Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics

Ida Sandsbraaten

Tissue Deformation Estimation With Deep Learning on Ultrasound Data

Master's thesis in Cybernetics and Robotics
Supervisor: Gabriel Kiss
June 2020

NTNU
Kunnskap for en bedre verden

Ida Sandsbraaten

Tissue Deformation Estimation With Deep Learning on Ultrasound Data

Master's thesis in Cybernetics and Robotics
Supervisor: Gabriel Kiss
June 2020

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics

NTNU
Kunnskap for en bedre verden

Abstract

Monitoring of patients undergoing surgery is essential for ensuring that the cardiac function restores the desired functionality and that no acute kidney failure occurs. As of today, the perioperative assessment is conducted manually by clinical observation and observing vital signs. Fully automatic monitoring of cardiac function, and the possibility of non-invasive measurement of the kidney, can reduce the risk of complications and decrease the intra-observer variability.

In recent years, deep learning has revolutionized solving complex tasks like image segmentation and object detection, and have been successful for medical applications. This thesis aims to contribute towards fully automatic monitoring, based on ultrasound recordings, by investigating the feasibility of tracking landmarks on the kidney and the myocardium in the heart with a deep learning model. By using a frame-to-frame displacement estimation, points in the tissue can be tracked. The displacements are estimated by a convolutional neural network (CNN), inspired by previous work in the field of image registration. The CNN is trained unsupervised and returns a low-resolution displacement field. By using cubic B-spline interpolation, the field is upsampled and can be used to track each pixel's motion from one frame to the next. In addition to the CNN, a state-of-the-art image registration method, Advanced Normalization Tools (ANTs), is used to estimate the tracking points. Based on the position of the points in the basal segments in the heart, the strain is estimated. The estimated strain is compared to reference values obtained by an echocardiographer.

The results show that deep learning is feasible to estimate the positions of the tracking points in most cases, and this method estimates faster than classic image registration methods such as ANTs. Estimations from ANTs are closer to the reference values on filtered ultrasound sequences than for those unfiltered.

Sammendrag

Overvåking av pasienter som opereres er avgjørende for å sikre at hjertet opprettholder ønsket funksjonalitet og at akutt nyresvikt ikke oppstår. Per i dag gjennomføres den perioperative vurderingen manuelt ved klinisk observasjon og overvåking av vitale tegn. Helautomatisk overvåking av hjertefunksjon, og muligheten for ikke-invasive målinger av nyren, kan senke risikoen for komplikasjoner og redusere intraobserver variabiliteten.

I løpet av de siste årene har dyp læring revolusjonert utførelsen av komplekse oppgaver som bildesegmentering og objektdeteksjon, og har også vært vellykket for medisinske applikasjoner. Denne oppgaven tar sikte på å bidra til helautomatisk monitorering basert på ultralydopptak, ved å undersøke muligheten for å spore gjenkjennbare punkter i nyrene og myokardium i hjertet med en dyp læringsmodell. Ved å bruke en bilde-til-bilde forflytningsestimering, kan punktene i vevet spores. Forflytningene er estimert av et konvolusjonelt nevralt nettverk, inspirert av tidligere arbeid innen bilderegistrering. Nettverket er trent ikke-veiledet (eng: unsupervised) og returnerer et lavtoppløselig forflytningsfelt. Deretter interpoleres feltet til et høyoppløselig felt som brukes for å finne forflytningsvektoren for hver piksel i bildene. I tillegg til det nevrale nettverket, brukes også Advanced Normalization Tools (ANTs) til å estimere forflytningsvektorene. Basert på posisjonen til punktene i basalsegmentene i hjertet, estimeres longitudinal strain. Disse estimerte verdiene sammenlignes med referanseverdier funnet av en ekkokardiograf.

Resultatene viser at dyp læring gjør det mulig å estimere posisjonene til gjenkjennbare punkter i de fleste tilfeller, og at denne metoden estimerer raskere enn klassiske bilderegistreringsmetoder som ANTs. Estimatene fra ANTs er nærmere referanseverdiene på filtrerte ultralydsekvenser, enn for ufiltrerte.

Preface

The task was proposed by The Operating Room of the Future, St. Olavs hospital, in the spring of 2019. During the autumn of 2019, a preliminary project [1] was conducted as preparation for the Master's project, and the work presented in this thesis is a continuation of the preliminary project. Section 1.1-1.2, Section 2.1 and 2.4.1-2.4.2, and Section 3.2.1 are based upon the work done in the project [1].

Due to COVID-19, the original plan of investigating the feasibility of supervised deep learning was not able to be carried out. As a result of this, the research goals in this thesis were changed to include a comparison between unsupervised deep learning and state-of-the-art medical imaging registration methods.

Acknowledgements

The materials needed for this thesis were acquired by echocardiographers and radiographers at St. Olavs University hospital in Trondheim and Oslo University hospital. Erik Andreas Rye Berg provided reference values of strain for the test samples.

Thank you, Idar Kirkeby-Garstad, for funding the online GPU service Floydhub related to deep learning training.

I would like to express my sincerest gratitude to my supervisor, Dr. Gabriel Hanssen Kiss, for his thorough feedback and guidance whenever needed. Throughout both the preliminary project and this thesis, he has always made time for answering all of my questions and provided helpful guidance.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and Background

Estimations show that over 300 million operations take place each year worldwide, and the number is increasing [2]. Impaired cardiac function and acute renal failure are two of the side effects that may occur as a result of major surgeries and interventions.

Surgical procedures like bypass surgery and valve replacements have been shown to cause decreased myocardial contractility and, in some cases, atrial fibrillation and myocardial infection [3][4]. Therefore, the heart of the patients undergoing major interventions is monitored before, during and after the surgery. This perioperative observation is done by clinical observations and evaluating vital signs, such as heart rate, blood pressure, and blood oxygen level. Also, echocardiographic evaluation based on ultrasound imaging of the heart is used [5]. Transthoracic echocardiography (TTE) is the most common type of echocardiography used to assess cardiac function in the control room. During surgery, transesophageal echocardiography (TEE) is more used as the probe is placed in the patient's esophagus undergoing surgery. Both these tools are used to assess diastolic function and left ventricular systolic, and other cardiac parameters such as strain and ejection fraction. These measurements can provide important information about cardiac function.

It has been reported that 30% of the patients that undergo cardiac surgery experience acute renal failure [6][7], which can lead to a higher risk of infections, post-operative complications and higher mortality rates [8]. Studies [9] [10] [11] have identified some of the risk factors associated with acute renal failure in patients undergoing a cardiopulmonary bypass. They aimed to propose and validate a predictive scoring system capable of identifying patients at risk of developing acute renal failure and adopt strategies that would offer renal protection for these patients. The kidneys receive about 20% of the overall blood flow in the body, and constant perfusion is crucial for optimal kidney function. The kidney's auto-regulatory system is responsible for keeping the blood flow as constant as possible regardless of changes in blood pressure. As of today, there is no clinically non-invasive applicable method for monitoring the kidney's auto-regulatory function in real-time. In order to characterize this function, real-time continuous measurements of blood pressure and kidney perfusion are required. While blood pressure can be measured non-invasively with a finger cuff or invasively with a pressure catheter, the kidney perfusion is more challenging. A possible non-invasive measurement is Doppler based blood flow measurement. However, because the changes in the auto-regulatory mechanism are slow and have a low frequency, the Doppler sequences need to be at least three minutes long. The challenge is that the kidneys are located close to the diaphragm and their position changes in sync with the respiratory cycle. Thus, the position of the Doppler sample constantly has to be updated.

Fully automatic methods for cardiac and renal assessment are still in the research stage and currently not in use in the operating room. Automatization of monitoring renal and cardiac function can potentially lead to significant advantages like faster and more reliable results such that it can be used in real-time and still provide a complete assessment. It may also potentially detect changes in the kidney's auto-regulatory system and cardiac function earlier than today and reduce the risk of complications. Another advantage of automatic assessment, is the reduced intra-observer variability.

Over the last few years, machine learning has revolutionized several fields such as speech recognition, natural language processing and computer vision [12]. These approaches take advantage of layered structures of artificial neurons to enable the computer to build complex concepts out of simpler ones. Deep learning breaks complicated mapping down to a series of nested easier mappings, each described by a different layer in the deep learning model [12]. Object detection and classification has also been useful in medical applications, and these methods are often based on convolutional neural networks (CNN).

1.2 Previous Work

Image registration is a vast field and finds applications in various disciplines such as remote sensing, computer vision, and medicine. New image registration methods have been continuously developed for a long time, and only some main contributions related to this project will be presented briefly in this section.

Image registration has earlier been used in kidney-related tasks with success. A non-rigid registration algorithm was introduced by Sance et al. [13] in 2006. Their goal was to develop an automated method to correct the kidneys' deformation and motion in 3D DCE-MRI recordings. This method is classic image registration, where the calculations are too slow to use this approach when continuously monitoring the kidney in real-time.

Methods for automatic assessment of cardiac function have also been developed. Heimdal et al. [14] used tissue Doppler imaging (TDI) to visualize the regional function of the left ventricle in real-time by using a strain rate method. Speckle tracking echocardiography (STE) is an imaging technique used to analyze the tissue's motion in the heart based on the speckle pattern to the different regions of the myocardium. 2D STE was proposed by Reisner et al. in 2004 for estimating strain and strain rate [15] [16]. A limitation of the STE method is that the method relies on temporal resolution, and it does not handle rapid heart rates. STE also depends on good image quality [17].

de Vos et al. developed a deep learning approach for unsupervised affine and deformable image registration [18]. Here they proposed a Deep Learning Image Registration (DLIR) framework, which is an unsupervised training technique for convolutional neural networks. de Vos et al. used this method on cardiac sine MRIs and chest CTs. Using deep learning instead of the classic image registration method yields a much lower computation time, and the DLIR method is more applicable in predicting movements in real-time.

Based on the DLIR approach, a method for estimation of strain has been developed [19]. This contribution is based on myocardial landmark detection and frame-to-frame displacement estimation. The displacements are computed by using a convolutional neural network (CNN) and cubic B-spline interpolation, inspired by the work of de Vos et al. [18]. By using the displacements estimated, the landmarks are tracked, and the strain could be estimated.

1.3 Aim of Study

This study aims to, based on recent work in this field, continue the development of a fast deformable registration algorithm that is able to detect the movement of the kidneys and the left ventricle of the heart in ultrasound video. Predicted movements of the left ventricle will be used to estimate longitudinal strain. The result will be compared to state-of-the-art image registration methods. Different pre-processing methods, such as filtering, will be tested to see the influence on the prediction.

1.4 Structure of the Report

This thesis is a continuation of the work done in the project given in [1]. In this first chapter, the motivation behind this thesis is covered, as well as previous work in automatic assessment for medical applications. The theoretical background about the human heart and kidney, ultrasound imaging, and deep learning is presented in Chapter 2. Chapter 3 describes the preprocessing of the data and the different methods used in this project. Chapter 4 presents the results from the different methods. In Chapter 5, the results are discussed, along with recommendations for further work. A summary is found in Chapter 6.

