2 InceptionNet

InceptionV3 originated as a module for GoogLeNet[157], with the purpose of allowing for deeper networks without increasing too much the number of parameters. Figure 4.9 shows the structure of a single Inception module:

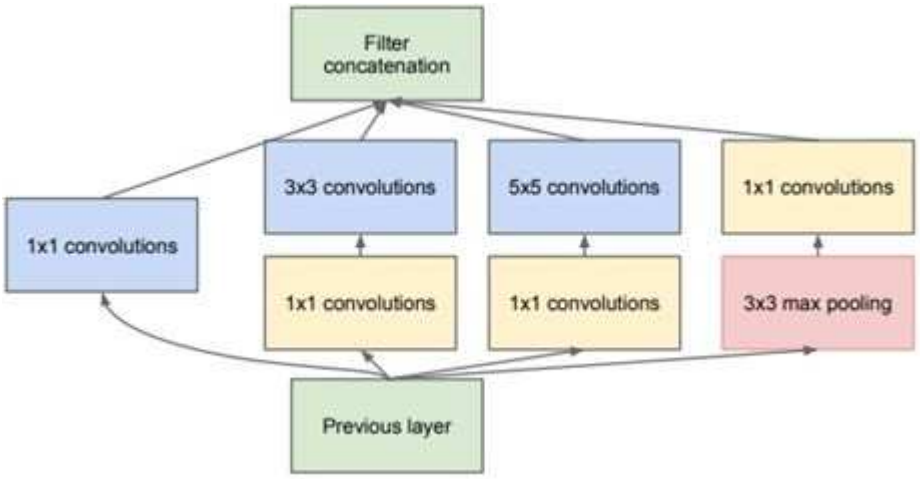


Figure 4.9 Inception Module

- 1x1 convolutional layers act as rectified linear activators as well, so their p urpos e is two-fold.
- 1x1 convolution blocks were introduced to reduce dimensionality.

The overall architecture is depicted in Figure 4.10:

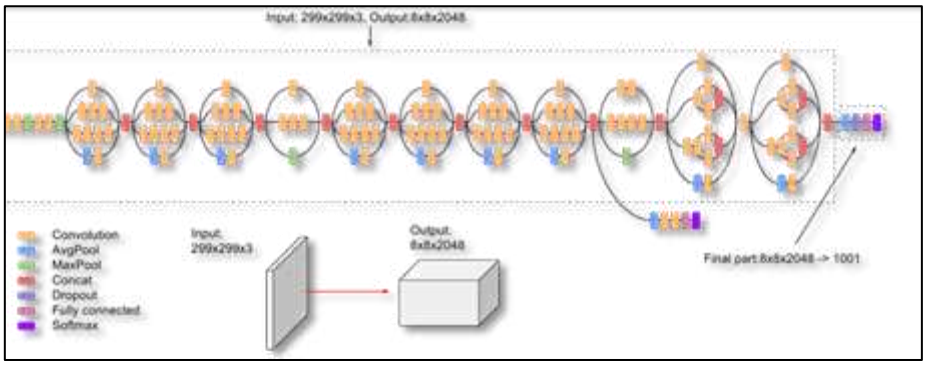


Figure 4.10 InceptionV3 architecture

4.4.3 XceptionNet

The next CNN chosen for analysis and comparison is Xception [158], which was described as an extreme version of InceptionV3 with the exploitation of the so-called depthwise separable convolution, and a subsequent redefinition of the Inception module, as shown in Figure 4.11.

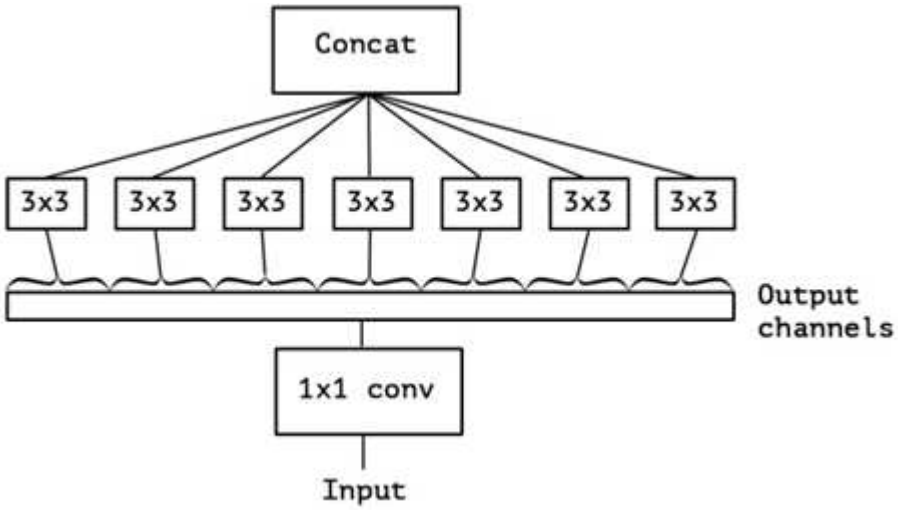


Figure 4.11 Xception Module

The basic underlying concept is the assumption that cross-channel correlation and patial correlation can be mapped separately. This leads to the idea of using a 1x1 convolution to map cross-channel correlations at first and apply 3x3 convolutions to map spatial correlations later on. This has been proved to slightly outperform InceptionV3. shows the Xception architecture:

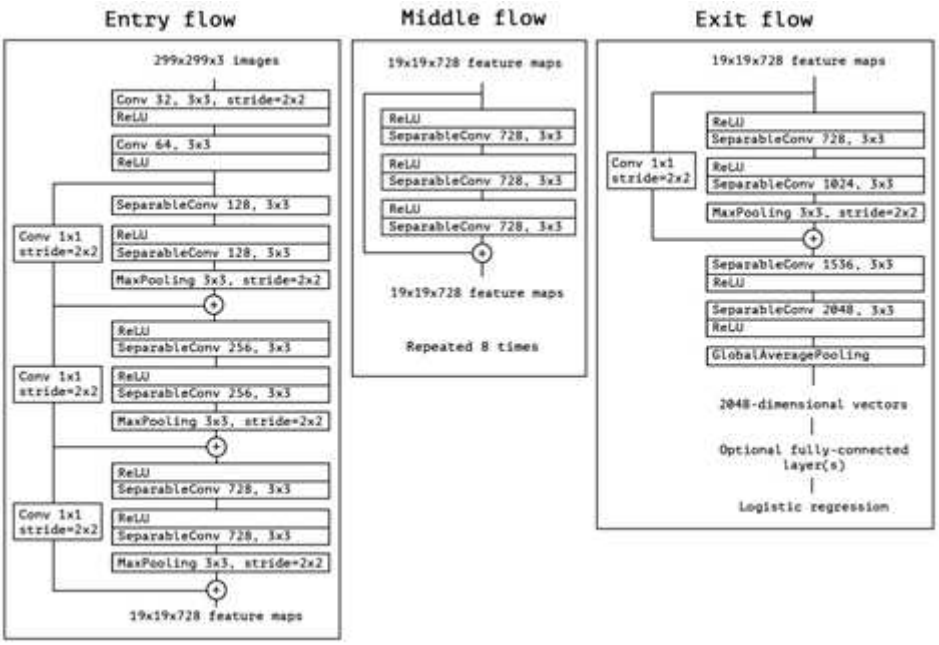


Figure 4.12 Xception architecture

4.4.4 ResNet50

Remarkably, the "middle flow" section of the network presents a skip connection, which is indeed the key element of the next CNN to be described, ResNet50.
ResNet had the peculiarity of using skip connections to tackle the problem of vanishing gradients[159]. This approach was successful and resulted in a number of variations of the original topology. ResNet50 is depicted in Figure 4.13:

