Politecnico di Torino

Politecnico di Torino

MSc in Biomedical Engineering

Ultrasound Assessment of Inferior Vena Cava Dynamics as a Predictor of Right Atrial Pressure

Candidate: Noemi Auria
Antonio Capitanio

Supervisors: Prof. Luca Mesin
Ing. Piero Policastro

Academic Year 2024/25

Abstract

Right Atrial Pressure (RAP) represents a fundamental hemodynamic parameter in the evaluation of cardiac pathologies. However, the gold standard for assessing this parameter remains an invasive measurement. The invasiveness of such techniques limits their routine use, highlighting the need for reliable non-invasive alternatives.The diameter and Caval Index (CI) of the inferior vena cava (IVC), derived from ultrasound imaging, allow an indirect estimation of RAP through semi-quantitative criteria defined by current international guidelines, which classify RAP into three levels (low, intermediate, and high). While these recommendations offer a practical and straightforward approach, their application is often limited by suboptimal precision and reproducibility. Considering these limitations, a dedicated software tool named "VIPER" has been developed. This semi-automatic algorithm is capable of accurately tracking the borders of the IVC, from which it extracts key parameters: vessel diameter, CI, Respiratory (RCI) and Cardiac (CCI) Caval Index.

In this thesis, two different algorithms of VIPER software, one developed on Matlab and one on Python, were employed to analyze and segment the IVC from ultrasound video recordings provided by expert clinicians at the Fondazione Toscana Gabriele Monasterio, Pisa. No statistically significant differences were observed between the two algorithm tools for either diameter or CI measurements, confirming the consistency and reproducibility of the automated software. The IVC diameters measured by the operator had a mean value of 17.24 ± 4.03 mm, while the automated estimates obtained from the Python and Matlab algorithms reported mean values of 15.89 ± 5.20 mm and 16.13 ± 5.44 mm, respectively. Similarly, the CI computed from the Python algorithm yielded a mean of 0.31 ± 0.16 mm, compared to 0.28 ± 0.16 mm from the Matlab-based version.

The study population consists of 37 patients (16 males and 21 females) with a mean age of 67.49 $\pm$ 16.81 years. For each subject, a comprehensive set of clinical, echocardiographic and invasive variables were available, along with additional features extracted by the two algorithm tools. Notably, invasive measurements were excluded from the dataset to preserve the non-invasive nature of the prediction.

A range of linear and non-linear classification models was implemented with the aim of assigning each patient to a RAP class. Among these, the best-performing approach was a meta-classifier trained using features extracted by the Matlab version of VIPER. This model integrates the predictions of two base classifiers: a linear Ridge model and a non-linear Support Vector Machine (SVM). The meta-classifier achieved an accuracy of 73%, as estimated through Leave-One-Out Cross-Validation (LOO-CV). The same meta-classifier architecture, when applied to features derived from the Python version of VIPER, achieved a LOO-CV accuracy of 65%. In comparison, when using the operator-measured diameter, the standard guideline-based assessment methods achieved a lower accuracy of 54%, suggesting that algorithmic approaches may offer a more reliable support for non-invasive RAP assessment and may represent a promising advancement in clinical practice.

Contents

List of Figures

List of Tables

List of Acronyms

AO Aorta.

ASE American Society of Echocardiography.

AV Atrioventricular.

CCI Cardiac Collapsibility Index.

CI Caval Index.

CNN Convolutional Neural Network.

CVP Central Venous Pressure.

DL Deep Learning.

FC Fully Connected.

HF Heart Failure.

IQR Interquartile Range.

IVC Inferior Vena Cava.

JVD Jugular Vein Distention.

LVEF Left Ventricular Ejection Fraction.

ML Machine Learning.

OLP Ordinary Least Product.

OLS Ordinary Least Square.

PAH Pulmonary Arterial Hypertension.

PCWP Pulmonary Capillary Wedge Pressure.

PH Pulmonary Hypertension.

PVR Pulmonar Vascular Resistance.

PZT Lead Zirconate Titanate.

RA Right Atrial.

RAP Right Atrial Pressure.

RCI Respiratory Collapsibility Index.

RHC Right Heart Catheterization.

RV Right Ventricular.

SA Sinoatrial.

SVC Superior Vena Cava.

SVM Support Vector Machine.

US UltraSound.

VIF Variance Inflation Factor.

Chapter 1

Cardiovascular System

In order to live, the human cells need to exchange substances with surrounding areas. Every cell by performing the diffusion process embeds the substances they need, such as oxygen and nutrients, and releases waste products like carbon dioxide. The diffusion process itself is pretty slow, so in order to make the transportation of substances faster, the cardiovascular system comes in help. This system is made up of three main elements:

- the heart whose job is to pump the blood into the blood vessels
- blood vessels which are ducts where the blood flows
- the blood, a fluid which circulates in the blood vessels towards the whole body.

The general scheme of blood flow in the cardiovascular system is shown in Figure 1.1. The cardiovascular system consist of two main circuits:

- the pulmonary circulation formed by the lung vessels and all the vessels connecting the lungs to the heart
- the systemic circulation formed by all the remaining vessels directed from the heart to the other body parts.

In both circuits, there is a dense network of capillaries where the exchange of nutrients and gas occurs. When blood flows inside the pulmonary capillaries, carbon dioxide is released while oxygen flows inside the blood and so it is called oxygenated blood. Whereas, when blood flows inside the systemic capillaries is rich in carbon dioxide due to the consumption of oxygen from the cells of human body: in this case the blood is called deoxygenated blood. [1]

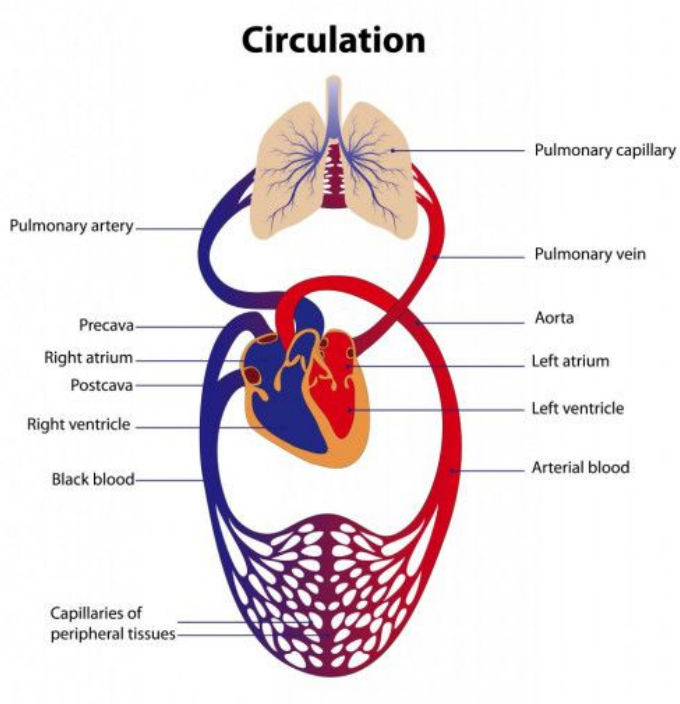


Figure 1.1: *Scheme of the blood system: blue deoxygentaed blood, red oxyganeted blood*

1.1 Heart

Heart pumps the oxygenated blood towards the organs through the blood vessels and removes waste products and carbon dioxide. It is formed by four chambers: atria are the upper chambers and they receive the blood coming from the veins, while ventricles are the two lower chambers and they receive the blood from atria and they generate a pressure big enough to push the blood towards the big arteries. [1] It is located in the center of the chest, near the lungs and it is made of three layers of tissues. The inner one is the endocardium, a thin epithelial layer which also forms the surface of the valves followed by the myocardium, a thick layer of muscular tissue that allows heart chambers to contract and relax to pump blood to the whole body. Pericardium is the sac that surrounds the heart and it is made of thin layers of connective tissue, it holds the heart in place and protects it. Heart chambers are four and are separated by heart valves, which make sure that the blood keeps flowing in the right direction. Heart valves help control the direction the blood flows and they prevent blood from flowing backward.The heart has four valves:

- the tricuspid valve separates the right atrium and right ventricle
- the mitral valve separates the left atrium and left ventricle
- the pulmonary valve separates the right ventricle and the pulmonary artery
- the aortic valve separates the left ventricle and aorta.

The valves open and shut in time with the pumping action of heart's chambers. Oxygen-poor blood from the body enters the heart through two large veins called

the Superior and Inferior Vena Cava (IVC). The blood enters the heart's right atrium and is pumped to the right ventricle, which in turn pumps the blood to lungs. Lungs add oxygen to your blood which returns to your heart through the pulmonary veins (Figure 1.2). [2]

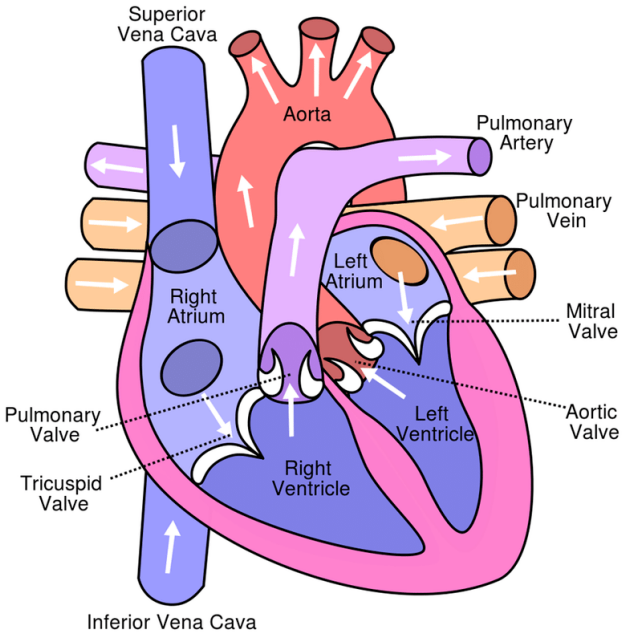


Figure 1.2: *Representation of heart and its parts*

1.1.1 Electrical activities of heart

In order for the heart to pump the blood into the blood system, the cardiac muscle needs to contract synchronously: first atria both contract and then ventricle both contract. Contractions of cardiac muscle are generated by signals coming from the muscle itself. There are two types of cells that are able to start and coordinate the cardiac contractions:

- pacemaker cells that are able to start action potentials and control heart rate,they are mainly located in two specific zones of myocardium which are the sinoatrial node (SA) and atrioventricolar node (AV).

- conduction fibers lead action potentials generated by pacemaker cells and they propagate them inside the heart. [1]

The impulse starts in the Sinoatrial (SA) node in the right atrium, then the electrical impulse travels through the atria, causing them to contract and force blood into the ventricles. The electrical impulse is then transmitted down to the pacemaker cells in the Atrioventricular (AV) node, which is located between the atria and the ventricles. The impulse is delayed here so that the ventricles have time to finish filling with blood. The AV node fires another impulse that travels along the walls of the ventricles, causing them to contract and push the blood out of the heart. The ventricles relax, and a heartbeat process starts again in the SA node (Figure 1.3).

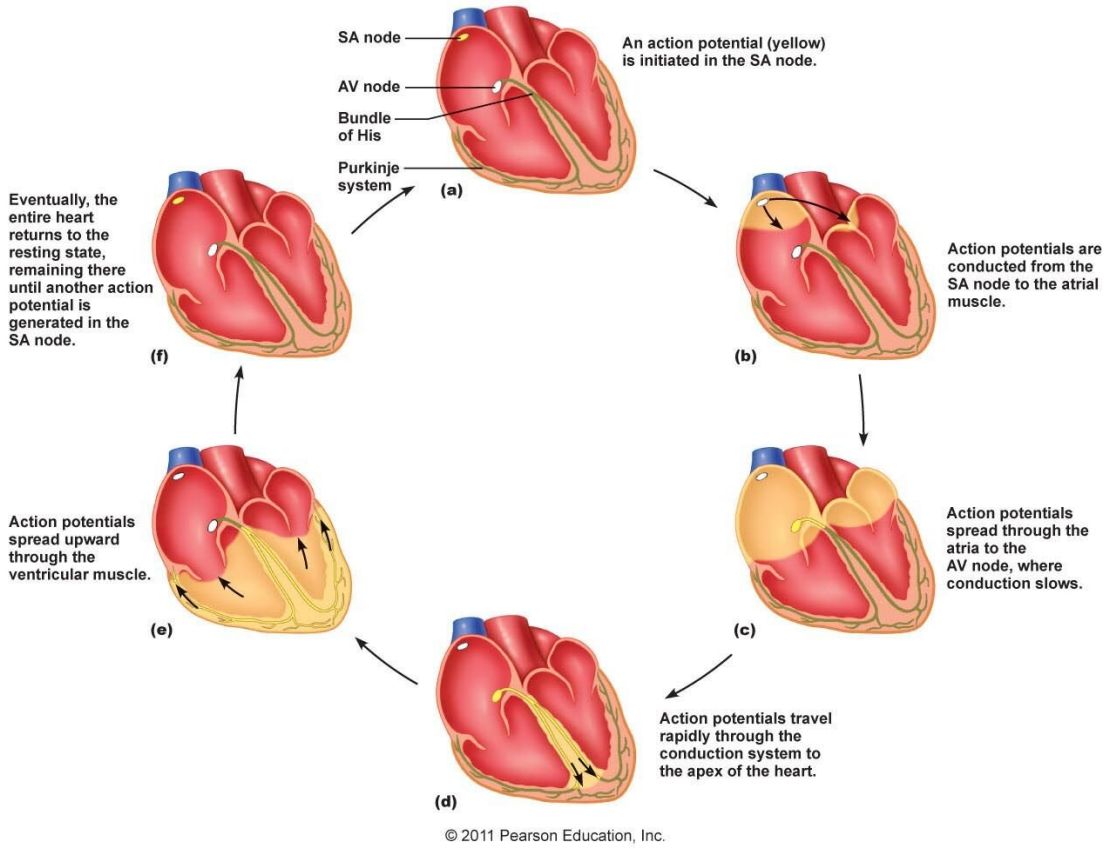


Figure 1.3: *Representation of action potentials propagation inside the heart*

The heart's electrical activity can be monitored through electrodes placed on the skin surface. The resulting signal, known as the Electrocardiogram (ECG), is composed of positive and negative waves separated by isoelectric segments. This waveform follows a repetitive pattern that recurs consistently with each cardiac cycle, and its frequency is determined by the individual's heart rate. ECG signals is usually characterized by three types of characteristics waves (Figure 1.4):

- P wave, a positive deflection due to the atrial depolarization
- PQ interval which estimates the conduction time through the AV node
- QRS complex which includes a small negative deflection followed by a peak positive deflection, it indicates the ventricular depolarization
- QT interval which estimates the releasing time of the ventricle, called ventricular diastole
- T wave, a positive deflection due to the repolarization of the ventricle

The R-R interval indicates the time between two successive peaks of QRS complex and it represents the time between a heartbeat and the following one. The repolarization of the atrial chambers is not visible on the ECG trace since they are hidden by the QRS complex. ECG signal amplitude varies from 0.5 to 4 mV with a bandwidth between 0.01 and 250 Hz. [1]

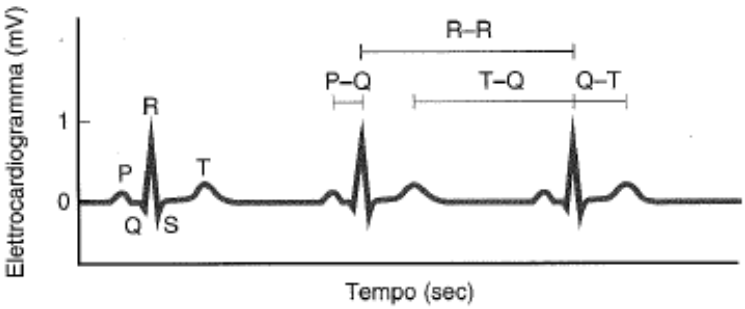


Figure 1.4: *Normal ECG waveform*

1.1.2 Cardiac cycle

The cardiac cycle is the period that begins with atrial contraction and ends with ventricular relaxation (Figure 1.5). The heart goes through two broad phases during this cycle:

- systole, the phase of contraction during which the blood enters the circulation
- diastole, the phase of relaxation by which the chambers of the heart get refilled with blood.

Both the atria and the ventricles go through systole and diastole, and coordination between them is required for efficient ejection of blood and refilling.

Early in the cardiac cycle, atria and ventricles are in diastole: blood enters the right atrium through the superior and inferior vena cava and the coronary sinus while the left atrium simultaneously receives oxygenated blood via the four pulmonary veins. The atrioventricular valves (mitral and tricuspid) are open, allowing unimpeded flow of blood from the atria into the ventricles.The semilunar valves (pulmonary and aortic) are closed, preventing backflow from the great arteries to the ventricles.

Atrial contraction follows atrial depolarization, which is reflected by the P wave in the ECG. Early in atrial systole, ventricles are already 70–80% filled with blood passively during diastole. The additional 20–30% of ventricular filling comes from the atrial contraction, or "atrial kick." This lasts approximately 100 milliseconds and ends before the onset of ventricular systole, when the atrial myocardium resumes diastole.

Ventricular systole follows ventricular depolarization, as marked by the QRS complex of the electrocardiogram (ECG), and has two phases (total duration of about 270 ms):

- isovolumetric contraction where ventricular pressure rises, but is not high enough to cause the semilunar valves to open. Since no blood is being expelled during this early phase, ventricular volume also doesn't change.
- Ventricular ejection phase: ventricular contraction keep going, the pressure within the chambers is higher than that within the aorta and pulmonary trunk, forcing the semilunar valves open and blood out of the heart.

Finally, ventricular diastole begins with ventricular repolarization, represented by the T wave of the ECG, the onset of muscle relaxation and beginning of the next filling phase. [3]

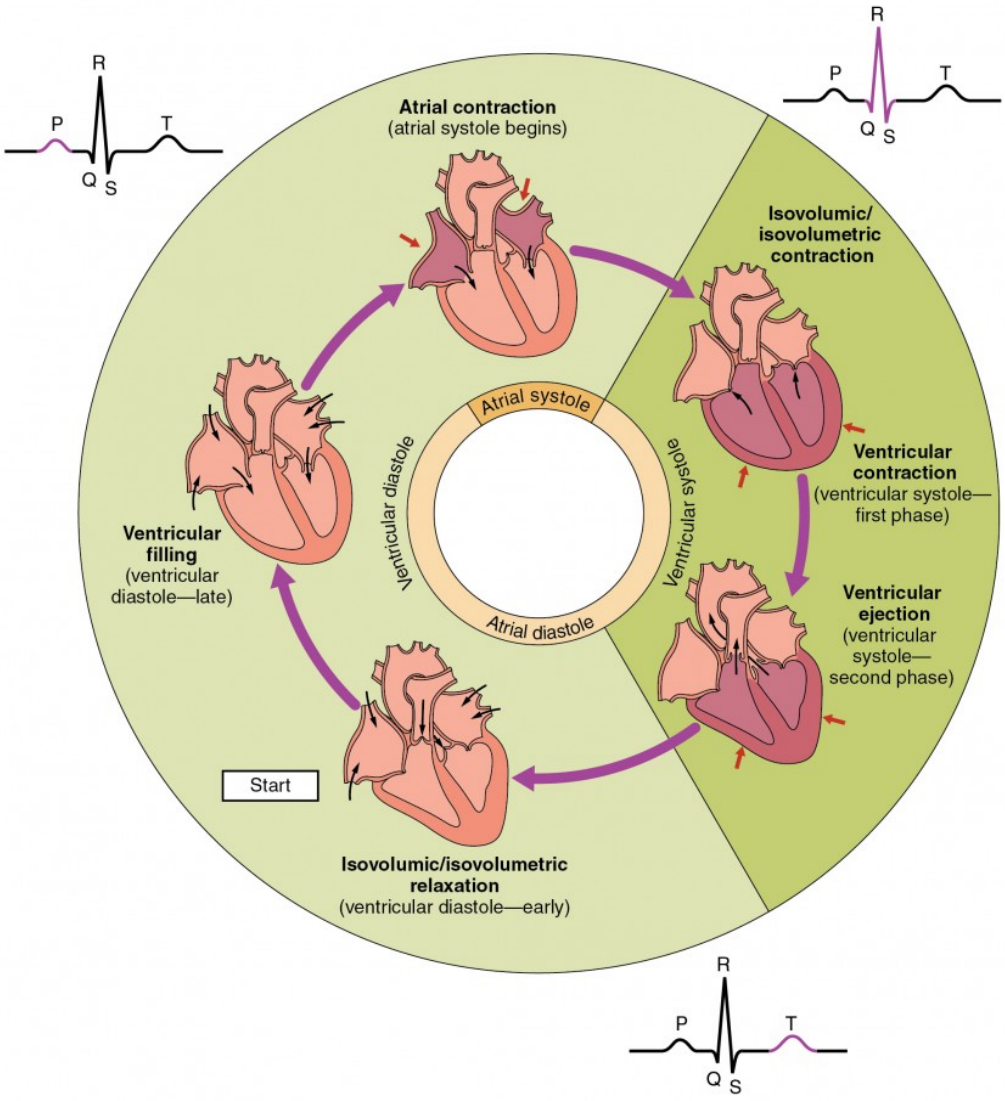


Figure 1.5: *Phases of cardiac cycle*

1.2 Blood vessels

Blood vessels can be classified based on their direction from the heart to the tissues or vice versa and based on their caliber. Arteries transport oxygenated blood from the heart to the capillaries, while veins bring back deoxygenated blood to the heart (Figure 1.6). Every blood vessel has a cavity in which the blood flow, called lumen, is surrounded by an epithelial layer, called endothelium. The wall of all vessels is formed by endothelial cells, smooth muscle, and fibrous / elastic connective tissue which allows the vessel to distend under pressure without breaking, thanks to an extracellular protein called elastin.

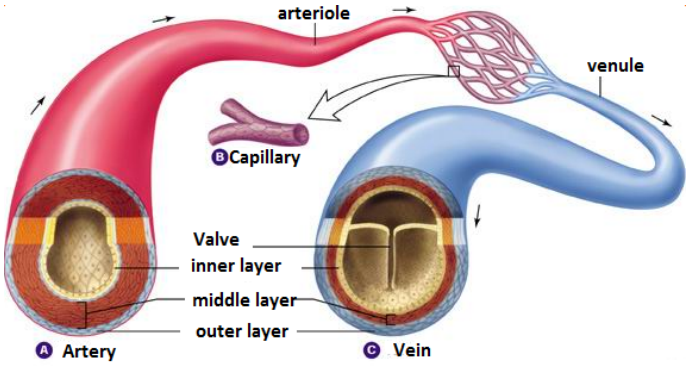


Figure 1.6: *Section of a blood vessel*

1.2.1 Arteries and veins

The largest size artery is the aorta whose diameter is 12.5 mm and thickness 2 mm, while smaller arteries have a diameter between 2 and 6 mm and a thickness of 1 mm. Bigger arteries offer small resistance to the flow and their main job is to carry the blood: this function of the arteries is possible thanks to their elasticity and rigidity. Arteries can be considered as blood reservoirs (Figure 1.7). During systole it occurs an expansion of the arteries wall due to the flowing of the blood; part of the elastic energy produced will be used to let the blood keep flowing during diastole. The

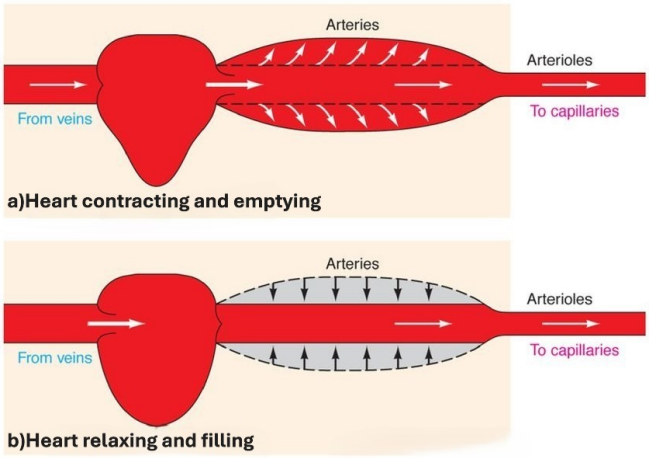


Figure 1.7: *Arteries as blood reservoirs: a) during systole the entrance of the blood through the aorta increase the volume, causing en expansion of the walls; b) during diastole, the elastic component of the aorta's walls keep pushing the blood ahead.*

pressure inside the aorta is called arterial pressure whose value changes during the cardiac cycle: the maximum value is achieved during systole and is called systolic pressure while the minimum value is achieved during diastole and is called diastolic pressure. [1]

While arteries have thicker walls, in order to sustain high blood pressure, veins have thinner walls and they can distend much more easily. Even though the venous pressure is much lower than the arterial pressure, veins contain a quantity of blood much higher than the one contained in the arteries. In fact, in normal condition al-

most 60% of the blood volume is contained inside the veins. However, this blood can be moved towards the arteries when needed and for this reason veins are considered as a volume reservoir.