Chapter 2

Theory

2.1 The Kidney

The kidneys are the organs in the body that produce urine. The pair of kidneys are located at each side of the spine, in the retroperitoneal space. This means that the kidneys' position will move in sync with the respiratory cycle, as illustrated in Figure 2.1. They are responsible for several essential functions in the body, such as filtering waste products from the blood and controlling the pH and potassium levels. The blood is received through the renal artery and leaves through the renal vein. Renal blood flow is between 1.0 and 1.2 liters per minute per 1.73 m^2 of body surface area [20].

The kidneys can regulate the blood flow over a wide range of blood pressures. The renal circulation is controlled by intrinsic mechanisms, and renal blood flow tends to remain constant despite changes in blood pressure [21]. Regulation of renal blood flow is controlled via tubuloglomerular feedback and the myogenic mechanism, as illustrated in Fig. 2.2. These two internal autoregulatory mechanisms will operate without influence from the outside.

It is proven in animal models [23] [24], which mimicked clinical situations, that changes in the kidney's auto-regulatory system occur early on during the disease, before changes in overall renal function are observable. Therefore, it has been hypothesized that monitoring kidney perfusion and detecting changes in the auto-regulation

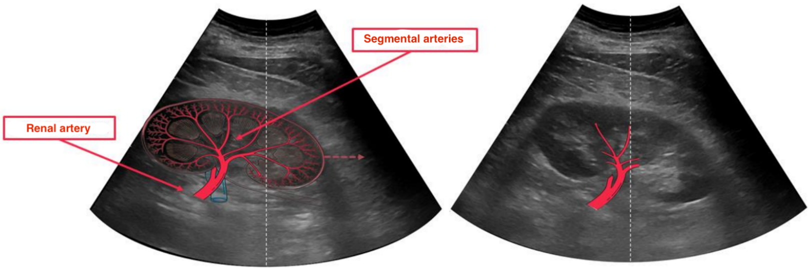


Figure 2.1: The renal vascular system overlaid on an ultrasound image. Both the kidney and associated arteries move significant as a result of respiration. Illustration: Andreas Østvik.

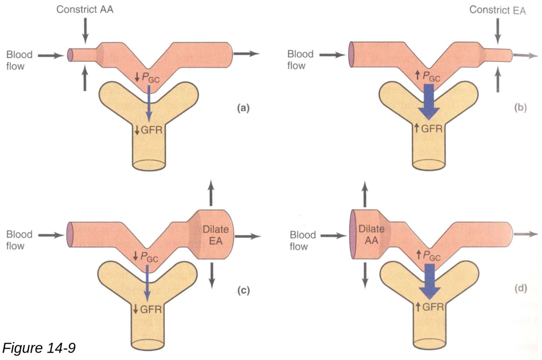


Figure 2.2: Diagram illustrating the kidney's auto-regulation system, including the two different mechanisms for keeping a constant blood flow in response to changes in blood pressure [22].

system can offer early diagnostic information and can be used for treatment of acute renal failure. Acute renal failure is said to happen when there is a fast decline in the glomerular filtration rate (GFR) resulting in nitrogenous wastes, such as creatinine and blood urea nitrogen, do not get filtered out by the kidneys [25].

2.2 The Heart

The human heart is located between the lungs in the area called thoracic cavity. It is divided into four chambers; upper left and right atria and lower left and right ventricles. Both the left and right sections have one atrium and one ventricle each and are separated by the septum. The two upper chambers are the receiving chambers, while the two lower ones act as the discharging chambers. The septum between the chambers includes four valves that allow the blood to flow from the atria to the ventricles and from the ventricles into the pulmonary trunk and aorta. These valves ensure that the blood runs only one way. The mitral valve is between the left atrium and the left ventricle. It opens and closes as a result of differences in the pressure in the left atrium and left ventricle [26].

Deoxygenated blood from the body is received through the venae cavae into the right atrium of the heart. From the right atrium, the blood will flow through the tricuspid valve into the right ventricle. Hereafter, the blood leaves through the pulmonary valve and into two arteries leading to each of the lungs. Gas exchange occurs, and carbon dioxide is removed from the blood, while oxygen is added. The oxygenated blood flows to the left atrium via the pulmonary veins. From here, the blood is pumped from the left atrium through the mitral valve and into the left ventricle. Then the blood flows through the aortic valve and into the aorta, which is the largest artery in the body. The aorta is branched into several arteries, which will supply the body with oxygenated blood [26].

Cardiac function is a cyclic process, commonly divided into two phases - diastole and systole. Systole is the part where the heart is contracted, and the blood is pumped into circulation. The systolic phase begins when the mitral valve closes. This marks the ending of the diastolic phase and is referred to as end-diastole (ED). Diastole is

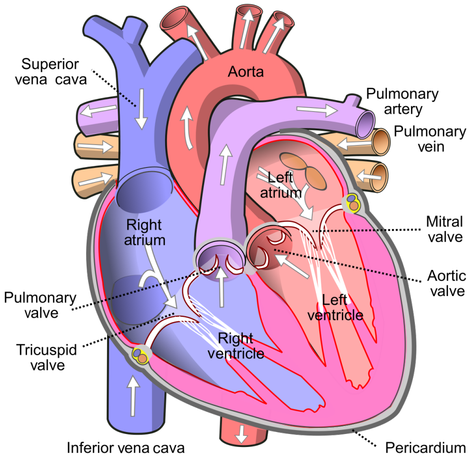


Figure 2.3: Illustration of the cardiac structure. Illustrated by Wikipedia user Wapcaplet. URL: `https://commons.wikimedia.org/wiki/File:Diagram_of_the_human_heart_(cropped).svg`.

the phase where the heart muscle relaxes, and the heart is refilled with blood after the emptying done during systole. The closing of the aorta marks the ending of the systolic phase and is referred to as end-systole (ES). In Figure 2.4, a Wiggers diagram shows how the pressure changes during the cardiac cycle.

2.2.1 Assessment of Cardiac Function

In order to gain knowledge about the cardiac function, the left ventricle (LV) is very important. The function of the LV is related to almost all cardiac diseases. LV mechanics consist of circumferential and longitudinal shortening and lengthening, radial thickening and thinning. By inspecting the LV deformation, essential information about the global and regional function can be extracted. Assessment of the regional cardiac function can be obtained by looking at the deformation of segments in the myocardium. This is done with visual inspection by a physician which makes the both the intra- and inter-observer variability high. The American Heart Association has

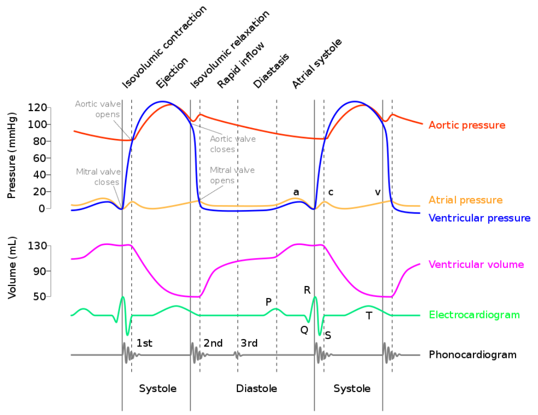
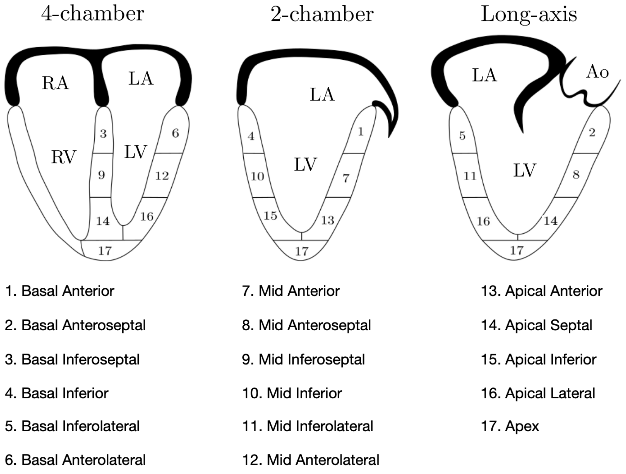


Figure 2.4: Wiggers diagram illustrating the aortic, ventricular and atrial pressures along with ventricular volume in the cardiac cycle. Illustration by Wikipedia user adh30 who revised work by DanielChangMD who revised original work of DestinyQx. Redrawn as SVG by xavax. `https://en.wikipedia.org/wiki/Wiggers_diagram#/media/File:Wiggers_Diagram_2.svg`

recommended a 17-segment model to standardize the definitions of the regions and is visualized in Figure 2.5.

Strain is a measure of how much an object has been deformed, and has been adopted in cardiology to standardize the measurement [27]. Strain describes the deformation relative to baseline length, and the calculation of the Lagrangian strain is defined in Equation (2.1) where L is the distance between two points on the myocardium [28]. L_0 is the initial length and typically at ED. By setting $t = ES$, the end-systolic strain can be calculated.

$$\epsilon(t) = \frac{L(t) - L_0}{L_0} \tag{2.1}$$



1. Basal Anterior
2. Basal Anteroseptal
3. Basal Inferoseptal
4. Basal Inferior
5. Basal Inferolateral
6. Basal Anterolateral
7. Mid Anterior
8. Mid Anteroseptal
9. Mid Inferoseptal
10. Mid Inferior
11. Mid Inferolateral
12. Mid Anterolateral
13. Apical Anterior
14. Apical Septal
15. Apical Inferior
16. Apical Lateral
17. Apex

Figure 2.5: The standardized 17-segment model proposed by The American Heart Association (AHA). Figure by Torjus Haukom [19].

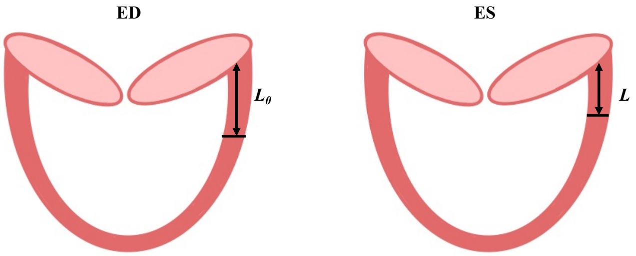


Figure 2.6: Figure illustrating the distances L and L_0 that are involved in the strain calculation.

2.3 Diagnostic Ultrasound

Ultrasound for medical diagnostics has been continuously developed since it was first used as a tool to image the brain in 1942 [29]. Today, ultrasound is one of the most used imaging technology in medicine as it is portable, free of radiation risk, and less expensive than other imaging principles such as magnetic resonance and computed tomography. Ultrasound can be used to show detailed tissue structures in 3D, measure blood flow, and screening of fetal abnormalities [30].

2.3.1 The Principles of Ultrasound

A pulse-echo approach is used to obtain medical ultrasound as we know it. The transducer transmits small pulses of ultrasound echo into the patient. These high-frequency sound waves will propagate through the tissue of the body where they are reflected and scattered as a result of tissue boundaries and small irregularities in the tissue. Echos are generated, and some echos will travel back to the transducer and are used to form the ultrasound B-mode image. The brightness of the image corresponds to the amplitude of the received echo, which is the reason it is called B-mode (brightness mode) [31].

