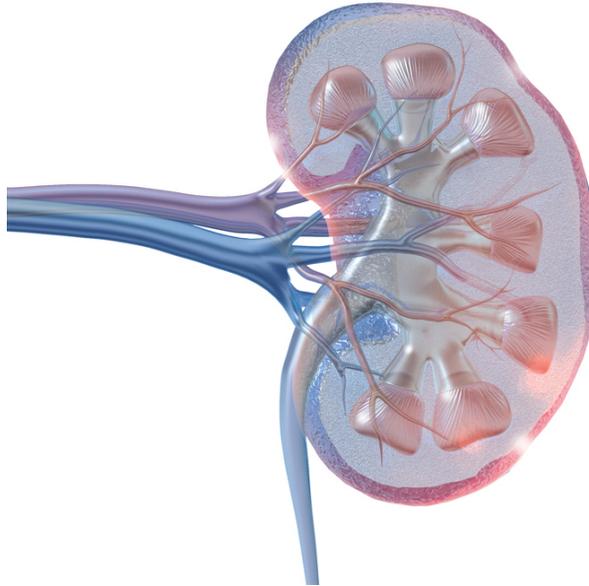




**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

---



# Investigation of methods for intensity based kidney registration with MRI

Master of Science Thesis in Electrical Engineering

Lovisa Ekberg

Jackie Johansson



MASTER'S THESIS 2022

# Investigation of methods for intensity based kidney registration with MRI

Lovisa Ekberg  
Jackie Johansson



**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering  
CHALMERS UNIVERSITY OF TECHNOLOGY  
Gothenburg, Sweden 2022

Investigation of methods for intensity based kidney registration with MRI

Lovisa Ekberg  
Jackie Johansson

© LOVISA EKBERG, 2022.  
© JACKIE JOHANSSON, 2022.

Supervisor: Simon Ekström, Antaros Medical  
Supervisor: Bettina Selig, Antaros Medical  
Supervisor: Torbjörn Lundh, Department of Mathematical Sciences  
Examiner: Torbjörn Lundh, Department of Mathematical Sciences

Master thesis 2022  
Department of Electrical Engineering  
Chalmers University of Technology  
SE-412 96 Gothenburg

Typeset in LaTeX

Image on front page show an illustration of a kidney. The image is showing the kidneys internal structures as well as contours. The images is borrowed from Britannica image Quest [1].

## Abstract

**Background:** The kidneys are an important part of the body and an absence of the functions a kidney provide is be lifethretening. To be able to provide correct treatment, data needs to be collected. A safe and efficient way of collecting this data is the use of MRI which can be used to asses both function, structure and pathophysiological changes. But to be able to access this data, images need to be collected over different time instances. This results in a issue: the movement of the kidneys due to breathing and heartbeats introduced during this time period of scanning complicates the analysis of the data. To compensate for these motions, image registration can be used. This is the process of aligning the images so that they have a pixel to pixel correspondance.

**Aim:** The purpose of this study was to assist in the search for an automated registration method that can compensate for motions in kidney MRI images.

**Method:** The method used in this project is literature study and performance study. The literature study was used to get a broad understanding of the topic and to gather information about promising techniques. These techniques were then implemented in the performance study. The main parts investigated was: i) the registration process, ii) choice of transform, iii) choice of similarity measure.

**Results and conclusion:** Kidney registration is a complex task including many possible combinations of settings. Mutual information (MI) outperformed Cross-correlation (CC) as a similarity measure. When it comes to different combinations of registration processes and transformation models, the result does not show a clear optimal combination. Thus, this study urges the complexity of selecting registration settings and there are many areas where further research is needed.

Key words: “Registration”, “Segmentation”, “Kidneys”, “MRI”, “Imaging”, “Mutual Information”, “Groupwise registration”

## Acknowledgments

We would like to thank Antaros Medical, for having us as interns and your very warm welcome. Thanks for giving us an interesting topic to investigate and supporting us throughout this masters thesis. A special thanks to Simon Ekström and Bettina Selig for being our supervisors, giving us support and guidance during the whole process. Many thanks also to Torbjörn Lundh, for your support and new perspective on the topic.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Aim . . . . .	2
1.2	Scope . . . . .	2
<b>2</b>	<b>Background of Image Registration</b>	<b>3</b>
2.1	Geometric transformation models . . . . .	4
2.1.1	Rigid Transformation model . . . . .	5
2.1.2	Similarity Transformation model . . . . .	5
2.1.3	Affine Transformation model . . . . .	6
2.1.4	B-spline Transformation model . . . . .	7
2.2	Similarity metrics . . . . .	9
2.2.1	Sum of squared differences (SSD) . . . . .	10
2.2.2	Mutual information (MI) . . . . .	10
2.2.3	Intensity cross-correlation (CC) . . . . .	12
2.2.4	Principal component analysis (PCA) . . . . .	12
2.3	Optimisation . . . . .	15
2.3.1	Gradient Descent Optimisation . . . . .	15
2.3.2	Stochastic Gradient Descent Optimisation . . . . .	16
2.3.3	Pyramids and Sampling . . . . .	16
2.4	Registration process for multiple images . . . . .	17
2.5	Validation of Registration Method . . . . .	18