1.2.2 Compliance

Every blood vessel tends to expand when the pressure inside the vessel increases and to contract when the pressure decreases. To be more precise,the difference between the internal and external pressure (P_{in}- P_{ext}) of the vessel determines the expansion or contraction of the vessel: this pressure is called transmural pressure P_{tm}. [1] When $P_{tm} > 0$ the volume of the vessel increases, while when $P_{tm} < 0$ the volume of the vessel decreases. Vessel compliance is defined as

$$C = \frac{\Delta V}{\Delta P_{\mathrm{tm}}} \tag{1.1}$$

where ΔV represent volume variation of the vessel, P_{tm} represent transmural pressure. [1] The relation between ΔV and P_{tm} is generally represented by a volume-pressure or capacitance curve (Figure 1.8) and shows an increase of the volume vessel with increasing transmural pressure.The same change in P_{tm} will produce different changes in vessel size, depending on the resting value of P_{tm} and volume vessel V: if average P_{tm} is low size changes will be large, while at a higher P_{tm} value the vessel size will be larger and smaller phasic changes will be detected. [4]

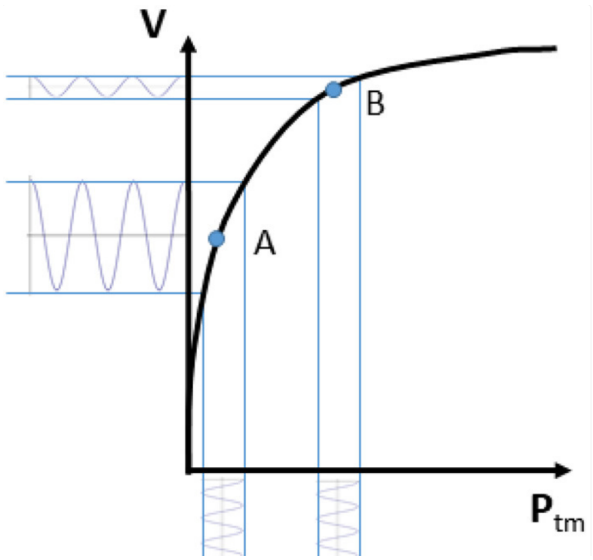


Figure 1.8: *Volume-pressure curve of a venous blood vessel: point A high vessel compliance, point B low vessel compliance*

The primary distinctions between veins and arteries regarding compliance are their mechanical properties: arteries are stiffer and subjected to greater pressure fluctuations than veins. Pressure oscillations within the arteries are mostly due to the pulsatile nature of the heart's action, and also to peripheral vascular resistance.; while P_{tm} changes in the veins are largely influenced by a variation of the external pressure. [4] In order to perform as pressure reservoirs, compliance in the arteries needs to be low: a high variation of the pressure cause a small volume variation. [1] Veins,therefore,as said before are a volume reservoirs due to their high compliance:

a small increase of pressure inside the vessel cause a big increase of volume. In other words, veins can adapt themselves to huge increase in blood volume and so they are particularly efficient when it comes to storing blood.

1.2.3 IVC: Inferior Vena Cava

The vena cava is divided into IVC and Superior Vena Cava (SVC). They are the body's two largest veins (Figure 1.9), which drain deoxygenated blood to the right atrium of the heart. Specifically, the IVC drains blood from the lower part of the body (lower limbs and abdomen) to the heart, and the SVC drains blood from the upper part of the body to the heart.

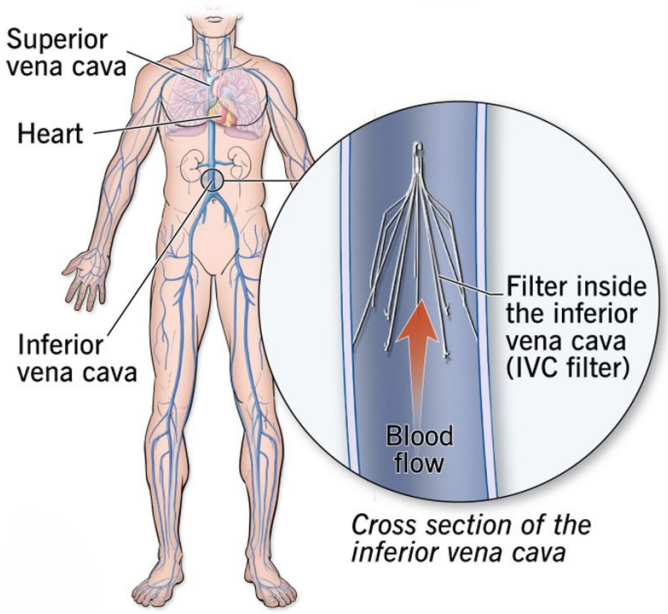


Figure 1.9: *Representative scheme of the veins flowing in the human body*

The IVC runs along the posterior abdominal wall, to the right of the aorta. It arises from the convergence of the two common iliac veins at the sacral promontory. It then connects to the posterior aspect of the heart and, after passing through the hepatic diaphragm, empties into the right atrium. Since IVC represents two-thirds of the venous return system, it can be clinically used as a marker of body volume status, which is crucial for clinical evaluation. [5]

1.3 Right Atrial Pressure

RAP is referred to as the pressure of the blood in the right atrium of the heart. This pressure is related to the volume of blood returning to the heart and to the heart's capacity to pump the blood into the arterial system. The RAP can often correspond to the Central Venous Pressure (CVP). [6] RAP is related to Right Ventricular (RV) diastolic function, volume status, and Right Atrial (RA) compliance. Increased RAP is a predictor of mortality in patients with Heart Failure (HF) due to acquired heart disease. [7] The RAP is measured conventionally after exhaling, through invasive techniques (e.g., right atrium catheterisation) but some studies have shown that the agreement between the RAP measured invasively and echocardiography is modest, even though there are issues related to the reproducibility of IVC measurement by US. [6] In accordance with the recommendations of the American Society of Echocardiography (ASE), the correlation between RAP and echocardiographic indices of right ventricular (RV) diastolic function was assessed. The indices considered were:

- the ratio of tricuspid inflow early diastolic velocity to late diastolic velocity (E/A)
- the ratio of tricuspid inflow early diastolic velocity to tricuspid annular early diastolic tissue Doppler velocity (E/E′)
- the tricuspid inflow deceleration time (DT)
- the diameter of the IVC [7]

ASE also partitioned RAP into 3 categories:

Range	**Value [mmHg]**
Normal	0-5
Intermediate	5-10
High	10-20

Table 1.1: Range of RAP according ASE guidelines

1.3.1 Techniques for RAP monitoring in clinical practice

Currently, pulmonary artery catheterization represents the standard method for the estimation of RAP in clinical practice. Pulmonary artery catheterization (Swan-Ganz or Right Heart Catheterization (RHC)) is an invasive diagnostic procedure that involves the insertion of a catheter through a central vein, advancing it into the pulmonary artery (Figure 1.10). This technique allows for the measurement of right atrial, right ventricular, and pulmonary artery pressures, the estimation of cardiac output, the detection of intracardiac shunts, and the calculation of pulmonary vascular resistance. The hemodynamic data obtained are essential for a deeper understanding of the pathophysiology of heart failure and pulmonary hypertension. The RHC procedure is usually a low risk procedure, but complications can occur and can be fatal.[8]

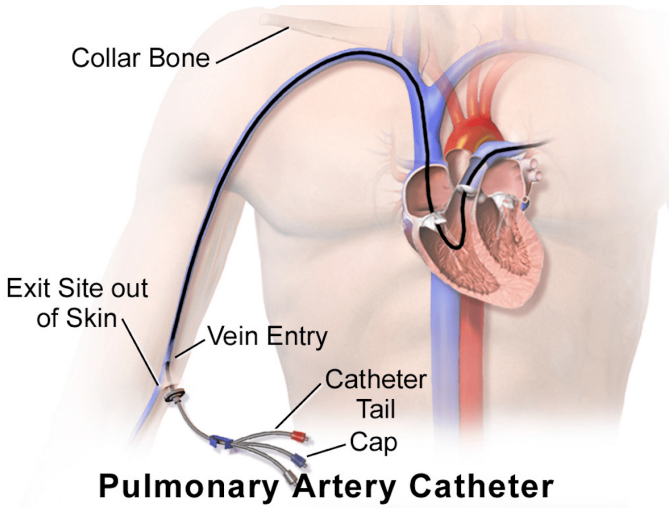


Figure 1.10: *Swan-ganz catheter representation*

Due to the limitations associated with invasive procedures, several non-invasive methods have been developed to estimate RAP, with particular focus on the ultrasound assessment of the IVC. The dynamic changes in the diameter of the IVC provide valuable information about the patient's volume status and RAP. Among the parameters used in clinical practice, the Caval Index (CI) derived from ultrasound recordings, is the standard indicator of IVC pulsatility. Nevertheless, its measurement lacks a standardized protocol and is often affected by artifacts, mainly due to the physiological movements of the IVC during respiration.[9]

1.3.2 Heart Failure and RAP

In the context of HF, congestion refers to the accumulation of fluid within the intravascular space and the interstitial tissues. This condition arises due to elevated cardiac filling pressures, primarily driven by the kidneys maladaptive retention of sodium and water. Congestion is the predominant cause of hospitalization in patients with acute HF (Figure 1.11). [10]

Given that fluid overload plays a central role in the pathophysiology and clinical presentation of HF, therapeutic strategies are largely aimed at achieving effective decongestion. Diuretics represent the primary pharmacological approach for promoting sodium and fluid excretion in patients who are not candidates for renal replacement therapies. Their primary function is to stimulate the kidneys to excrete excess sodium and water. However, when renal response is inadequate, the condition is referred to as diuretic resistance.

Despite the limited evidence supporting its impact on long-term outcomes, this remains the first-line approach approved by cardiology guidelines. When conventional diuretic strategies fail to relieve congestion, renal replacement techniques such as ultrafiltration or hemofiltration offer an alternative method for rapid fluid removal and may restore responsiveness to diuretics. One of the advantages of ultrafiltration in the management of HF is its ability to remove excess fluid predominantly from the extravascular compartment, while preserving the circulating blood volume. This selective mechanism helps to avoid neurohormonal activation and the development of renal dysfunction—adverse effects commonly associated with aggressive diuretic

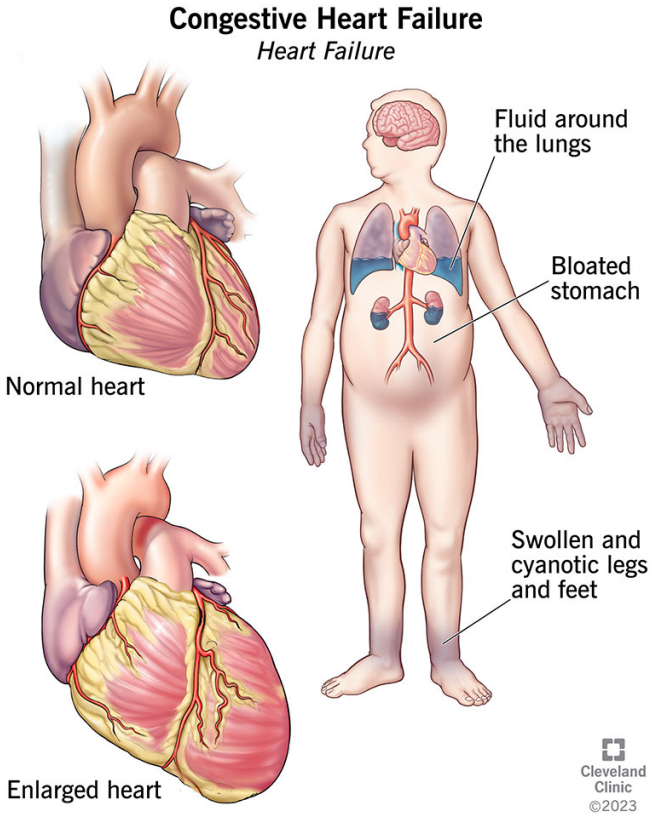


Figure 1.11: *Impact of congestion on human body*

therapy.[11]

Currently, most clinical markers of volume overload (e.g., Jugular Vein Distention (JVD), peripheral edema, or ascites) reflect elevated RAP. Although it is true that an elevation in RAP mirrors an elevation in PCWP in most patients with HF, approximately 25% to 30% of patients have discordance between right- and left-sided filling pressures, with an isolated elevation on either side.Discordance can occur in those with preserved or reduced ejection fraction. [12] HF is further classified by Left Ventricular Ejection Fraction (LVEF).The prognosis and response to treatment of patients with HF differs significantly when patients are stratified based on LVEF. HF can be classified based on LVEF:

- HF with Reduced Ejection Fraction (HFrEF): patients with an LVEF $\leq$40%
- HF with Preserved Ejection Fraction (HFpEF): patients with an LVEF $\geq$50% more commonly affecting older women, with a history of hypertension and, less commonly, coronary disease, than patients with HFrEF. [13]

Patterns of elevated filling pressures, right-sided only, left-sided only, or both, tend to remain relatively stable over time. To enhance clinical assessment of congestion, patients can be classified based on whether right-sided, left-sided, or biventricular filling pressures are elevated (Figure 1.12). Although this classification may better reflect the nature of volume overload, it remains unclear whether patients with isolated pressure elevation have different treatment responses or prognoses compared to those with concordant pressures. A disproportionately elevated

RAP/PCWP ≥ 0.67, referred to as the "right-left equalizer" pattern, has been associated with impaired renal function and worse clinical outcomes, offering preliminary support for this hemodynamic classification. This pattern should be considered when a patient with persistently elevated JVD experiences worsening renal function during diuresis. Further investigation is needed to fully understand the characteristics and prognostic significance of these hemodynamic profiles. [12]

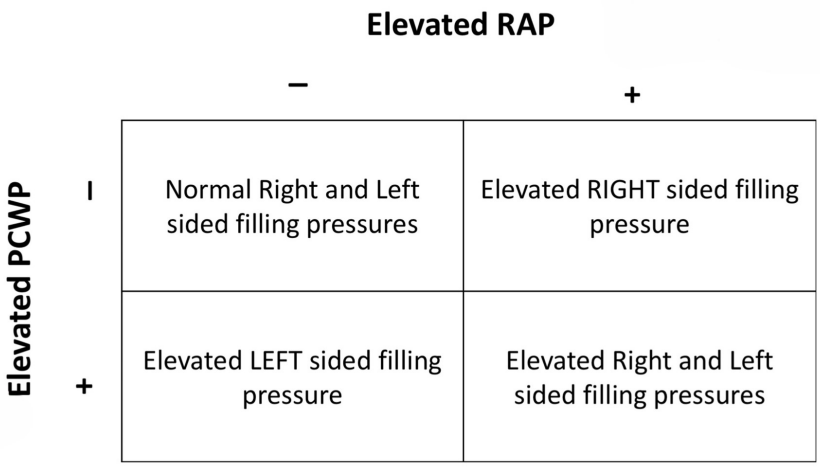


Figure 1.12: *Classification of Congestion Based on Clinical Assessment of Elevated Right- or Left-Sided Filling Pressure*

1.3.3 Pulmonary Hypertension and RAP

RAP is a crucial marker of prognosis in Pulmonary Hypertension (PH), whether the condition arises from pulmonary vascular pathology, such as Pulmonary Arterial Hypertension (PAH) or from cardiac causes like HF. [14] PAH is a progressive condition characterized by proliferative and fibrotic remodeling of the pulmonary vasculature. This leads to increased Pulmonar Vascular Resistance (PVR), reduced pulmonary vascular compliance, right ventricular dysfunction, and ultimately death. Diagnosis of PAH relies on precise hemodynamic assessment of the right heart and pulmonary circulation through RHC. [15] As the disease advances, clinical signs of right ventricular failure emerge, including elevated filling pressures, diastolic dysfunction, and reduced cardiac output. In parallel, the IVC, which drains into the right atrium, also adapts to these hemodynamic shifts, typically becoming dilated and exhibiting reduced inspiratory collapse due to elevated pressure in the right sided. [16]

1.3.4 IVC as a predictor of RAP

The current gold standard for evaluating intravascular congestion involves measuring RAP, typically ranging from 2 to 6 mmHg, and Pulmonary Capillary Wedge Pressure (PCWP), usually between 3 and 8 mmHg, using RHC. However, due to its invasive nature and the fact that it is not routinely performed, RHC is not suitable for the long-term monitoring of patients. As a non-invasive alternative, the evaluation of the IVC through ultrasonography provides a simple and effective method for estimating

RAP. An increase in the diameter of the IVC has been shown to predict a higher risk of hospitalization and mortality in patients with HF.[10] Nevertheless, the accuracy of this method may be affected by several limitations, such as:

- variability in the anatomical site of measurement
- inspiratory movement of the IVC, potentially causing misalignment with the original ultrasound imaging
- interindividual differences in respiratory maneuver
- the lack of integration of other hemodynamic parameters from both ventricles in the current estimation approach

In clinical practice, RAP estimation via ultrasonography is commonly performed by measuring the IVC diameter and calculating its percentage collapse during a brief inspiratory effort(sniff), known as the collapsibility index or caval index. These measurements are interpreted according to established guideline-based criteria, which define threshold combinations of IVC diameter and collapsibility for estimating RAP. [17]

Spontaneous breathing and mechanical ventilation are based on distinct physiological mechanisms. During spontaneous inspiration, negative intrathoracic pressure develops, which facilitates venous return to the heart and leads to a temporary reduction in the diameter of the IVC. At end-expiration, intrathoracic pressure rises toward atmospheric levels, decreasing venous return and resulting in an increase in IVC diameter. This respiratory-driven variation in venous dimensions is commonly assessed through the CI, which quantifies the degree of IVC narrowing during the breathing cycle. [18]

However, certain physiological and pathological conditions may limit the accuracy of this method. In healthy young athletes, for instance, a dilated IVC may be observed despite normal RAP, representing a benign physiological variant. Furthermore, in patients undergoing mechanical ventilation, the IVC is often persistently dilated and may show minimal or absent respiratory collapse. These alterations reduce the reliability of IVC measurements as indirect indicators of RAP in such populations. The classification of RAP levels based on these parameters is summarized in Table 1.2. [19]

RAP	**Normal (0–5 mmHg)**	**Intermediate (5–10 mmHg)**	**High (10–20 mmHg)**
IVC diameter	≤ 21 mm	≤ 21 mm , > 21 mm	> 21 mm
Collapsibility index	$> 50\%$	$< 50\%$, $> 50\%$	$< 50\%$

Table 1.2: Classification of RAP levels based on IVC diameter and collapsibility index, according to the ASE guidelines

Chapter 2

Ultrasonography

UltraSound (US) is based on high-frequency sound waves and can be considered a safe, generally non-invasive medical imaging and diagnostic tool. The frequencies of these waves are between 1 to 20 MHz, higher than the band of human hearing, i.e. above 20 kHz. Sound waves are longitudinal mechanical waves, which can propagate through various materials such as fluids, soft tissues, and solids. The importance of ultrasound techniques in medicine is due to three factors:

- complete safety for the patient due to the absence of ionizing radiation, so this technique is reliable to perform long-term measurements, and it is applicable to patients at risk
- high temporal resolution, US imaging is fast enough to document rapid phenomena such as cardiac contraction
- reflection and diffusion based operation generated by acoustic interfaces [20]

2.1 Physics of ultrasound

The operating principle of ultrasound devices is based on the piezoelectric effect, a property of materials such as Lead Zirconate Titanate (PZT), which can convert mechanical stress into electrical signals and, conversely, deform when subjected to an electric field.

Inside the ultrasound probe, a PZT piezoelectric crystal is exposed to high-frequency electrical pulses from the generator that cause deformation of the crystal and emit pressure waves into the biological tissues. The waves reflected back from the tissue strike the crystal for the second time and get converted into electrical signals by the same piezoelectric effect so that they aid in producing the ultrasound image.

Once generated, these ultrasound waves travel through the body, and their speed of propagation depends on the physical characteristics of the medium they pass through. In general, the propagation speed of a wave is described by the following relationship:

$$v = \lambda \cdot f \tag{2.1}$$

where v is the velocity of sound waves in a specific tissue, λ is the wavelength and f is the frequency of the wave. The wavelength determines the minimum spatial resolution required in an image. Higher frequencies result in shorter wavelengths, which in turn provide better image quality. In ultrasonic devices, the speed of sound is typically set at 1540 m/s. This value is considered an average for soft tissues in the human body and is used as a standard reference.

Ultrasound waves interact with body tissues on the basis of their acoustic impedance (Z), which represents the resistance that a sound wave encounters as it travels through different types of tissue. The acoustic impedance is defined as the product between the density of the medium (ρ) and the velocity of the US (v):

$$Z = \rho \cdot v \tag{2.2}$$

The unit of measurement for acoustic impedance is the Rayl, defined as $\frac{\text{kg}}{\text{m}^2 \cdot \text{s}} \cdot 10^6$. Within the human body, the acoustic impedance of biological tissues typically ranges from approximately 1.38 Rayl, as observed in adipose tissue, to about 1.62 Rayl in kidney tissue. Significant deviations from this range are found in the lungs (0.26 Rayl), which exhibit very low impedance due to their air content, and in the bones (7.80 Rayl), which demonstrate considerably higher impedance as a result of their higher density.

When the wavefront meets an interface, that is a discontinuity in the value of acoustic impedance, in the passage from a medium having a value Z_1 to one with a value Z_2, part of the energy is transmitted and part reflected back or diffused due to irregularities of the encountered interface. The phenomenon can be described by the Snell's law (Figure 2.1):

$$\frac{\sin \theta_i}{\sin \theta_r} = \frac{v_1}{v_2} = \frac{\eta_2}{\eta_1} \tag{2.3}$$

where the subscript i refers to the incident wave and r to the reflected one, while v_1 and v_2 are the propagation speeds and η_1 η_2 are refractive indices in the two media. If the velocities in the two media are equal, there is no refraction and the US continues in a straight line without any deviation. If the acoustic impedance of the crossed medium is equal, there is no reflection, and all the energy of the incident wave is transmitted to the second medium. If the impedances are different, part of the US energy is reflected.

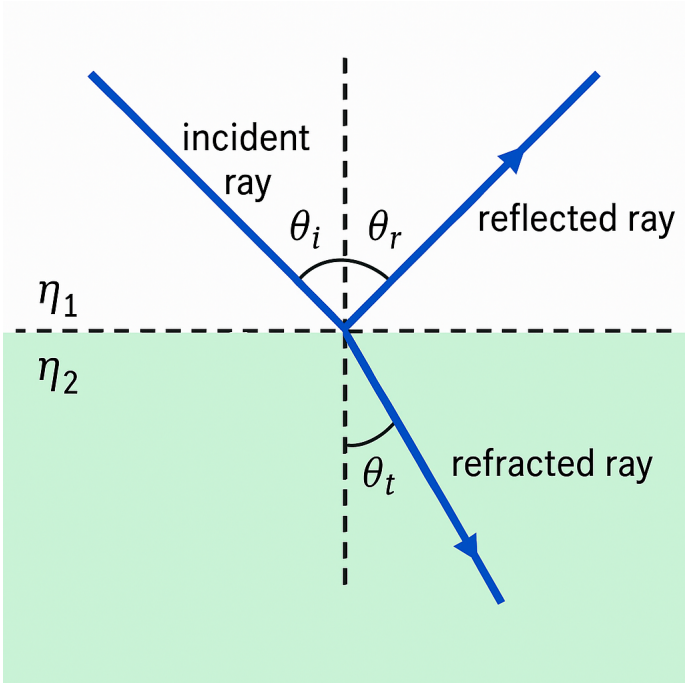


Figure 2.1: *Snell's law*

In order to understand how much light is reflected and how much is refracted, we need to focus on two components: the reflected (R) and the transmitted (T) coefficients, defined as follows:

$$R = \left(\frac{Z_1 - Z_2}{Z_1 + Z_2} \right)^2 \tag{2.4}$$

with $0 \leq R \leq 1$

$$T = 1 - R \tag{2.5}$$

If $Z_1 \gg Z_2$ or $Z_2 \gg Z_1$, as in the case of the muscle-bone interface, all the incident energy is reflected back. [20]

2.2 Ultrasound imaging

An ultrasound imaging system typically consists of a piezoelectric transducer, a signal amplification circuit and a display device. The transducer is responsible for generating US pulses and, immediately after transmission, switches to the receive mode to capture the echoes reflected by acoustic discontinuities encountered as the pulse propagates through the patient's body. These discontinuities occur at interfaces between tissues with different acoustic impedances, giving rise to reflections whose return times are directly related to their depth. The flowchart of a typical echo pulse system is shown in Figure 2.2, and illustrates the signal path from pulse generation to image display.