2.3.2 Echocardiography

Echocardiography refers to the use of ultrasound of the heart, and is a commonly used tool in cardiology for assessing cardiac function. The most used types of echocardiography are transthoracic echocardiogram (TTE) and transesophageal echocardiogram (TEE). TTE is the standard echocardiogram where the probe is placed on the exterior of the patient's chest. This gives a noninvasive and quick assessment of the heart. TEE is, on the other hand, an invasive procedure where the probe is inserted into the patient's esophagus, as shown in Figure 2.8. The proximity of the probe in the esophagus to the heart and great vessels gives additional information to the TTE and is often used in the operating room. TEE can provide high-resolution images in real-time and be used for perioperative use in order to assess cardiac function [32].

Figure 2.7: Example of a B-mode ultrasound image.

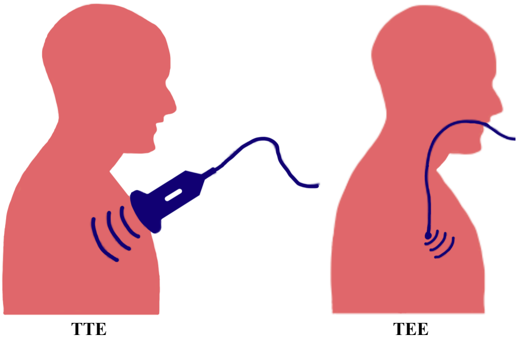
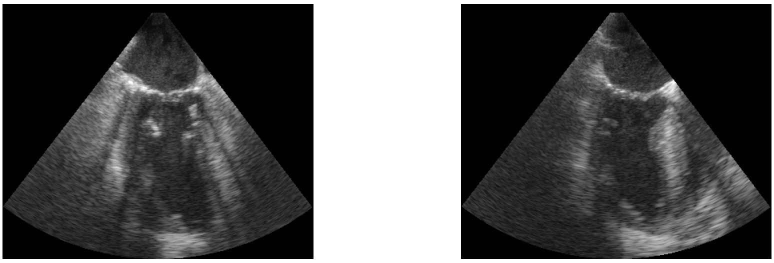


Figure 2.8: Illustration of the conceptual difference between TTE and TEE.

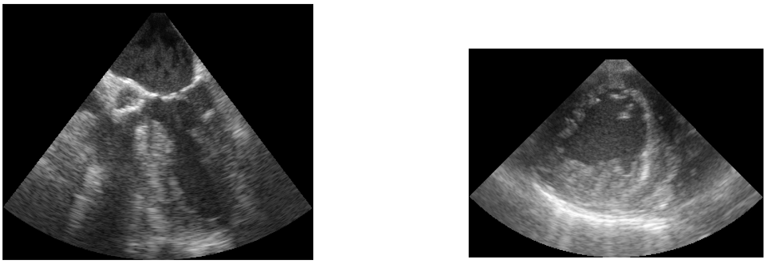
2.3.3 Echocardiographic Tomographic Views

When using TEE to examine the heart, different views can be obtained to give a detailed assessment. These views are defined by the transducer position and the orientation of the tomographic plane through the heart [33]. Four essential views for TEE recordings are four-chamber (4C), two-chamber (2C), apical long-axis (ALAX) and short-axis (PSAX). Figure 2.9 shows the different images for the four views.



(a) Two-chamber

(b) Four-chamber



(c) Apical long-axis

(d) Short-axis

Figure 2.9: The four essential views for TEE recordings.

Figure 2.9a shows the left ventricle and parts of the left atrium. Figure 2.9a also shows the complete left ventricle as well as some of the left and right atrium and right ventricle. Figure 2.9c shows the left ventricle, parts of the left atrium, the aortic valve, and the start of the aorta. In Figure 2.9d, the aorta is in the center of the image, surrounded by the left atrium, the right atrium, the right ventricle, and the pulmonary artery. This view is also great for observing the aortic valve [34].

2.4 Deep Learning

Deep learning is a subset of machine learning methods which is based on neural networks. The neural networks consist of a collection of artificial neurons inspired by the biological neural networks in brains. Deep learning is the key to solve many problems and has been proven useful in fields such as speech and audio recognition, natural language processing, online advertising, robotics and image segmentation [12]. It has become even more useful as the amount of available data for training has expanded and the computer infrastructure has become more powerful over the recent years.

Machine learning methods can be divided into three categories: supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, the model is given examples of input-output pairs and is asked to learn from these. Unsupervised learning techniques extract patterns directly from the given data set without any prior information given. Some main applications in unsupervised learning are clustering, generative models, and feature learning. Reinforcement learning uses feedback from actions taken in an unknown environment, and from this, the agent will learn suitable actions to maximize the rewards.

Figure 2.10 shows a neural net with one input layer, one hidden layer, and one output layer. The computation performed starts with an input vector from the left and flows to the right. There is no feedback loop, and the neural network is called a feed-forward neural network, which is the simplest form of an artificial neural network (ANN). Several neurons form a layer, and many layers can be stacked to obtain deep networks. In a neural network, each node in the hidden layer is connected to all the nodes in the previous layer via a weighted edge. The edge from node i in layer $(l-1)$ to the node j in layer l is denoted as $w_{i,j}^{(l)}$. The output at the node j in layer l is denoted

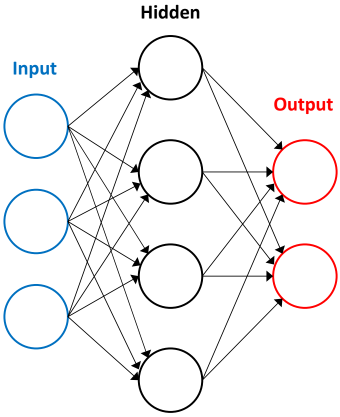


Figure 2.10: Illustration of a simple neural network with one input layer, one hidden layer and one output layer.

by x_j and is given by Equation (2.2).

$$x_j^{(l)} = \phi\left(\sum w_{i,j}^{(l)} x_i^{(l-1)} + b_j^{(l)}\right) \tag{2.2}$$

The function ϕ denotes an activation function, and there are several possibilities for choosing this. The most popular activation functions are sigmoid, softmax, rectified linear units (ReLUs), and tanh. An important property of activation functions is to make the neuron non-linear, by applying one of these functions [12]. The constant term $b_j^{(l)}$ represents the bias and is a parameter like any of the weights. The learning part of the neural network training consists of finding all these optimal parameters.

2.4.1 Convolutional Neural Networks

Convolutional neural networks (CNNs) are a subset of neural networks where at least one layer performs a convolutional computation of its input. The mathematical operation convolution is the substitute of the general matrix multiplication, which is the traditional computation. When using convolution, a filter is applied to one section of the input and calculates the value within the section of the filter, and moves

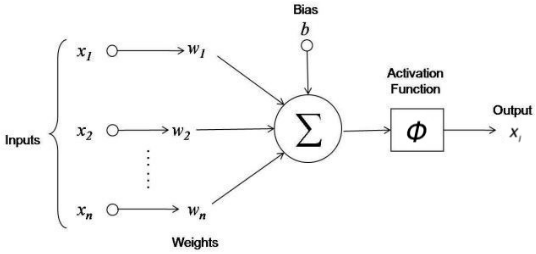


Figure 2.11: Illustration of the calculation performed by a neuron. Illustration:`https://medium.com/shallow-thoughts-about-deep-learning/how-would-we-find-a-better-activation-function-than-relu-4409df217a5c`

to the next position. This goes on until the filter has reached the end of the input. The size of the filter can vary, but its height and width are usually smaller than the input dimensions. This gives the CNN sparse connectivity, contrary to traditional neural networks, where every output unit is connected to every input unit. Sparse connections decrease the number of parameters the network has to learn.

In addition to the sparse connectivity, CNNs also exploit another principle of convolutional networks - parameter sharing. By sharing the weights, the number of parameters to learn is reduced even further. Figure 2.12 illustrates these two principles. Each output node will have exactly the same number of input edges as the total filter size. The color of the weights shows that the parameters are shared between the different nodes.

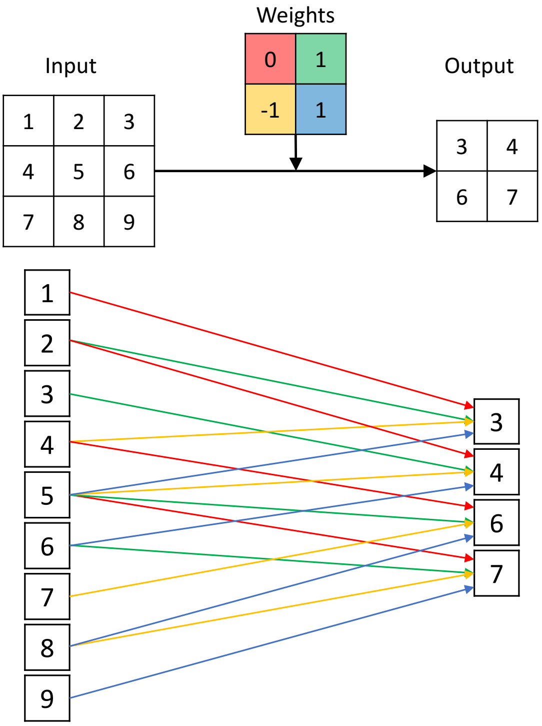


Figure 2.12: A convolutional layer illustrated like an ordinary neural layer.

After the convolutional layer, there is often a ReLU activation function, followed by a pooling layer [12]. Pooling is a way of non-linear downsampling, and the most common method is max-pooling. Here, the layer takes the highest value inside the region of pooling and pass it further to the output. After the final layer of convolution, average pooling is often used in order to make the layer fully connected. The final pooling layer will also make the network return the correct dimensions of the output regardless of the size of the input to the network.

2.4.2 Training Neural Networks

When training neural networks, a cost function is used to find out what kind of parameters θ yield the best result. The cost function $J(\theta)$ quantifies how good the model is doing, and the goal of the training is to minimize this cost. $J(\theta)$ is defined as the expected value over the training set x of the loss function $L(\hat{f}(x,\theta)) : J(\theta) = \mathbb{E}\left[L(\hat{f}(x,\theta))\right]$ where $\hat{f}(x,\theta)$ is the output of the network. The most used technique for training neural nets is stochastic gradient descent, where the parameters are updated with a learning rate α, as shown in the update rule defined in Equation (2.3).

$$\theta_i = \theta_{(i-1)} - \alpha \nabla J(\theta_{(i-1)}) \tag{2.3}$$

The algorithm to compute all the derivatives jointly is called backpropagation, which uses the chain rule of calculus to calculate the loss of a network with N layers. That means it starts with the computation of the last layer with respect to the second last layer's output. After that, the second last layer's gradient is computed with respect to the third last layer's output and so on. This gives us the following method to compute the gradients

$$\frac{\partial}{\partial\theta} L(\theta) = \frac{\partial L}{\partial \hat{f}(x,\theta)} \frac{\partial \hat{f}(x,\theta)}{\partial\theta}$$

$$\frac{\partial \hat{f}(x,\theta)}{\partial\theta} = \frac{\partial \hat{f}^{(N)}(\hat{f}^{(N-1)}(\cdots \hat{f}^{(2)}(\hat{f}^{(1)}(x))))}{\partial \hat{f}^{(N-1)}(\hat{f}^{(N-2)}(\cdots \hat{f}^{(2)}(\hat{f}^{(1)}(x))))} \cdots \frac{\partial \hat{f}^{(2)}(\hat{f}^{(1)}(x))}{\partial \hat{f}^{(1)}(x)} \frac{\partial \hat{f}^{(1)}(x)}{\partial\theta}$$

The calculation of the derivatives gets computationally heavy when the network consists of many hidden layers. Training is then quite a time-consuming operation and requires powerful machines in order to optimize all these large amounts of data. When the training is done, using the network to predict new samples is done in a split second.