2.5.1	Spatial Overlapping Metrics . . . . .	18
2.5.2	Distance Measures metrics . . . . .	19
2.5.3	Volumetric metrics . . . . .	20
<b>3</b>	<b>Methodology</b>	<b>20</b>
3.1	Data used in study . . . . .	21
3.2	Image Pre-processing . . . . .	22
3.3	Performance study . . . . .	23
3.3.1	Selection of reference image . . . . .	24
3.3.2	Similarity metrics to investigate . . . . .	24
3.3.3	Geometric transformation models to investigate . . . . .	24
3.3.4	Optimisation method to investigate . . . . .	25
3.3.5	Selection of registration process and transformation model . . . . .	25
3.3.6	Validation of methods . . . . .	26
<b>4</b>	<b>Results</b>	<b>27</b>
4.1	Selection of reference image . . . . .	27
4.2	Investigation of different similarity metrics . . . . .	28
4.3	Investigation of B-spline grid spacing . . . . .	31
4.4	Selection of registration process and transformation model . . . . .	33
<b>5</b>	<b>Discussion</b>	<b>36</b>
5.1	Selection of reference image . . . . .	36
5.2	Investigation of different similarity metrics . . . . .	37

5.3	Investigation of B-spline grid spacing . . . . .	37
5.4	Selection of registration process and transformation model . . . . .	38
5.5	Validation of methods . . . . .	39
5.6	Suggestions for future work . . . . .	40
<b>6</b>	<b>Conclusion</b>	<b>41</b>

# 1 Introduction

The kidneys play an essential role in the function of the body. The primary purpose is to filter the blood to remove waste and water but also to maintain a balance of salts and minerals. The filtered excesses are either reabsorbed by the blood or excreted from the body with the urine. Consequently, the absence of these functions leads to life-threatening conditions [2]. Kidney disease is a global problem where one out of ten people worldwide suffer from chronic kidney disease (CKD) [3]. The causes of CDK are several, and treatments for prevention and control of the disease are essential.

MR imaging biomarkers have great potential in the clinical constellation of CKD. This method allows assessment of the renal structure, function, and pathophysiological changes, making MRI a good substitute for biopsy. There are benefits in avoiding invasive measurement methods since they can imply complications due to the great amount of blood flowing through the kidneys. Therefore, MRI is not only offering accurate diagnostic and prognostic tools but also a safer alternative [4].

In conventional MRI, a qualitative MR image is acquired by using specific values of the imaging parameters, for example, time of repetition, flip angle and echo time [5]. The intensities in the resulting image depend on these parameter values and may vary between MRI scanners and instants. As a result, the pixel values can not be compared between different data sets [6]. In contrast to qualitative MRI, quantitative MRI is a method aiming to provide measures in physical units. This method requires multiple images acquired at different time instances during one scanning [5]. This process results in images independent of operator-specific imaging parameters so that they can be compared between scans and different data sets [5]. Another feature of quantitative MRI is the extra tissue characterisation information that is provided with the technique [5]. In this project, the data used is quantitative MR images.

When capturing images of the kidneys over time, it takes up to a few minutes resulting in several images with different information and contrast [7]. During the imaging acquisition, movement of the body caused by, for example, breathing or the heart pulse leads to displacement and distortion of the kidneys between the images [8]. These movements in addition with the different contrasts between images hinder the analysis of the kidneys. For example, to quantify the MR images, the movements need to be compensated for, and the kidneys in the images need almost a pixel-to-pixel correspondence [8].

Examples of images acquired with different imaging parameters are modified look locker inversion recovery (MOLLI) and diffusion weighted imaging (DWI) images. MOLLI images are acquired with respect to the subjects heartbeats [9]. The result is images with less motion caused by the cardiac pulse [9]. Multiple MOLLI images can be acquired during one scanning, resulting in images with different inversion times [10]. On the other hand, DWI

images show information about the diffusion of water in cells and add a functional view to the MRI images [11]. The DWI images are attenuated with different b-values. The b-value is a factor dependent on the magnitude, duration and time interval of the gradients used for the MRI [12]. For example, a more substantial diffusion effect is represented by a higher b-value [12].

Regardless of the method to acquire MR images of kidneys, the miss-alignment is still a problem that needs to be solved for a more effective analysis. Image registration is a way of compensating for the movements induced during imaging by aligning the images taken at different time instances [7]. Today, registration of medical images is mainly achieved manually, which is an expensive and time-consuming process done by experts. With this project, the goal is to investigate methods for automated registration of kidneys in MR images.

## 1.1 Aim

The project aims to assist in the search for an automated registration method that can compensate for motions in kidney MR images. The work first focuses on literature and previous studies to identify promising intensity-based registration methods. Then, the focus is on implementing selected methods to evaluate and compare their performance on the data used in this project. Key questions that the thesis attempts to answer include:

- What are the current promising registration techniques used for kidney MR images?
- What element within the registration process will be prioritised to investigate?
- How does different options affect intra-modality registration of MOLLI and DWI images individually, as well as inter-modality registration of MOLLI and DWI images together?