The time span between the pulse emission and the reception of the corresponding echo is used to compute the spatial location (depth) of the reflecting surface. However, since ultrasound signals are attenuated during their travels through biological tissues, echoes from deeper structures are inherently weaker than those that are reflected from more superficial interfaces. To attempt to compensate for this effect, the system employs time-dependent signal amplification as a function of the echo delay, which is the same as the depth of the target. This action is accomplished by the Time Gain Compensation (TGC) circuit that adds a time-increasing gain. This action discriminately amplifies the backscattered echoes based on their time-of-flight, which is related to the depth of the reflecting object. The TGC generally functions as a logarithmic amplifier where the factor of amplification is a function of the echo arrival time. As a result, the depth-dependent attenuation is mitigated, allowing the final image to primarily reflect variations in acoustic impedance, which are then encoded using grayscale or color to enhance visual interpretation.[20]

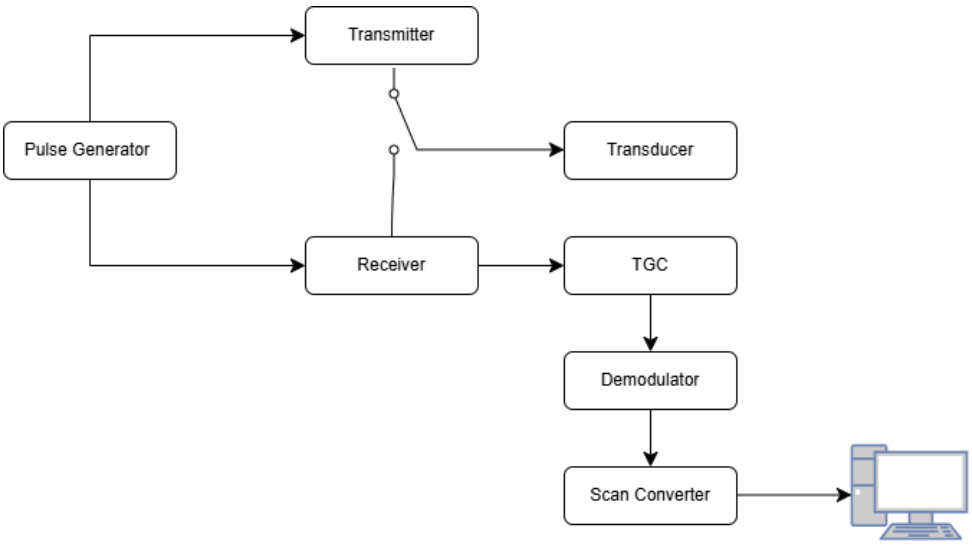


Figure 2.2: *Flowchart of an echo pulse system*

2.3 Ultrasound probes and display modes

In ultrasound imaging, the interface between the skin and the ultrasound probe is crucial to the effective transmission of sound waves. When tissues are examined using ultrasound, microbubbles can form between the skin and the probe, leading to a high reflection coefficient, severely reducing the quality of the image. The presence of air disrupts the smooth transmission of sound waves, causing them to scatter instead of penetrate the tissue effectively. To avoid this issue, a layer of gel is applied between the probe and the skin. This gel, typically made of a water-based substance, eliminates the air gap, ensuring that sound waves are efficiently transmitted into the body, allowing for accurate imaging and better visualization of internal structures.

Three main types of probes are commonly used (Figure 2.3):

- *Linear* are typically designed with higher frequencies and feature a rectangular footprint. These probes are ideal for imaging superficial structures because they provide high-resolution images
- *Curvilinear* are arranged along a convex face and generally operate at lower frequencies compared to linear probes. The lower frequency allows these probes to penetrate deeper into the body, making them suitable for imaging deeper structures.
- *Phased-array* use an electronically steered beam within a compact array to generate images. The beam is directed from a single point, which makes these probes excellent for imaging in confined spaces. [21]

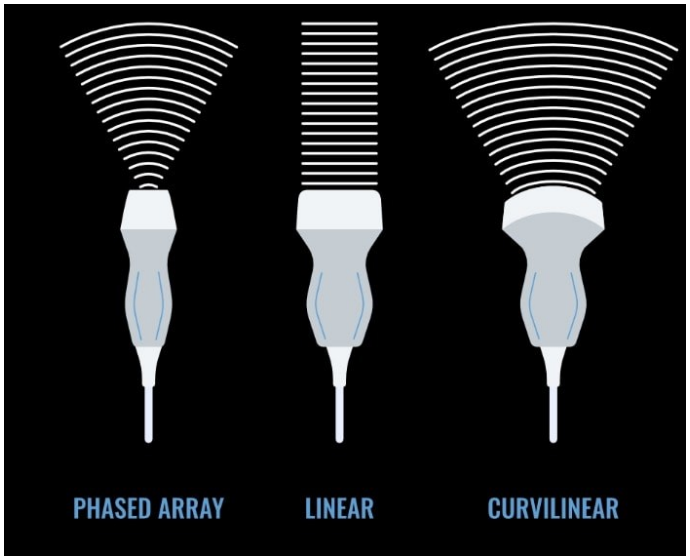


Figure 2.3: *Ultrasound probes type*

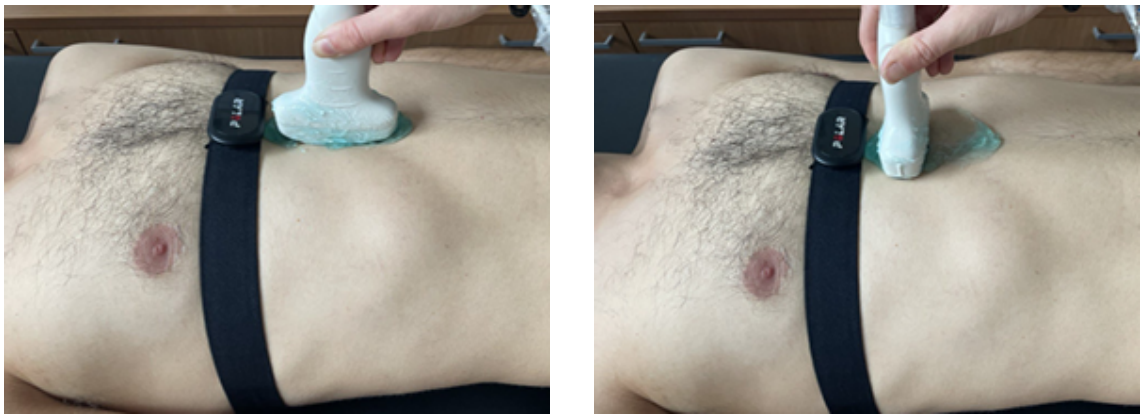
Four different modes of ultrasound are used in medical imaging:

- *A-Mode* (Amplitude) is the most basic form of ultrasound, where a single transducer scans through the body along a single line, and the resulting echoes are displayed on a screen based on their depth.
- *B-Mode* (Brightness) uses a linear array of transducers to scan a plane through the body simultaneously. A two-dimensional image is created on the screen from the data collected, guaranteeing a detailed view of internal structures. This mode is the most common used for general diagnostic purposes.
- *M-Mode* (Time Motion) captures a rapid succession of B-mode images that are displayed in sequence, enabling the observation and measurement of movement over time. This technique is particularly useful for assessing the motion of organ boundaries. [22]

2.4 IVC echography

The assessment of the IVC by ultrasound is typically performed with the patient lying in the supine position. The subcostal window is commonly used to obtain both transverse and longitudinal views of the IVC (Figure 2.4). For the transverse view, the ultrasound probe is placed just inferior to the xiphoid process, with the marker oriented towards the patient's right side. This orientation produces images where anatomical structures on the patient's right appear on the left side of the screen and vice versa. In this configuration, both the IVC and the Aorta (AO) can typically be visualized and differentiated.

To obtain a longitudinal view of the IVC, the probe is rotated 90 degrees from the transverse orientation, aligning the marker cranially.



a) *b)*

Figure 2.4: *a)Probe positioning for longitudinal view, b)Probe positioning for transversal view*

The probe is then shifted approximately 1–2 cm to the right of the midline while maintaining visualization of the RA to view the IVC in its long axis. It is important to sweep the probe carefully to the left of the midline in every patient in order to clearly identify the AO and avoid confusing it with the IVC. There are very few instances in which the IVC can be seen but not the AO, while if only the AO is visualized, the IVC may be totally collapsed.[18]

Chapter 3

Material and methods

3.1 Objectives and Expected Impact of the Study

The aim of this thesis is to estimate RAP using a non-invasive method. RAP is a key hemodynamic parameter that provides valuable information about a patient's cardiovascular function and volume status. In this work, the focus is placed on the IVC, whose diameter and respiratory variations have shown significant correlation with RAP. By combining specific echocardiographic features with others related to IVC morphology and dynamics, this study explores their predictive value in estimating RAP. The proposed method is designed to be simple, repeatable, and applicable in routine clinical practice, including in emergency and intensive care settings where rapid, non-invasive hemodynamic assessment is crucial. The use of VIPER software improves this approach by automating the extraction of key parameters such as RCI, CCI, CI and IVC diameter, making the process faster, more intuitive, and easily accessible to clinicians.

3.2 VIPER

VIPER is a semi-automated tracking algorithm for the IVC, specifically developed to detect and track its boundaries in both longitudinal and transverse ultrasound views, and it can operate in both real-time and offline modes. Its primary objective is to provide more accurate and reliable estimations of CI and IVC diameters, which are crucial parameters for the clinical assessment of a patient's hemodynamic status. The software offers several practical benefits for healthcare professionals:

- It eliminates time-consuming manual measurements, providing accurate and easily interpretable data for assessing a patient's hydration status
- It delivers quantitative values that can be readily included in fluid therapy reports
- It serves as objective evidence to support clinical decision-making in fluid management [23]

Among the most widely adopted non-invasive techniques in clinical settings for assessing a patient's volume status is the ultrasound-based evaluation of the IVC and the subsequent estimation of the CI. The maximum and minimum IVC diameters ($min(D)$ and $max(D)$ respectively) are measured over a respiration cycle and used to compute the CI:

$$CI = \frac{max(D) - min(D)}{max(D)} \tag{3.1}$$

The CI provides information about the collapsibility of the investigated blood vessel, reflecting its mechanical properties, compliance of the vessel and of the surrounding tissues, transmural pressure. [6]

Despite its popularity, the ultrasound method has significant limitations. It is time-consuming, subject to coarse approximations, and highly operator-dependent.

Figure 3.1: *VIPER logo*

The result is that often the operator renounces to use this method by defining the patient "full" or "empty", obtaining from ultrasound a measure only qualitative. This leads to a substantial loss of clinically relevant information, potentially compromising the accuracy of fluid therapy. To address these limitations, VIPER was specifically designed to provide a more robust and automated alternative. By performing real-time semi-automatic segmentation of the IVC boundaries, the software enables continuous and accurate tracking of the vessel throughout ultrasound video sequences. This allows for the extraction of detailed time series, such as diameter for each frame, offering a dynamic and quantitative view of the IVC behavior. [23]

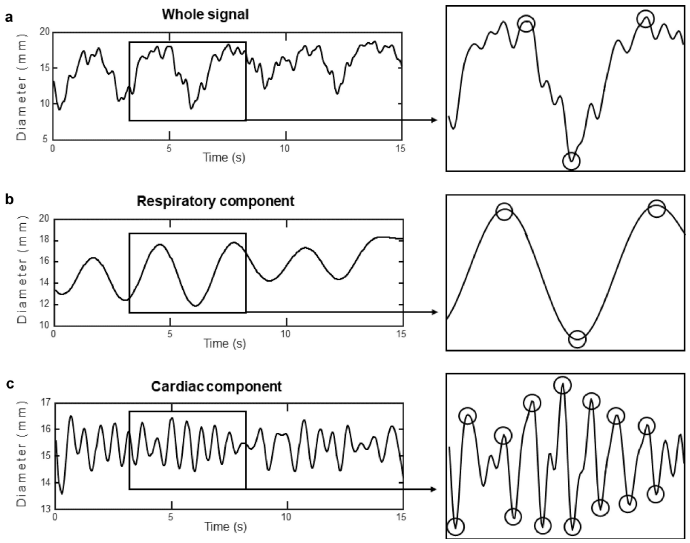


Figure 3.2: *a) Whole diameter signal and the extraction of CI b) Respiratory component and extraction of RCI c) Cardiac component and extraction of CCI.*

The dynamics of the IVC are influenced by two primary sources of physiological oscillation (Figure 3.2) : low-frequency fluctuations due to spontaneous respiratory activity and higher-frequency variations induced by cardiac pulsatility. By applying frequency-based filtering techniques to the IVC diameter signal extracted from ultrasound videos, VIPER enables the effective separation of the respiratory and cardiac components of IVC pulsatility. This, in turn, allows for the computation of two distinct indices:

- Respiratory Collapsibility Index (RCI) computed by low-pass filtering the whole diameter time series with a cutoff frequency of 0.4 Hz

- Cardiac Collapsibility Index (CCI) computed by high-pass filtering the whole diameter time series with a cutoff frequency of 0.8 Hz [6]

In this thesis, two different algorithms of VIPER software, one developed in Matlab and one in Python, were employed to analyze and segment the IVC from ultrasound video recordings provided by expert clinicians.

3.2.1 Python

In the main interface of the VIPER software developed in Python, the user can select the desired acquisition mode: either offline or live, as well as the ultrasound view, longitudinal or transverse (Figure 3.3). For the purpose of this thesis, only the longitudinal view in offline mode was utilized.

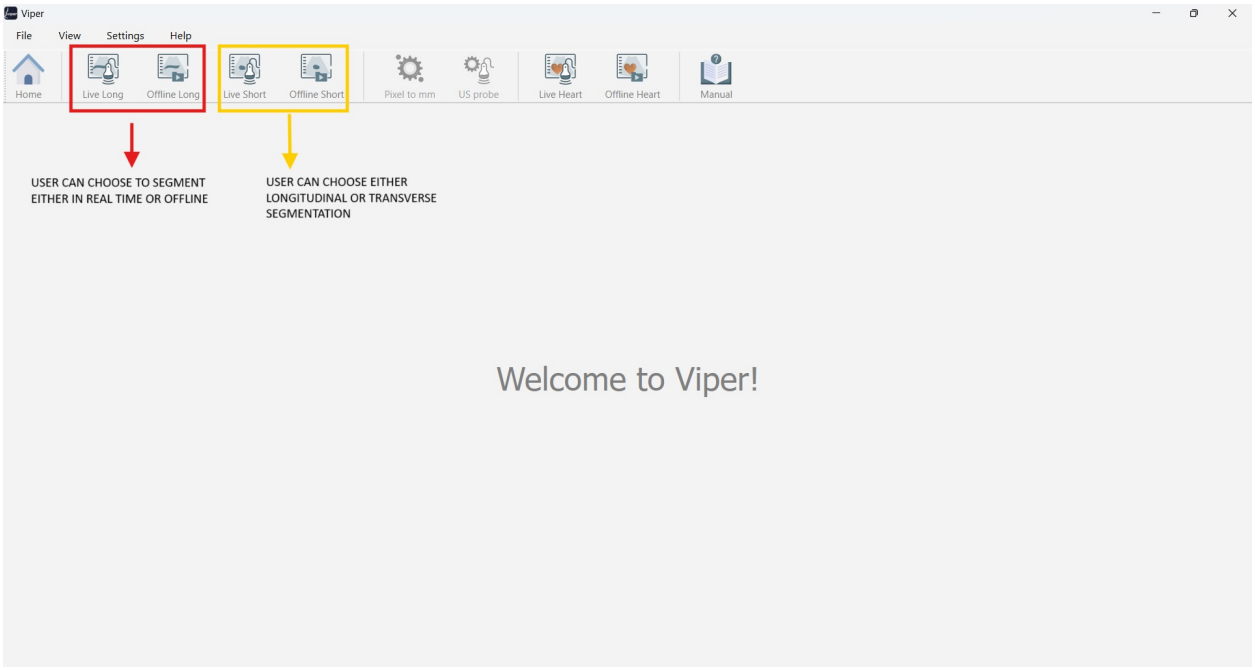


Figure 3.3: *Initial Screen of VIPER Software*

After selecting the preferred configuration, the software requests the operator to define a reference scale by selecting two points on the graduated scale adjacent to the ultrasound image, corresponding to a 5 cm distance (Figure 3.4). This step is essential for converting pixel-based measurements to real-world units (centimeters). The operator is then asked to manually select the center of the vessel to be analyzed. Immediately after this input, VIPER performs automatic segmentation of the selected vessel and displays the result on the screen. On the left side of the interface, the software presents key parameters extracted from the ultrasound data: mean diameter of the IVC, CI, RCI and CCI, along with a plot showing the temporal variation of the vessel diameter, expressed in seconds (Figure 3.5).

The software also includes several auxiliary features to enhance usability and flexibility during analysis. For instance, if the initial segmentation is unsatisfactory, the user can reinitiate the process by clicking the "Restart" button.

Additionally, VIPER allows the placement of custom markers along the ultrasound timeline to annotate specific events or features of interest. Once the desired

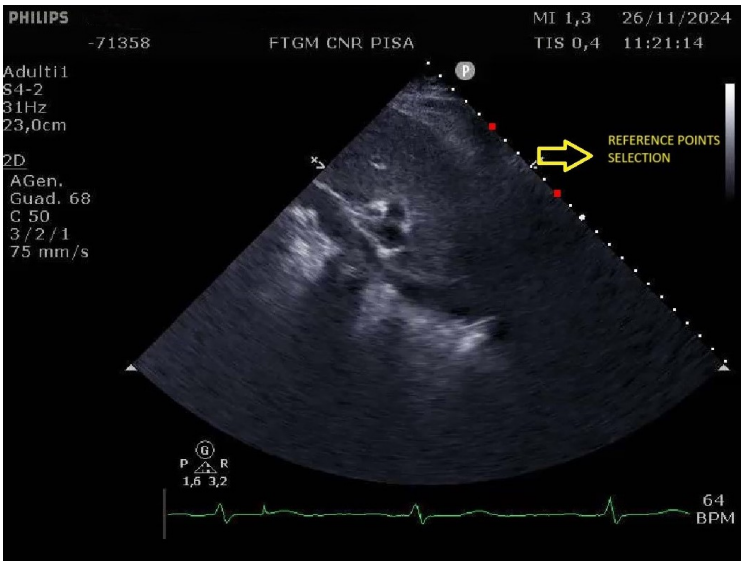


Figure 3.4: *Selection of Reference Points on Python algortithm*

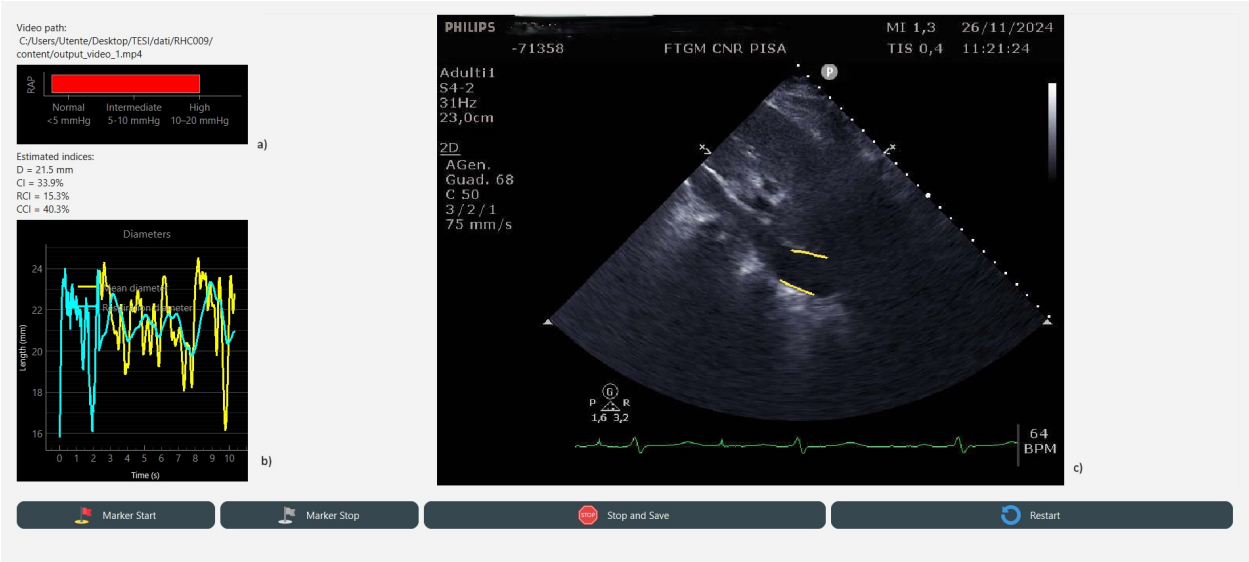


Figure 3.5: *a)RAP level of the patient based on IVC diameter b)Temporal variations of diameters c) Segmentation of IVC vessel on Echo Images*

segmentation is achieved, VIPER automatically generates and saves an Excel file containing the extracted data. This file includes 11 columns and a number of rows corresponding to the video duration, sampled according to the acquisition frequency. For each sampled time point, the Excel sheet records the following parameters: time, marker, diameter, RCI and CCI. It also reports several statistical summaries computed over the entire video, such as mean, maximum and minimum diameter, along with the overall CI, RCI and CCI values.

3.2.2 Matlab

Unlike the Python implementation, the Matlab version of the software does not support real-time segmentation of the IVC; instead, it requires the ultrasound video to be uploaded and processed offline. The segmentation process begins with the man-

ual selection of two high-contrast reference points in the first frame of the video. These points serve as reference markers for the subsequent tracking algorithm, which is essential to compensate for positional changes of the vein that may occur during recording. This step enhances the accuracy of diameter estimation by allowing the algorithm to follow the vein's movement throughout the video. Furthermore, the user defines the upper and lower boundaries of the vein to ensure accurate identification of its anatomical structure. A final step involves drawing a line perpendicular to the vessel, which sets the distal limit of the region of interest and specifies the IVC segment to be tracked and analyzed in subsequent processing. The mean diameter was then computed by averaging across different diameters and frames, in order to obtain a robust estimation of the IVC size.[17] As a result, a video is generated showing the detailed tracking of the selected segment of the IVC throughout the analysis (Figure 3.6).

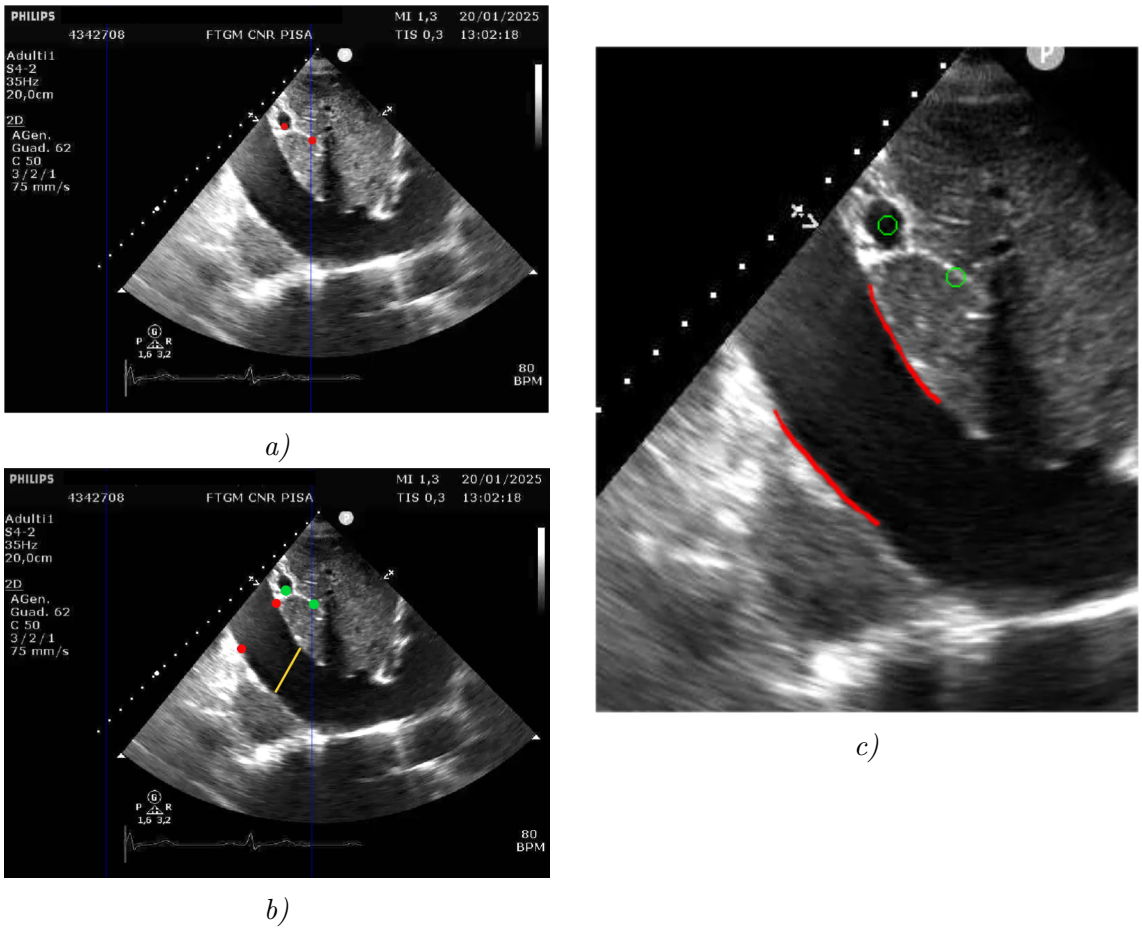


Figure 3.6: *a)Selection of reference points on Matlab algortithm, b)Selection of the portion of the vessel to be tracked, c)Tracking of the selected segment*

3.3 Echography

The ultrasound videos were acquired using the Philips iE33 xMATRIX echocardiography system (Philips Healthcare, Andover, MA, USA) in longitudinal view of the IVC (Figure 3.7). All acquisitions were performed by using a curvilinear probe in B-mode (brightness mode), which is the standard imaging modality used in clinical

ultrasound for anatomical visualization. In addition to the IVC ultrasound acquisition, a complete echocardiographic examination was performed for each subject. From these comprehensive cardiac assessments, additional clinical and functional features were extracted, which were subsequently included in the dataset. These features provide complementary information on cardiovascular function and enable a more integrated analysis in conjunction with the IVC measurements.