2.4.3 Cross-Validation

In order to estimate the model's generalization error in the most accurate way, the test examples can not also be used for training. Therefore, no examples in the test set can be used in the training nor validation sets. The validation set is constructed from the training data, which gives two disjoint subsets. One of the subsets, the training set, is used to learn the parameters, while the other subset, the validation set, is used to estimate the generalization error during the training. The hyperparameters will be updated accordingly. The ratio of the training data size to the validation data size is often 80/20 [12]. Because the validation set is included in the training of the model, the error it estimates is underestimated compared to the true generalization error. If the available data is limited, and the validation set is small, the statistical uncertainty of the estimated error will make it difficult to claim that one model performs better than others on a given task. By performing training and validation repeatedly on randomly chosen subsets of the original dataset, all examples can be used to estimate the mean error. This increases the computational cost compared to no-repeat training. The most common technique is k-fold cross-validation, where the dataset is split into k non-overlapping subsets. The error is estimated by taking the mean of the error across k trials. On the first trial, the first subset is used as the validation set, and the rest of the subsets are used for training. This is repeated k times [12].

Chapter 3

Data and Method

3.1 Data

The renal data set used in this project consists of ultrasound recordings of the kidney gathered from 30 healthy volunteers with a Vivid E95 scanner and a C1-6-D probe (GE Vingmed Ultrasound, Horten, Norway). The acquisitions are done with the patient lying on the left side when scanning the right kidney, and only the 2D B-mode images are used in this project. The length and frame rate of the recordings varied, but all included at least one breathing cycle. As a result of different methods used when scanning, 16 of the recordings had a frame rate of 33 fps, while the remaining 14 had 9.5 fps. Due to different body fat compositions in the patients, the size and placement of the kidney varied between the recordings. This lead to varying sector width and acquisition depth.

For the cardiac data set, 2C, 4C and ALAX recordings from 88 patients are collected. Only the 2C and 4C sequences are used in this project. Cardiologists with echocardiographic expertise have scanned the patients with Vivid E95 and E9 systems with a 6VT-D probe (GE Vingmed Ultrasound, Horten, Norway). The patients were examined for diagnostic purposes, while five of these were scanned under surgery (coronary artery bypass grafting in four cases and mitral valve clipping for one patient). All of the cardiac recordings last for at least three cardiac cycles and are taken from three different views: 4C, 2C, and ALAX. The frame rate varies in the range between 30 to

60 frames per second, and the pixel brightness was recorded in the range $[0, 255]$.

In order to evaluate estimated strain, reference values for basal longitudinal strain were provided by an expert echocardiographer. The reference values were acquired by manually annotating the images and tracking the myocardium using the EchoPAC (GE Vingmed Ultrasound, Horten, Norway) speckle tracking software.

Three recordings with PSAX view are all taken from the same patient undergoing surgery and are taken with the same equipment as the 2C, 4C and ALAX recordings. The PSAX recordings are taken perioperatively, at The Intervention Centre on Oslo.

All samples from both the renal and cardiac data set were used. After recording, the samples were anonymized before exported to proprietary DICOM files. Every recording was then converted from DICOM to 2D images by applying a polar-Cartesian transform on the raw B-mode lines. Finally, the image data and geometric information were exported to HDF5 files. In order to be able to read the HDF5 files in this project, the open-source library $h5py$ has been used.

3.1.1 Preprocessing

Both raw and filtered images are tested in this project. When the images were filtered, a bilateral filter and a non-linear correction are applied. The bilateral filter will smooth the images while preserving the edges, and this is done by computing a non-linear combination of the nearby pixel values. It combines domain and range filtering, which means it takes both the coordinates and value of the pixel into consideration when computing the filtered pixel value [35]. A simple non-linear correction was then added to map the input pixel value to a given output. The non-linear function is shown in Figure 3.1. This method will suppress noise around the edges and preserve the edges further.

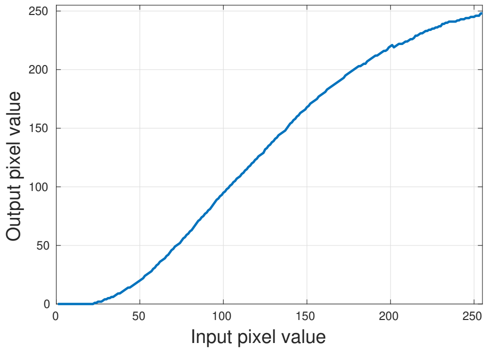


Figure 3.1: Non-linear function to map input value to an output value in a given pixel.

An example of a raw image which is filtered with bilateral filter and then corrected with a non-linear function is illustrated in Figure 3.2.

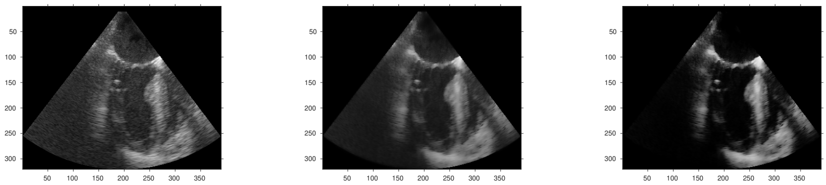


Figure 3.2: The original raw image, bilateral filtered image, and bilateral filtered and non-linear corrected image.

3.1.2 Data Preparation for Deep Learning

Each cardiac recording was split into multiple sequences such that one new file contained one ED-ED cycle. Before the learning of the model could begin, the data was divided into three separate datasets. Twenty-three patients and their corresponding ED-ED cycles had known strain collected by the physician and were allocated to the test set. The remaining 65 patients were divided randomly into five different groups in order to perform 5-fold cross-validation. For each training of the network, four groups were used as the test set, while the last one was the validation set. The one group chosen for validation changed for each training. Both the training and validation sets are organized into HDF5-files containing pairs of consecutive frames, as can be seen in Figure 3.3. The order of these pairs is randomly shuffled. All pixels in the datasets were scaled to $[0, 1]$, and all images were zero-padded in order to have the same size before training of the neural network could start.

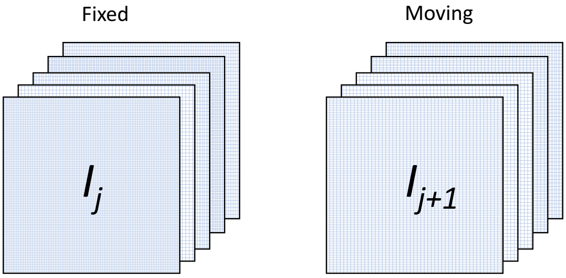


Figure 3.3: Overview of how the training and validation sets are organized into fixed and moving images, where frame I_j is in the fixed set and the consecutive frame I_{j+1} is located in the moving set with the same index as the fixed frame.

3.2 Method

In order to estimate strain from TEE images, the location of at least two points on each basal segment and a method to track these points are needed. The initial location of the landmarks is in this project assumed to be known, and corresponds to the landmark detection part of the proposed pipeline Figure 3.7. Two methods for motion estimation are proposed - Deep learning and Advanced Normalization Tools (ANTs). These two different methods are presented in this section. In the strain estimation step, the distance between the two points in each basal segment are found. The smallest distance is defined as the time of ES, and the distance at this point is used as $L(t)$ in the calculation of the Lagrangian strain in Equation (2.1). Ideally, if ECG data were available, this would have been used to precisely find the time of ES.

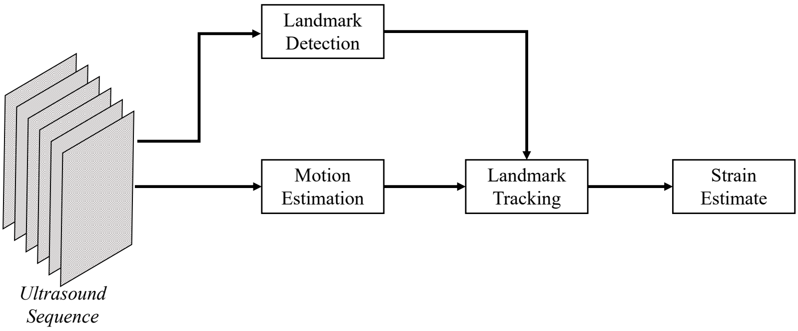


Figure 3.4: Proposed pipeline for strain estimation.

3.2.1 The Deep Learning Model

The main part of the method, which is based on the work by de Vos et Al [18], is a convolutional neural network which takes in two consecutive frames I_i and I_{i+1} and returns a displacement field $\overrightarrow{D}$ describing the two-dimensional motion between the two frames. This displacement field has a low resolution, so after the CNN, it is upsampled by using cubic B-spline such that each pixel corresponds to a motion vector. These vectors can then be used for point tracking from one frame to the next one.

To obtain the $\overrightarrow{D}$, the CNN architecture takes in one fixed and one moving frame and concatenates them. After concatenation, alternating layers of convolutions and downsampling are applied. The number of downsampling layers is specified by the user when calling the network to train. Two more layers of convolution are applied before two final 1 x 1 convolutions are used in order to obtain the estimated displacements. See Figure 3.5 for the illustration of one network.

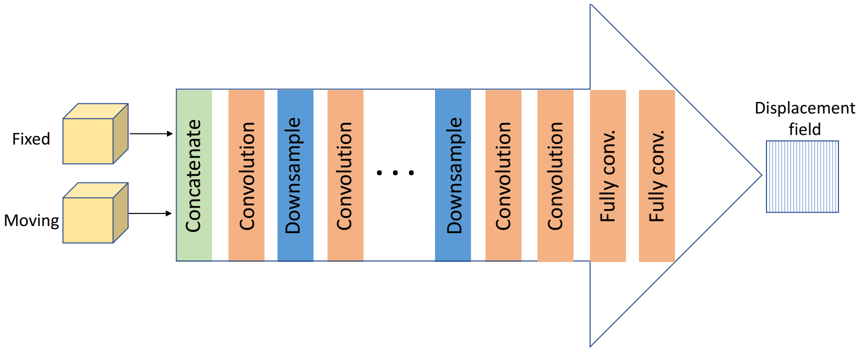


Figure 3.5: Illustration of the different layers in one CNN.

A multi-stage strategy has been introduced [18] in order to predict the displacements by using several networks in a daisy-chain with a decreasing number of downsampling rates. The warped image from one network is propagated to the next one and is compared to frame I as illustrated in Figure 3.6. By using this multi-stage approach, the model is less sensitive to local optima and image folding [18], and each CNN is trained sequentially.