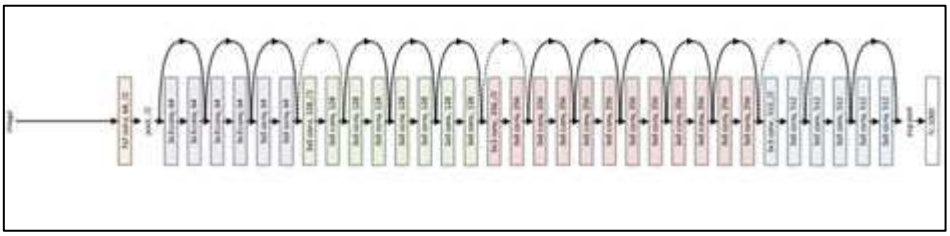


Figure 4.13 ResNet50 architecture

4.4.5 Vision Trasformer

CNNs do exhibit some issues, such as the inability to retain information about the composition and position of specific elements within an image, and to pass such information on to subsequent layers. For this reason, several architectures were developed and presented in recent years. Specifically, Transformers have aroused great interest, especially in NLP applications [160]. In this work the focus is centered on what is probably the most popular version of the Transformer

architecture for image classification, the Vision Transformer, or ViT [161]. The peculiar structure and basic working principles of this network are represented in Figure 4.14:

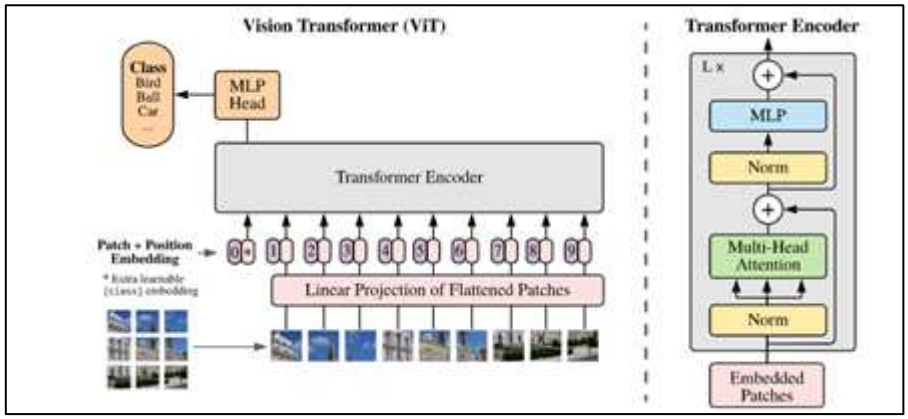


Figure 4.14 Transformer Architecture

4.5 Experimental Setup

The first step of the process consisted in training and testing the three aforementioned CNNs over the chest X-ray dataset. All of the networks were trained by exploiting the transfer learning technique, which allows to retain and freeze network weights derived from previous training sessions over specific datasets. In this case, weights obtained over the ImageNet21k database were used. Subsequently, only some of the top layer weights were set to be trainable. Indeed, the differences among the chosen architectures caused the number of trainable layers to change from one neural network to another. The reduced number of unfrozen layer weights were trained and tested over the chest X-Ray database. This method allowed to significantly reduce the overall amount of time dedicated to the training and testing phases for all convolutional networks.

All networks were trained and tested on a PC with a CPU@3.70GHz with TensorFlow 2.5.0 and Keras. Hyperparameters were configured in the very same fashion for all architectures: initial learning rate was set to 0.0001; fine-tuning learning rate to 0.00001; the chosen optimizer was Adam; batch size was set to 32; dropout coefficient equal to 0.5; loss function of choice was the Sparse Categorical Cross Entropy function. The same number of 60 epochs was set both for the initial training epochs and for fine-tuning epochs, for a total of 120 epochs. The hyperparameters selection criteria are the same as Chapter 1. However, in this case the maximum number of epochs instead of the stop criterion was used. The choice of the drop out coefficient was empirical, and the Sparse Categorical Cross Entropy function was chosen because the classes of the experiment are mutually exclusive, i.e. each sample belongs to exactly one class.

A difference was set in the layer selected to start the fine-tuning process, just as previously mentioned, as follows: the fine-tune was set at layer 308 (over 311 total layers) for InceptionV3; at layer 128 (over 132 total layers) for Xception; at layer 172 (over 175 total layers) for ResNet50.
The next step was the deployment of the Vision Transformer architecture, or ViT, which is the equivalent of the BERT Transformer applied to vision and image classification. Once again, transfer learning was exploited in order to reduce the total amount of time spent on training, validating and testing.
The following hyperparameters were used: the base architecture is ViT-B_32; batch size was set to 32; learning rate was set to 0.00001; the selected optimizer was Rectified Adam; loss function of choice was the Categorical Cross Entropy function; label smoothing was set to 0.2; overall number of epochs was 30. All settings and hyperparameters were chosen in order to make a fair comparison against CNNs, and to take into account the significant architectural differences between CNNs and the Vision Transformer.
Data augmentation with random horizontal flipping and random rotation (set to 0.2), as well as Mitchell-Netravali filtering were applied to the database images before feeding them to the CNNs. On the other hand, no operation of any kind was performed before feeding the images to the Vision Transformer, with the exception of resizing them from an initial resolution of 229x229x3 to 224x224x3 in order to fit the ViT input layer.

4.6 Results

Table 4.1 shows a comparison of the results of the Vision Transformer versus convolutional networks:

Table 4.1 Covid-19 Classification Accuracy

Network Architecture	Test Accuracy
InceptionV3	0.7936
Xception	0.8362
ResNet50	0.8558
ViT	**0.9930**

ResNet50 exhibits the best performance among the convolutional neural networks of choice, with a test accuracy of about 86%. However, the Vision Transformer architecture clearly outperforms the selected CNNs on this specific image classification task with an outstanding accuracy of 99.3%. A few more indicators are shown in Table 4.2 and Figure 4.15 to further describe the performance of the ViT architecture over the four classes of the database:

Table 4.2 Classification indicators for each class

	Precision	Recall	F1-score	Support
COVID (Class: 0)	0.97	0.94	0.96	353

Lung Opacity (Class: 1)	0.87	0.93	0.90	602
Normal (Class: 2)	0.95	0.92	0.94	1019
Viral Pneumonia (Class: 3)	0.96	0.98	0.97	135