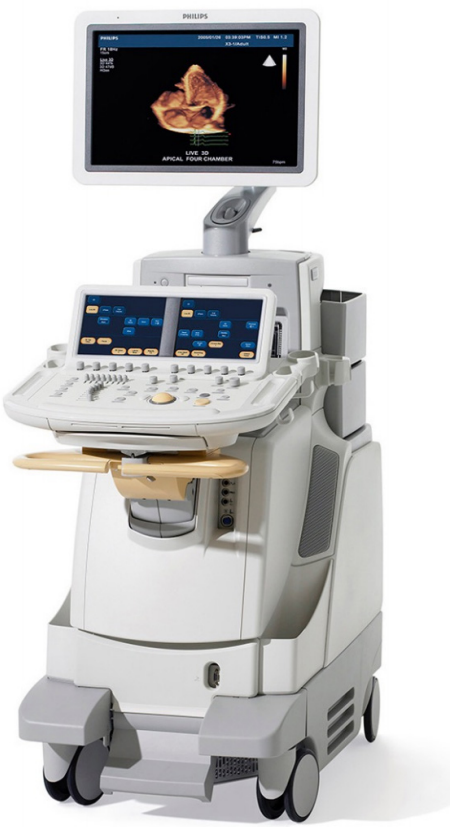


Figure 3.7: *Representation of Philips iE33 xMATRIX echocardiography system*

3.4 Dataset

The final dataset consists of 37 subjects (16 males and 21 females; mean age 67.49 ± 16.81 years), each associated with an ultrasound video of the IVC at Fondazione Toscana G. Monasterio (FTGM, Pisa, Italy). The analyzed ultrasound videos have durations ranging from 0.900 to 19.433 seconds, with an average duration of 3.859 ± 4.131 seconds (mean ± standard deviation) (Figure 3.8). However, due to the short duration of several recordings, particularly those under 3 seconds, VIPER was unable to compute the RCI, which requires a sufficient temporal window to capture multiple respiratory cycles. As a result, RCI could only be calculated for a limited number of videos and was therefore excluded from the analysis across the entire dataset to ensure consistency. The dataset is moderately balanced among the three RAP level classes, consisting of 15 subjects belonging to class 1, 13 to class 2 and 9 to class 3.

For each video, IVC segmentation was performed using the VIPER software, implemented both in Matlab and Python, in order to enable a comparative analysis

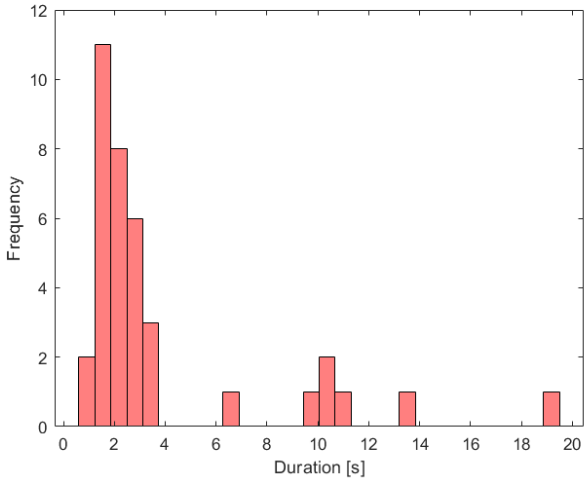


Figure 3.8: *Histogram of video duration*

between the two environments. In addition to the video data, each subject is characterized by a set of complementary features, organized into three main categories:

- clinical variables:

Acronym	**Name**	**Meaning**
SEX	Sex	male or female
AGE	Age	categorical variable
BSA	Body Surface Area	indicator of metabolic mass and it is calculated or measured considering weight and height
TR	Tricuspid Regurgitation	the tricuspid valve, between the right atrium and ventricle, doesn't close properly, causing blood to leak backward
AR	Aortic Regurgitation	the tricuspid valve, between the lower left heart chamber and the aorta, doesn't close properly, causing blood to leak backward
MR	Mitral Valve Regurgitation	the valve between the left heart chambers doesn't close fully
DD	Diastolic Dysfunction	the heart's ventricles (lower chambers) have trouble relaxing properly between beats, preventing them from filling with enough blood

Table 3.1: Clinical Variables explained

- invasive measurements obtained via catheterization:

Acronym	**Name**	**Meaning**
sPAP	Systolic Pulmonary Artery Pressure	measure of blood pressure in the pulmonary arteries during the heart's contraction
dPAP	Diastolic Pulmonary Arterial Pressure	pressure in the pulmonary arteries during the heart's relaxation phase (diastole)
mPAP	Mean Pulmonary Artery Pressure	measurement of the average pressure within the pulmonary arteries
PCWP	Pulmonary Capillary Wedge Pressure	pressure within the pulmonary capillaries, which is an indirect indicator of the left atrial pressure and left ventricular filling
PVR	Pulmonary Vascular Resistance	measure of the resistance that blood encounters as it passes through the pulmonary vasculature
RAP	Right Atrial Pressure	RAP indicates the amount of blood returning to the heart and the capability of the heart to pump the blood into the arterial system

Table 3.2: Catheterization Variables explained

These parameters were both extracted via catheterization and echo.

- echocardiographic parameters:

Acronym	Name	Meaning
HR	Heart Rate	heart pulsing during echocardiograms
LAVi	Left Atrial Volume index	measurement of the volume of the left atrium, one of the heart's chambers, adjusted for body surface area
LVEF	Left Ventricular Ejection Fraction	measurement of the percentage of blood ejected from the left ventricle with each heartbeat
LVMi	Left Ventricular Mass index	measurement that reflects the mass of the left ventricle of the heart, adjusted for the patient's body surface area
E/E'	E/E' ratio	ratio between early mitral inflow velocity (E) and early diastolic mitral annular velocity (E'), it is used to estimate left ventricular filling pressures
CI	Cardiac Input	volume of blood entering the heart during diastole, or the "filling" phase
TRV	Tricuspid Regurgitation Velocity	speed at which blood flows back into the right atrium from the right ventricle through the tricuspid valve
RV FAC	Right Ventricular Fractional Area Change	percentage change in the right ventricle's area during the heart cycle

Table 3.3: Echo Variables explained

These features were further enriched by additional measurements extracted directly by the VIPER software: IVC diameter, RCI, CCI and CI.

3.4.1 Preprocessing

In order to make the dataset correctly structured and valid before model development, a systematic data cleaning process was adopted. The important steps taken are presented below:

1. Handling Missing Values: in each of the variables, the presence of missing values (NaN) was tested. Any variable with more than 50% missing data with respect to the total number of subjects was excluded from subsequent analysis because their inclusion could compromise the robustness of the models.

2. Assessment of Distribution Symmetry:the skewness of each of the variables was calculated to evaluate the symmetry of its distribution. Skewness is a measure of the tendency of the distribution to be asymmetric relative to the mean. A near zero skewness shows that the distribution is symmetric; positive and negative values, however, represent right-skewed and left-skewed distributions, respectively.

$$\text{Skewness} = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s} \right)^3 \tag{3.2}$$

where:

- n is the total number of observations in the dataset
- x_i is the i-th data point
- $\bar{x}$ is the sample mean, defined as $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
- s is the sample standard deviation, given by $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

3. Imputation of Missing Values: based on the previously computed skewness, missing values were imputed using either the mean or the median. Variables with approximately symmetric distributions (i.e., skewness near zero) were imputed with the mean, while those exhibiting asymmetry were imputed using the median, to better preserve the distribution's characteristics.

4. Outlier Detection: outliers were identified using the Interquartile Range (IQR) method, which creates a cutoff based on the 25th (Q1) and 75th (Q3) percentiles (Figure 3.9). Data points falling outside the range

$$[Q_1 - 1.6 \cdot \text{IQR},\ Q_3 + 1.6 \cdot \text{IQR}] \tag{3.3}$$

were considered outliers.

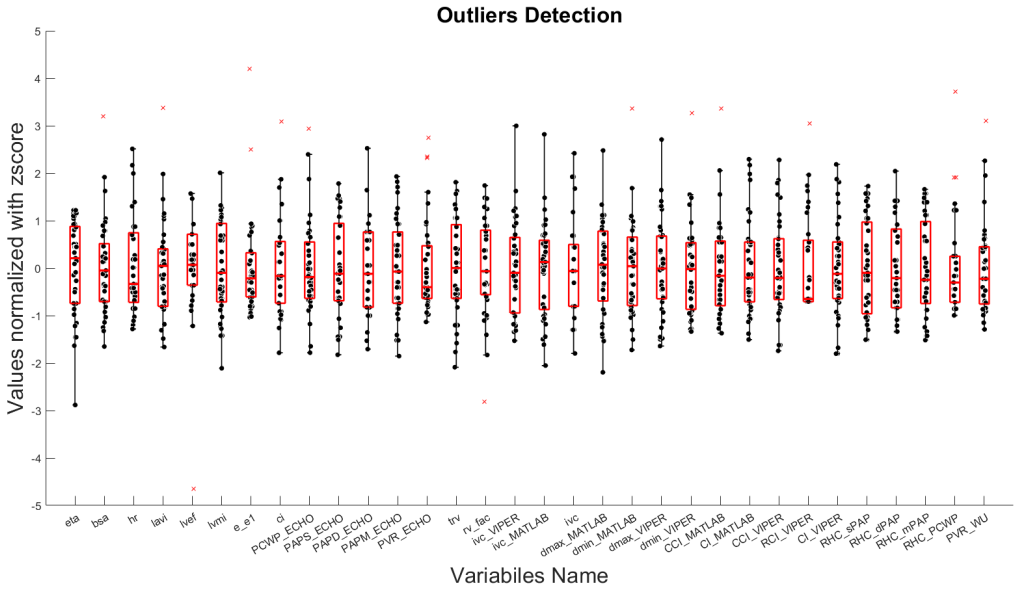


Figure 3.9: *Outlier Detection*

5. Outlier Treatment: outliers identified were not excluded but instead imputed using the median of the individual variable. This approach reduces the influence of extreme values while maintaining the overall structure of the dataset.

3.5 Statistical Analysis

Statistics is a discipline that combines elements of mathematics and logical reasoning to analyze and interpret data. While statistics is grounded in mathematical theory, many of its key concepts are conceptually simple and based on a few essential assumptions. Broadly, statistics can be divided into two main branches: descriptive statistics, which are used to give a brief idea and description of the main features of a dataset, and inferential statistics, which allow us to draw conclusions and make generalizations about a larger population based on data obtained from a representative sample. [24] In this thesis, descriptive statistics were employed as an initial step to explore and better understand the structure and characteristics of the dataset. Given the large number of variables, this approach was essential to gain a deeper understanding of the distribution, central tendency, and variability of each feature. In addition, descriptive analyses facilitated the identification of potential patterns and relationships within the data, including the presence of correlations between variables. These preliminary investigations not only guided subsequent data preprocessing steps but also provided a foundation for the development of more advanced models.

3.5.1 Shapiro-Wilk test

Shapiro-Wilk test is used to test whether a sample is from a normal distribution. The null hypothesis provides for normality of the sample: in case the resulting p-value is below the level of significance needed (usually $\alpha = 0.05$), then the null hypothesis is not true, i.e., the sample is not likely to follow a normal distribution. [25]

3.5.2 Ordinary Least Product Regression

Ordinary Least Product (OLP) regression, also known as standard major axis regression, is a statistical method used to assess systematic disagreement between two measurements when both are affected by random error. Unlike Ordinary Least Square (OLS) regression, in which one variable is taken to be free from error, OLP treats both variables equally and is therefore particularly well suited for method comparison studies where there is not a clear dependent variable. OLP regression minimizes the product of vertical and horizontal deviations of data points from the regression line, rather than OLS where only the vertical deviations are minimized (Figure 3.10).

The loss function is:

$$\varepsilon = \sum_{i=1}^{n} \delta x \cdot \delta y = \sum_{i=1}^{n} (x_i - \hat{x}_i)(y_i - \hat{y}_i) \tag{3.4}$$

where

- x_i and y_i are the observed values for i-th subject
- $\hat{x}_i$ and $\hat{y}_i$ are the estimated values obtained with OLP regression

The general formula for the OLP regression line is:

$$\hat{y} = a + bx \tag{3.5}$$

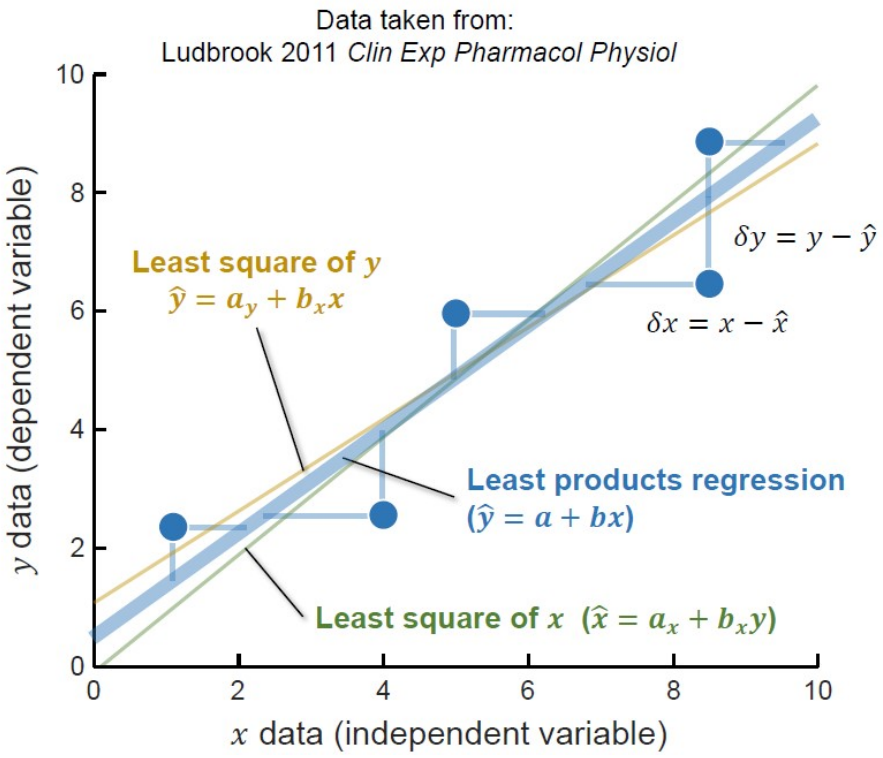


Figure 3.10: *OLP Regression, minimisation of the product of x and y residuals [26].*

where a represent the intercept and b represent the slope, but firstly it is necessary to compute these two parameters. The simplest way to compute the slope b is to use the ratio of the standard deviation of y and x

$$b = \frac{s_y}{s_x} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}} \tag{3.6}$$

where

- s_y and s_x represent the standard deviation of y and x
- $\bar{y}$ and $\bar{x}$ represent the mean values of y and x

Finally, since the OLP regression line passes through the point $\bar{y}$ and $\bar{x}$, the intercept a, can be obtained from the formula:

$$a = \bar{y} - b\bar{x} \tag{3.7}$$

OLP is especially useful when the goal is to detect fixed bias (a constant difference) and proportional bias (a difference that changes with magnitude) between two methods. By analyzing the confidence intervals of the intercept and slope, researchers can infer the presence of fixed or proportional bias: if the 95% Confidence Interval of the intercept does not include 0, this indicates a fixed bias. If the 95% Confidence Interval of the slope does not include 1, this indicates a proportional bias. In our study confidence intervals of a and b were calculated considering Fisher's distribution, the Pearson correlation coefficient r between the two methods and the number of subjects. [27]

OLP regression has been performed in order to confront the estimated diameter extracted with both Python and Matlab algorithm and the measured diameter of the operator. Even if the measured diameter of the operator represent the gold standard of our study, it may be affected by error so that's why OLP regression has been chosen.

Finally, Pearson correlation coefficient r was calculated for all models: it is defined as the product-moment correlation between two variables and quantifies the strength and direction of their linear relationship. It is calculated using the formula:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3.8}$$

3.5.3 Bland Altman Plot

The Bland-Altman plot is a well-established graphical method for evaluating the agreement between two quantitative measurement techniques. It was originally proposed to quantify the difference between paired observations and has since become one of the most commonly used tools in method comparison studies. In this plot, the differences between paired measurements are plotted against their average, providing a clear visual representation of systematic disagreement. [27]

Even in the presence of a fixed bias between the software-derived and operator-derived measurements, the Bland-Altman plot remains a valuable tool for evaluating measurement agreement. Its primary role lies in the visualization of the limits of agreement (LOA), defined as the interval where 95% of the differences between methods are expected to fall. By plotting the individual differences against the mean of the two methods, researchers can visually assess the consistency and spread of the measurement differences. In this study LOA are computed by first calculating the mean of the differences represented by the grey line in the middle:

$$\bar{d_{xy}} = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i) = \frac{1}{n}\sum_{i=1}^{n} d_{x_i y_i} \tag{3.9}$$

Superior and inferior LOA are then calculated:

$$\bar{d_{xy}} + 2s_{xy} = \bar{d_{xy}} + 2\sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_{x_i y_i} - \bar{d_{xy}})^2} \tag{3.10}$$

$$\bar{d_{xy}} - 2s_{xy} = \bar{d_{xy}} - 2\sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_{x_i y_i} - \bar{d_{xy}})^2} \tag{3.11}$$

Thus, despite a known fixed bias, the Bland-Altman plot contributes important insights into the reliability and clinical interpretability of the agreement between two measurement methods.

3.5.4 Wilcoxon Signed Rank Test

The Wilcoxon signed rank test, which is also known as the Wilcoxon signed rank sum test and the Wilcoxon matched pairs test, is a non-parametric statistical test used

to compare two dependent samples (in other words, two groups consisting of data points that are matched or paired). As with other non-parametric tests, this test assumes no specific distribution of the data being analyzed. Non-parametric tests, also known as distribution-free tests, make no assumptions about the shape of the distributions of your data. They are used to test hypotheses when the assumptions for the normality of the data are not met. [28]

Paired data arise when observations from one independent sample are uniquely matched or related to observations in another independent sample. In this study a confront has been performed between the parameters extracted with Matlab algorithm and the ones extracted with Phyton algorithm. The hypotheses for the Wilcoxon signed rank test for paired data are as follows:

- the null hypothesis (H0) is that the difference between the paired observations in the population is zero
- the alternative hypothesis (H1) is that the difference between the paired observations is not equal to zero

Based on the p-value obtained from the statistical test and using a significance level of 0.05, it was determined whether the observed differences are statistically significant. If the p-value was less than 0.05, this indicates that the null hypothesis can be rejected, suggesting that a statistically significant difference exists between the compared groups or conditions. Conversely, if the p-value was greater than or equal to 0.05, the result is not statistically significant, and there is insufficient evidence to reject the null hypothesis.

3.6 Models

Following the statistical analysis, Machine Learning (ML) models were developed with the objective of non-invasively predicting RAP. To achieve this, all invasive variables obtained through catheterization were excluded from the dataset (Table 3.2). Subsequently, two distinct datasets were created, differing in the set of features used: one dataset included variables extracted using Matlab-based algorithm, while the other included variables extracted using Python-based tools. Prior to model training, both datasets were standardized using Z-score normalization in order to ensure that each feature contributed equally to the learning process. This transformation centers the data around zero with unit variance, improving convergence and performance in many ML algorithms.

$$z_i = \frac{x_i - \mu}{\sigma} \tag{3.12}$$

where:

- x_i is the observed value for $i-$th subject,
- μ is the sample mean,
- σ is the sample standard deviation.

3.6.1 Analysis of correlation between features

An initial exploration of the dataset was carried out to reduce redundancy among the input features and improve model generalization. Pairwise correlations between variables were calculated using the Spearman correlation coefficient, rather than Pearson's, because not all features were normally distributed, violating one of the key assumptions of Pearson correlation. Spearman, being a non-parametric measure based on rank-order, is more appropriate for this context, as it does not require the data to follow a normal distribution and is more robust to outliers.

Following this, a high correlation threshold (approximately 0.8–0.9) was established to identify strongly correlated feature pairs. For each pair exceeding this threshold, one of the two variables was removed, based on their correlation with the target variable (RAP). Specifically, the variable with the weaker association with the target was discarded, while the more informative feature was retained. This method helped ensure that the final feature set preserved the most predictive information while reducing multicollinearity. This procedure was performed for almost all the models described in the following sections.

3.6.2 Linear Model

The initial modeling approach was based on the use of the $fitlm$ function, which fits a standard ordinary least squares (OLS) linear regression. This method estimates the relationship between a set of predictor variables and a response by minimizing the sum of squared residuals:

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} (y_i - \mathbf{x}_i \boldsymbol{\beta})^2 \tag{3.13}$$

where:

- y_i is the observed value of the target (dependent) variable for the i-th instance,
- x_i is the row vector of input features (independent variables) corresponding to the i-th observation,
- β is the column vector of model coefficients (also called weights or parameters) associated with the input features. These are learned during the training process and define the linear relationship between the predictors and the response variable,
- n is the total number of observations (samples).

While $fitlm$ provides a simple and interpretable baseline, it relies on strong assumptions(such as linearity, lack of regularization and independence of errors) and can perform poorly in the presence of multicollinearity or irrelevant features.Although a careful feature selection process was conducted, including pairwise correlation analysis and the removal of highly collinear variables based on the Variance Inflation Factor (VIF), the linear model implemented with Matlab's $fitlm$ function yielded limited predictive performance likely due to residual multicollinearity and the inability of linear regression to capture complex, non-linear relationships.

To address these limitations,regularized linear models were implemented: Ridge, Lasso, and Elastic Net regression.These techniques modify the loss function of standard linear regression by adding a penalty term that discourages overly complex models, promoting better generalization.

- **Ridge Regression (L2 Regularization)** Ridge regression adds a penalty equivalent to the square of the magnitude of beta coefficients:

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} (y_i - \mathbf{x}_i \boldsymbol{\beta})^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \tag{3.14}$$

  where:

  - λ is the regularization parameter, which controls the strength of the penalty term. A higher λ imposes more regularization,
  - β_j is the coefficient corresponding to the j-th feature,
  - p is the number of features (predictors) in the dataset.

  This technique does not set any coefficients exactly to zero, making it suitable for situations where all variables are believed to have some effect on the variable that wants to be predicted.

- **Lasso Regression (L1 Regularization)** Lasso regression employs L1 regularization, adding a penalty equal to the absolute value of the magnitude of coefficients:

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} (y_i - \mathbf{x}_i \boldsymbol{\beta})^2 + \lambda \sum_{j=1}^{p} |\beta_j| \tag{3.15}$$

  This approach can shrink some coefficients to exactly zero, effectively performing variable selection and yielding sparse models (it is used when only a few predictors actually influence the response).

- **Elastic Net Regression (L1 + L2 Regularization)** Elastic Net combines both L1 and L2 regularization penalties:

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} (y_i - \mathbf{x}_i \boldsymbol{\beta})^2 + \lambda \left[\alpha \sum_{j=1}^{p} |\beta_j| + (1 - \alpha) \sum_{j=1}^{p} \beta_j^2 \right] \tag{3.16}$$

  where α is a mixing parameter that balances L1 and L2 penalties:

  - α=0 equivalent to Ridge (pure L2)
  - α=1 equivalent to Lasso (pure L1)
  - $0 < \alpha < 1$ combination of both

  This hybrid approach retains the variable selection benefits of Lasso and the coefficient shrinkage properties of Ridge, making it particularly useful when dealing with datasets containing highly correlated predictors. [29]

3.6.3 Random Forest

A Random Forest is an ensemble ML algorithm that merges several decision trees. Each tree in the forest is trained on a random subset of the data (bootstrap sampling) and only a random subset of features is considered when making splits (feature randomization).

For classification tasks, the forest predicts by majority vote of trees, while for regression tasks, it averages the predictions (Figure 3.11).

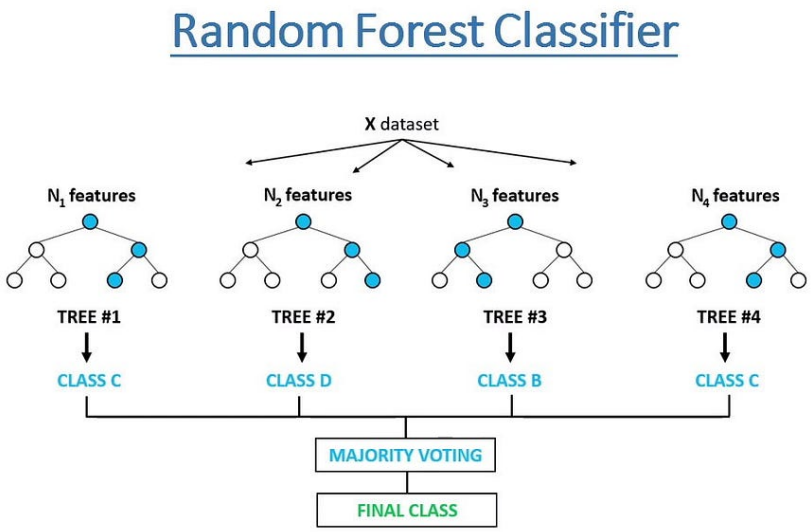


Figure 3.11: *Random Forest Classifier*

The model's strength comes from its "wisdom of crowds" approach: while individual trees might make errors, the collective decision-making process tends to average out these errors and arrive at more stable predictions. [30]

Feature selection was performed using Matlab's *predictorImportance* function, which is specifically designed for ensemble models such as Random Forests. This function quantifies feature importance by analyzing how much each feature improves the quality of the splits across all decision trees in the ensemble. The core idea is based on measuring the reduction in node impurity (commonly using the Gini index for classification or MSE for regression) whenever a feature is used to split a node in a decision tree. Each split that involves a particular feature leads to a certain gain in "purity," meaning that the resulting groups are more homogeneous with respect to the target classes. This gain is weighted by the number of observations affected by the split, reflecting its overall impact. By summing these weighted gains across all trees and all relevant splits, the algorithm assigns an importance score to each feature. Features that consistently contribute to purer and more informative splits are considered more important. [31]

3.6.4 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a powerful supervised learning algorithm used for both classification and regression tasks, with a primary focus on classification.