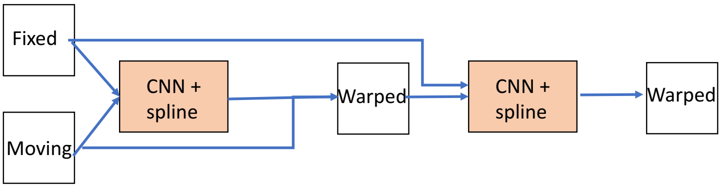


Figure 3.6: The multi-stage approach, where the warped image from one CNN is propagated to the next one.

3.2.1.1 Loss Function

The loss function used when training the neural network proposed by de Vos et al [18] is defined as

$$L = L_{NCC} + \alpha P,$$

where L_{NCC} is the negative normalized cross-correlation (NCC) and P is the bending energy penalty. α is defined by the user. NCC is a statistical measure of the tendency between two signals to vary linearly with each other and is often used for finding image correspondences. The bending energy will provide smooth deformations because it will minimize the second-order derivatives of the local transformations of the displacement vector field.

3.2.1.2 Implementation and Training

After preprocessing the data as described in Section 3.1.1, the training started. The model is implemented in Python 3.6.9 using Tensorflow version 1.14, and the original model is developed by Haukom T. [19] inspired by de Vos et al [18]. The source code is then adapted to train on the cardiac and renal data and can be found in the GitHub repository for this project [1]. To improve the speed of training, a Tesla K80 GPU was rented at Floydhub cloud service.

From earlier project [1][19][18], a daisy-chain approach has been showing the most promising results. A model with three networks of four, two and one downsampling layers respectively and fixed hyperparameters with a learning rate for the Adam optimizer equal to $\alpha = 10^{-4}$ and regularization parameter at $\lambda = 10^{-6}$ was tested. 5-fold cross-validation was also implemented in order to estimate the most accurate generalization error.

---

[1] `https://github.com/idasand/Tissue_deformation_estimation`

3.2.2 ANTs

Advanced Normalization Tools (ANTs) is software for image analysis and can be used for applications such as image registration, segmentation, geometric quantification, and statistics. ANTS is often considered as one of the state-of-the-art tools for biomedical image registration and segmentation [36]. A lot of the ANTs' functionality is based on Insight ToolKit (ITK), and the code is open source.

The registration method in ANTs takes in a pair of fixed and moving images and returns the transform to move from the moving to the fixed image. It is important that the input images are in the correct file format, more specifically an ANTsImage [2]. This conversion needs to be done before the registration. Different types of transforms can be used, and this is defined in the input by the user. Some of the different transforms that can be chosen are scaling, rotation, translation, affine, and rigid. The transformation method used in this project is SyNCC, which is symmetric normalization based on affine and deformable transformation with cross-correlation as the optimization metric. After the type of transform is given, ANTs will output the transform to map the fixed image to the moving one. This transform is then applied to the set of tracking points to estimate the position of the points in the fixed image.

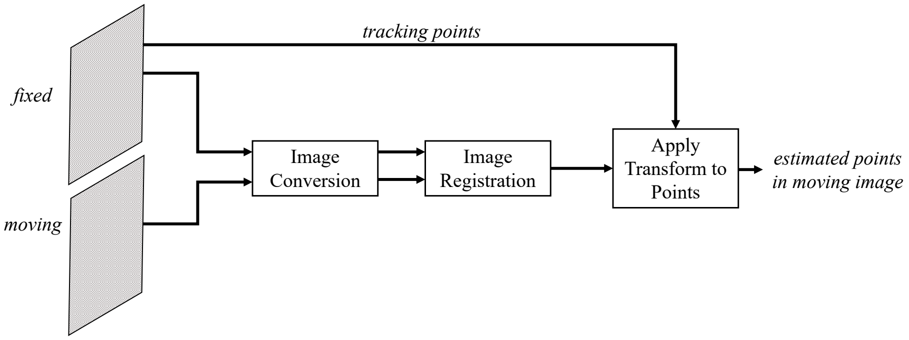


Figure 3.7: Pipeline that illustrates how the ANTs method estimated the position of the tracking points in the moving image.

---

[2] `https://antspy.readthedocs.io/en/latest/core.html#ants.core.ants_image.ANTsImage`

3.3 The Experiments

This project consists of two parts. The first one is landmark estimation used on ultrasound videos of kidney and cardiac short axis view of the heart. The second part is strain estimation based on two- and four-chamber view. Both parts use the same two methods. These methods are ANTs, both with the raw, unfiltered B-mode images and with images filtered with bilateral filter and non-linear correction, and Deep Learning network trained on images filtered with bilateral filter and non-linear correction.

3.3.1 Landmark Tracking

Initial tracking points are placed on the first frame in the ultrasound sequence. Then different methods are used to calculate position for the same points in the next frame based on the deformations between the first and second frame. This estimation of the points' position continues throughout the sequence. As no true position for the points is available, reference points are found manually by the author throughout the cycle. This method is performed on both renal and short-axis videos.

3.3.2 Strain Estimation

Initial points are placed in the first frame also for this part of the project. Two points are placed on each of the flaps in the mitral valve and the two remaining are placed further down on the myocardium. These four tracking points are estimated throughout one ED-ED cycle for the test recordings with the different methods. By using the positions, the time of ES is estimated, and the Lagrangian strain is computed. This estimated strain is compared to the reference strain from an expert echocardiographer, obtained in EchoPAC (GE, Vingmed, Horten, Norway).

Chapter 4

Results

4.1 Deep Learning Model Training

During the training of the networks, the normalized cross-correlation (NCC) was monitored for both the training and validation set. The NCC was calculated each step for the training set and every 100 steps for the validation set. Training stopped when the NCC converged.

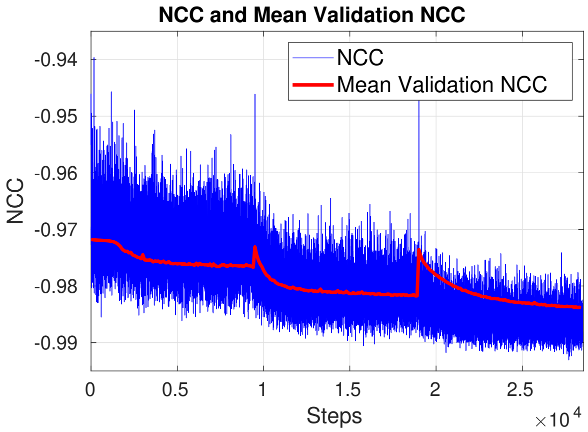


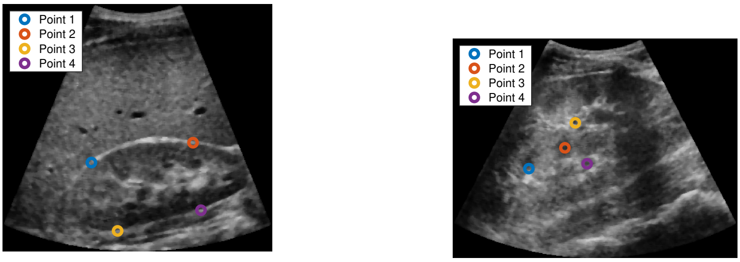
Figure 4.1: NCC and Mean NCC for the daisy-chain network consisting of 4, 2 and 1 downsampling layers.

5-fold cross-validation was performed to ensure the estimated error is as close to the real generalization error as possible, and the validation results are averaged over the five rounds in order to give an estimate of the model's predictive performance. The Mean Validation NCC for the five different validation sets gave average Mean Validation normalized cross-correlation $-0,982 \pm 0,0021$.

4.2 Landmark Tracking

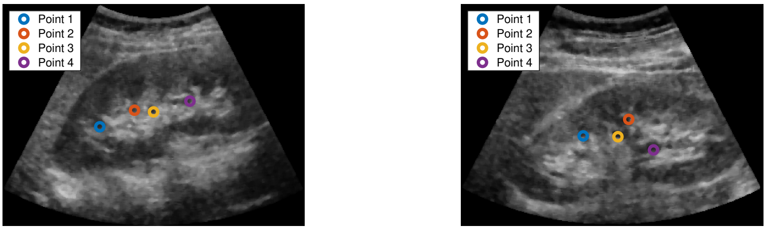
4.2.1 Renal Results

Four sample recordings were extracted from the kidney data set and used for testing. In each recording, four initial tracking points were marked, as seen in Figure 4.2. Both the renal cortex and the veins area are included. These points were tracked throughout each recording with different image registration methods and filtering. Manually reference points were marked by a non-clinician and compared to the estimated new positions for the points. The error between the estimated and reference points can be found in Figure 4.3-6.



(a) Patient 1

(b) Patient 2



(c) Patient 3

(d) Patient 4

Figure 4.2: Initial points for the renal test images.

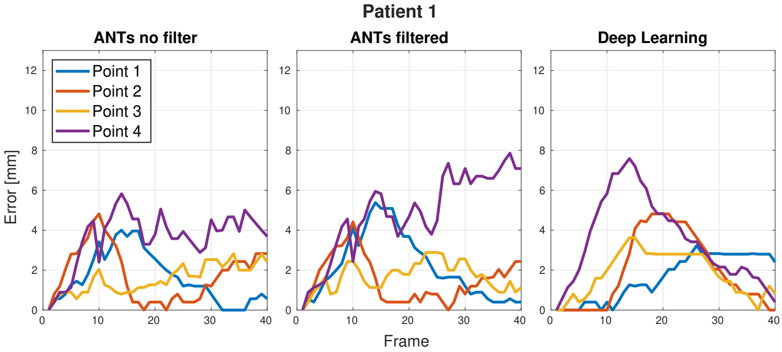


Figure 4.3: Patient 1

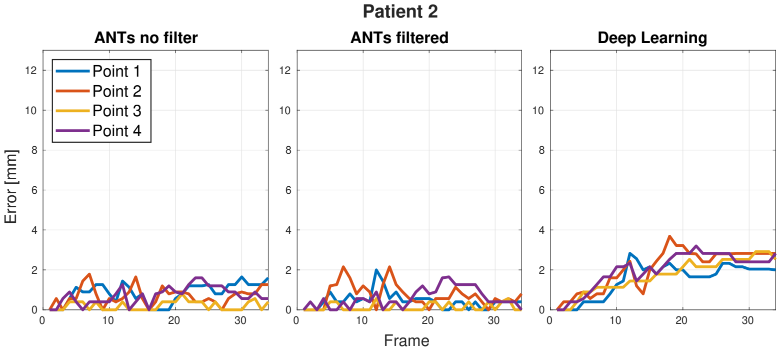


Figure 4.4: Patient 2

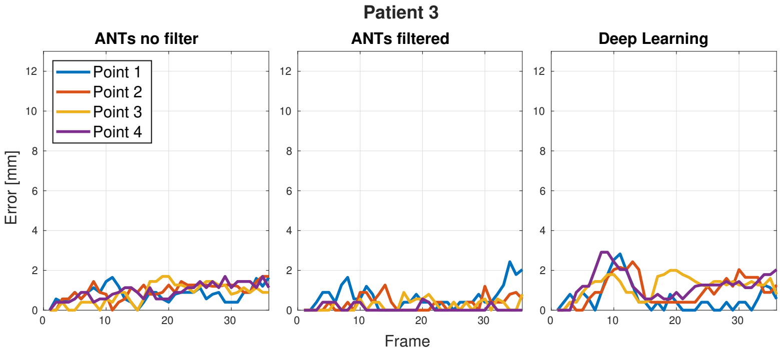


Figure 4.5: Patient 3

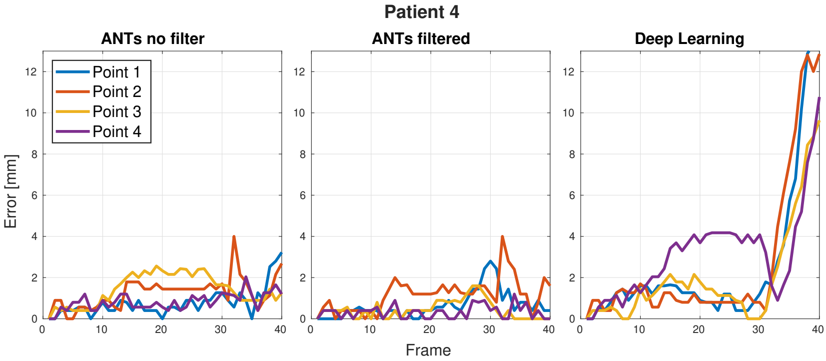


Figure 4.6: Patient 4

Tracking points at the renal cortex and in the renal veins area are tested, and the mean error in each of these two categories can be seen in Table 4.1.