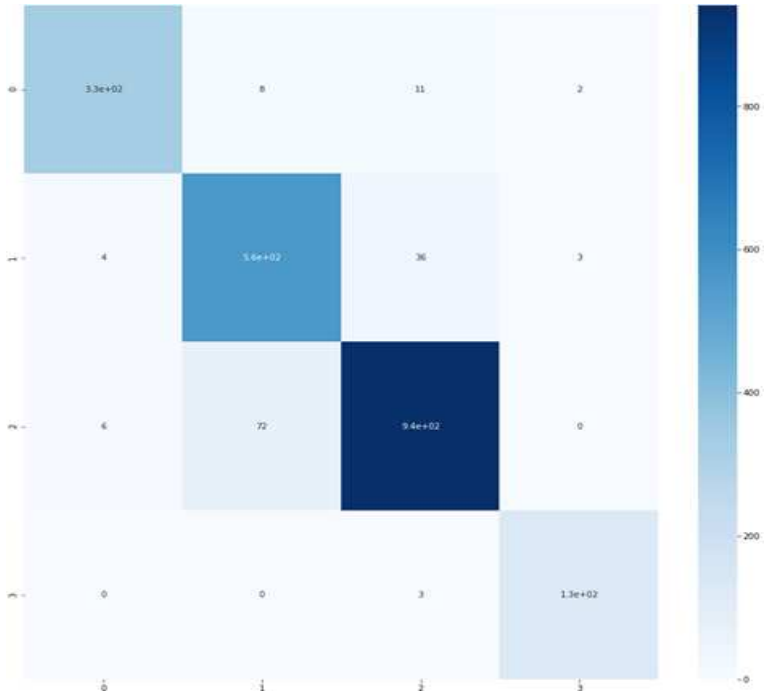


Figure 4.15 Vision Transformer Confusion Matrix

These results shows that the Vision Transformer is highly capable to correctly classify images in each category. This is verified when evaluating the ratio of correctly classified over total classified images per category (precision), the ratio of correctly classified over total actual images per category (recall), and a sort of balance in between the two (F1-score). The poorest results are associated to the Lung Opacity class, which might be due to often the small spots that indicate the presence of the disease are often indistinguishable from the cloud that characterizes the other diseases. Indeed, it is worth highlighting that the Vision Transformer is able to reach a significantly higher accuracy with respect to Convolutional Neural Networks after iterating for only 30 epochs, as opposed to 60 + 60 = 120 overall epochs to train and fine tune the other architectures.

4.7 Conclusion

The Deep Transfer Learning technique is used to build a COVID-19 infected patient classification model. Experimental results demonstrated that the proposed deep transfer learning-based COVID-19 classification model achieves efficient outcomes compared to other supervised learning models. In particular, three convolutional networks (InceptionNet, XceptionNet and ResNet50) and the

Vision Transformer were used. The ViT (best performing model) achieves a training and test accuracy of 99.30%.
The main research insights are based on the detection of Covid-19 infection without the use of antigenic swabs (the most used method to detect the infection). This could represent an alternative method because it is based on the detection of X-ray images of the lungs.

Chapter 5

Conclusions

The goal of the thesis consists of a double aspect. On the one hand, the focus of the thesis is of a medical nature; indeed, the topics, addressed with recent intelligent techniques, are aimed at solving important medical issues that have not yet been answered in the medical literature. On the other hand, a neural thesis based on the search for nuances of machine learning from shallow learning to deep learning is applied to different fields and different applications concerning the medical field.

From a medical point of view, various issues were treated and for each one, methods and approaches customized ad hoc for the problem were chosen.

Depending on the fields of biomedical application, several networks were treated considering the inputs, i.e., the data available to us (images, signals, and time series).

The first field of application concerns cardiovascular problems, and in particular hypertension. Currently, the wearable devices for health-monitoring control the heart activity by means of the electrocardiogram signal and the blood oxygenation by means of the photoplethysmogram signal. However, to keep cardiovascular problems under control, the monitoring of arterial blood pressure becomes essential. For this reason, a *forecasting* problem is addressed to trace the blood pressure signal and detect any anomalies.

The second field of application is based on the *stratification* of the risk of death in patients suffering from channelopathies, that is rare diseases involving the heart system. In this case, by means of a digitization algorithm, the ECG images were transformed into signals and then through the fiducial parameters of the ECG a neural network with a hidden layer was trained.

Another task, considered in this thesis, relates to the *classification* of nanomaterials used for biocompatible sensors for tissues. The images of the nanomaterials are placed at the entrance of the neural network, which classifies the nanomaterials without anomalies by means of a defect detection model.

The last topic concerns *diagnostic* using chest X-Ray images. Radiographic images represent an essential resource for the diagnosis of pathologies (here only Covid-19 is analysed). For this reason, artificial intelligence will play an important role as a diagnostic support tool in the medical field because it is very fast and economical.
From a neural point of view, a wide range of neural networks have been studied and applied.
First of all, shallow learning systems were applied to medical problem, i.e., with a single hidden layer neural network.
Before testing a deep learning approach to predict the blood pressure signal, a network with an input regression vector was trained to understand the dynamics of the system. However, due to the large amount of samples that a signal contains, good performance cannot be achieved by using the neural network. A case in which the network with the single hidden layer has given promising results is related to the classification for the risk stratification of channelopathies: since the approach for the classification was based on features extraction, the results of the system are optimized with a shallow approach.
Deep learning approaches are widely used to solve problems that exploit the logic of artificial intelligence. In the case of the diagnosis of Covid-19 infection, various deep learning networks were examined and compared. In particular, convolutional neural networks (InceptionNet, XceptionNet and ResNet50) and very recent networks called transformers (ViT- Vision Transformer). The latter represent the new frontier of artificial intelligence; indeed, even in the proposed application, transformers have obtained interesting results.
Finally, neural networks may be combined to generate custom neural models that acquire characteristics of two or more networks.
In the proposed thesis work several hybrid approaches were examined. In the case of pressure signal detection, the best approach requires a cascade of a convolutional and recurrent networks, the former to perform feature extraction automatically from the signal and the latter to predict the entire signal.
Another hybrid approach was used to build a defect detection classification system for the evaluation of the goodness of nanomaterials. In this case, an AutoEncoder (unsupervised system) was trained to compress an image in an encoding and a multilayer perceptron (with two hidden layers) was used for the classification of the encodings (associated with each image of the nanomaterial). This hybrid approach turns out to be the most performing among those studied in literature for the same purpose.
Although there are two different and apparently opposite aspects, these two aspects integrate harmoniously since the neural aspect proves to be efficient and effective in a field that highlights its capabilities; instead, from a medical point of view the most relevant contribution stems from the fact that artificial intelligence represents the best way to solve medical problems, which cannot be solved with classical methods. Recently the importance of investing in health technologies enabled by artificial intelligence has taken on considerable interest for the health system. On the medical level, technology is at the service of health for purposes of

real-time monitoring and telemedicine, storing electronic records to maintain data quality, disease prediction and making early and accurate diagnoses.
The technological support based on artificial intelligence for medicine represents an evolving process that will acquire more and more value and credibility.

References

[1] "Cardiovascular diseases." https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1 (accessed Aug. 24, 2021).

[2] M. Kelly and C. Semsarian, "Multiple Mutations in Genetic Cardiovascular Disease," *Circ. Cardiovasc. Genet.*, vol. 2, no. 2, pp. 182–190, Apr. 2009, doi: 10.1161/CIRCGENETICS.108.836478.

[3] J. E. Rossouw, "Hormones, genetic factors, and gender differences in cardiovascular disease," *Cardiovasc. Res.*, vol. 53, no. 3, pp. 550–557, Feb. 2002, doi: 10.1016/S0008-6363(01)00478-3.

[4] M. E. Mendelsohn and R. H. Karas, "The Protective Effects of Estrogen on the Cardiovascular System," *http://dx.doi.org/10.1056/NEJM199906103402306*, vol. 340, no. 23, pp. 1801–1811, Oct. 2008, doi: 10.1056/NEJM199906103402306.