The core idea of SVM is to find the optimal hyperplane that best separates data points belonging to different classes. This hyperplane is chosen in such a way that it maximizes the margin between the nearest points of the two classes, which are called support vectors (Figure 3.12*c*).

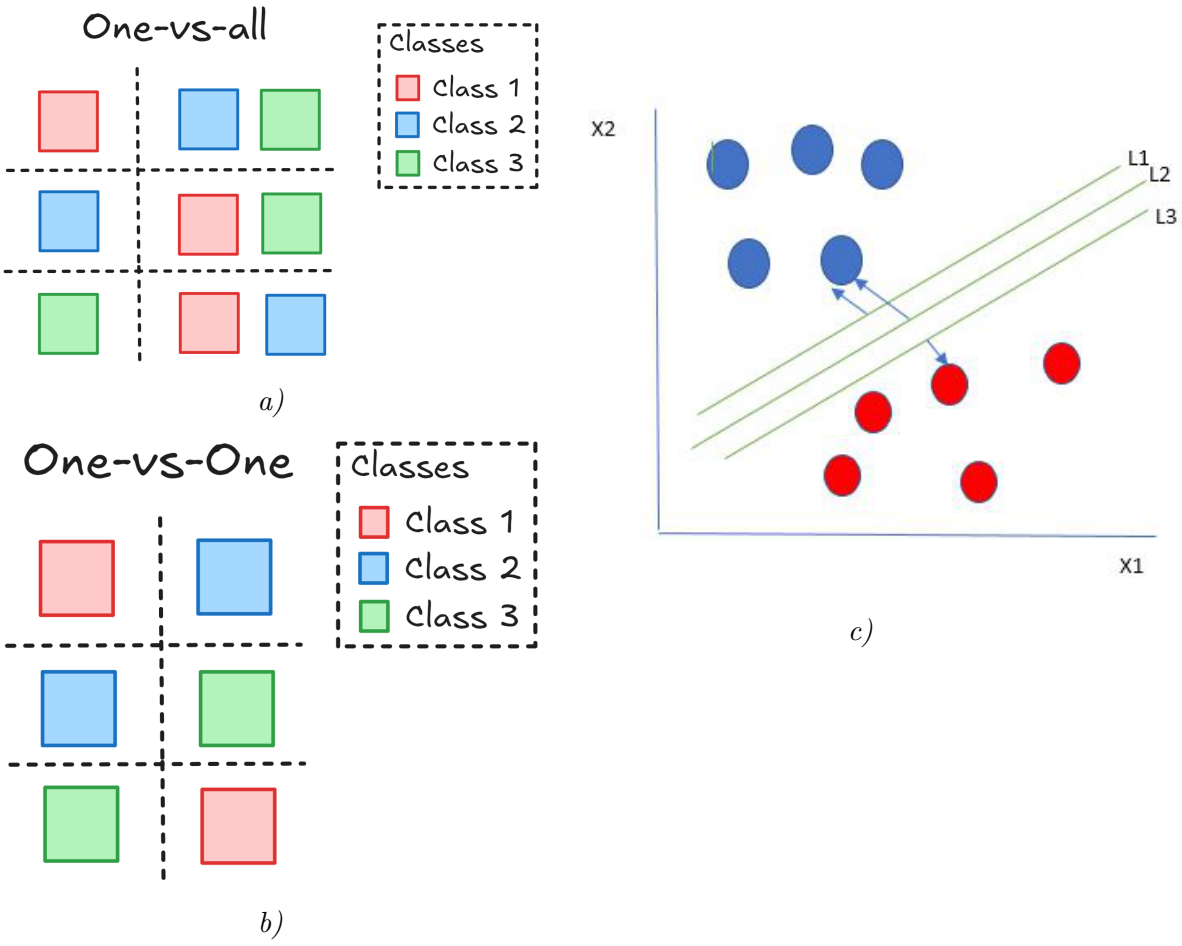


Figure 3.12: *a)One-vs-All: training of n classifiers, one for each class against all others, b) One-vs-One: training of* $\frac{n(n-1)}{2}$ *classifiers,one for each pair of classes, c)SVM for classification task*

A larger margin generally leads to better generalization performance for unseen data. SVM is useful when it comes to handle high-dimensional spaces and can be extended to many types of different data because the use of kernel functions allows to map non-linearly separable data into a higher dimension where one can linearly separate it. Linear, polynomial, and radial basis function (RBF) kernels are popular kernel functions. [32] Even-though Support Vector Machines are meant to perform binary classification, they can be modified to handle multi-class classification problems using two common strategies: One-vs-All (OvA) and One-vs-One (OvO). With the application of the One-vs-All method, for every class, an SVM model is trained with that class as the positive class and all the rest as negative (Figure 3.12*a*).During prediction, the maximum trust score determines the class label. Contrarily, the One-vs-One method will train a single SVM per possible pair of classes and thus result $\frac{n(n-1)}{2}$ classifiers for n classes.Each classifier votes for a single class out of its two classes, and the overall prediction is achieved by majority voting(Figure 3.12*b*). Al-

though One-vs-One can be computationally more intensive due to the number of classifiers required, it often provides higher accuracy by focusing on class-specific boundaries. [33]

3.6.5 MetaClassifier

Stacking is an ensemble learning technique that combines the outputs of multiple individual models to enhance predictive performance. Rather than relying on a single classifier, stacking uses the predictions from several base models as inputs for a final model that learns how to best combine them, known as the meta-learner.

In this work, a Support Vector Machine (SVM) and a Random Forest were trained using as input the predictions produced by two independently trained models (Ridge Model and SVM Model).The idea is that each base model captures different patterns or characteristics in the data, and the final classifier learns to interpret and integrate their outputs in a more informed way. This approach can help mitigate the limitations of each individual model and often leads to improved accuracy and generalization. [34]

3.6.6 1D CNN

Convolutional Neural Network (CNN) is one of the most popular and used Deep Learning (DL) networks. The advantage related to CNN is that it is capable to automatically detects the most relevant features without any human supervision, which makes it the most used. The structure of CNNs takes inspiration from neurons in human and animal brains. A commonly used type of CNN consists of numerous convolution layers preceding pooling layers, while the ending layers are Fully Connected (FC) layers. [35]

The general architecture of a CNN, as shown in Figure 3.13, can be divided into two main phases: feature extraction and classification. During the feature extraction phase, the network automatically learns the most relevant features from the input through a series of layers, including convolutional layers, ReLU activation layers and max pooling layers. These layers work together to identify spatial structures and recognize patterns within the data. The second phase, classification, uses the extracted features to assign the input to a specific category. This stage typically involves fully connected layers that interpret the high-level features and produce the final output.

Here's a breakdown of the main layers in CNN architecture:

- convolutional layer: this layer performs a convolution operation, which is to apply filters (kernel) to the input image to generate a feature map useful to detect basic patterns, like edges, corners, and textures.

- pooling layer: this layer decreases the spatial dimensions of the feature map while preserving the essential features. This reduces the risk of overfitting by summarizing the features extracted in the convolutional layer and also lowers computational costs.

- fully connected layer: this layer connects every neuron in one layer to every neuron in the next. The flattening process is used to convert all the feature maps into a one-dimensional vector.

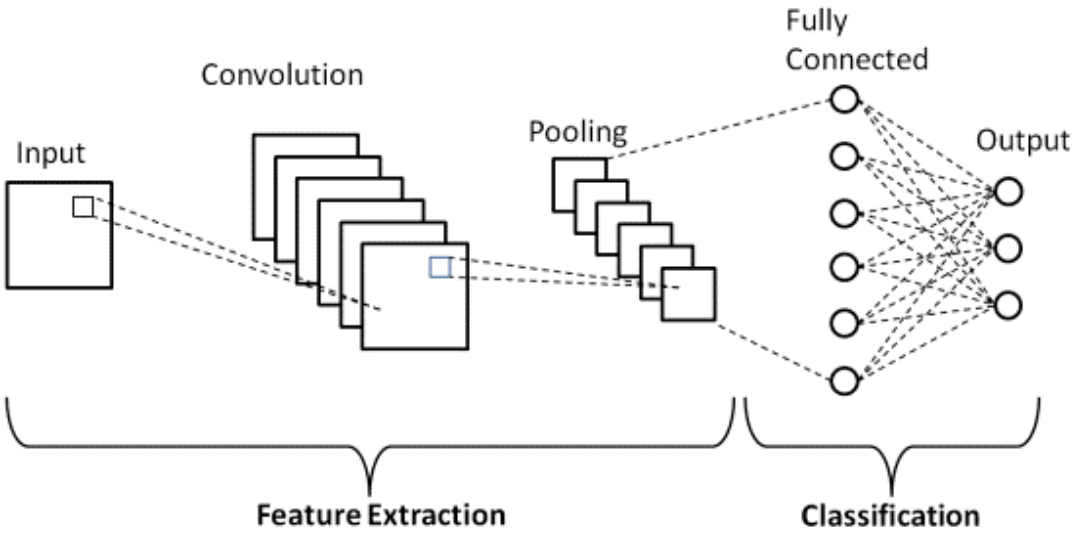


Figure 3.13: *General architecture of a CNN*

- dropout layer: this layer randomly deactivates a fraction of neurons during training to avoid overfitting. This prevents the model from relying too much on specific neurons, helping it generalize better on unseen data.

- activation function: this layer introduce non-linearity into the model,enabling it to capture complex relationships in the data.The most common activation function are ReLu,sigmoid,softmax. [36]

CNN can be distinguished based on the dimension of the convolutional layer. While 2D convolutional layers are widely used in image processing, 1D convolutional layers are network layer that performs convolution operations on one-dimensional data. 1D convolutions are particularly useful for extracting local patterns from sequential data, like peaks or valleys in time-series signals. [37]

In this thesis, a 1D CNN was employed to model the temporal dynamics of the IVC diameter. The input data to the network consisted of time-series sequences of IVC diameters automatically extracted using the VIPER software. This architecture was chosen due to its ability to efficiently capture local patterns and variations over time, making it particularly suitable for sequential medical data. To further improve performance, a hybrid neural network architecture was developed. This model combined two parallel input branches: one processing the temporal IVC diameter signals via the 1D CNN, and the other receiving numerical features extracted from the ultrasound data.

3.7 Metrics

To evaluate which model offers the most accurate and reliable prediction of RAP, a thorough performance assessment was conducted using a range of classification metrics. As a first step, confusion matrices were generated for each model in order to analyze the distribution of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). This allowed for a detailed understanding of the types of errors made by each model. Subsequently, a set of standard evaluation metrics was applied to quantify the predictive performance. These included:

- Accuracy

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.17}$$

This metric represents the proportion of correctly classified instances (both positive and negative) over the total number of cases. It provides a general measure of overall model performance but can be inaccurate in imbalanced datasets where one class predominates.

- Precision

$$PR = \frac{TP}{TP + FP} \tag{3.18}$$

Precision measures the proportion of true positive predictions among all instances classified as positive by the model.

- Recall (or sensitivity)

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.19}$$

Recall indicates the ability of a model to correctly identify all relevant positive cases. A high recall indicates that the model is able to detect most of the positive instances, minimizing missed cases.

- Specificity

$$\text{specificity} = \frac{TN}{TN + FP} \tag{3.20}$$

Specificity indicates the ability of the model to correctly identify negative instances, thus minimizing false positives which implies the reduction of false alarms.

- F1 score

$$\text{FS} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.21}$$

The F1 score represents the harmonic mean of precision and recall, offering a balanced metric that accounts for both false positives and false negatives. This metric is especially useful when working with datasets that are not perfectly balanced, since it provides a more balanced evaluation of model performance and it mitigates the informations given by precision or recall alone.

Since the classification problem faced in this study is a multi-classification problem (three classes in consideration), evaluation metrics such as precision, recall, specificity, and F1-score were computed with a One-vs-All strategy. Here, every class is individually considered to be the positive class while the other two are considered to be negative, thus allowing the computation of the respective metrics for every class individually. To provide an overall estimate of the performance of the model across the tree classes,the macro-average was employed, which computes the unweighted mean of the metrics across all classes. In this manner, each class contributes equally to the final assessment. Also, since the dataset employed in this analysis is relatively well balanced over the three classes, the application of macro-averaging is appropriate and does not introduce large bias.

Chapter 4

Results

This chapter reports the results of the present study, which investigates a non-invasive alternative for estimating RAP. The analysis focuses on identifying correlations between RAP and echocardiographic parameters, particularly the diameter of the IVC. The aim is to evaluate whether ultrasound-based measurements can reliably reflect central venous pressure, thereby offering a potential substitute for traditional invasive methods.

4.1 Statistical Analysis

4.1.1 Test

- Shapiro-Wilk test

  This test has been performed on our final dataset to verify the normal distribution of each variable (using a significance level of $\alpha = 0.05$) and showed that out of 32 variables, 22 of them have a normal distribution and the 10 left have a NON normal distribution (Figure 4.1 and Figure 4.2).

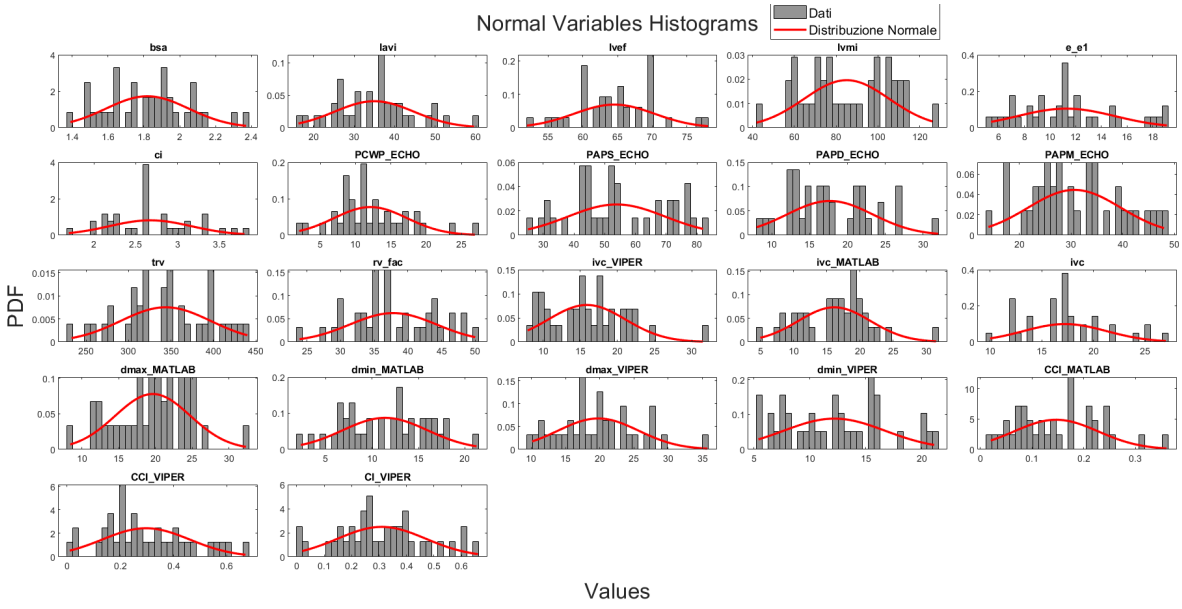


Figure 4.1: *Histogram of Normal Variables*

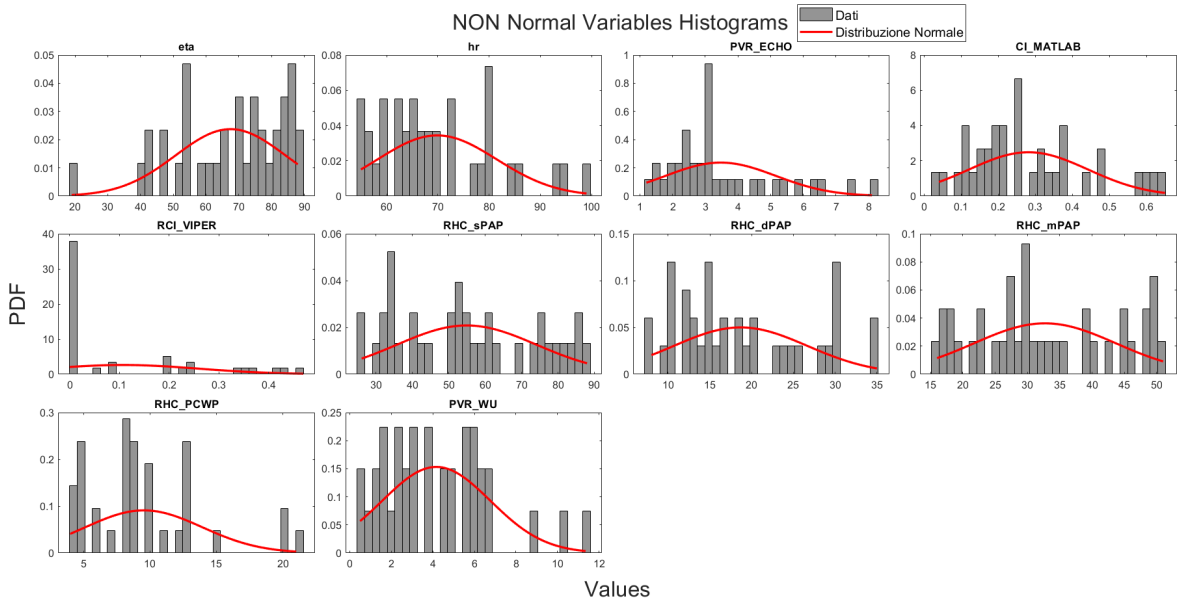


Figure 4.2: *Histogram of NON Normal Variables*

- Wilcoxon Signed Rank test

  To assess whether the measurements obtained from the Matlab and Python algorithm differed significantly, a Wilcoxon signed-rank test was conducted for each parameter. The resulting p-values are reported in Table 4.1.

Feature	**p-value**
IVC	0.1561
Max Diameter	0.8034
Min Diameter	0.0801
Caval Index	0.1184
Cardiac Caval Index	2.02×10^{-5}

Table 4.1: *p-values obtained from the Wilcoxon signed-rank test comparing the measurements estimated by the Matlab and Python algorithm for each parameter.*

4.1.2 Graphs

- OLP

  Figures 4.3*b* and 4.3*c* illustrate the comparison between automated and manual measurements of the IVC diameter. In both plots, each point represents a paired observation, with the x-axis indicating the automatic estimate and the y-axis showing the corresponding manual measurement performed by the operator. Figure 4.3*a* further illustrates a direct comparison between the two algorithm tools, with Matlab algorithm measurements plotted on the x-axis and Python algorithm measurements on the y-axis. The red line represents the best-fit OLP regression line that minimizes the loss function, while the dashed line corresponds to the identity line (y = x), which denotes perfect agreement between the two methods. The regression analysis yielded the following 95%

confidence intervals, which are summarized in Table 4.2. The corresponding Pearson correlation coefficients are reported in Table 4.3.

Algorithm	**Slope CI**	**Intercept CI**
Python	0.58 - 1.04	0.77 - 8.02
Matlab	0.55 - 1.00	1.17 - 8.35
Matlab vs Python	0.86 - 1.06	-1.25 - 2.03

Table 4.2: *95% confidence intervals*

Algorithm	**Pearson Correlation Coefficients** r
Python	0.5149
Matlab	0.4839
Matlab vs Python	0.9507

Table 4.3: *Pearson Correlation Coefficients r performed on each model*

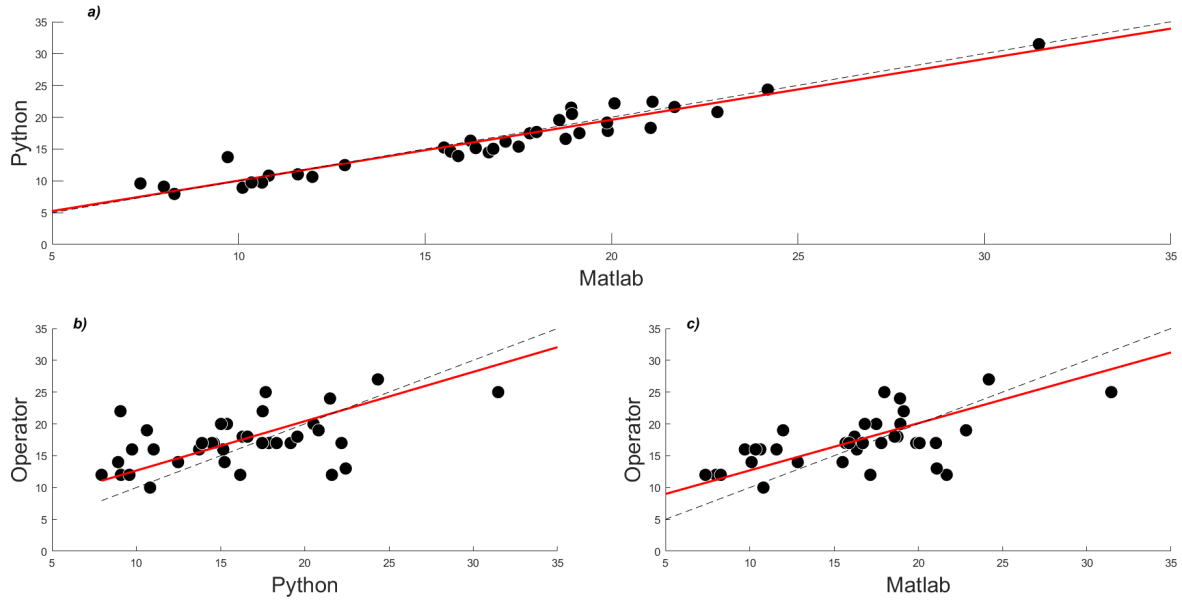


Figure 4.3: *a) OLP Regression of diameter estimated with Matlab vs diameter estimated with Python algorithm , b)OLP Regression of diameter measured from the operator vs diameter estimated with Python algorithm, c) OLP Regression of diameter measured from the operator vs diameter estimated with Matlab algorithm*

- Bland Altman plot

  Figures 4.4 display the Bland-Altman plots used to assess the agreement between the reference diameters and those estimated using the Python and Matlab algorithm, respectively. For each pair of measurements, the mean of the two values is reported on the x-axis, while the corresponding difference (reference – estimated) is reported on the y-axis. In both figures, the solid grey line in the center represents the bias, i.e., the mean of the differences between the reference and estimated values. This line indicates the presence of any systematic offset between the two measurement methods. The two dashed lines, located above and below the bias line, represent the limits of agreement.