Table 4.1: Error between the estimated and manual reference points for the whole respiratory cycle in the two categories in all four patients.

Category	**ANTs no filter**	**ANTs filtered**	**Deep Learning**
Renal cortex	2.1 ± 1.5 mm	2.6 ± 2.0 mm	2.2 ± 1.8 mm
Renal veins area	0.9 ± 0.6 mm	0.5 ± 0.5 mm	1.8 ± 2.1 mm

4.2.2 Cardiac Results

In three different ultrasound recordings, all from the same patient undergoing surgery, six tracking points in the myocardium were placed and tracked through a cardiac cycle with different methods and filtering. The initial points can be seen in Figure 4.7. No exact position of the points was available, leading to the manual reference points are found by a non-clinician. The error between the estimated and manual reference points is plotted in Figure 4.8-10.

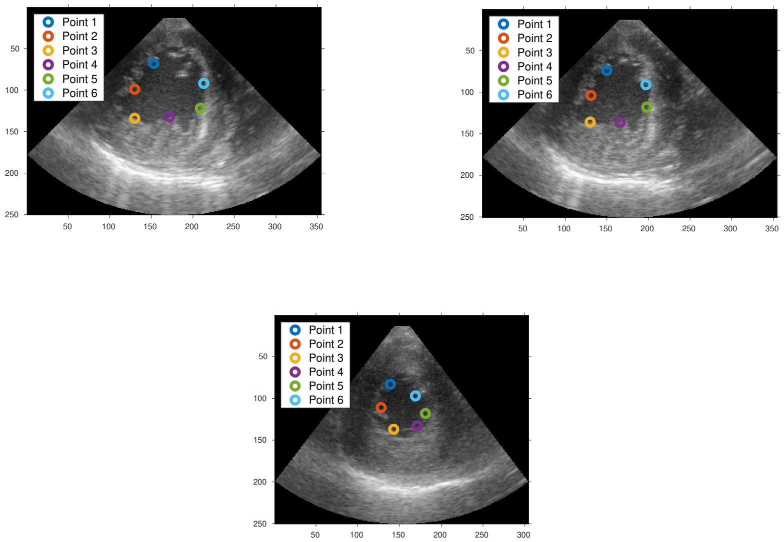


Figure 4.7: Initial tracking points for the PSAX images. Corresponding to sequence 1, 2 and 3

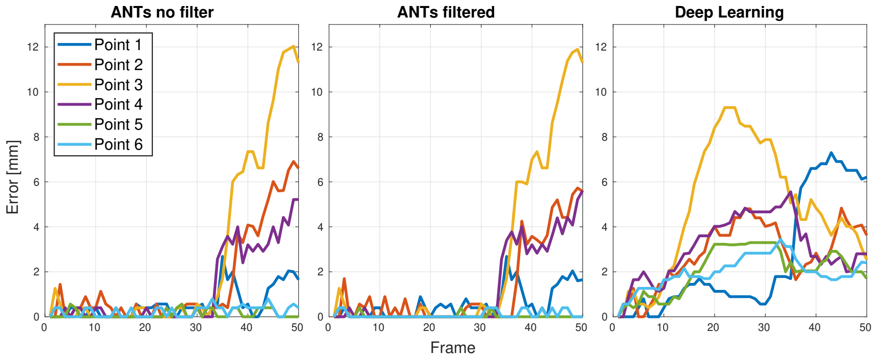


Figure 4.8: Sequence 1

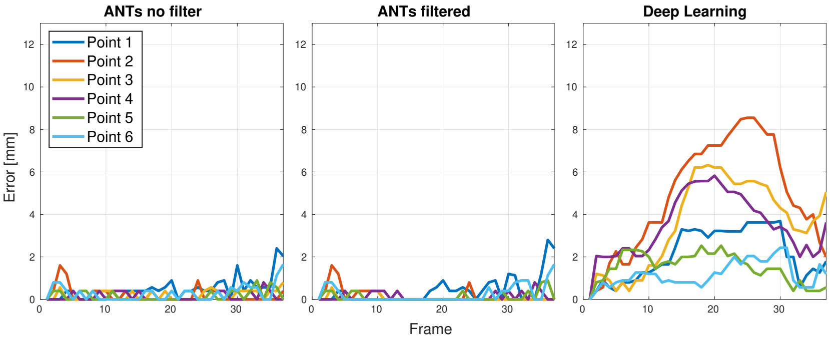


Figure 4.9: Sequence 2

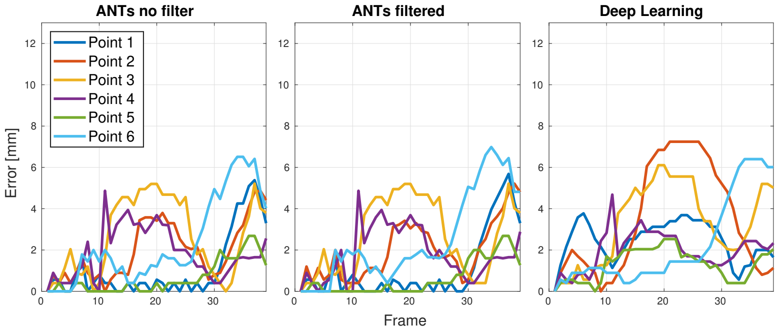


Figure 4.10: Sequence 3

Table 4.2: Error between estimated and manual reference points for the three PSAX recordings with three different image registration methods.

ANTs no filter	ANTs filtered	Deep Learning
1.1 ± 1.8 mm	1.0 ± 1.8 mm	2.7 ± 2.0 mm

4.3 Strain Estimation

Landmark tracking was performed on 2C and 4C test images with the three different image registration methods. Based on these positions, Lagrangian strain was calculated with Eq. (2.1). As each video is one ED-ED cycle, the time of ES is in this project defined when the length between the two points is the smallest in the cardiac cycle. The initial length and the smallest length are then used to calculate the strain for the given video and is then compared to the reference value strain. Prediction error is defined as

$$Prediction\ error = reference\ strain - predicted\ strain,$$

such that a positive error tells us that the absolute value of the predicted strain is bigger than the reference value because both values are negative.