[5] G. A. Colditz, W. C. Willett, M. J. Stampfer, B. Rosner, F. E. Speizer, and C. H. Hennekens, "Menopause and the Risk of Coronary Heart Disease in Women," *http://dx.doi.org/10.1056/NEJM198704303161801*, vol. 316, no. 18, pp. 1105–1110, Jan. 2010, doi: 10.1056/NEJM198704303161801.

[6] D. Mozaffarian, P. W. F. Wilson, and W. B. Kannel, "Beyond Established and Novel Risk Factors," *Circulation*, vol. 117, no. 23, pp. 3031–3038, Jun. 2008, doi: 10.1161/CIRCULATIONAHA.107.738732.

[7] E. M. Cherry and F. H. Fenton, "Heart Structure, Function and Arrhythmias."

[8] J. R. Bradford and H. P. Dean, "The Pulmonary Circulation1," *J. Physiol.*, vol. 16, no. 1–2, pp. 34–158, Mar. 1894, doi: 10.1113/JPHYSIOL.1894.SP000493.

[9] J. Segen, "Concise Dictionary of Modern Medicine," *eweb:292340*.

[10] "The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC 7)." https://web.archive.org/web/20070824210934/http://www.nhlbi.nih.gov/guidelines/hypertension/ (accessed Aug. 30, 2021).

[11] R. E. Klabunde, "Neurohumoral Control of the Heart and Circulation," *Cardiovasc. Physiol. Concepts*, pp. 117–140, 2005.

[12] "Low Blood Pressure | NHLBI, NIH." https://www.nhlbi.nih.gov/health-topics/low-blood-pressure (accessed Aug. 30, 2021).

[13] “High Blood Pressure | NHLBI, NIH.” https://www.nhlbi.nih.gov/health-topics/high-blood-pressure (accessed Aug. 30, 2021).

[14] J. R. Chiong *et al.*, “Secondary hypertension: Current diagnosis and treatment,” *Int. J. Cardiol.*, vol. 124, no. 1, pp. 6–21, Feb. 2008, doi: 10.1016/J.IJCARD.2007.01.119.

[15] B. Williams *et al.*, “2018 ESC/ESH Guidelines for the management of arterial hypertensionThe Task Force for the management of arterial hypertension of the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH),” *Eur. Heart J.*, vol. 39, no. 33, pp. 3021–3104, Sep. 2018, doi: 10.1093/EURHEARTJ/EHY339.

[16] B. N. Van Vliet, L. L. Chafe, V. Antic, S. Schnyder-Candrian, and J. P. Montani, “Direct and indirect methods used to study arterial blood pressure,” *J. Pharmacol. Toxicol. Methods*, vol. 44, no. 2, pp. 361–373, Sep. 2000, doi: 10.1016/S1056-8719(00)00126-X.

[17] A. Roguin, “Scipione Riva-Rocci and the men behind the mercury sphygmomanometer,” *Int. J. Clin. Pract.*, vol. 60, no. 1, pp. 73–79, Jan. 2006, doi: 10.1111/J.1742-1241.2005.00548.X.

[18] T. Tamura and W. Chen, “Seamless healthcare monitoring: Advancements in wearable, attachable, and invisible devices,” *Seamless Healthc. Monit. Adv. Wearable, Attach. Invis. Devices*, pp. 1–469, Nov. 2017, doi: 10.1007/978-3-319-69362-0.

[19] T. Dieterle, “Blood pressure measurement - an overview,” *Swiss Med. Wkly. 2012 3*, vol. 142, no. 3, Jan. 2012, doi: 10.4414/SMW.2012.13517.

[20] G. P. Molhoek *et al.*, “Evaluation of the Penàz servo-plethysmo-manometer for the continuous, non-invasive measurement of finger blood pressure,” *Basic Res. Cardiol. 1984 795*, vol. 79, no. 5, pp. 598–609, Sep. 1984, doi: 10.1007/BF01910489.

[21] K. Tegtmeyer, G. Brady, S. Lai, R. Hodo, and D. Braner, “Placement of an Arterial Line,” 2006.

[22] P. Verdecchia *et al.*, “White coat hypertension and white coat effect similarities and differences,” *Am. J. Hypertens.*, vol. 8, no. 8, pp. 790–798, Aug. 1995, doi: 10.1016/0895-7061(95)00151-E.

[23] F. J. Callaghan, C. F. Babbs, J. D. Bourland, and L. A. Geddes, “The relationship between arterial pulse-wave velocity and pulse frequency at different pressures,” *http://dx.doi.org/10.3109/03091908409032067*, vol. 8, no. 1, pp. 15–18, 2009, doi: 10.3109/03091908409032067.

[24] R. Wang, W. Jia, Z. H. Mao, R. J. Sclabassi, and M. Sun, “Cuff-free blood pressure estimation using pulse transit time and heart rate,” *Int. Conf. Signal Process. Proceedings, ICSP*, vol. 2015-January, no. October, pp. 115–118, 2014, doi: 10.1109/ICOSP.2014.7014980.

[25] A. Hennig and · A Patzak, “Continuous blood pressure measurement using

pulse transit time," *Somnologie*, vol. 17, pp. 104–110, 2013, doi: 10.1007/s11818-013-0617-x.

[26] P. Fung, G. Dumont, C. Ries, C. Mott, and M. Ansermino, "Continuous noninvasive blood pressure measurement by pulse transit time," *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, vol. 26 I, pp. 738–741, 2004, doi: 10.1109/IEMBS.2004.1403264.

[27] M. Hosanee *et al.*, "Cuffless Single-Site Photoplethysmography for Blood Pressure Monitoring," *J. Clin. Med.*, vol. 9, no. 3, p. 723, Mar. 2020, doi: 10.3390/jcm9030723.

[28] V. Randazzo, J. Ferretti, and E. Pasero, "ECG WATCH: A real time wireless wearable ECG," Jun. 2019, doi: 10.1109/MeMeA.2019.8802210.

[29] Randazzo, Ferretti, and Pasero, "A Wearable Smart Device to Monitor Multiple Vital Parameters—VITAL ECG," *Electronics*, vol. 9, no. 2, p. 300, Feb. 2020, doi: 10.3390/electronics9020300.

[30] V. Randazzo, E. Pasero, and S. Navaretti, "VITAL-ECG: A portable wearable hospital," in *2018 IEEE Sensors Applications Symposium, SAS 2018 - Proceedings*, Apr. 2018, vol. 2018-January, pp. 1–6, doi: 10.1109/SAS.2018.8336776.

[31] Y. Kurylyak, F. Lamonaca, and D. Grimaldi, "A Neural Network-based method for continuous blood pressure estimation from a PPG signal," *Conf. Rec. - IEEE Instrum. Meas. Technol. Conf.*, pp. 280–283, 2013, doi: 10.1109/I2MTC.2013.6555424.

[32] U. Senturk, I. Yucedag, and K. Polat, "Repetitive neural network (RNN) based blood pressure estimation using PPG and ECG signals," Dec. 2018, doi: 10.1109/ISMSIT.2018.8567071.

[33] "MIMIC II Databases." https://archive.physionet.org/mimic2/ (accessed Aug. 31, 2020).