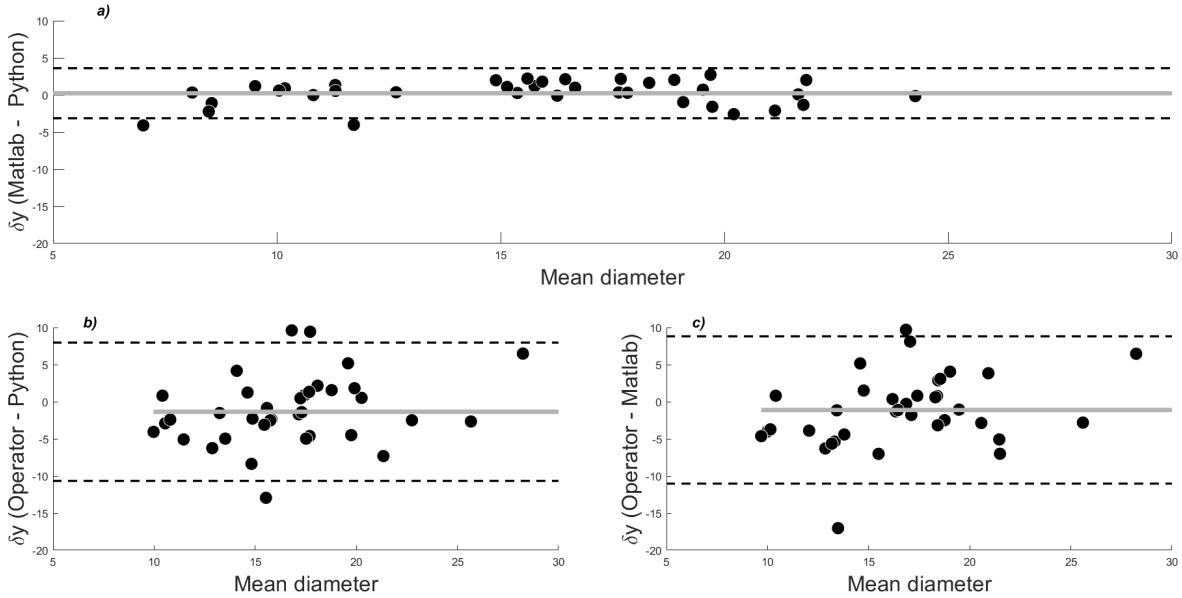


Figure 4.4: *a)Bland-Altman plot of Matlab and Python algorithm, b)Bland-Altman plot of Operator and Python algorithm, c) Bland-Altman plot of Operator and Matlab algorithm*

- Histograms and Boxplots

  The figures below present a comparative analysis of the variables obtained from the two algorithm tools, Matlab and Python algorithm. The left panel displays overlaid histograms of the measured variables. The red bars represent the values derived from Python, while the black bars correspond to those computed with Matlab. The right panel shows boxplots of the values normalized using z-score transformation. Each dot represents an individual measurement, and the gray connecting lines link the values from the same subject across the two methods. The red boxes illustrate the interquartile range, the median, and potential outliers, marked as isolated points (e.g., the red asterisk)

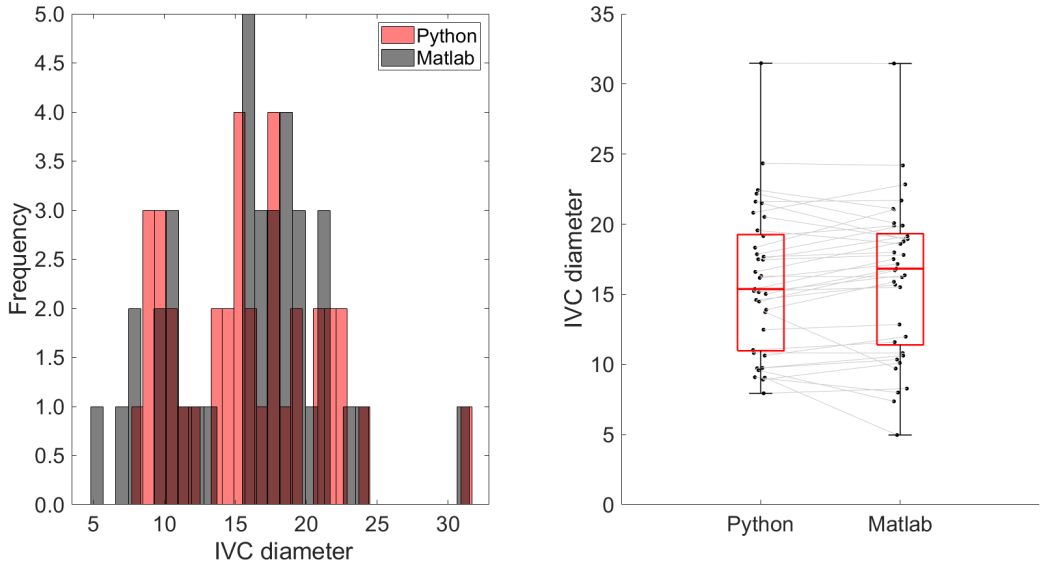


Figure 4.5: *Comparison of IVC diameter measurements using Python and Matlab. Left: overlaid histograms of diameter distributions. Right: boxplots with connections per subject*

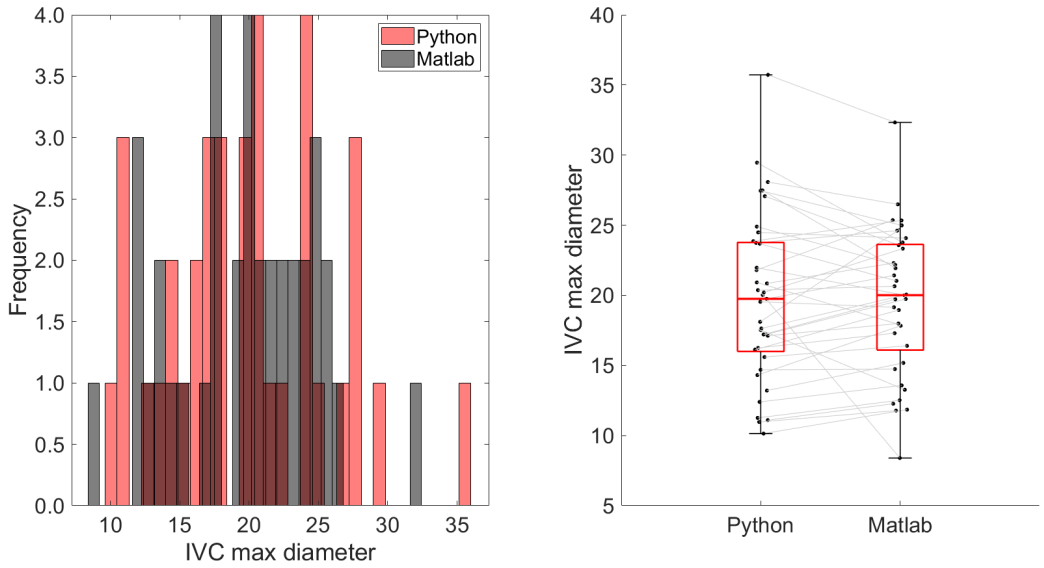


Figure 4.6: *Comparison of IVC max diameter measurements using Python and Matlab. Left: overlaid histograms of diameter distributions. Right: boxplots with connections per subject*

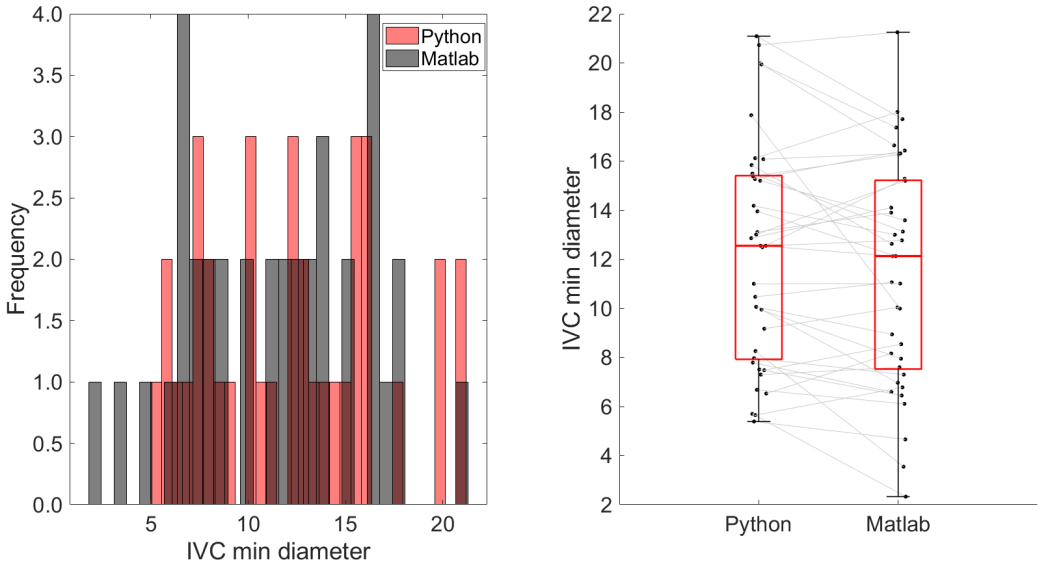


Figure 4.7: *Comparison of IVC min diameter measurements using Python and Matlab. Left: overlaid histograms of diameter distributions. Right: boxplots with connections per subject*

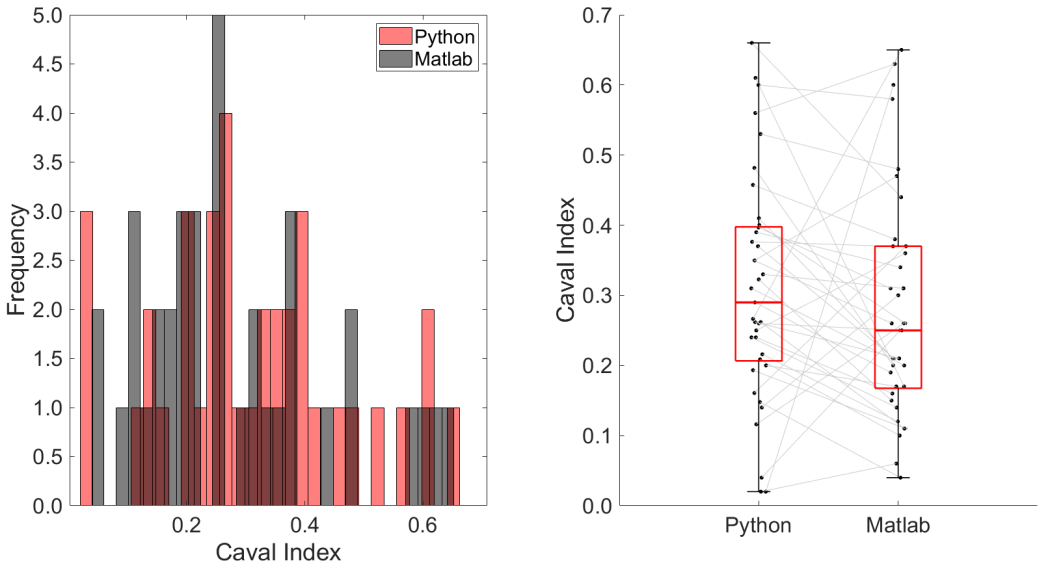


Figure 4.8: *Comparison of CI measurements using Python and Matlab. Left: overlaid histograms of diameter distributions. Right: boxplots with connections per subject*

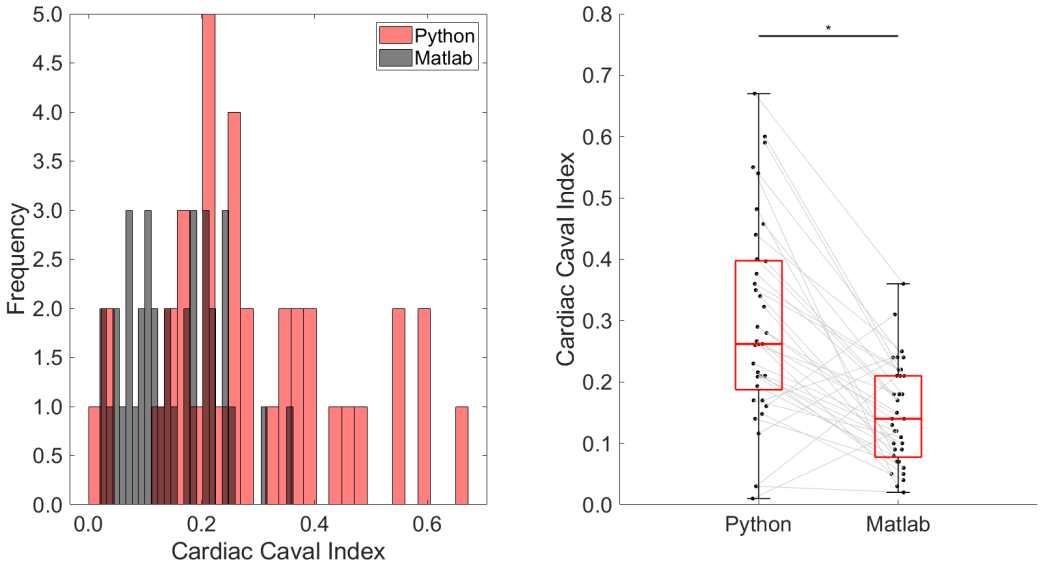


Figure 4.9: *Comparison of CCI measurements using Python and Matlab. Left: overlaid histograms of diameter distributions. Right: boxplots with connections per subject*

- Correlation table

  To assess the consistency between the two implementations, a Spearman correlation matrix was computed, as shown in Figure 4.10. This figure presents the correlation coefficients between corresponding variables obtained from Matlab and Python. High correlation values indicate a strong agreement between the two computational environments.

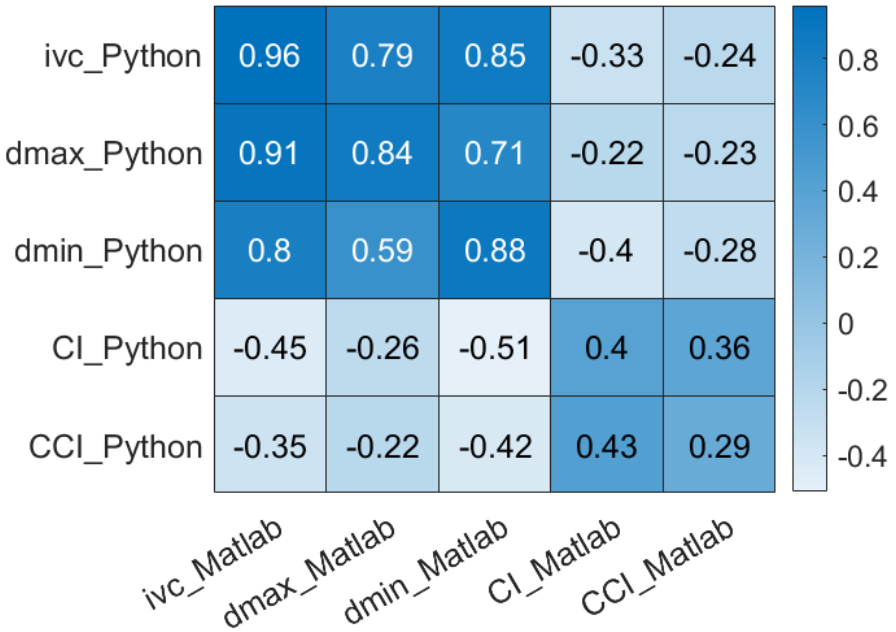


Figure 4.10: *Spearman correlation matrix between the variables of Matlab and Python*

4.2 Model Comparison

4.2.1 Guidelines

Before performing the comparison between the developed models, it is shown how the clinical guidelines traditionally used to estimate RAP based on IVC diameter and collapsibility index (Table 1.2) are employed to assess the model performance. Confusion Matrix (Figure 4.11) and metrics (Table 4.4) are shown below. This step allows to validate the relevance of the models obtained with Matlab and Python algorithm in the clinical context and to provide a baseline for comparing their predictive capabilities.

Metrics	**Value**
Accuracy	0.54
Precision	0.58
Recall	0.52
Specificity	0.75
F1-Score	0.52

Table 4.4: Metrics of RAP guidelines

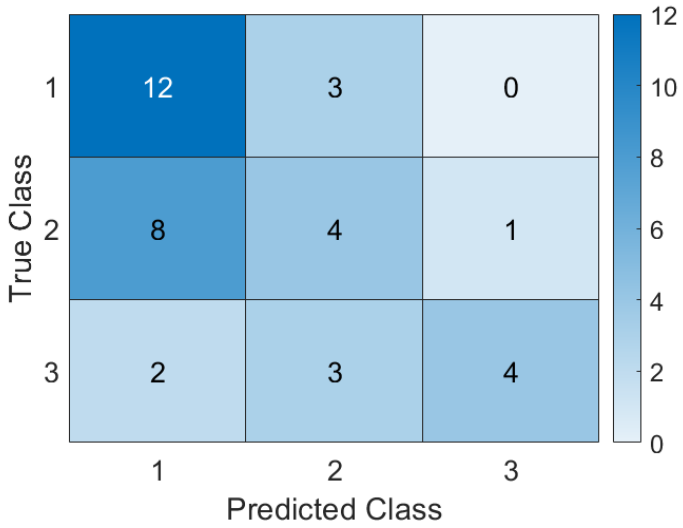


Figure 4.11: *Confusion Matrix of RAP guidelines*

4.2.2 Matlab algorithm

- **Linear Model**

  In this section are shown the results of each regularized linear model:

  – **Ridge Regression**: to determine the optimal value of the regularization factor λ , a range of candidate values is explored using a grid search approach on a logarithmic scale. For each candidate λ, a ridge regression model is trained using leave-one-out cross-validation (LOOCV), and the corresponding mean squared error (MSE) is computed.The value of λ that results in the lowest average MSE across all folds is then selected as the optimal regularization parameter (Figure 4.12).

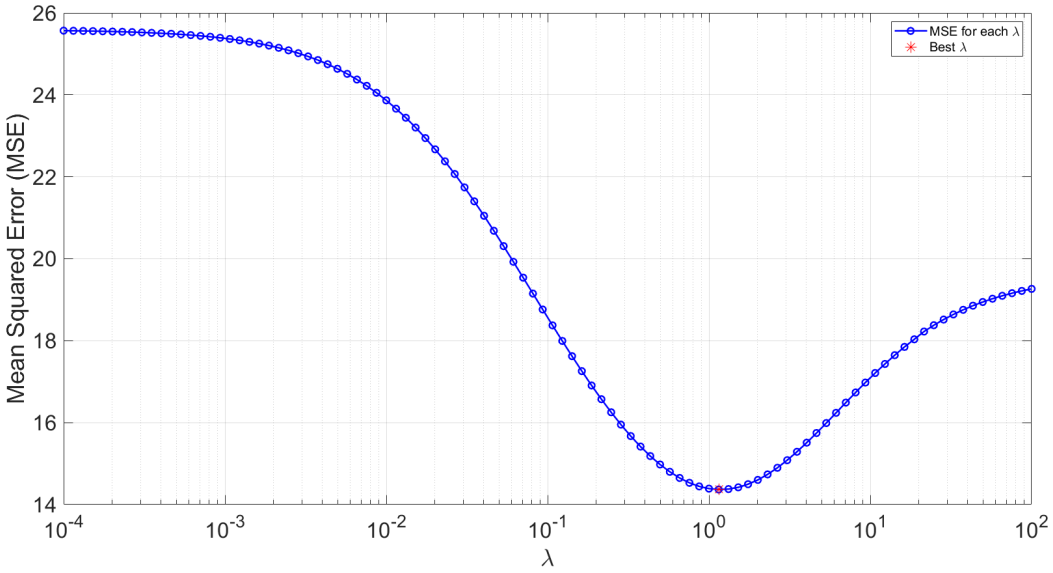


Figure 4.12: *Cross-Validated MSE of Ridge Regression*

The final model was then trained using this optimal λ and validated again using LOOCV. Finally, confusion matrix (Figure 4.13) and classification performance metrics such as accuracy,precision, recall, specificity, and F1 score were calculated to evaluate the model (Table 4.5). The features selected for the model and their weights are also shown below in Table 4.6.

Metrics	Value
Accuracy	0.65
Precision	0.72
Recall	0.67
Specificity	0.82
F1-Score	0.67

Table 4.5: Metrics of Ridge Model

Features	Weight
PCWP_echo	2.0693
CI_Matlab	1.3442
ci	1.2490
IVC_Matlab	1.2382
Age	0.7042
PAPM_echo	0.5399
Lavi	0.5110
rv_fac	0.4919
hr	0.3844
Dmin_Matlab	0.3445
BSA	0.2701
lvmi	0.2491
PVR_echo	0.1908
lvef	0.1845
e_e'	0.0463
CCI_Matlab	0.0108

Table 4.6: Feature ranking with the Ridge model

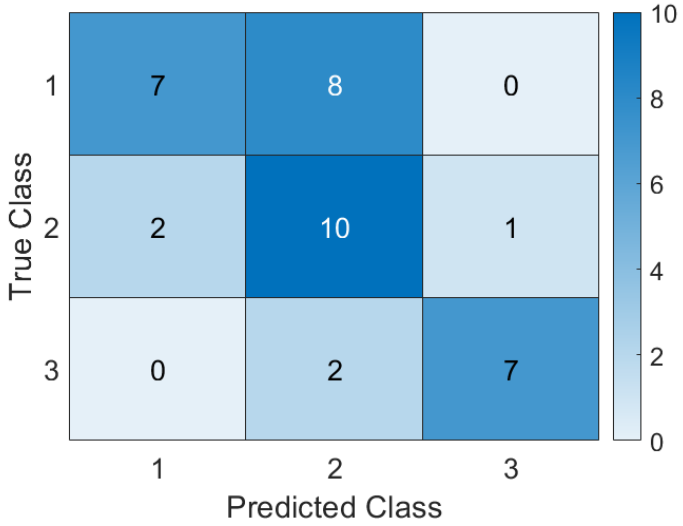


Figure 4.13: *Confusion Matrix for Ridge Model*

– **Lasso Regression**: as in RIDGE regression, the optimal λ value was the one that minimized MSE during cross-validation. The final model was then computed by using the optimal λ value and validated with a LOOCV. The feature selection results in the identification of only 6 features from the initial dataset (Table 4.8), in comparison to the 16 features retained by the Ridge regression model. Finally, Confusion Matrix (Figure 4.14) and metrics (Table 4.7) were computed.

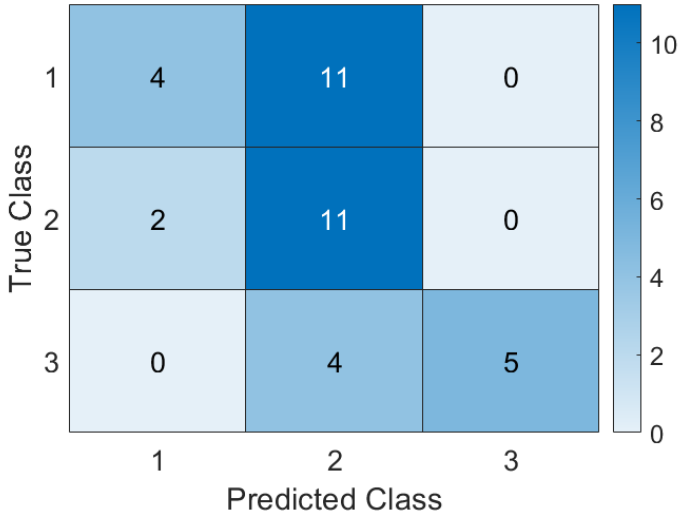


Figure 4.14: *Confusion Matrix for Lasso Model*

Metrics	**Value**
Accuracy	0.54
Precision	0.70
Recall	0.56
Specificity	0.76
F1-Score	0.55

Table 4.7: Metrics of Lasso Model

Features	**Weight**
PCWP_echo	1.8347
ci	0.7598
IVC_Matlab	0.3357
Age	0.2927
PAPM_echo	0.0682
e_e'	0.0424

Table 4.8: Feature ranking with the Lasso model

– **Elastic Net Regression**: a grid search was performed over a predefined range of alpha values to identify the optimal alpha that minimizes the mean squared error (MSE) in the Elastic Net regression. The final model was then built using this optimal alpha (α= 0.1) along with the λ value that achieved the minimum MSE for that alpha. The feature ranking obtained through this model constitutes a middle ground between the results of the two preceding models, with a total of 13 features being selected, as detailed in Table 4.10. Confusion Matrix (Figure 4.15) and metrics (Table 4.9) are shown in the images below.

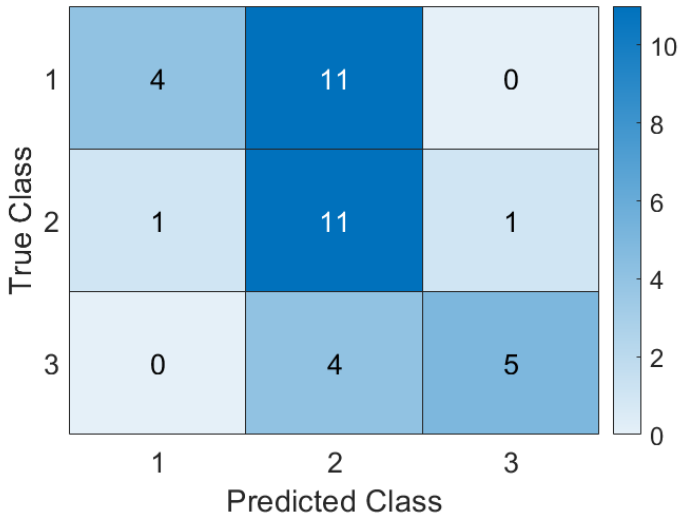


Figure 4.15: *Confusion Matrix for Elastic Net Model*

Metrics	Value
Accuracy	0.54
Precision	0.69
Recall	0.56
Specificity	0.76
F1-Score	0.54

Table 4.9: Metrics of Elastic Net Model

Features	Weight
PCWP_echo	0.8946
ci	0.6336
PAPM_echo	0.4870
e_e'	0.4475
Age	0.4420
IVC_Matlab	0.3904
CI_Matlab	0.3029
rv_fac	0.2455
lvmi	0.2318
Dmin_Matlab	0.1996
lvef	0.0463
PVR_echo	0.0445
lavi	0.0335

Table 4.10: Feature ranking with the Elastic Net model

- **Random Forest**: feature selection was evaluated by how much each feature improves the quality of the splits across all decision trees (Figure 4.16). Only the features with an importance greater than a predefined threshold (the mean of all importance values) were retained, as shown in Table 4.11.

  These selected features were then used to train the final classification model using a Random Forest with 1000 trees. The model's performance was evaluated using the Out-of-Bag (OOB) prediction method, which provides an unbiased estimate of the classification error by using the samples not included in the bootstrap training sets. In the figure below is shown how OOB Error changes by increasing the number of trees (Figure 4.17).