The project is in collaboration with Antaros Medical and is intended to assist in their process of finding an automated method that eventually can be used for clinical studies of MR images of kidneys.

## 1.2 Scope

There are several things limiting the project. The images used in this study are Modified look-locker inversion recovery (MOLLI) and Diffusion-weighted imaging (DWI). The images are in 2D and the project is restricted to only look at one slice from the axial plane at a time. Furthermore, the datasets contains more MOLLI slices than DWI. The datasets of MRI

images are provided by Antaros Medical and not collected during the project. Therefore, the thesis does not regard the imaging acquisition, and no time is spent on improvements in reducing motions in this field. The project only investigate intensity based registration methods and do not consider feature based ones. All implementation is conducted using the image registration toolbox Elastix. Lastly, machine learning methods are not considered, and a time frame limits the project from January 2022 to June 2022.

## 2 Background of Image Registration

Image registration is a technique frequently used in medical image processing to find a spatial relationship between images and to align them. A flow chart of the image registration process can be seen in Figure 1. The underlying idea is to map an image  $I_M$ , here called the moving image, to a fixed image,  $I_F$ . Registration is an iterative process where  $I_M$  is initially transformed, resulting in  $I_{M'}$ . Then, a similarity metric measures how well  $I_{M'}$  aligns with the fixed image  $I_F$ . Lastly, the transformation parameters are updated and reapplied to  $I_M$  and the same process is carried out until a set of optimal parameters is found. It can be assumed that for the first iteration,  $I_M$  is transformed by a identity transformation. In conclusion, the goal of the registration is to find the transformation that maximises the similarities between the fixed and the moving image, alternatively minimising their dissimilarities. This process is illustrated in 1.

Another part of image registration is how to determine the quality of the registration methods and specific parameters. Due to a lack of ground truth images to compare the registered image to, this is challenging. One solution to this is using manually segmented images, created by experts. The segmentation of the fixed image  $I_F$  is then used as the gold standard to which the other transformed mask is compared against.

The image registration process is dependent on the specific application, rather than a universal method, since the diversity of images to register is so broad. [13]. Even in the area of medical images there are several decisions to make depending on the modality (MRI, CT, etc.), organ (kidney, brain, etc.) and dimensions (2D, 3D). The different parts of image registration in focus for this study can be seen in Figure 1 and in the following sections the choice for each of the parts are described.

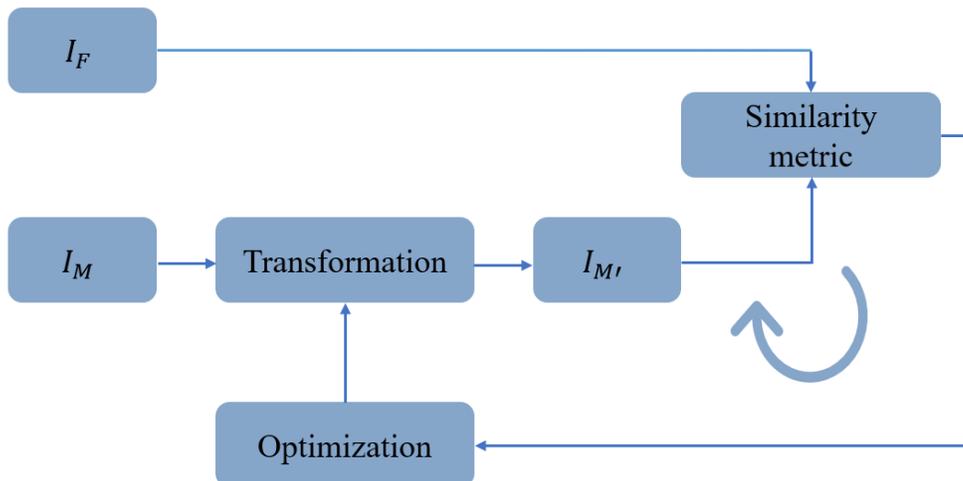


Figure 1: A flow chart of the optimisation process of image registration, with the steps in focus for this study. The goal is to align the moving image,  $I_M$ , to the fixed image,  $I_F$ . First,  $I_M$  is transformed, resulting in  $I_{M'}$ , which is then compared with  $I_F$  by a similarity metric. The transformation model is updated accordingly and reapplied to  $I_M$ , and the same process is carried out. It can be assumed that for the first iteration, the transform is a identity matrix.

## 2.1 Geometric transformation models

The choice of a transformation model in image registration is essential, as the model needs to represent the geometric differences between  $I_M$  and  $I_F$  [7]. These differences are introduced by movements during scanning [7]. For instance, structures like bone often undergo translation and rotation during image acquisition and can therefore be represented by a rigid model [14]. At the same time, more deformable organs might also experience shearing, scaling, or other deformations between images. Images of such organs require other transformation models, for instance, non-rigid ones [14]. Each transformation model contains a set of parameters depending on the specific model. Therefore, after deciding on a proper transformation model to use, a search for the specific model parameters that best align  $I_M$  and  $I_F$  starts. Example of parameters to determine are angle or translations between the moving and the fixed image.