[34] C. P. Chua and C. Heneghan, "Continuous blood pressure monitoring using ECG and finger photoplethysmogram," in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2006, pp. 5117–5120, doi: 10.1109/IEMBS.2006.259612.

[35] G. Parati and M. Valentini, "Prognostic relevance of blood pressure variability," *Hypertension*, vol. 47, no. 2. Lippincott Williams & Wilkins , pp. 137–138, Feb. 01, 2006, doi: 10.1161/01.HYP.0000198542.51471.c4.

[36] "Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine - Leonard S. Lilly, Eugene Braunwald - Google Libri." https://books.google.it/books?hl=it&lr=&id=blq42K8AY5AC&oi=fnd&pg=PP2&dq=E.+Braunwald,+5ª+ed.,+Philadelphia,+W.B.+Saunders+Co&ots=_g6t3akJaq&sig=1-LwXov_Q8WMilwn4dxd0PUIdSc&redir_esc=y#v=onepage&q&f=false (accessed Sep. 02, 2021).

[37] “Rapid ECG Interpretation - M. Gabriel Khan - Google Libri.” https://books.google.it/books?hl=it&lr=&id=Id6xAn4g8nEC&oi=fnd&pg=PP2&dq=ecg+interpretation&ots=RrihEoxWzD&sig=3f4Vix7BmtOTgy91sPMXt81tLoU&redir_esc=y#v=onepage&q=ecg interpretation&f=false (accessed Sep. 08, 2021).

[38] H. P. Selker *et al.*, “Standard ECG,” *Ann. Emerg. Med.*, vol. 29, no. 1, pp. 17–20, Jan. 1997, doi: 10.1016/S0196-0644(97)70299-3.

[39] K. H. Shelley, “Pulse Oximeter Waveform: Photoelectric Plethysmography,” 2001.

[40] A. AA and S. KH, “Photoplethysmography,” *Best Pract. Res. Clin. Anaesthesiol.*, vol. 28, no. 4, pp. 395–406, 2014, doi: 10.1016/J.BPA.2014.08.006.

[41] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiol. Meas.*, vol. 28, no. 3, p. R1, Feb. 2007, doi: 10.1088/0967-3334/28/3/R01.

[42] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.,” *Circulation*, vol. 101, no. 23, Jun. 2000, doi: 10.1161/01.cir.101.23.e215.

[43] G. B. Moody and R. G. Mark, “A database to support development and evaluation of intelligent intensive care monitoring,” *Comput. Cardiol.*, vol. 0, no. 0, pp. 657–660, 1996, doi: 10.1109/cic.1996.542622.

[44] “MIMIC Database v1.0.0.” https://physionet.org/content/mimicdb/1.0.0/ (accessed Apr. 06, 2021).

[45] G. Slapničar, N. Mlakar, and M. Luštrek, “Blood Pressure Estimation from Photoplethysmogram Using a Spectro-Temporal Deep Neural Network,” *Sensors 2019, Vol. 19, Page 3420*, vol. 19, no. 15, p. 3420, Aug. 2019, doi: 10.3390/S19153420.

[46] A. Paviglianiti, V. Randazzo, E. Pasero, and A. Vallan, “Noninvasive Arterial Blood Pressure Estimation using ABPNet and VITAL-ECG,” May 2020, doi: 10.1109/I2MTC43012.2020.9129361.

[47] A. Paviglianiti, V. Randazzo, G. Cirrincione, and E. Pasero, “Neural Recurrent Approches to Noninvasive Blood Pressure Estimation,” 2020.

[48] A. Paviglianiti, V. Randazzo, G. Cirrincione, and E. Pasero, “Double Channel Neural Non Invasive Blood Pressure Prediction,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12463 LNCS, pp. 160–171, Oct. 2020, doi: 10.1007/978-3-030-60799-9_14.

[49] M. Enqvist, “Linear Models of Nonlinear Systems,” 2005.

[50] M. Nørgaard, *Neural networks for modelling and control of dynamic systems : a practitioner’s handbook*. Springer, 2000.

[51] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, Oct. 2006, doi: 10.1016/J.IJFORECAST.2006.03.001.

[52] C. J. Willmott and K. Matsuura, “Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance,” *Clim. Res.*, vol. 30, no. 1, pp. 79–82, Dec. 2005, doi: 10.3354/CR030079.

[53] X. He and H. Asada, “New method for identifying orders of input-output models for nonlinear dynamic systems,” *Am. Control Conf.*, pp. 2520–2523, 1993, doi: 10.23919/ACC.1993.4793346.

[54] K. Levenberg1 and F. Arsenal, “NOTES-A METHOD FOR THE SOLUTION OF CERTAIN NON-LINEAR PROBLEMS IN LEAST SQUARES*.”

[55] H. P. Gavin, “The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems,” 2020.

[56] A. Paviglianiti, V. Randazzo, S. Villata, G. Cirrincione, and E. Pasero, “A Comparison of Deep Learning Techniques for Arterial Blood Pressure Prediction,” *Cogn. Comput. 2021*, vol. 1, pp. 1–22, Aug. 2021, doi: 10.1007/S12559-021-09910-0.

[57] A. Géron, *Hands-On Machine Learning With Scikit-Learn, Keras, And TensorFlow: Concepts, Tools, And Techniques To Build Intelligent Systems* . 2019.

[58] D. Ciregan, U. Meier, and J. Schmidhuber, “Multi-column deep neural networks for image classification,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 3642–3649, 2012, doi: 10.1109/CVPR.2012.6248110.

[59] “4. Fully Connected Deep Networks - TensorFlow for Deep Learning [Book].” https://www.oreilly.com/library/view/tensorflow-for-deep/9781491980446/ch04.html (accessed Sep. 15, 2021).

[60] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” 2016.

[61] A. van den Oord *et al.*, “WaveNet: A Generative Model for Raw Audio,” Sep. 2016.

[62] H. H. Sak, A. Senior, and B. Google, “Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling.”

[63] F. A. Gers, N. N. Schraudolph, and J. Schmidhuber, “Learning Precise Timing with LSTM Recurrent Networks,” *J. Mach. Learn. Res.*, vol. 3, pp. 115–143, 2002.

[64] C. P. Chua and C. Heneghan, “Continuous blood pressure monitoring using

ECG and finger photoplethysmogram," in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 2006, pp. 5117–5120, doi: 10.1109/IEMBS.2006.259612.

[65] R. Maronna, R. Martin, V. Yohai, and M. Salibián-Barrera, "Robust statistics: theory and methods (with R)," 2019.

[66] Ü. Şentürk, I. Yücedağ, K. P.-2018 2Nd international, and undefined 2018, "Repetitive neural network (RNN) based blood pressure estimation using PPG and ECG signals," *ieeexplore.ieee.org*.

[67] S. Mishra *et al.*, "ECG Paper Record Digitization and Diagnosis Using Deep Learning," *J. Med. Biol. Eng.*, vol. 41, pp. 422–432, 1234, doi: 10.1007/s40846-021-00632-0.

[68] R. Patil and R. Karandikar, "Image digitization of discontinuous and degraded electrocardiogram paper records using an entropy-based bit plane slicing algorithm," *J. Electrocardiol.*, vol. 51, no. 4, pp. 707–713, Jul. 2018, doi: 10.1016/J.JELECTROCARD.2018.05.003.