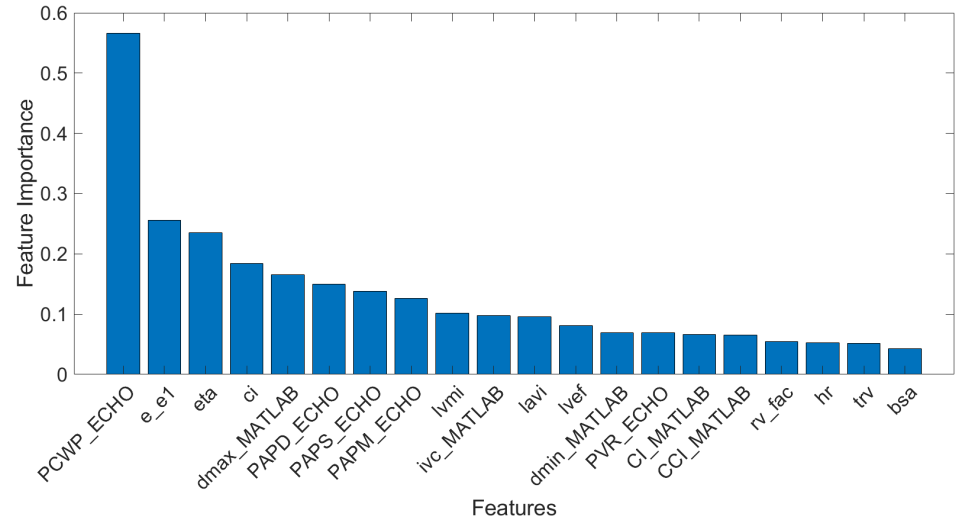


Figure 4.16: *Histogram of Feature Importance*

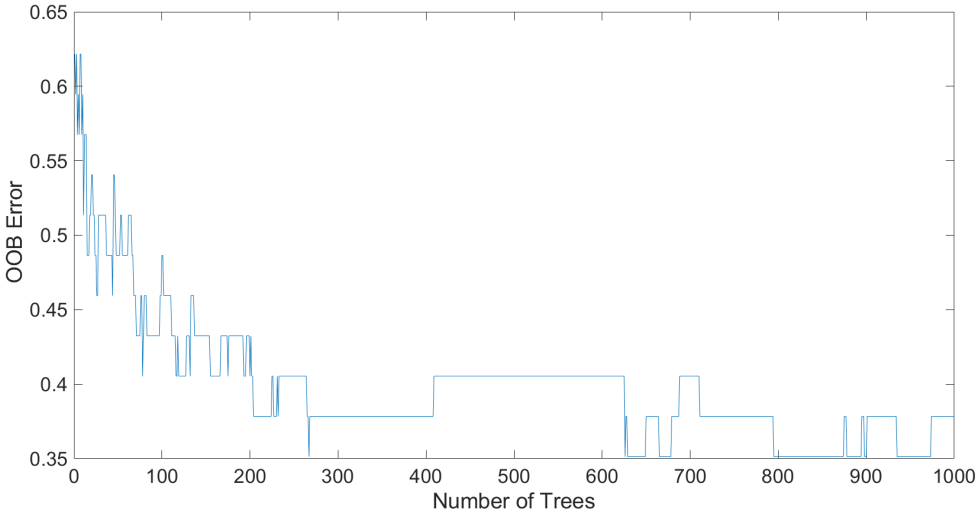


Figure 4.17: *OOB Error in function of tree number*

These unused observations, referred to as "out-of-bag" samples, can then be used to evaluate the tree's performance. The OOB predictions were used to calculate the confusion matrix (Figure 4.18) and the overall classification metrics (Table 4.12).

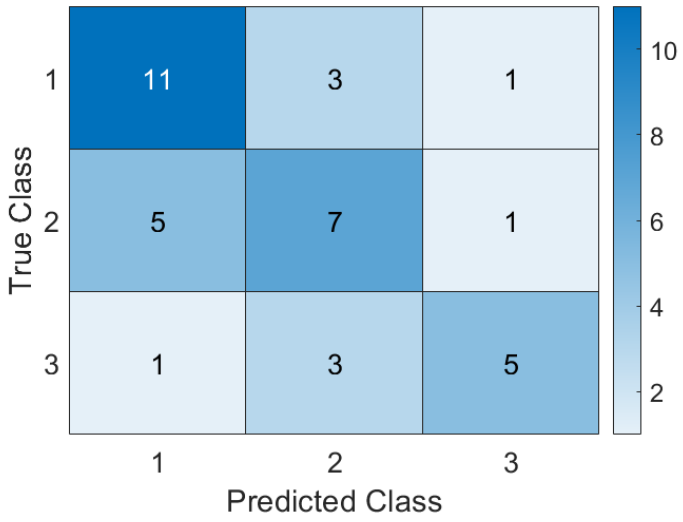


Figure 4.18: *Confusion Matrix of Random Forest Model*

Features	**Weight**
PCWP_echo	0.5657
e_e'	0.2555
Age	0.2349
ci	0.1840
Dmax_Matlab	0.1654
dPAP_echo	0.1493
sPAP_echo	0.1389

Table 4.11: Feature ranking with the Random Forest model

Metrics	**Value**
Accuracy	0.62
Precision	0.63
Recall	0.61
Specificity	0.80
F1-Score	0.62

Table 4.12: Metrics of Random Forest Model

- **Support Vector Machine**: feature selection has been performed with backward sequential feature selection strategy, using a non-linear Support Vector Machine (SVM) with a radial basis function (RBF) kernel as the evaluation model. In this approach, features were progressively removed one by one based on their impact on the classification error: at each step, the feature whose removal caused the smallest increase or the greatest reduction in classification error was discarded. After identifying the optimal subset of features (Table 4.14), the final classification model was trained using only the selected variables. Model performance was then evaluated using a Leave-One-Out cross-validation strategy. In the figure below are shown the Confusion Matrix (Figure 4.19) and metrics (Table 4.13).

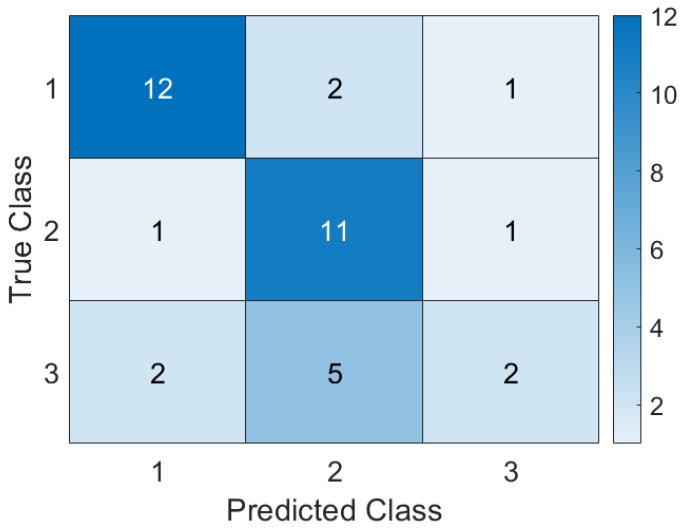


Figure 4.19: *Confusion Matrix of SVM model*

Metrics	Value
Accuracy	0.68
Precision	0.64
Recall	0.62
Specificity	0.83
F1-Score	0.61

Table 4.13: Metrics of SVM Model

Features
PCWP_echo
PVR_echo
CI_Matlab

Table 4.14: Feature selected with the SVM model

- **MetaClassifier** To further improve classification performance, a stacking strategy was implemented by combining the predictions of two base models: a Ridge regression model and a Support Vector Machine (SVM) classifier. The continuous predictions from the Ridge model were discretized into the same three classes used for the SVM. These outputs were then stacked to form a new feature set, which served as input for a meta-classifier (SVM model) . The meta-classifier was trained using a leave-one-out cross-validation (LOO-CV) approach. Confusion Matrix (Figure 4.20) and metrics are shown below (Table 4.15).

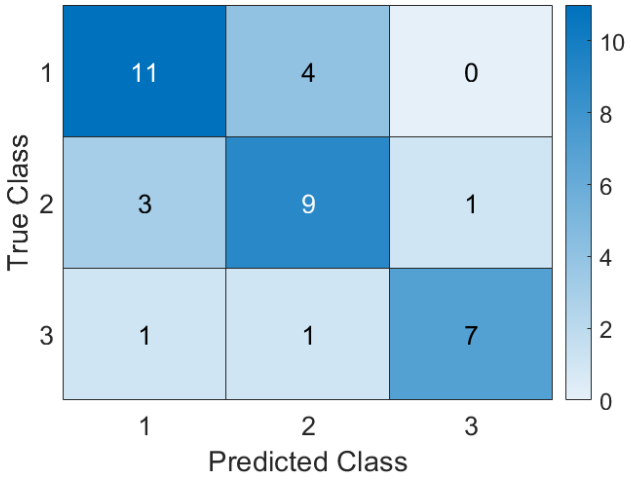


Figure 4.20: *Confusion Matrix of MetaClassifier*

Metrics	Value
Accuracy	0.73
Precision	0.75
Recall	0.73
Specificity	0.86
F1-Score	0.74

Table 4.15: Metrics of Meta-Classifier

- **1D CNN**

  To ensure data quality and consistency, ultrasound videos with a duration shorter than 1.5 seconds were excluded from the dataset, resulting in the removal of 7 samples. Model evaluation was conducted using a 5-fold cross-validation strategy, with stratification constraints applied to guarantee that each fold contained at least one subject from each of the three target classes. The architecture of the CNN adopted in this study is depicted in Figure 4.21. The left branch is designed to process the temporal sequences corresponding to the 30 subjects and consists of 11 layers. Notably, both max pooling and average pooling operations are implemented in parallel within this branch, with the objective of enriching the feature representation and maximizing the informational content derived from each subject. The output of this branch is subsequently concatenated with that of the right branch, which receives as input the numerical features extracted from the original dataset for each subject.

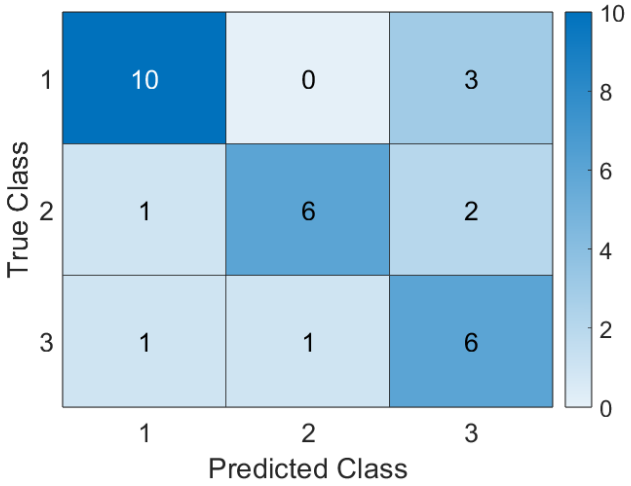


Figure 4.21: *Confusion Matrix of 1D CNN*

Metrics	Value
Accuracy	0.73
Precision	0.75
Recall	0.73
Specificity	0.87
F1-Score	0.73

Table 4.16: Metrics of 1D CNN

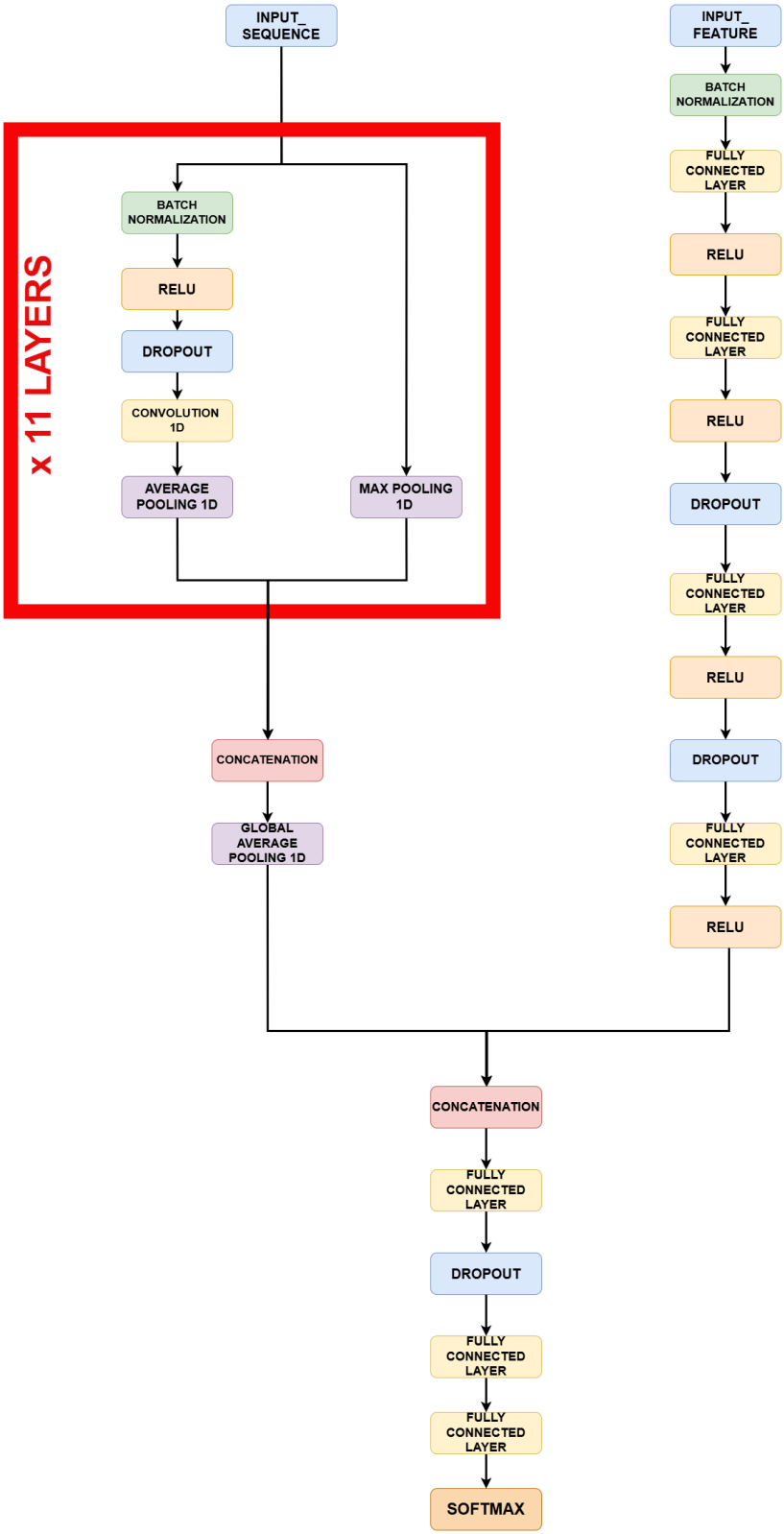


Figure 4.22: *Architecture of hybrid CNN*

4.2.3 Python algorithm

All considerations and procedures described for the Matlab dataset were also applied to the Python dataset. This includes data preprocessing, feature selection, and model training and evaluation strategies. The only difference between the two pipelines lies in the specific set of variables provided by each algorithm.

- **Linear Model**:in this section are shown the results of each regularized linear model:
  - **Ridge Regression**

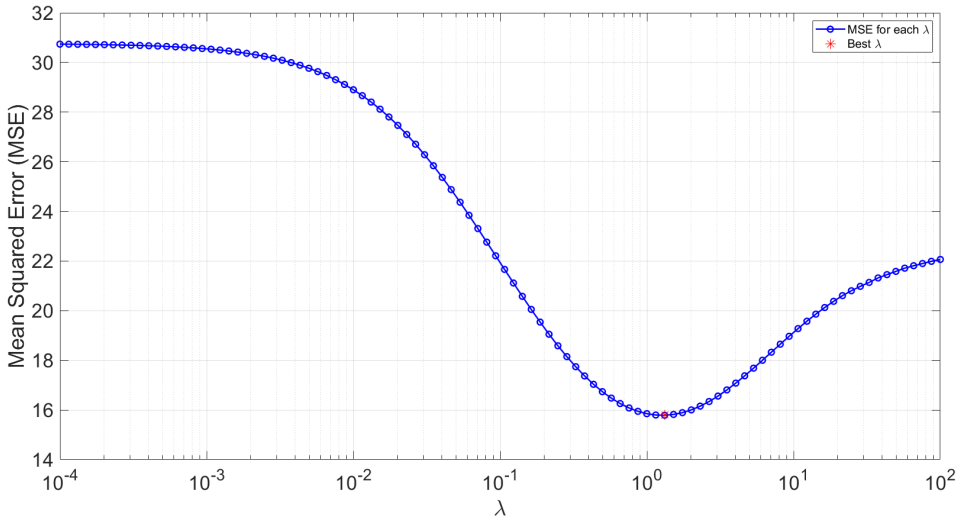


Figure 4.23: *Cross-Validated MSE of RIDGE Regression*

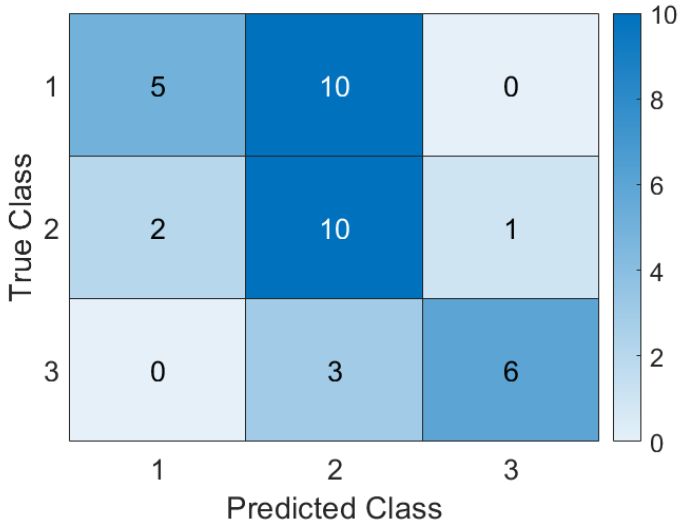


Figure 4.24: *Confusion Matrix of RIDGE Model*

Metrics	Value
Accuracy	0.57
Precision	0.67
Recall	0.59
Specificity	0.78
F1-Score	0.59

Table 4.17: Metrics of RIDGE Model

Features	Weight
PCWP_echo	1.8141
ci	1.4471
CI_Python	1.1305
mPAP_echo	0.7762
Lavi	0.7690
Age	0.7062
e_e'	0.6764
lvmi	0.6451
rv_fac	0.5628
IVC_Python	0.5080
Dmax_Python	0.3488
Dmin_Python	0.2387
PVR_echo	0.2280
hr	0.1719
BSA	0.1702
CCI_Python	0.0774
lvef	0.0297

Table 4.18: Feature ranking with the Ridge model

– **Lasso Regression**

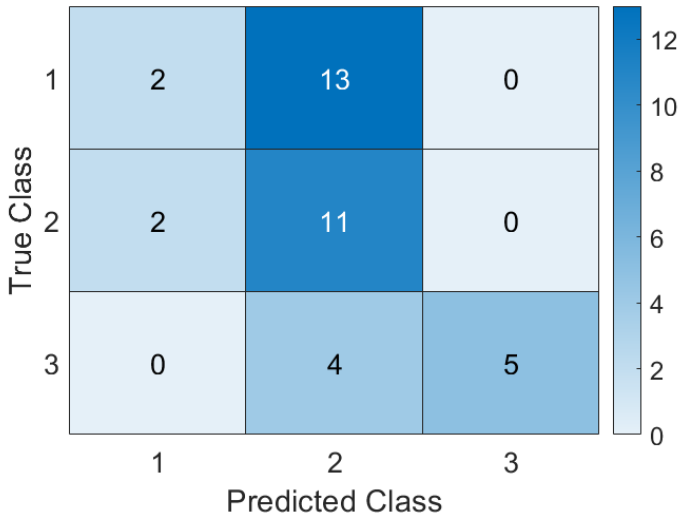


Figure 4.25: *Confusion Matrix of LASSO Model*

Metrics	Value
Accuracy	0.49
Precision	0.63
Recall	0.51
Specificity	0.73
F1-Score	0.49

Table 4.19: Metrics of LASSO Model

Features	Weight
PCWP_echo	1.8588
ci	0.6405
Age	0.2182
Dmin_Python	0.2099
PAPM_echo	0.1032
e_e'	0.0643

Table 4.20: Feature ranking with the Lasso model

– **Elastic Net Regression**: the optimal value of α that minimizes the MSE is 0.3 (α=0.3).

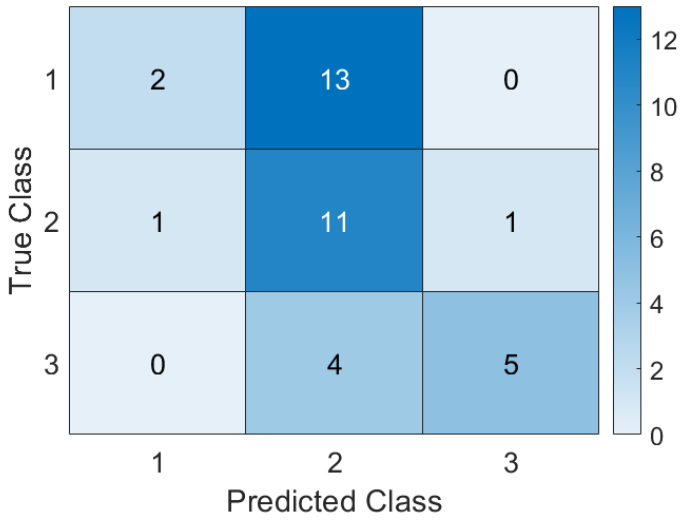


Figure 4.26: *Confusion Matrix of ELASTIC NET Model*

Metrics	Value
Accuracy	0.49
Precision	0.63
Recall	0.51
Specificity	0.74
F1-Score	0.48

Table 4.21: Metrics of ELASTIC NET Model

Features	Weight
PCWP_echo	1.0585
ci	0.7052
e_e'	0.4705
mPAP_echo	0.4371
Age	0.4212
Dmin_Python	0.2407
lvmi	0.2269
rv_fac	0.2171
CI_Python	0.2066
Dmax_Python	0.0594
lvef	0.0034

Table 4.22: Feature ranking with the Elastic Net model

- **Random Forest**

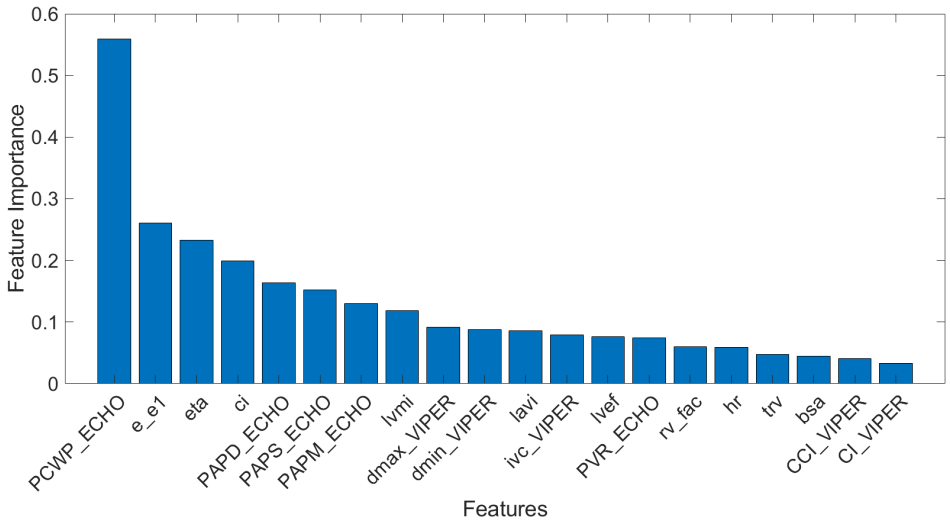


Figure 4.27: *Histogram of Feature Importance*

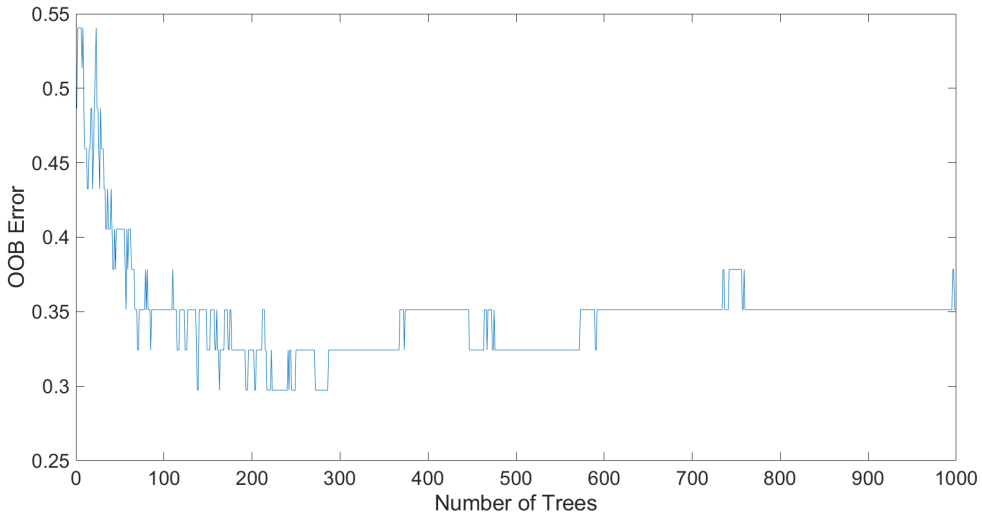


Figure 4.28: *OOB Error in function of tree number*

Metrics	Value
Accuracy	0.65
Precision	0.66
Recall	0.63
Specificity	0.81
F1-Score	0.64