The transformation models are in this paper divided into rigid and non-rigid transforms. The difference between the models is the degrees of freedom, which specify the deformations that can be recovered [15]. The following sections describe a rigid transformation, followed by three non-rigid ones: Similarity, Affine, and B-spline transform.

### 2.1.1 Rigid Transformation model

A common transformation model is a rigid transformation. This model is most often used for the registration of rigid structures [14]. Such a transformation model allows for rotation and translation, corresponding to three degrees of freedom in a two-dimensional image [15]. Seen in Figure 2 is an example of a rigidly transformed image. Recall that, mathematically the expression for a rigid transform is

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}. \quad (1)$$

where  $x$  and  $y$  are coordinates in the moving image and  $\tilde{x}$  and  $\tilde{y}$  are coordinates in the fixed image. The transformation parameters for this model are  $\varphi$ , the rotation angle and  $t_x$  and  $t_y$ , the translation parameters.

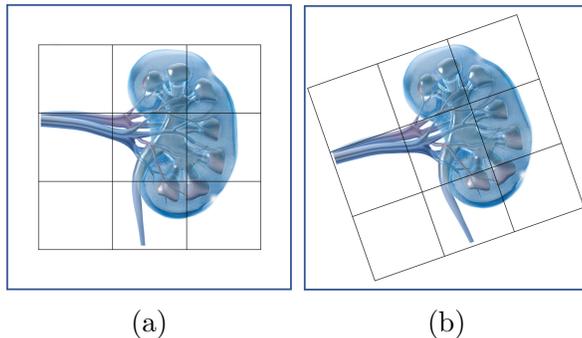


Figure 2: Example of a rigid transformation. (a) Original image before any registration is applied. (b) A rigidly transformed image, rotated counterclockwise.

### 2.1.2 Similarity Transformation model

Another transformation model is the Similarity transform. This transformation is similar to the rigid one. However, except for rotation and translation, it can also scale the images for better alignment [13]. As a result, the similarity transform allows four degrees of freedom in a two-dimensional image [15] and is classified as a non-rigid model [16]. An illustration of the transform is seen in Figure 3 and the formula is

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = s \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}. \quad (2)$$

The transformation parameters for a similarity transform are the same as the rigid one in Chapter 2.1.1. However, here a scaling parameter  $s$  has also been added.

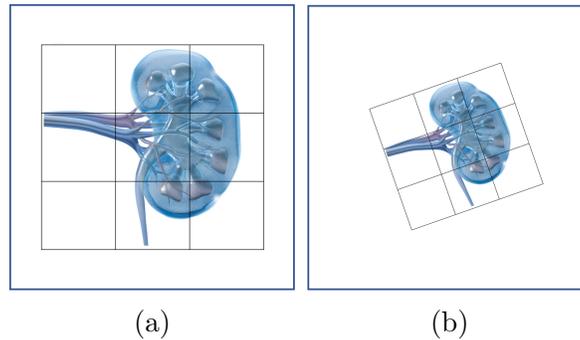


Figure 3: Example of a similarity transformation. (a) Original image before any registration is applied. (b) A similarity transformed image, rotated and scaled.

### 2.1.3 Affine Transformation model

An Affine transformation is another example of a non-rigid transformation model [16]. The Affine model is similar to both the rigid and similarity transform, and allows for rotation, translation, scaling and shearing of the image [17]. Owing to this, the degrees of freedom in an Affine transform and a two-dimensional image is six degrees. An example of an Affine transformation can be seen in Figure 4, and the equation for the model is

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \quad (3)$$

In the expression above,  $\tilde{x}$  and  $\tilde{y}$  are coordinates in the fixed image,  $x$  and  $y$  are coordinates in the moving image and  $a - f$  are constants [18].

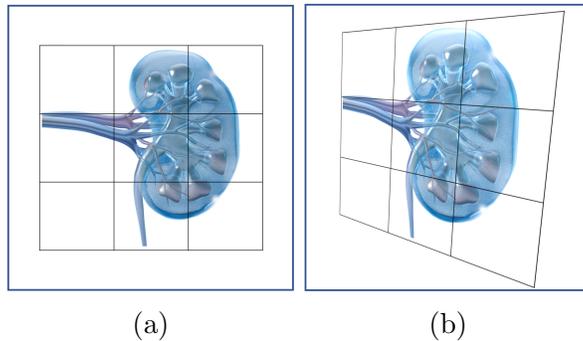


Figure 4: Example of an Affine transformation. (a) Original image before any registration is applied. (b) An Affine transformed image, showing a sheared kidney.