[69] J. Chebil, J. Al-Nabulsi, and M. Al-Maitah, "A novel method for digitizing standard ECG papers," *Proc. Int. Conf. Comput. Commun. Eng. 2008, ICCCE08 Glob. Links Hum. Dev.*, pp. 1308–1312, 2008, doi: 10.1109/ICCCE.2008.4580816.

[70] P. R. K. Shrivastava, S. Panbude, and G. Narayanan, "Digitization of ECG Paper Records using MATLAB," *Int. J. Innov. Technol. Explor. Eng.*, no. 6, pp. 2278–3075, 2014.

[71] M. Woś, G. Kłosowski, D. Wójcik, and T. Rymarczyk, "A complete system for an automated ECG diagnosis," doi: 10.15199/48.2021.01.32.

[72] "ecg-kit - File Exchange - MATLAB Central." https://it.mathworks.com/matlabcentral/fileexchange/50769-ecg-kit (accessed Sep. 30, 2021).

[73] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson Correlation Coefficient," *Springer Top. Signal Process.*, vol. 2, pp. 1–4, 2009, doi: 10.1007/978-3-642-00296-0_5.

[74] J. M. Bland and D. G. Altman, "Measurement error and correlation coefficients.," *BMJ Br. Med. J.*, vol. 313, no. 7048, p. 41, Jul. 1996, doi: 10.1136/BMJ.313.7048.41.

[75] R. Khattree, "Ch. 22. Repeatability, reproducibility and interlaboratory studies," *Handb. Stat.*, vol. 22, pp. 795–822, Jan. 2003, doi: 10.1016/S0169-7161(03)22024-1.

[76] J. Adler and I. Parmryd, "Quantifying colocalization by correlation: The Pearson correlation coefficient is superior to the Mander's overlap coefficient," *Cytom. Part A*, vol. 77A, no. 8, pp. 733–742, Aug. 2010, doi: 10.1002/CYTO.A.20896.

[77] O. Campuzano, G. Sarquella-Brugada, E. Arbelo, J. Brugada, and R. Brugada, “Short QT syndrome,” *Horizons World Cardiovasc. Res.*, vol. 8, pp. 137–147, Apr. 2015, doi: 10.1161/01.CIR.0000085071.28695.C4.

[78] C. Giustetto *et al.*, “Short QT syndrome: Clinical findings and diagnostic-therapeutic implications,” *Eur. Heart J.*, vol. 27, no. 20, pp. 2440–2447, Oct. 2006, doi: 10.1093/eurheartj/ehl185.

[79] C. Giustetto *et al.*, “Long-Term Follow-Up of Patients With Short QT Syndrome,” *J. Am. Coll. Cardiol.*, vol. 58, no. 6, pp. 587–595, Aug. 2011, doi: 10.1016/J.JACC.2011.03.038.

[80] A. Algra, J. G. Tijssen, J. R. Roelandt, J. Pool, and J. Lubsen, “QT interval variables from 24 hour electrocardiography and the two year risk of sudden death.,” *Heart*, vol. 70, no. 1, pp. 43–48, Jul. 1993, doi: 10.1136/HRT.70.1.43.

[81] Y. Sun, X. Quan, S. Fromme, R. Cox, … P. Z.-J. of molecular and, and undefined 2011, “A novel mutation in the KCNH2 gene associated with short QT syndrome,” *Elsevier*.

[82] C. Giustetto *et al.*, “Long-term follow-up of patients with short QT syndrome,” *J. Am. Coll. Cardiol.*, vol. 58, no. 6, pp. 587–595, Aug. 2011, doi: 10.1016/j.jacc.2011.03.038.

[83] A. Mazzanti *et al.*, “Novel insight into the natural history of short QT syndrome,” *J. Am. Coll. Cardiol.*, vol. 63, no. 13, pp. 1300–1308, Apr. 2014, doi: 10.1016/j.jacc.2013.09.078.

[84] Y. Sun *et al.*, “A novel mutation in the KCNH2 gene associated with short QT syndrome,” *J. Mol. Cell. Cardiol.*, vol. 50, no. 3, pp. 433–441, Mar. 2011, doi: 10.1016/J.YJMCC.2010.11.017.

[85] B. Yaman, E. Açıkgöz, S. K. Açıkgöz, and A. Abacı, “Early Repolarization and Short QT Interval Correlation in Healthy Population,” *J. Am. Coll. Cardiol.*, vol. 62, no. 18, p. C220, Oct. 2013, doi: 10.1016/J.JACC.2013.08.622.

[86] E. J. Topol, “High-performance medicine: the convergence of human and artificial intelligence,” *Nature Medicine*, vol. 25, no. 1. Nature Publishing Group, pp. 44–56, Jan. 2019, doi: 10.1038/s41591-018-0300-7.

[87] A. Van de Water, J. Verheyen, R. Xhonneux, and R. S. Reneman, “An improved method to correct the QT interval of the electrocardiogram for changes in heart rate,” *J. Pharmacol. Methods*, vol. 22, no. 3, pp. 207–217, Nov. 1989, doi: 10.1016/0160-5402(89)90015-6.

[88] A. Towfighi, L. Zheng, and B. Ovbiagele, “Sex-specific trends in midlife coronary heart disease risk and prevalence,” *Arch. Intern. Med.*, vol. 169, no. 19, pp. 1762–1766, Oct. 2009, doi: 10.1001/archinternmed.2009.318.

[89] Z. Z.-A. of translational medicine and undefined 2016, “A gentle introduction to artificial neural networks,” *ncbi.nlm.nih.gov*.

[90] M. Svensén and C. M. Bishop, “Pattern Recognition and Machine Learning Solutions to the Exercises: Web-Edition,” 2002.

[91] “MathWorks - Creatori di MATLAB e Simulink - MATLAB e Simulink - MATLAB & Simulink.” https://it.mathworks.com/?s_tid=gn_logo (accessed Sep. 21, 2021).

[92] “Systems of Units. Some Important Conversion Factors.”

[93] S. RAI and A. RAI, “Review: Nanotechnology- The secret of fifth industrial revolution and the future of next generation,” *Nusant. Biosci.*, vol. 7, no. 2, 2015.

[94] A. Cappy, D. Stievenard, and D. Vuillaume, “Nanotechnology : the Next Industrial Revolution ?”

[95] R. Vasita and D. S. Katti, “Nanofibers and their applications in tissue engineering,” *International Journal of Nanomedicine*, vol. 1, no. 1. Dove Press, pp. 15–30, 2006, doi: 10.2147/nano.2006.1.1.15.

[96] A. Rakovich, “Nanomaterials for biosensing and phototherapy applications,” in *Proceedings - International Conference Laser Optics 2018, ICLO 2018*, Aug. 2018, p. 540, doi: 10.1109/LO.2018.8435651.

[97] M. S. Islam, B. C. Ang, A. Andriyana, and A. M. Afifi, “A review on fabrication of nanofibers via electrospinning and their applications,” *SN Appl. Sci.*, vol. 1, no. 10, Oct. 2019, doi: 10.1007/s42452-019-1288-4.

[98] J. Fang, H. T. Niu, T. Lin, and X. G. Wang, “Applications of electrospun nanofibers,” *Chinese Science Bulletin*, vol. 53, no. 15. pp. 2265–2286, Aug. 2008, doi: 10.1007/s11434-008-0319-0.