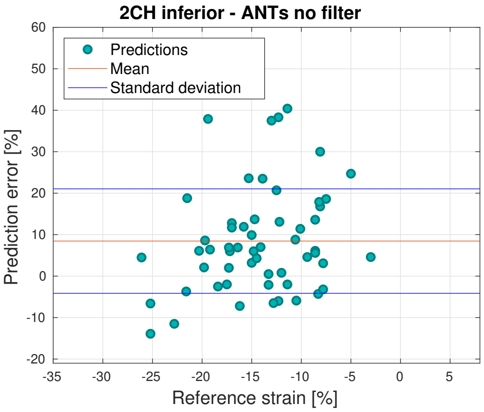


Figure 4.11: Mean error 8, 5 ± 12, 6%

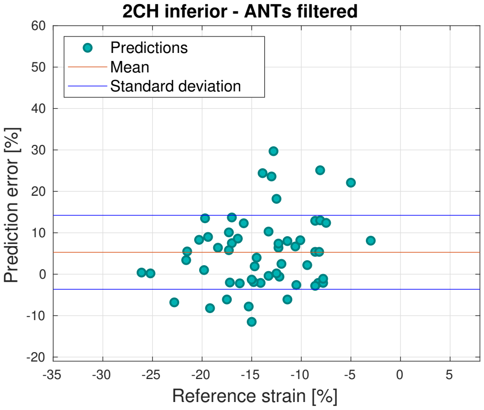


Figure 4.12: Mean error 5, 3 ± 8, 9%

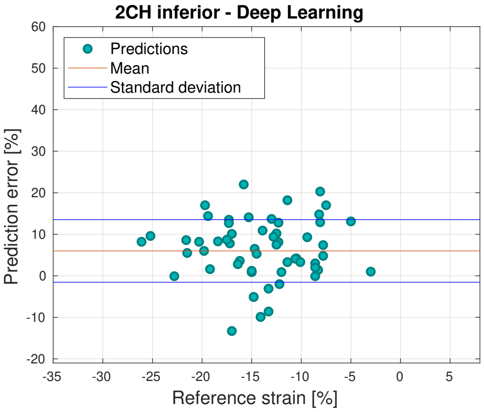


Figure 4.13: Mean error 6, 0 ± 7, 5%

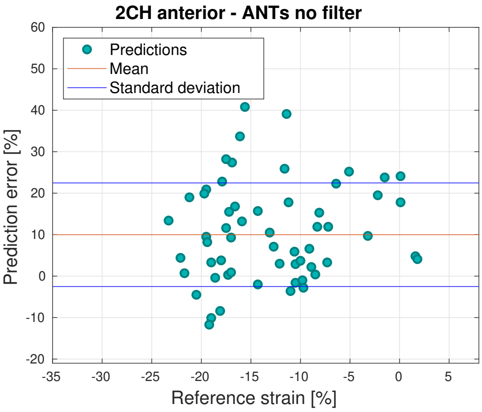


Figure 4.14: Mean error 10, 0 ± 13, 5%

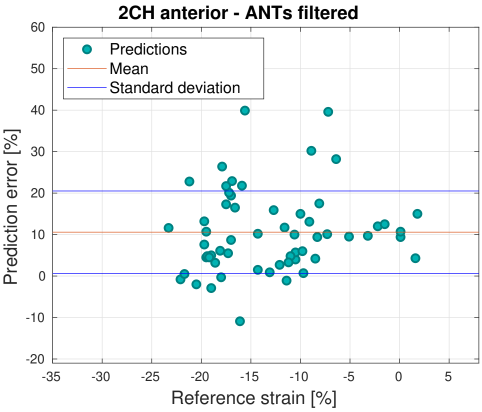


Figure 4.15: Mean error 10, 6 ± 9, 9%

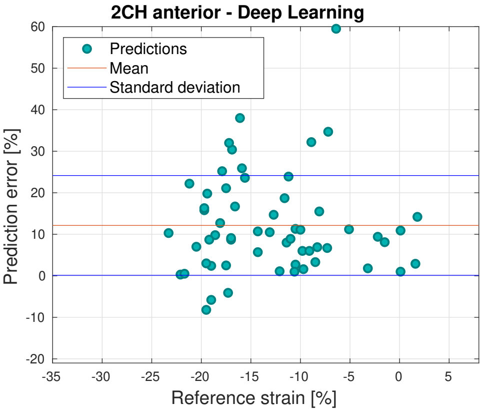


Figure 4.16: Mean error 12, 1 ± 12, 0%

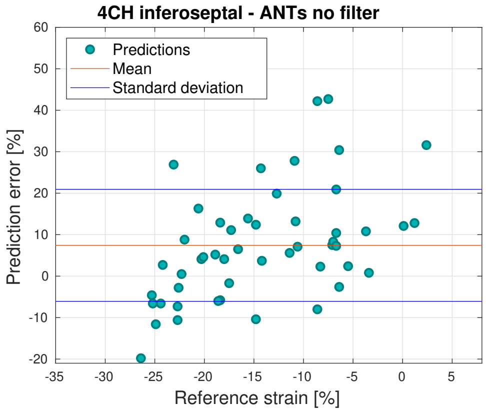


Figure 4.17: Mean error 7, 4 ± 13, 5%

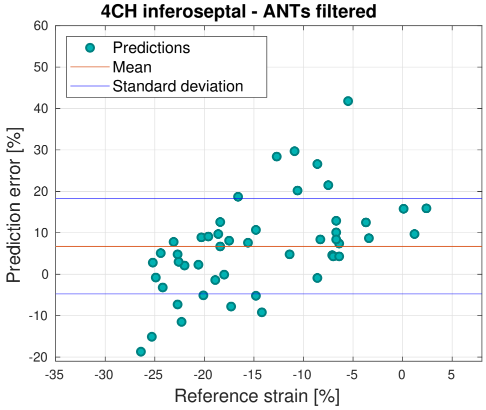


Figure 4.18: Mean error 6, 7 ± 11, 5%

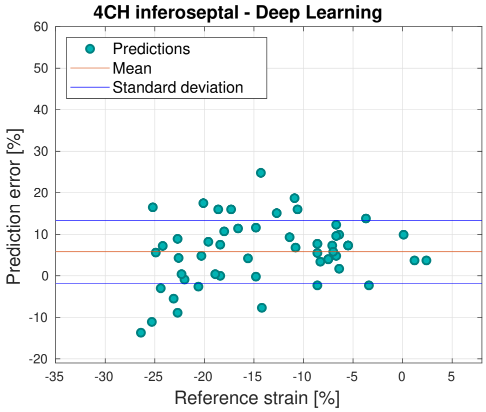


Figure 4.19: Mean error 5, 8 ± 7, 6%

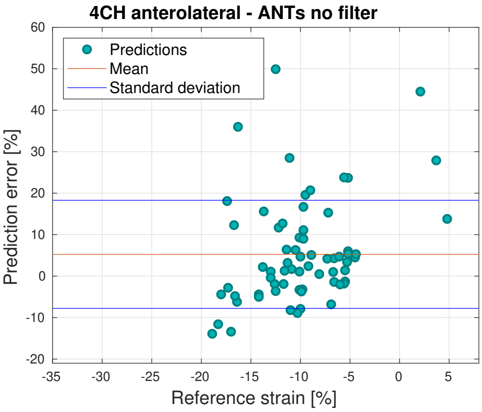


Figure 4.20: Mean error 5, 2 ± 13, 0%

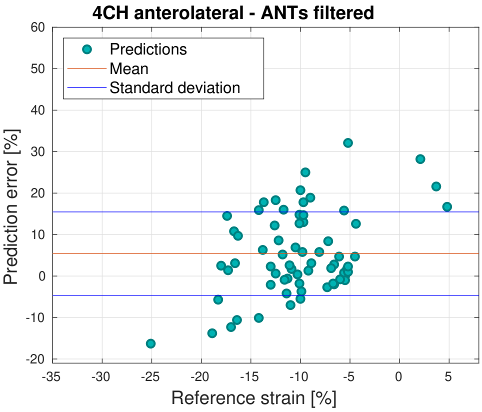


Figure 4.21: Mean error 5, 4 ± 10, 0%

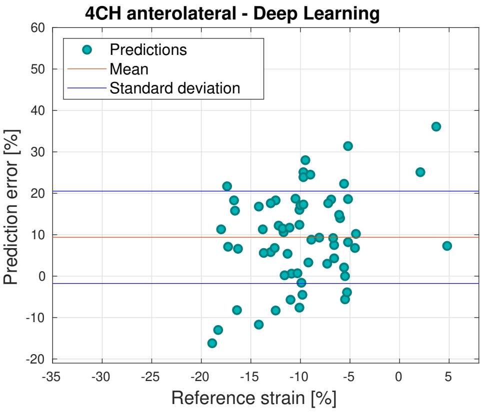


Figure 4.22: Mean error 9, 4 ± 11, 0%

All three methods and their estimation error for 2C and 4C strain estimation are summarized in Table 4.3.

	2 chamber		4 chamber	
	Inferior	Anterior	Infero-septal	Antero-lateral
ANTs no filter	$8,5 \pm 12,6\%$	$10,0 \pm 13,5\%$	$7,4 \pm 13,5\%$	$5,2 \pm 13,0\%$
ANTs filtered	$5,3 \pm 8,9\%$	$10,6 \pm 9,9\%$	$6,7 \pm 11,5\%$	$5,4 \pm 10,0\%$
Deep Learning	$6,0 \pm 7,5\%$	$12,1 \pm 12,0\%$	$5,8 \pm 7,6\%$	$9,4 \pm 11,0\%$

Table 4.3: Error between estimated strain in the different views with different methods.

4.4 Time Performance

A comparison of the time needed to estimate the new position of the landmarks is done on a Intel® Core™ i7-8700 CPU @ 3.20GHz × 12. The time for each method per frame is presented in Table 4.4, and the test is done on the same recording for both methods. The time needed to convert the image into ANTsImage in the ANTs method is not taken into consideration, only the image registration and applying the transform to the tracking points.

Table 4.4: Table of how long the different methods take in order to compute the position of the tracking points for one frame in a ultrasound sequence.

ANTs	**Deep Learning**
14.51 ± 3.95 seconds	0.12 ± 0.011 seconds

Chapter 5

Discussion

5.1 Training of the Deep Learning Model

With fixed hyperparameter values based on previous results, the average mean normalized cross-correlation was $-0,982 \pm 0,0021$. The different folds that were allocated to the validation set in turn, contained random samples from the available training data. There was not much difference between them, as they all contained an equal amount of 2C and 4C sequences. Further tests with different hyperparameter values could be tested in order to see if this would improve the results. In addition, a different number of downsampling layers could be tested, but it takes about 24 hours to train the network with a daisy-chain setup on a Tesla K80 GPU server at Floydhub, so a trial-and-error approach is not recommended for finding the hyperparameter values.

5.2 Landmark Tracking

5.2.1 Renal Results

Points in both the renal cortex and the veins area are tested for landmark tracking through a respiratory cycle. For patient 1, the tracking points were only marked at the cortex, and in Figure 4.3, it can be seen that the error is larger than in the rest of the recordings, which only had renal veins area tracking points. This is a result of the points drifting along the cortex because normalized cross-correlation cannot distinguish between the different points along the line. It will only take into account the similarity between images. The remaining recordings in Figure 4.4-Figure 4.6, have a smaller error for all three methods, except for the predictions made by the Deep learning model for patient 4. In this recording, the model is not able to keep up with the fast movement of the kidney towards the end of the respiratory cycle. For constant movements, all three methods are able to track points in the renal veins area, with the ANTs method with filtered frames estimating closest to the reference position.

5.2.2 Cardiac Results

In the PSAX recordings, some points are affected by the valve leaflets which move out and in again. This results in the tracking point being at the wrong place when the leaflets are back towards the myocardium. An example of this can be seen in Figure 4.8, where the ANTs predictions are completely off after halfway in the cycle. The error for each point is lower in the deep learning method, but the average error is more prominent than for the ANTs. ANTs with filtered frames performed the best also for the PSAX recordings.

The deep learning method is trained on 2C and 4C sequences, and not PSAX. Using deep learning on PSAX sequences is suboptimal. After training, the model's parameters are tweaked and fine-tuned for estimations in 2C and 4C views. A new deep learning model should be trained on PSAX before it can be used for the same view. This would have yielded better results than the one estimated within this project because the new models are tuned for PSAX sequences.

5.3 Strain Estimation

There was no data selection in this project, meaning all recordings were used. In some cases, parts of the myocardium were missing or not completely visible. Strain estimation is then infeasible, like in Figure 5.1. Some recordings also had unclear boundaries between the ventricle and myocardium or noise even after filtering.

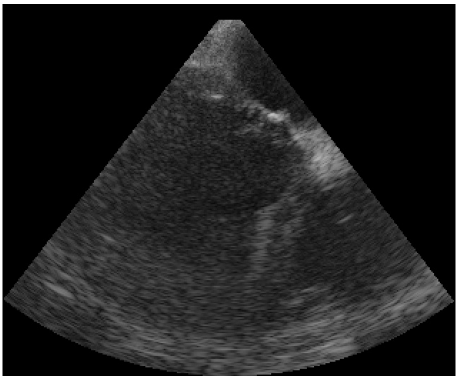


Figure 5.1: The left myocardium segment is not visible, and calculation of the strain is infeasible.

For this project, no method of annotating ES in the cardiac cycle was available such as ECG data. Thus, the smallest distance between the two points in a segment defined the time of ES in this project. For most cases, this worked well. However, if the tracking is not working properly, and the estimated position of the points are drifting towards each other for the whole cycle, the distance gets too small. An example of a sequence where this happened, is illustrated in Figure 5.2. Using Equation (2.1) for calculating he Lagrangian strain, the strain is sensitive to small errors in the distance between the two points, especially when the distance L is small. This resulted in errors around 40% for several cases. In addition, an error of 49.9% in 4CH anterolateral for ANTs no filter and 59.5% in 2CH anterior with deep learning.

It is also worth mentioning that the data annotation on the initial frame is done by a non-clinician with no previous experience in echocardiography or cardiac assessment. Thus, the initial tracking points may not be placed completely at the requested positions.