#### 2.1.4 B-spline Transformation model

The last category of transformation models is a cubic B-spline model. In a B-spline model, a uniform spaced grid overlay the image region, containing a control point at corner of the grid [19]. An example of a grid with corresponding grid knots can be seen in Figure 5. The idea behind a B-spline model is to construct a curve between a set of knots, estimating these control points. An illustration of a B-spline curve is seen in Figure 7.

The grid is divided into several smaller sections, reaching from one control point to another, between which curves are defined, known as segments [20]. Each segment can then be represented as a linear combination of basis functions. There are four basis functions used to define a cubic B-spline curve,  $B_n(t)$  for  $n = 0, 1, 2, 3$ , these are:

$$B_0(t) = (1 - t)^3/6, \quad (4)$$

$$B_1(t) = (3t^3 - 6t^2 + 4)/6, \quad (5)$$

$$B_2(t) = (-3t^3 + 3t^2 + 3t + 1)/6, \quad (6)$$

$$B_3(t) = t^3/6. \quad (7)$$

A final curve can then be constructed by attaching adjacent segments at the knots. The attachment is done by fulfilling a constraint that the curve should have  $C^2$  continuity [19]. In other words, the segments should share the same second derivative at each of the knots. This results in a smooth curve, not always passing through all of the control points. A weighted sum of these basis functions, placed on the grid, can then be used to model the total transformation [15]. The approximation of the transformation  $f(x, y)$  can be calculated by

$$f(x, y) = \sum_{k=0}^3 \sum_{l=0}^3 B_k(s) B_l(t) \phi_{(i+k)(j+l)}. \quad (8)$$

Whereas  $i = \lfloor x/n_x \rfloor - 1$ ,  $j = \lfloor y/n_y \rfloor - 1$ ,  $s = x/n_x - \lfloor x/n_x \rfloor$  and  $t = y/n_y - \lfloor y/n_y \rfloor$ , and  $n_x$  and  $n_y$  represents the mesh containing the control points [19].

The spacing between the control points specifies the flexibility of the deformation [15]. For example, an increasing number of grid points results in a transformation model with increased degrees of freedom [16]. A smaller grid spacing can increase the accuracy of the transform, but then it also increases the computational time [19]. One way to compromise between the flexibility and time is to use a hierarchical approach of grid sizes, starting with a large grid that decreases to a final resolution [19]. In addition, a B-spline transform also has local support which implies that moving one knot will only affect a part and not the whole curve [20]. This is also one of the reasons for the flexibility of the transformation.

Even though a B-spline can align images both globally and locally, pre-alignment of the images is often required [16]. The pre-alignment is often performed by applying a Rigid or Affine transformation before the B-spline model is used.

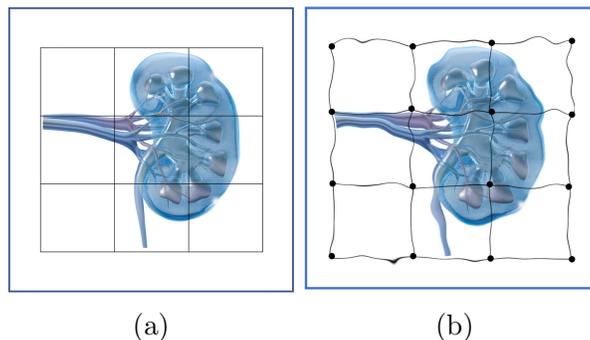


Figure 5: Example of a B-spline transformation. (a) Original image before any registration is applied. (b) An example of a grid overlaying the kidney with control points in each knot.

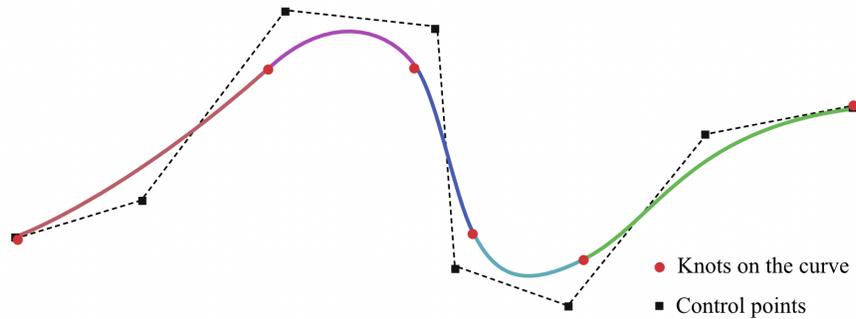


Figure 6

Figure 7: Illustration of a B-spline curve. Each colour representing one segment.

## 2.2 Similarity metrics

Similarity metrics, also known as similarity objective functions or cost functions, measure how well the transformed moving image,  $I_{M'}$ , resemble the fixed image,  $I_F$ . This step is a crucial part of the registration process, yielding the measure of what the iterative process is to optimise. There exist several types of similarity metrics [21] [16], and the choice of which to use is based on the object to be registered and on what modality and transformation method to use [16]. Similarity metrics are commonly divided into two categories; feature-based and intensity based.