[99] A. K. Gaharwar, S. Sant, M. J. Hancock, and S. A. Hacking, “Nanomaterials in tissue engineering: Fabrication and applications,” *Nanomater. Tissue Eng. Fabr. Appl.*, pp. 1–444, 2013.

[100] R. Bhattarai, R. Bachu, S. Boddu, and S. Bhaduri, “Biomedical Applications of Electrospun Nanofibers: Drug and Nanoparticle Delivery,” *Pharmaceutics*, vol. 11, no. 1, p. 5, Dec. 2018, doi: 10.3390/pharmaceutics11010005.

[101] L. Wang and G. Healey, “Using Zernike moments for the illumination and geometry invariant classification of multispectral texture,” *IEEE Trans. Image Process.*, vol. 7, no. 2, pp. 196–203, 1998, doi: 10.1109/83.660996.

[102] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002, doi: 10.1109/TPAMI.2002.1017623.

[103] N. Otsu, “THRESHOLD SELECTION METHOD FROM GRAY-LEVEL HISTOGRAMS.,” *IEEE Trans Syst Man Cybern*, vol. SMC-9, no. 1, pp. 62–66, 1979, doi: 10.1109/tsmc.1979.4310076.

[104] X. C. Yuan, L. S. Wu, and Q. Peng, “An improved Otsu method using the weighted object variance for defect detection,” *Appl. Surf. Sci.*, vol. 349, pp. 472–484, Sep. 2015, doi: 10.1016/j.apsusc.2015.05.033.

[105] F. Zhou, G. Liu, F. Xu, and H. Deng, “A generic automated surface defect detection based on a bilinear model,” *Appl. Sci.*, vol. 9, no. 15, Aug. 2019, doi: 10.3390/app9153159.

[106] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, May 27, 2015, doi: 10.1038/nature14539.

[107] Y. Bengio, “Deep Learning of Representations for Unsupervised and Transfer Learning,” 2012.

[108] C. Ieracitano, A. Adeel, F. C. Morabito, and A. Hussain, “A novel statistical analysis and autoencoder driven intelligent intrusion detection approach,” *Neurocomputing*, vol. 387, pp. 51–62, Apr. 2020, doi: 10.1016/J.NEUCOM.2019.11.016.

[109] C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, and F. C. Morabito, “A Convolutional Neural Network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings,” *Neurocomputing*, vol. 323, pp. 96–107, Jan. 2019, doi: 10.1016/j.neucom.2018.09.071.

[110] E. Ragusa, P. Gastaldo, R. Zunino, M. J. Ferrarotti, W. Rocchia, and S. Decherchi, “Cognitive Insights into Sentic Spaces Using Principal Paths,” *Cogn. Comput. 2019 115*, vol. 11, no. 5, pp. 656–675, Jul. 2019, doi: 10.1007/S12559-019-09651-1.

[111] F. Gao, T. Huang, J. Sun, J. Wang, A. Hussain, and E. Yang, “A New Algorithm for SAR Image Target Recognition Based on an Improved Deep Convolutional Neural Network,” *Cogn. Comput. 2018 116*, vol. 11, no. 6, pp. 809–824, Jun. 2018, doi: 10.1007/S12559-018-9563-Z.

[112] Q. Lian, W. Yan, X. Zhang, and S. Chen, “Single image rain removal using image decomposition and a dense network,” *IEEE/CAA J. Autom. Sin.*, vol. 6, no. 6, pp. 1428–1437, Nov. 2019, doi: 10.1109/JAS.2019.1911441.

[113] E. Principi, D. Rossetti, S. Squartini, and F. Piazza, “Unsupervised electric motor fault detection by using deep autoencoders,” *IEEE/CAA J. Autom. Sin.*, vol. 6, no. 2, pp. 441–451, Mar. 2019, doi: 10.1109/JAS.2019.1911393.

[114] G. Boracchi, D. Carrera, and B. Wohlberg, “Novelty detection in images by sparse representations,” in *IEEE SSCI 2014 - 2014 IEEE Symposium Series on Computational Intelligence - IES 2014: 2014 IEEE Symposium on Intelligent Embedded Systems, Proceedings*, Jan. 2014, pp. 47–54, doi: 10.1109/INTELES.2014.7008985.

[115] D. Carrera, F. Manganini, G. Boracchi, and E. Lanzarone, “Defect detection in SEM images of nanofibrous materials,” *IEEE Trans. Ind.*

Informatics, vol. 13, no. 2, pp. 551–561, Apr. 2017, doi: 10.1109/TII.2016.2641472.

[116] P. Napoletano, F. Piccoli, and R. Schettini, "Anomaly detection in nanofibrous materials by CNN-based self-similarity," *Sensors (Switzerland)*, vol. 18, no. 1, Jan. 2018, doi: 10.3390/s18010209.

[117] C. Ieracitano, F. Pantó, N. Mammone, A. Paviglianiti, P. Frontera, and F. C. Morabito, "Toward an Automatic Classification of SEM Images of Nanomaterials via a Deep Learning Approach," in *Smart Innovation, Systems and Technologies*, vol. 151, Springer Science and Business Media Deutschland GmbH, 2020, pp. 61–72.

[118] C. Ieracitano, A. Paviglianiti, N. Mammone, M. Versaci, E. Pasero, and F. C. Morabito, "SoCNNet: An Optimized Sobel Filter Based Convolutional Neural Network for SEM Images Classification of Nanomaterials," *Smart Innov. Syst. Technol.*, vol. 184, pp. 103–113, 2021, doi: 10.1007/978-981-15-5093-5_10.

[119] C. Ieracitano, N. Mammone, A. Paviglianiti, and F. C. Morabito, "Toward an Augmented and Explainable Machine Learning Approach for Classification of Defective Nanomaterial Patches," pp. 244–255, Jun. 2021, doi: 10.1007/978-3-030-80568-5_21.

[120] Z. M. Huang, Y. Z. Zhang, S. Ramakrishna, and C. T. Lim, "Electrospinning and mechanical characterization of gelatin nanofibers," *Polymer (Guildf).*, vol. 45, no. 15, pp. 5361–5368, Jul. 2004, doi: 10.1016/J.POLYMER.2004.04.005.

[121] S. A. Theron, E. Zussman, and A. L. Yarin, "Experimental investigation of the governing parameters in the electrospinning of polymer solutions," *Polymer (Guildf).*, vol. 45, no. 6, pp. 2017–2030, Mar. 2004, doi: 10.1016/j.polymer.2004.01.024.

[122] A. Abutaleb, D. Lolla, A. Aljuhani, H. U. Shin, J. W. Rajala, and G. G. Chase, "Effects of surfactants on the morphology and properties of electrospun polyetherimide fibers," *Fibers*, vol. 5, no. 3, Sep. 2017, doi: 10.3390/fib5030033.

[123] T. Chaira and A. Ray, "Fuzzy image processing and applications with MATLAB," 2017.

[124] M. Versaci, F. C. Morabito, and G. Angiulli, "Adaptive Image Contrast Enhancement by Computing Distances into a 4-Dimensional Fuzzy Unit Hypercube," *IEEE Access*, vol. 5, pp. 26922–26931, Nov. 2017, doi: 10.1109/ACCESS.2017.2776349.