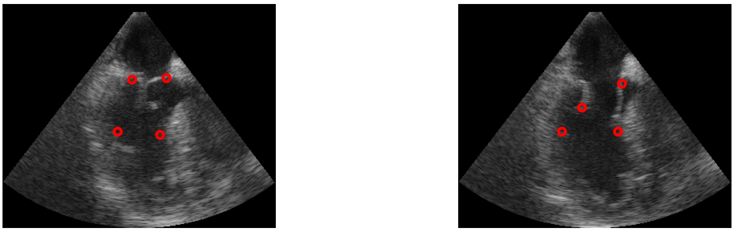


Figure 5.2: The image to the left shows the initial frame with tracking points, while the right shows the position for the points at ES. The distance between the two estimated points on the left segment, is closer than the real distance, because the left upper point follow the leaflet during the cardiac cycle

5.4 Time Performance

The deep learning method is significantly faster in estimating the new position of the tracking points than the ANTs method. A calculation time of $0,12$ seconds for each frame will make the Deep Learning method almost fast enough to use it in real-time, depending on the frame rate for the given ultrasound sequence. The training of the deep learning models takes a lot of time, but when it is finished training, the model can estimate points in under a second. The ANTs method, on the other hand, is too slow in order to be used in real-time. ANTs is a classic image registration method, in which the computation takes time. In this time performance test, the file format conversion is not taken into account, and the total time needed to use this model is larger than the estimated $14,51$ seconds, as shown in Table 4.4.

5.5 Future Work

The results presented in Chapter 4 shows that the model of estimating strain with a deep learning model is feasible, but the results can be improved. In order to enhance the results, more data could be added. The data set used is limited, so augmentation of the data set could be useful. Like most deep learning approaches, this framework requires a large set of training data. The amount of available training data is an important element in the proposed pipeline for estimating strain.

Another limitation of the proposed pipeline, is that the position of the initial landmark points are assumed to be known. In this project, this is done manually before the tracking started. In order to be able to actually estimate longitudinal strain in TEE images, a trained physician or a method to place the points should be utilized. A method for annotating ES correctly should also be used, in order to find the correct distance that should be included in the calculation of strain.

A supervised learning method with the ground truth of what the position of the tracking points in the training set could also improve the results. Such method will require manual labelling of a decent amount of data, but in return, a precise predicting of the kidneys position in the recordings could be achieved.

A specific PSAX deep learning model should be developed in order to use on PSAX images, such that each task has a designated model trained on the specific problem.

Chapter 6

Conclusion

In this project, the feasibility of tissue deformation estimation is demonstrated. The method uses an adapted version of the Deep Learning Framework for Unsupervised Affine and Deformable Image Registrated presented by de Vos et al. [18] and implementation by Haukom in Basal Strain Estimation in Transesophageal Echocardiography using Unsupervised Deep Learning [19]. Advanced Normalization Tools (ANTs) is another approach for estimating the position and is also covered in this thesis.

Ultrasound sequences were filtered using a bilateral filter and a non-linear correction. From previous projects, a deep learning model with a multi-stage approach was trained on 2C and 4C cardiac data, and 5-fold cross-validation was performed. In addition, ANTs with symmetric normalization based on affine and deformable transformation with cross-correlation as optimization metric was utilized to estimate deformations in the tissue. The results for renal data and PSAX recordings were compared to manual reference points, while estimated strain in 2C and 4C sequences was compared to reference values obtained by an expert echocardiographer.

In conclusion, calculation of the strain is feasible for most cases where the segment of the myocardium is visible, and the quality of the ultrasound sequence is sufficient. However, further improvements are recommended to increase the precision of locating the points.

References

[1] Sandsbraaten, I., “Kidney and renal artery movement detection and tracking in ultrasound images using deep learning,” Dec. 2019.

[2] Weiser, T., Haynes, A., Molina, G., Lipsitz, S., Esquivel, M., Uribe-Leitz, T., Fu, R., Azad, T., Chao, T., Berry, W. and Gawande, A., “Size and distribution of the global volume of surgery in 2012,” *Bulletin of the World Health Organization*, vol. 94, 201–209F, Mar. 2016.

[3] Warltier, D.C., Laffey, J.G., Boylan, J.F.. and Cheng, D.C., “The Systemic Inflammatory Response to Cardiac Surgery: Implications for the Anesthesiologist,” *Anesthesiology: The Journal of the American Society of Anesthesiologists*, vol. 97, no. 1, pp. 215–252, Jul. 2002. eprint: `jasa/content\_public/journal/jasa/931215/0000542-200207000-00030.pdf`.

[4] Zaunseder, S., Riedl, M., Kurths, J., Malberg, H., Nauernschmitt, R. and Wessel, N., “Impact of cardiac surgery on the autonomic cardiovascular function,” *Journal of Computational Surgery*, vol. 1, Apr. 2014.

[5] Vincent, J.-L-, Pelosi, P., Pearse, R., Payen, D., Perel, A., Hoeft, A., Romagnoli, S., Ranieri, V., Ichai, C., Forget, P., Rocca, G. and Rhodes, A., “Perioperative cardiovascular monitoring of high-risk patients: A consensus of 12,” *Critical Care*, vol. 19, May 2015.

[6] M. H. Rosner and M. D. Okusa, “Acute kidney injury associated with cardiac surgery,” *Clinical journal of the American Society of Nephrology*, vol. 1, pp. 19–32, 1 2006.

[7] M. H. Rosner, “Acute kidney injury associated with cardiac surgery,” *Redo Cardiac Surgery in Adults*, pp. 37–52, 2012.

[8] Thakar, C. V., Yared, J.-P., Worley, S., Cotman, K., and Paganini, E. P., “Renal dysfunction and serious infections after open-heart surgery,” *Kidney international*, vol. 64, pp. 239–246, 1 2003.

[9] Hilberman, M., Myers, B., Carrie, B., Derby, G., Jamison, R., and Stinson, E., “Acute renal failure following cardiac surgery,” *The Journal of thoracic and cardiovascular surgery*, vol. 77, pp. 880–888, 6 1979.

[10] Zanardo, G., Michielon, P., Paccagnella, A., Rosi, P., Caló, M., Salandin, V., Da Ros, A., Michieletto, F., and Simini, G., “Acute renal failure in the patient undergoing cardiac operation,” *The Journal of thoracic and cardiovascular surgery*, vol. 107, pp. 1489–1495, 6 1994.

[11] Mangano, C. M., Diamondstone, L. S., Ramsay, J. G., Aggarwal, A., Herskowitz, A., and Mangano, D. T., “Renal dysfunction after myocardial revascularization: Risk factors, adverse outcomes, and hospital resource utilization,” *Annals of internal medicine*, vol. 128, pp. 194–203, 3 1998.

[12] Goodfellow, I., Bengio, Y. and Courville A., *Deep Learning*. MIT Press, 2016, `http://www.deeplearningbook.org`.

[13] Sance, R., Ledesma-Carbayo, M., Lundervold, A. and Santos, A., “Image registration for quantitative analysis of kidney function using mri,” *AIP Conference Proceedings*, vol. 860, Jan. 2006.

[14] Heimdal, A., Støylen, A., Torp, H. and Skjærpe, T., “Real-time strain rate imaging of the left ventricle by ultrasound,” pp. 1013–1019, 11 Nov. 1998.

[15] Reisner, S., Lysyansky, P., Agmon, Y., Mutlak, D., Lessick, J. and Friedman, Z., "Global longitudinal strain: A novel index of left ventricular systolic function," *Journal of the American Society of Echocardiography : official publication of the American Society of Echocardiography*, vol. 17, pp. 630–3, Jun. 2004.

[16] Blessberger, H. and Binder, T., "Two dimensional speckle tracking echocardiography: Basic principles," *Heart*, vol. 96, no. 9, pp. 716–722, 2010. eprint: `https://heart.bmj.com/content/96/9/716.full.pdf`.

[17] Goldstein, St.A., Kronzon, I., Khandheria, B.K. and Mor-Avi, V., *ASE's Comprehensive Echocardiography*, 2nd ed. Elsevier, 2016.

[18] De Vos, B., Berendsen, F., Viergever, M., Sokooti, H., Staring, M. and Išgum, I., "A deep learning framework for unsupervised affine and deformable image registration," *Medical Image Analysis*, vol. 52, Dec. 2018.

[19] Haukom, T., "Basal strain estimation in transesophageal echocardiography using unsupervised deep learning," Jun. 2019.

[20] Field, M.J., Harris, D.C. and Pollock, C.A., *The Renal system.* 2010, vol. 2, pp. 57–67.

[21] Pappano, A.J and Wier, W.G., *Cardiovascular Physiology.* 2013, vol. 10.

[22] Hansen, E., Sjaastad, Ø.V. and Sand, O., *Menneskets fysiologi.* Gyldendal akademisk, 2014.

[23] Conger, J. D., Schultz, M. F., Miller, F., and Robinette, J. B., "Responses to hemorrhagic arterial pressure reduction in different ischemic renal failure models," *Kidney international*, vol. 46, pp. 318–323, 1994.

[24] Verbeke, M., Smöllich, B., van de Voorde, J., de Ridder, L., and Lameire, N., "Beneficial influence of ketanserin on autoregulation of blood flow in post-

ischemic kidneys," *Journal of the American Society of Nephrology : JASN*, vol. 7, pp. 621–627, 1996.

[25] Basile, D. P., Anderson, M. D. and Sutton, T. A., "Pathophysiology of acute kidney injury," *Comprehensive Physiology*, vol. 2, pp. 1303–1353, 2 2012.

[26] Biga, L.M., Dawson, S., Harwell, A., Hopkins, R., Kaufman, J., LeMaster, M., Matern, P., Morrison-Graham, K., Quick, D. and Runyeon, J., *Anatomy Physiology*. Open Oregon State, Oregon State University, 2019, pp. 1840–1850.

[27] J. Gorcsan and H. Tanaka, "Echocardiographic assessment of myocardial strain," *Journal of the American College of Cardiology*, vol. 58, no. 14, pp. 1401–1413, 2011. eprint: `https://www.onlinejacc.org/content/58/14/1401.full.pdf`.

[28] A. Støylen, "Strain rate imaging of the left ventricle by ultrasound. feasibility, clinical validation and physiological aspects.," 2001.

[29] Shampo, M. A. and Kyle, R. A., "Karl theodore dussik - pioneer in ultrasound," *Mayo Clinic Proceedings*, vol. 70, 12 1995.

[30] Chan, V. and Perlas, A., "Basics of ultrasound imaging," *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management*, pp. 13–19, 2011.

[31] Hoskins, P., Martin, K. and Thrush, A., *Diagnostic Ultrasound Physics and Equipment*. Cambridge University Press, 2010.

[32] Hahn, R. T. et al., "Guidelines for performing a comprehensive transesophageal echocardiographic examination: Recommendations from the american society of echocardiography and the society of cardiovascular anesthesiologists," *Journal of the American Society of Echocardiography*, vol. 2, pp. 921–964, 2013.

[33] M. e. a. Cerqueira, "American heart association writing group on myocardial segmentation and registration for cardiac imaging : Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart : A state-

ment for healthcare professionals from the cardiac imaging committee of the council on clinical cardiology of the american heart association," *Circulation*, vol. 105, pp. 539–542, 2002.

[34] Brennan, P., Standring, S. and Wiseman, S., *Gray's Surgical Anatomy E-Book*. Elsevier Health Sciences, 2019.

[35] Tomasi, C and Manduchi, R., "Bilateral filtering for gray and color images," pp. 839–846, 1998.

[36] Avants, B.B., `http://stnava.github.io/ANTs/`, accessed May 18, 2020.

NTNU
Kunnskap for en bedre verden