Feature based registration is based on the geometrical structures extracted from the images [22], such as edges, contours, line intersections and corners [13]. The similarity metric for a feature based method are often the measured distance between corresponding features between the moving and fixed images [16]. This method requires a previous step; automated feature detection [13], or manually defined landmarks. MR images of kidneys usually lack distinctive features, making this kind of image challenging to extract features from automatically. Even landmarks, which are placed manually, faces many difficulties [23]. Amongst other problems, this placement can be time-consuming and difficult to set correctly [23].

Intensity based registration refers to the scalar values in pixel or voxel intensity and aims to maximise the similarities between two images [16]. Due to the difficulties found in feature based registration of kidney images, intensity based is the most common approach [21] [16]. For the same reason, intensity based metrics were investigated in this paper.

### 2.2.1 Sum of squared differences (SSD)

The sum of squared differences (SSD) is an intensity difference measurement. The measurement is based on the assumption that corresponding structures in the two images have the same intensities. The underlying idea is to compare the pixel intensities between the moving and the fixed image by subtraction.

Given images  $I_F$  and  $I'_M$  with pixels  $i$ , the definition of SSD is given by

$$\text{SSD} = \frac{1}{N} \sum_i^N |I_F(i) - I_{M'}(i)|^2 \quad \forall i \in I_F \cap I_{M'}. \quad (9)$$

The value of the SSD decreases as the alignment of the images is improved. For instance, when the SSD is zero, the alignment between the images is perfect. Due to the assumption that the images have equal intensity, the use of SSD is limited to the registration of images of the same modality and matching intensities [21]. However, matching intensities between images is not the case for the ones concerned in this project.

### 2.2.2 Mutual information (MI)

Mutual information (MI), or normalised mutual information (nMI), uses entropy as a measure of how well one image explains the other [16]. Registration with MI has received much attention and is by many studies and literature [13, 21, 24] regarded as the most common similarity metric for registration of medical images with differing intensities.

MI is a measure of the statistical dependency between two data sets, in this case, between the fixed image  $I_F$  and the transformed moving image  $I_{M'}$ . The definition of MI and entropy can be found in [25]. Given a set of two images, the measure of MI between two images is given by

$$MI(I_F, I_{M'}) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X, Y) \quad (10)$$

where  $X$  and  $Y$  are two discrete random variables,  $X$  related to the first image and  $Y$  to the second.  $H$  represents the entropy, a measure of uncertainty of a random variable given by

$$H(X) = - \sum_i p_X(i) \log(p_X(i)), \quad (11)$$

$$H(Y) = - \sum_j p_Y(j) \log(p_Y(j)), \quad (12)$$

$$H(X, Y) = - \sum_i \sum_j p_{X,Y}(i, j) \log(p_{X,Y}(i, j)). \quad (13)$$

The base of the logarithm depends on the unit in which the entropy is measured. For example, the logarithm with the base two is used when the entropy is expressed in bits. The probability mass functions  $p_X(i)$  and  $p_Y(j)$  are evaluated at intensity  $i, j$  respectively.  $p_{X,Y}(i, j)$  is the joint probability distribution. These functions are obtained from

$$p_X(i) = \sum_j p_{X,Y}(i, j), \quad (14)$$

$$p_Y(j) = \sum_i p_{X,Y}(i, j), \quad (15)$$

$$p_{X,Y}(i, j) = \frac{h(i, j)}{\sum_{i,j} h(i, j)}. \quad (16)$$

The joint histogram  $h(i, j)$  gives the relationship between pixel values of corresponding intensities in two images [26]. The value in each histogram location  $h(i, j)$  corresponds to the number of pixels in image  $I_M$  having the intensity  $i$  and the intensity  $j$  in image  $I_F$ .

The number of intensities  $i, j$  examined is rarely the same as the total number of intensities in the image. Instead, the intensities are divided into a number of intensity bins (also called joint histogram bins), where each bin is represented by one intensity [27]. Assuming  $n$  intensity bins for both images,  $i \in \{i_1, \dots, i_n\}$ , and respectively,  $j \in \{j_1, \dots, j_n\}$ , see Figure 8. The integer  $n$  can be tuned to affect the result [15].

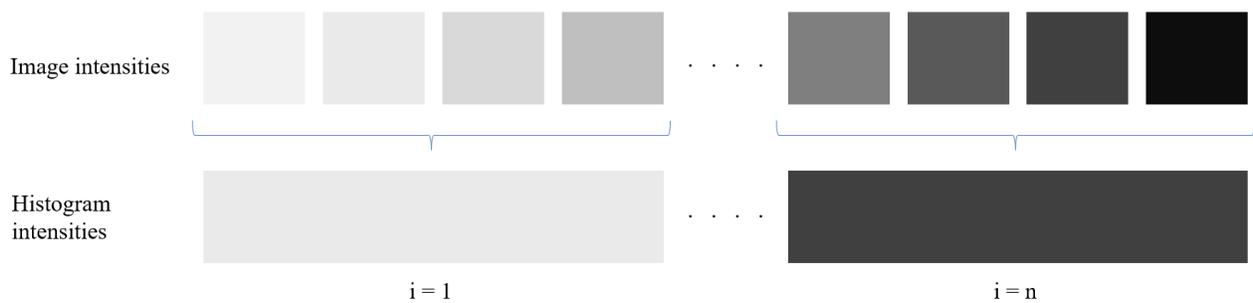


Figure 8: Illustration of how the intensities in an image is dividend into histogram bins.