[125] M. Versaci, S. Calcagno, and F. C. Morabito, "Image contrast enhancement by distances among points in fuzzy hyper-cubes," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9257, pp. 494–505, 2015, doi: 10.1007/978-3-319-23117-4_43.

[126] M. Versaci, … S. C.-2015 I. I., and undefined 2015, "Fuzzy geometrical

approach based on unit hyper-cubes for image contrast enhancement," *ieeexplore.ieee.org*.

[127] B. Du, W. Xiong, J. Wu, L. Zhang, L. Zhang, and D. Tao, "Stacked Convolutional Denoising Auto-Encoders for Feature Representation," *IEEE Trans. Cybern.*, vol. 47, no. 4, pp. 1017–1027, Apr. 2017, doi: 10.1109/TCYB.2016.2536638.

[128] P. Baldi, "Autoencoders, Unsupervised Learning, and Deep Architectures," 2012.

[129] Y. Ollivier, "Auto-encoders: reconstruction versus compression," Mar. 2014.

[130] P. Liu, P. Zheng, and Z. Chen, "Deep Learning with Stacked Denoising Auto-Encoder for Short-Term Electric Load Forecasting," *Energies*, vol. 12, no. 12, p. 2445, Jun. 2019, doi: 10.3390/en12122445.

[131] P. S. Bradley and O. L. Mangasarian, "Massive data discrimination via linear support vector machines," *Optim. Methods Softw.*, vol. 13, no. 1, pp. 1–10, 2000, doi: 10.1080/10556780008805771.

[132] S. Balakrishnama, A. G. information Processing, and undefined 1998, "Linear discriminant analysis-a brief tutorial," *music.mcgill.ca*.

[133] M. Ojala, G. G.-J. of M. L. Research, and undefined 2010, "Permutation tests for studying classifier performance.," *jmlr.org*, vol. 11, pp. 1833–1863, 2010.

[134] C. Ieracitano, F. Pantó, N. Mammone, A. Paviglianiti, P. Frontera, and F. C. Morabito, "Toward an Automatic Classification of SEM Images of Nanomaterials via a Deep Learning Approach," *Smart Innov. Syst. Technol.*, vol. 151, pp. 61–72, 2020, doi: 10.1007/978-981-13-8950-4_7.

[135] H. Jiang, K. Huang, R. Zhang, and A. Hussain, "Style-Neutralized Pattern Classification Based on Adversarially Trained Upgraded U-Net," *Cogn. Comput. 2019 134*, vol. 13, no. 4, pp. 845–858, Sep. 2019, doi: 10.1007/S12559-019-09660-0.

[136] "Generalized adversarial training in riemannian space," *ieeexplore.ieee.org*.

[137] M. E. H. Chowdhury *et al.*, "Can AI Help in Screening Viral and COVID-19 Pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020, doi: 10.1109/ACCESS.2020.3010287.

[138] T. Rahman *et al.*, "Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-rays Images," *Comput. Biol. Med.*, p. 104319, Mar. 2021, doi: 10.1016/j.compbiomed.2021.104319.

[139] L. EJ, D. DM, H. RC, O. RJ, S. SG, and S. DB, "Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV)," *Nucleic Acids Res.*, vol. 46, no. D1, pp. D708–D717, Jan. 2018, doi: 10.1093/NAR/GKX932.

[140] “COVID-19: polmonite, infiammazione, rischio tromboembolico ed effetti sul sistema cardiovascolare - Gradenigo.” https://www.gradenigo.it/news/covid-19-polmonite-infiammazione-rischio-tromboembolico/ (accessed Oct. 29, 2021).

[141] “Le statistiche del coronavirus nel mondo.” https://statistichecoronavirus.it/ (accessed Oct. 29, 2021).

[142] “Coronavirus disease (COVID-19).” https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed Jul. 02, 2020).

[143] “Covid-19 PCR test: how does it work? are there any alternatives? | Auxologico.” https://www.auxologico.com/covid-19-pcr-test-how-does-it-work-are-there-any-alternatives (accessed Oct. 29, 2021).

[144] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, “Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network,” *Appl. Intell. 2020 512*, vol. 51, no. 2, pp. 854–864, Sep. 2020, doi: 10.1007/S10489-020-01829-7.

[145] D. D *et al.*, “The Role of Imaging in the Detection and Management of COVID-19: A Review,” *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 16–29, 2021, doi: 10.1109/RBME.2020.2990959.

[146] “Deep learning approaches for COVID-19 detection based on chest X-ray images | Elsevier Enhanced Reader.” https://reader.elsevier.com/reader/sd/pii/S0957417420308198?token=8A7F6CAC26D35E27E9FE1D731673FC47A4D885B43A53A556D428619885FE0D1DD7FAEF6162B7B244059B7266CCA06413&originRegion=eu-west-1&originCreation=20211029145645 (accessed Oct. 29, 2021).

[147] I. Katsamenis, E. Protopapadakis, A. Voulodimos, A. Doulamis, and N. Doulamis, “Transfer Learning for COVID-19 Pneumonia Detection and Classification in Chest X-ray Images,” doi: 10.1101/2020.12.14.20248158.

[148] A. M. Freeman and J. Townes R. Leigh, “Viral Pneumonia,” *Encycl. Respir. Med. Four-Volume Set*, pp. 456–466, Jul. 2021.

[149] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *J. Big Data 2016 31*, vol. 3, no. 1, pp. 1–40, May 2016, doi: 10.1186/S40537-016-0043-6.

[150] H. Witten, E. Frank, M. Kaufmann, and J. Geller, “Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations Review by,” 2000.

[151] H. S.-J. of statistical planning and inference and undefined 2000, “Improving predictive inference under covariate shift by weighting the log-likelihood function,” *Elsevier*.

[152] C. Wang, S. M.-T. international joint conference on, and undefined 2011, “Heterogeneous domain adaptation using manifold alignment,” *aaai.org*.

[153] Y. Zhu *et al.*, “Heterogeneous transfer learning for image classification,” *aaai.org*.

[154] M. Harel and S. Mannor, “Learning from multiple outlooks,” *Proc. 28th Int. Conf. Mach. Learn. ICML 2011*, pp. 401–408, 2011.

[155] J. Nam, W. Fu, S. Kim, T. M.-… on S. Engineering, and undefined 2017, “Heterogeneous defect prediction,” *ieeexplore.ieee.org*.

[156] P. Prettenhofer, B. S.-P. of the 48th annual meeting of the, and undefined 2010, “Cross-language text classification using structural correspondence learning,” *aclweb.org*, pp. 11–16, 2010.

[157] B. Zhao, J. Feng, X. Wu, and S. Yan, “A Survey on Deep Learning-based Fine-grained Object Classification and Semantic Segmentation,” *Int. J. Autom. Comput.*, vol. 14, no. 2, pp. 119–135, 2017, doi: 10.1007/s11633-017-1053-3.

[158] F. Chollet, “Xception: Deep Learning With Depthwise Separable Convolutions.” pp. 1251–1258, 2017.

[159] Y. Tai, J. Yang, and X. Liu, “Image Super-Resolution via Deep Recursive Residual Network.” pp. 3147–3155, 2017.

[160] T. Wolf *et al.*, “Transformers: State-of-the-Art Natural Language Processing,” pp. 38–45, Nov. 2020, doi: 10.18653/V1/2020.EMNLP-DEMOS.6.

[161] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, “Transformers in Vision: A Survey,” Jan. 2021.