Table 4.23: Metrics of Random Forest Model

Features	Weight
PCWP_echo	0.5595
e_e'	0.2602
Age	0.2327
ci	0.1989
dPAP_echo	0.1638
sPAP_echo	0.1526
mPAP_echo	0.1297
lvmi	0.1189
Dmax_Python	0.0921

Table 4.24: Feature ranking with the Random Forest model

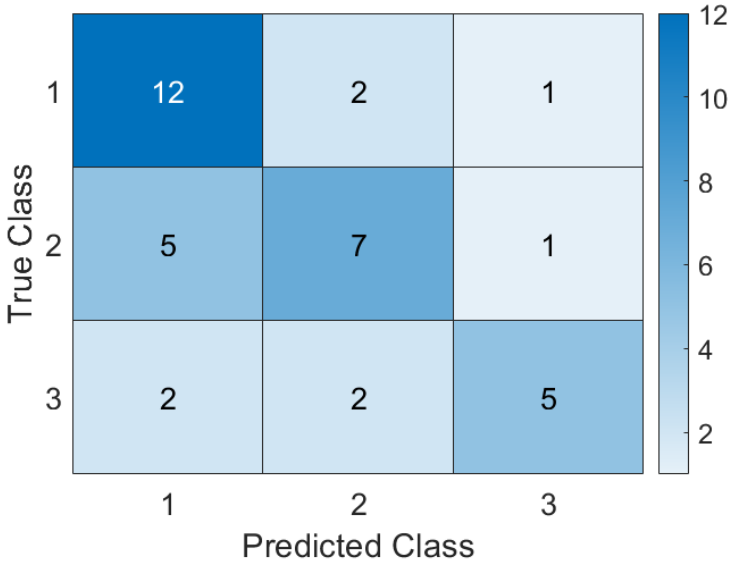


Figure 4.29: *Confusion Matrix of Random Forest Model*

- **Support Vector Machine**

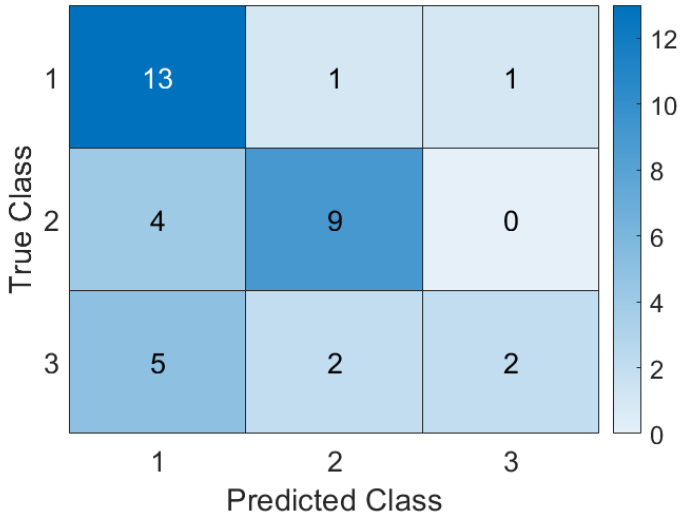


Figure 4.30: *Confusion Matrix of SVM Model*

Metrics	**Value**
Accuracy	0.65
Precision	0.67
Recall	0.59
Specificity	0.81
F1-Score	0.59

Table 4.25: Metrics of SVM Model

Features
PCWP_echo
PVR_echo

Table 4.26: Feature ranking with the SVM model

- **MetaClassifier** : a meta-classifier (RF Model) was trained using as input the prediction of 2 base models (RIDGE Model and SVM model)

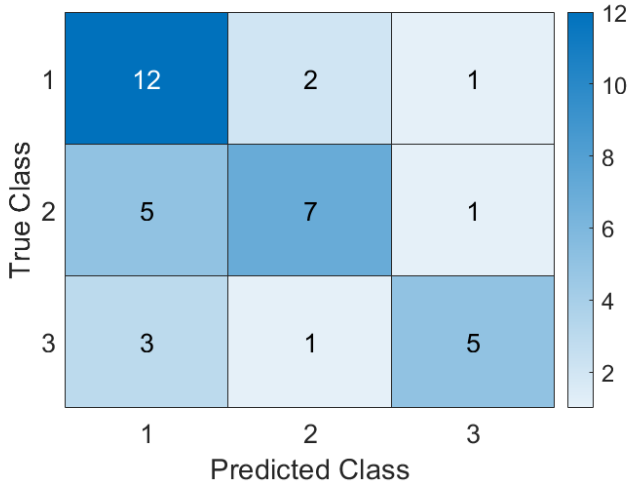


Figure 4.31: *Confusion Matrix of MetaClassifier*

Metrics	Value
Accuracy	0.65
Precision	0.67
Recall	0.63
Specificity	0.81
F1-Score	0.64

Figure 4.32: Metrics of Meta-Classifier

- **1D CNN**: the architecture implemented was the same as the one shown in Figure 4.21.

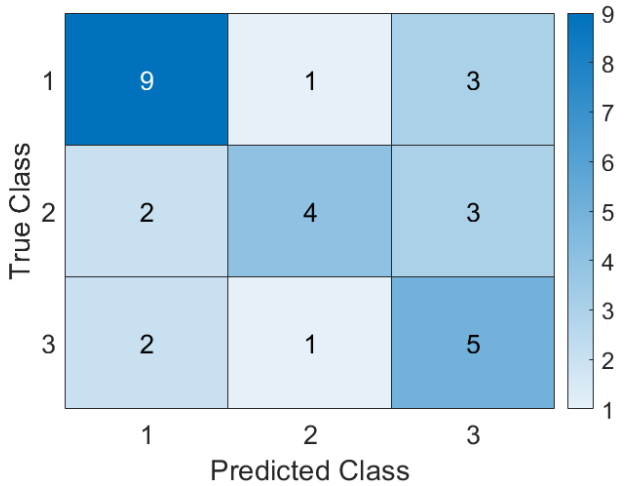


Figure 4.33: *Confusion Matrix of 1D CNN*

Metrics	Value
Accuracy	0.60
Precision	0.60
Recall	0.59
Specificity	0.80
F1-Score	0.58

Figure 4.34: Metrics of 1D CNN

Chapter 5

Discussion

5.1 Statistical Analysis

To provide a comprehensive understanding of the comparisons between Python and Matlab, multiple statistical analyses and graphical representations were employed. These complementary approaches help to elucidate the nature of the agreement, biases, and variability between the two algorithm tools.

Starting with the OLP regression analysis, the results summarized in Table 4.2 of Chapter 4 indicate that for Python and Matlab, the slope CIs interval includes 1, while the positive intercept range does not include zero. Although the OLP regression line crosses the identity line (Figure 4.3) the statistical analysis did not confirm the presence of a proportional bias.

In the head-to-head comparison between Python and Matlab (Figure 4.3*a*), the slope CI includes 1, which confirms that the proportional relationship between the two algorithm outputs is strong and not significantly different from being in perfect agreement. The intercept CI includes zero, suggesting no significant fixed bias between the two algorithm measurements.

The regression analysis results give information about the nature of the biases in the measurements taken with Python and Matlab. For further clarification, the Pearson correlation coefficients in Table 4.3 provide a quantitative indicator of the linear relationship between the measurements taken by the two algorithm and the manual measurement.

For Python, the Pearson's correlation coefficient is 0.5149, while Matlab obtains a Pearson's correlation coefficient of 0.4839, showing a moderate positive correlation with the manual measurements. This suggests that both automated measurements are following the trends in the manual measurements with some fluctuations.

In addition, the direct comparison between Matlab and Python yields a far higher correlation coefficient of 0.9507. The high positive correlation indicates that the two algorithm tools provide strongly consistent measurements to each other, measuring extremely similar patterns. Slight differences are only detected when the two algorithm are compared separately to the manual measurement.

Building upon these insights, the Bland-Altman plots (Figures 4.4) provide a visual representation of the agreement between each algorithm and the manual reference. The data points are relatively scattered around the mean bias line, which appears to be slightly below zero, indicating a small systematic overestimation of the estimated diameters with software. The limits of agreement are relatively wide, suggesting a considerable variability between the algorithm estimates and the reference values. Some data points lie outside the LoA, indicating occasional large discrepancies.

On the other hand, when comparing the Bland-Altman plots of Matlab vs Python (Figure 4.4*a*) it is evident that the limits of agreement are narrower than the ones obtained before. This indicates a higher level of agreement between the two algorithm tools and the mean bias line lies very close to zero, denoting minimal systematic differences between the two automatic system.

Finally, the Wilcoxon Signed Rank test results in Table 4.1 shed light on specific parameter differences between the two methods. Notably, the CCI shows a statistically significant difference between the two methods, with a p-value of 2.02×10^{-5}, well below the threshold of 0.05. This result suggests that the estimates of this parameter differ systematically between the two approaches.

For all other parameters, no statistically significant differences were detected. In fact, the corresponding p-values exceeded 0.05, as shown below: p = 0.1561 for IVC mean diameter, p = 0.8034 for maximum diameter, p = 0.0801 for minimum diameter and p = 0.1184 for the Caval Index. This indicates that any observed differences between the two measurement methods could be attributed to random variation rather than systematic bias.

The results obtained from this statistical test are supported by the graphical analysis (Figure 4.5 to Figure 4.9), which highlights that only the CCI parameter exhibits minimally overlapping histograms and statistically distinct boxplots. To reinforce these results, a Spearman correlation matrix (Figure 4.10) was also computed, confirming a strong correlation among the mean, maximum and minimum diameters, in contrast to the weak correlation observed between CCI and CI.

5.2 Model comparison

The main objective of this study was to evaluate the effectiveness of various classification models for the non-invasive estimation of RAP, with a focus on each model's ability to accurately discriminate between the three different RAP classes.

Initially, the traditional guideline-based model currently adopted by clinicians was considered. It had an overall accuracy of 54%, with great performance in classifying class 1 while it was very poor in correctly classifying classes 2 and 3. Although this approach is non-invasive and easy, its low accuracy makes it unsuitable for precise RAP estimation, especially in more complex cases.

Linear models were then considered. In general, these models showed performance around 50%, with the Ridge model working marginally better in terms of accuracy. All linear models worked well to recognize classes 2 and 3, but struggled with the correct classification of class 1. The Ridge model assigned a weight to each variable, which was used to estimate the significance of each feature in model development. The Lasso model used a similar approach, but assigned a weight of zero to several features, thereby automatically selecting only those variables with non-zero weights. Finally, the Elastic Net model was similar to Lasso but selected a larger number of effective features, but still applied a level of regularization.

The Random Forest model obtained an overall accuracy above 60%. Although class 1 was better predicted than classes 2 and 3, the model was still able to generalize across all three RAP categories. An importance-based approach was used to perform feature selection: a histogram was generated, and the mean of the importance values was chosen as a threshold to retain only the most relevant features. This procedure resulted in the selection of a restraint number of features considered fundamental for model development.

The SVM algorithm outperformed previous models, with an accuracy above 65%. In particular, classes 1 and 2 were well recognized, while class 3 remained more challenging. The feature selection method used in this case identified only a small group of features as relevant, highlighting the model's ability to achieve

good performance even with a limited number of predictors. In the Matlab-based model, three variables were selected, one of which was extracted by the software. Conversely, in the Python-based model, only two variables were selected, none of which derived from software outputs, suggesting that predictive information can also be captured through alternative features.

The meta-classifier developed by combining the predictions achieved with the best performance among all models considered (Ridge and SVM models), with an overall accuracy around 65%. This approach aggregated the specific advantages of each individual model: on the one hand, Ridge was able to correct classify classes 2 and 3; on the other, SVM performed well for classes 1 and 2. The integration of the two allowed for a more balanced and productive classification across all three classes, suggesting that merging different models could represent a valuable solution for RAP estimation.

Ultimately, the deep learning model based on the CNN architecture achieved performance comparable to that of the meta-classifier, demonstrating a good generalization capability across the three RAP classes. This result was obtained despite the exclusion of seven subjects with video durations below 1.5 seconds, a limitation that appears to have been mitigated by the higher informational content provided by the temporal sequences of IVC diameters.

The considerations done above apply to both Matlab and Python. However, the models implemented in Matlab generally show slightly superior performance than those developed with Python. The only exception to this trend is the Random Forest model implemented with Python-extracted parameters which achieves marginally higher performance than Matlab.

5.3 Limits of the study

Despite the promising results, this study presents several limitations that should be acknowledged. A primary constraint lies in the relatively short duration of the ultrasound video recordings, which limited the possibility of extracting the RCI in a reliable manner. Moreover, the quality of the echographic acquisitions varied considerably across patients, with several recordings affected by low image clarity or suboptimal visualization of the IVC, potentially compromising the accuracy of automated segmentation. Another important limitation is that only longitudinal views of the IVC were provided: no transverse plane acquisitions were available. This prevented the use of an additional functionality offered by the VIPER software, namely the segmentation and analysis of the IVC in the transverse plane.This limitation reduced the scope of the investigation and prevented a more comprehensive assessment of the IVC geometry, which may hold additional diagnostic value.

5.4 Future Development

Future work should focus on improving the quality and duration of ultrasound acquisitions. Specifically, it is recommended that future recordings should last at least 10 seconds to ensure sufficient temporal information for the reliable extraction of dynamic indices such as the RCI. Longer acquisitions would also enhance the robustness of both manual and automated analyses.

A further objective is to implement VIPER for real-time use, enabling simultaneous ultrasound acquisition and IVC segmentation directly during the examination, thus improving efficiency and clinical applicability.

Another important direction for future work is to expand the current dataset. Increasing the number and diversity of subjects would improve the generalizability of the study and make the results more statistical significative. A bigger dataset would also support the development of more robust machine and deep learning models and allow for subgroup analyses across diverse patient cohorts.

Chapter 6

Conclusion

This thesis proposed an innovative approach for the non-invasive estimation of RAP through automated analysis of ultrasound imaging of the IVC. Two distinct algorithmic pipelines were implemented (one developed in Matlab and the other in Python), both designed to perform semi-automatic segmentation of the IVC and extract clinically relevant features, including IVC diameter, CI, CCI and RCI. These features were used as input for classification models aimed at assigning each patient to one of the three RAP categories. The results demonstrate that automated methods offer superior performance compared to traditional guideline-based approaches, which rely on manual diameter measurements. Notably, the level of agreement between the Matlab and Python algorithms was higher than that observed between either algorithm and the manual measurements performed by the operator. These results suggest that, despite being independently developed, the two automated pipelines exhibit greater reproducibility and systematic behaviour than human assessment, which may be subject to inter- and intra-operator variability. All ultrasound videos were processed offline. Consequently, the results obtained with the Matlab algorithm were slightly better than those achieved with the Python version. This difference can be attributed to the fact that the Matlab implementation was tailored for more accurate and refined segmentation of the IVC, while the Python algorithm was optimized for real-time execution, favoring speed and simplicity over maximum precision. Despite the slight differences in performance between the two software implementations, statistical analysis revealed no significant differences in the same features extracted by both algorithm , with the exception of the CCI, which was the only variable to show a statistically significant discrepancy.

These results underscore the value of automated algorithmic solutions to improve the reliability and consistency of RAP classification. By minimizing the subjectivity inherent in manual measurements, these methods may contribute to more robust and repeatable clinical evaluation.

However, achieving high segmentation accuracy still depends on the availability of long enough and high-quality ultrasound videos, as low quality images can compromise the algorithm's ability to accurately detect and track the IVC borders.This level of image quality, however, cannot always be guaranteed in clinical settings due to differences in patients' physical characteristics such as: body size, tissue composition, or how easily internal structures can be visualized during the scan.

Bibliography

[1] Cindy L. Stanfield. *Fisiologia 4a edizione*. Cindy L. Stanfield, 2012, pp. 360, 361, 368, 369, 370, 400, 401, 402, 418.

[2] Lung National Heart and Blood Institute. "How the heart works". In: *National Heart,Lung and Blood Institute* (2022).

[3] Lumen Learning. "Cardiac Cycle". In: *Anatomy and Physiology II - SUNY* (2024). URL: https://courses.lumenlearning.com/suny-ap2/chapter/cardiac-cycle/.

[4] Luca Mesin et al. "Assessment of Phasic Changes of Vascular Size by Automated Edge Tracking-State of the Art and Clinical Perspectives". In: *Frontiers in Cardiovascular Medicine 8* (2022).

[5] Ambrosi G. et al. *Anatomia dell'uomo*. Edi. Ermes, 2010, p. 240.

[6] Luca Mesin and Pietro Policastro. "Processing Ultrasound Scans of the Inferior Vena Cava: Techniques and Applications". In: *Bioengineering* (2023). DOI: 10.3390/bioengineering10091076. URL: https://www.mdpi.com/2306-5354/10/9/1076.

[7] Alexander C. Egbe et al. "Pathophysiologic and Prognostic Implications of Right Atrial Hypertension in Adults With Tetralogy of Fallot". In: *Journal of the American Heart Association* (2019). DOI: 10.1161/JAHA.119.014148. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6915294/.

[8] Erwin Argueta-Sosa Gaspar Del Rio-Pertuz Kenneth Nugent. "Right heart catheterization in clinical practice: a review of basic physiology and important issues relevant to interpretation". In: *American journal of cardiovascular disease* (2023). DOI: PMC10352814.

[9] Gianfranco Sinagra Luca Mesin Stefano Albani. "Non-invasive estimation of right atrial pressure using inferior vena cava echography". In: *Ultrasound in medicine and biology* (2019). DOI: 10.1016/j.ultrasmedbio.2018.12.013.

[10] Eva M. Boorsma Jozine M. ter Maaten Kevin Damman Wilfried Dinh Finn Gustafsson Steven Goldsmith Daniel Burkhoff Faiez Zannad James E. Udelson and Adriaan A. Voors. "Congestion in heart failure: a contemporary look at physiology, diagnosis and treatment". In: *Nature Reviews Cardiology* (2020). DOI: 10.1038/s41569-020-0379-7.

[11] Nicola Cosentino Giancarlo Marenzi Manuela Muratori Damiano Magrì Gaia Cattadori and Piergiuseppe Agostoni. "Fluid balance in heart failure". In: *European journal of preventive cardiology* (2023). DOI: 10.1093/eurjpc/zwad166.

[12] Mark H. Drazner. "The Role of the Clinical Examination in Patients With Heart Failure". In: *Heart Failure Clinics* (2019). DOI: 10.1016/j.hfc.2018.08.002. URL: https://www.sciencedirect.com/science/article/pii/S221317791830283X.

[13] Maheswara Satya G. Golla and Pirbhat Shams. *Heart Failure With Preserved Ejection Fraction (HFpEF)*. Accessed: 2025-05-04. 2024. URL: https://www.ncbi.nlm.nih.gov/books/NBK599960/.

[14] Matteo Toma Stefano Giovinazzo Gabriele Crimi Giovanni Masoero Manrico Balbi Fabrizio Montecucco Marco Canepa Italo Porto Pietro Ameri. "Multiparametric vs. Inferior Vena Cava–Based Estimation of Right Atrial Pressure". In: *Frontiers in Cardiovascular Medicine* (2021). DOI: 10.3389/fcvm.2021.632302.

[15] Jonathan D Rich et al. "Noninvasive assessment of pulmonary artery pressure with Doppler echocardiography: a simultaneous Doppler echocardiography and cardiac catheterization study". In: *Journal of the American Society of Echocardiography* (2011). DOI: 10.1016/j.echo.2011.06.021.

[16] Mustafa Cetin Musa Cakici Cemil Zencir Hakan Tasolar Ercan Cil Emrah Yıldız Mehmet Balli Sabri Abus Erdal Akturk. "Relationship between severity of pulmonary hypertension and coronary sinus diameter". In: *journal of the Portuguese Society of Cardiology* (2015). DOI: 10.1016/j.repc.2014.11.017.

[17] Luca Mesin Piero Policastro Stefano Albani Christina Petersen Paolo Sciarrone Claudia Taddei Alberto Giannoni. "Non-Invasive Estimation of Right Atrial Pressure Using a Semi-Automated Echocardiographic Tool for Inferior Vena Cava Edge-Tracking". In: *journal of clinical medicine* (2022). DOI: 10.3390/jcm11123257.

[18] Elaine M. Kaptein Matthew J. Kaptein. "Inferior Vena Cava Collapsibility Index: Clinical Validation and Application for Assessment of Relative Intravascular Volume". In: *Advances in chronic kidney disease* (2021). DOI: 10.1053/j.ackd.2021.02.003.

[19] Lawrence G. Rudski MD FASE Chaira Wyman W. Lai MD MPH FASEb Jonathan Afilalo MD Msca Lanqi Hua RDCS FASEc Mark D. Handschumacher BScc Krishnaswamy Chandrasekaran MD FASEd Scott D. Solomon MDe Eric K. Louie MDf Nelson B. Schiller MD. "Guidelines for the Echocardiographic Assessment of the Right Heart in Adults: A Report from the American Society of Echocardiography". In: *American Society of Echocardiography* (2010). DOI: 10.1016/j.echo.2010.05.010.

[20] G: Coppini S. Diciotti G. Valli. *Bioimmagini*. Pàtron, 2012, pp. 177, 178.

[21] Robert M. Fulton. *Focused—Basic Ultrasound Principles and Artifacts*. Gregory R. Lisciandro, 2013.

[22] D. Junuzovic A. Carovac F. Smajlovic. "Application of Ultrasound in Medicine". In: (2011). DOI: 10.5455/aim.2011.19.168-171.

[23] VIPER S.r.l. *VIPER Software*. 2025. URL: https://www.viper.srl/dissemination-en.html.

[24] Murray J. Fisher RN Andrea P. Marshall RN. "Understanding descriptive statistics". In: *Australian Critical Care* (2008). DOI: https://doi.org/10.1016/j.aucc.2008.11.003.

[25] Built In. *Shapiro-Wilk Test: Definition, Formula, Example*. Accessed: 2025-05-19. 2023. URL: https://builtin.com/data-science/shapiro-wilk-test.

[26] John Ludbrook. “Regression analysis for comparing two methods of measurement: A review of current recommendations”. In: *Clinical and Experimental Pharmacology and Physiology* (). DOI: 10.1111/j.1440-1681.2011.05643.x.

[27] Marcus V. Nascimento-Ferreira, Augusto C. F. De Moraes, and Heráclito B. Carvalho. “Ordinary Least Products Regression: A Robust Statistical Tool for Assessing Agreement Between Measures Attended by High Variability”. In: *Research Square* (2020). Preprint. DOI: 10.21203/rs.3.rs-71462/v1. URL: https://doi.org/10.21203/rs.3.rs-71462/v1.

[28] Ruairi J Mackenzie. *The Wilcoxon Signed-Rank Test*. Accessed: 21 May 2025. 2022. URL: https://www.technologynetworks.com/informatics/articles/the-wilcoxon-signed-rank-test-370384.

[29] Creatrohit Rohit. *Lasso, Ridge & Elastic Net Regression — A Complete Understanding*. Accessed: 26 May 2025. 2021. URL: https://medium.com/@creatrohit9/lasso-ridge-elastic-net-regression-a-complete-understanding-2021-b335d9e8ca3.

[30] Aakash Kapur. *Random Forest Explained: A Visual Guide with Code Examples*. Accessed: 2025-05-26. 2020. URL: https://medium.com/data-science/random-forest-explained-a-visual-guide-with-code-examples-9f736a6e1b3c.

[31] MathWorks. *predictorImportance - MATLAB*. https://it.mathworks.com/help/stats/classreg.learning.classif.compactclassificationensemble.predictorimportance.html. Accessed: 2025-05-26. n.d.

[32] GeeksforGeeks. *Support Vector Machine (SVM) Algorithm*. Accessed: 2025-05-28. 2021. URL: https://www.geeksforgeeks.org/support-vector-machine-algorithm/.

[33] GeeksforGeeks. *Multi-class Classification Using Support Vector Machines (SVM)*. Accessed: 2025-05-28. 2021. URL: https://www.geeksforgeeks.org/multi-class-classification-using-support-vector-machines-svm/.

[34] Sharma Saravanan. *Understanding Stacking Classifiers: A Comprehensive Guide*. Accessed: 2025-05-28. 2023. URL: https://sharmasaravanan.medium.com/understanding-stacking-classifiers-a-comprehensive-guide-195bfab58e48.

[35] Laith Alzubaidi Jinglan Zhang Amjad J. Humaidi Ayad Al-Dujaili Ye Duan Omran Al-Shamma J. Santamaría Mohammed A. Fadhel Muthana Al-Amidie and Laith Farhan. “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions”. In: *journal of big data* (2021). DOI: https://doi.org/10.1186/s40537-021-00444-8.

[36] UpGrad Blog. *Basic CNN Architecture*. UpGrad blog post. https://www.upgrad.com/blog/basic-cnn-architecture/.

[37] Abhishek Jain. *Understanding the 1D Convolutional Layer in Deep Learning*. Medium blog post. https://medium.com/@abhishekjainindore24/understanding-the-1d-convolutional-layer-in-deep-learning-7a4cb994c981. 2024.