Despite the popularity of MI, the measure has shown a lack of robustness for certain registration problems. Such problems can, for example, arise when the images are of low resolution. A possible cause is that MI considers the relationships between individual pixels but does not consider the original images' relevant spatial information. Plum et al. [28] suggest a possible solution where a combination of MI and gradient information will satisfy the desired spatial information. Another way to overcome this drawback is to apply MI for global registration as a first step and use another similarity measure for smaller patches [29].

### 2.2.3 Intensity cross-correlation (CC)

Cross-correlation (CC) is a standard method for measuring how two datasets are correlated [30], in this case, the intensities of the fixed image and transformed moving image.

For images  $I_F$  and  $I_{M'}$  with pixel values  $i$ , the normalised CC coefficient is defined as

$$\text{CC}(I_F, I_{M'}) = \frac{\sum_i (I_F(i) - \overline{I_F}) (I_{M'}(i) - \overline{I_{M'}})}{\sqrt{\sum_i (I_F(i) - \overline{I_F})^2 \sum_i (I_{M'}(i) - \overline{I_{M'}})^2}} \quad \forall i \in I_F \cap I_{M'}, \quad (17)$$

where  $\overline{I_F}$  is the mean intensity value of the fixed image, and  $\overline{I_{M'}}$  the transformed moving image [31]. When the images are well aligned, the value of the normalised CC will be one, and if the images have no linear relation, the ratio will be zero.

Like MI, CC is a widespread measure for similarities between images [32], but is often found in other applications than for medical images. Since a linear relationship is assumed between the intensities of the structures of the two images [16], CC is known for being sensitive to intensity changes, for instance, caused by multi-modal images [29, 31]. Still, CC has shown to be especially robust for images with a low resolution, where otherwise MI becomes unstable [29]. Therefore, a combination between MI and CC has been proposed as a successful method [29].

### 2.2.4 Principal component analysis (PCA)

Principal component analysis (PCA) is a technique for identifying patterns in high dimensional data by highlighting similarities and dissimilarities within the data [33]. In other words, the PCA finds what features best describe the data and transforms it according to these to reduce the number of dimensions. This is done by linearly transforming the data into a new coordinate system [34]. When used as a similarity metric, bases reflecting the source image information are compared instead of comparing the information directly [33].

To perform PCA, a set of images  $M_d$  is represented as columns in a matrix  $\mathbf{M}$ , see Figure 9.  $N$  is the number of pixels in the images, and  $G$  is the number of images in the set.

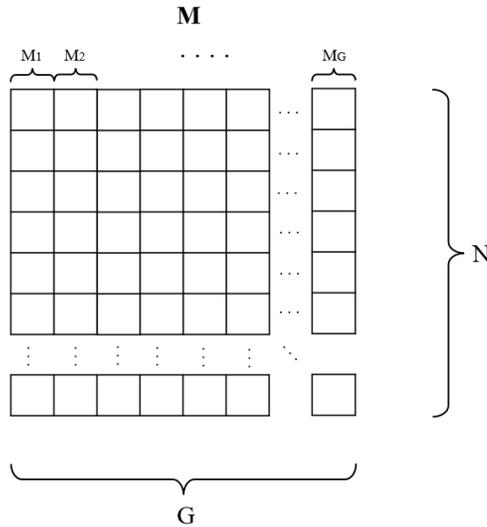


Figure 9: Image showing matrix  $\mathbf{M}$ . Each column holding the pixels of one flattened image and each row show a set of images.

Next, the correlation of the data points in  $\mathbf{M}$  is calculated, from which the PCA transformation is created such that the variation is as great as possible and simultaneously maximises the relation to the original features [35]. A simplified example of this principle is illustrated in Figure 10 where the blue dots represent some particular data and how the properties  $x$  and  $y$  of this data correlate. The red dots are projections of the blue dots along the purple line. The purple line is fitted such as the variance of the red dots is as big as possible at the same time as the total length of the red lines is minimised.

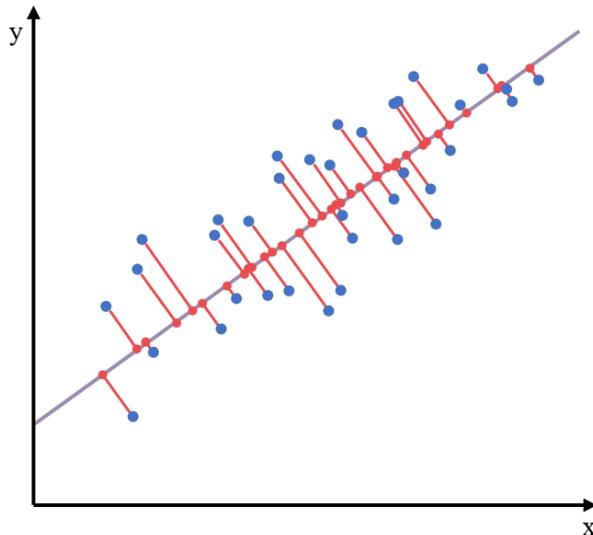


Figure 10: An illustration of how the PCA transform is chosen to represent as much variance as possible and still maximise the relationship of the original features. The axes are two arbitrary correlating properties, and the dots show how these correlate.

The new orthogonal coordinate system where the most variance is observed is given by the eigenvector associated with the largest eigenvalue. It should be noticed that different applications differ, some are based on the covariance of  $\mathbf{M}$  [35], while other are based on the correlation of  $\mathbf{M}$  [33].

W. Huizinga et al. [35] has proposed a couple of groupwise metrics based on this method. One method, denoted  $\mathcal{D}_{\text{PCA}}$ , represents a metric measuring the difference between the sum of all eigenvalues and the sum of the first few eigenvalues. Given a set of images  $I_{M_i}$  with corresponding eigenvalues  $\lambda_i$ , where  $i \in \{1, 2, \dots, n\}$ ,  $\mathcal{D}_{\text{PCA}}$  equals

$$\mathcal{D}_{\text{PCA}} = \sum_{j=1}^n \lambda_j - \sum_{i=1}^m \lambda_i = \sum_{i=m+1}^n \lambda_i,$$

where  $m$  is a selected cut-off constant  $0 \leq m \leq n$ . The eigenvalues are arranged according  $\lambda_i > \lambda_{i+1}$ , so that a eigenvalue with lower index has a larger value than the one with a larger index. An other method,  $\mathcal{D}_{\text{PCA2}}$  [35], is not requiring an user defined parameter  $m$ , and is designed to weight the last eigenvalues more than the first.  $\mathcal{D}_{\text{PCA2}}$  is given by

$$\mathcal{D}_{\text{PCA2}} = \sum_{i=1}^n i \lambda_i. \quad (18)$$

While MI is commonly used for pairwise registration, PCA is often used with groupwise registration because of its characteristics of finding patterns in the information from all

images together [36]. Both  $\mathcal{D}_{\text{PCA}}$  and  $\mathcal{D}_{\text{PCA}2}$  has shown great results on groupwise MRI registration compared to pairwise registration using MI or CC [35, 36].

## 2.3 Optimisation

Optimisation aims to find an optimal set of transform parameters by maximising or minimising the similarity metric [13]. For instance, a moving image  $I_M$  will best resemble a fixed one  $I_F$  when an optimal set of transform parameters is applied. The mathematical formulation for finding the optimal transform parameters  $\hat{\mu}$  is

$$\hat{\mu} = \underset{\mu}{\text{arg min}} C(\mu; I_F, I_M) \quad (19)$$

where  $C$  is the similarity metric to minimise, and  $\mu$  is the current parameters [37].

Finding the optimal set of transformation parameters is an iterative process. In every iteration  $k$ , the transform parameters  $\mu_k$  are updated by taking a step in a search direction with a specific step size [15]. An expression for the iterative approach is

$$\mu_{k+1} = \mu_k - a_k d_k \quad (20)$$

where  $d_k$  is the search direction and  $a_k$  the scalar gain factor determine the step size [15]. The iterative approach aims to choose a step size and search direction that results in a local minimum of the similarity metric [37]. Thus, the metric can result in several local minimums or maximums [37]. The one to select depends on the optimisation method, as well as the initial alignment of the images [37].

The difference between the optimisation methods are how  $a_k$  and  $d_k$  are computed [37]. In some approaches, the search direction is calculated based on exact gradients. In other cases, they are calculated based on approximations of gradients. However, the optimisation methods discussed in this paper are gradient descent and stochastic gradient descent. A final section also discusses other possibilities for improvement, such as pyramidal approaches and samples.

### 2.3.1 Gradient Descent Optimisation

A gradient descent optimisation method is based on gradients and assumes that these can be calculated precisely [37]. Furthermore, the search direction  $d_k$  in this approach takes steps

in the negative gradient direction of the similarity metric [37]. The formula for the search is  $d_k = -g(\mu_k)$  and is the derivative of the similarity metric evaluated at the current position  $\mu_k$  [37].

One approach is to adjust the step size  $a_k$  as a function of  $k$ , given the formula  $a_k = (a/(k + A))^\alpha$  [37]. The step size yields the user to choose values of three parameters:  $A$ ,  $a$  and  $\alpha$ . Adding the formulation from equation 20 and the ones for  $a_k$  and  $b_k$  it results in

$$\mu_{k+1} = \mu_k - \left(\frac{a}{k + A}\right)^\alpha g(\mu_k). \quad (21)$$

### 2.3.2 Stochastic Gradient Descent Optimisation

In some cases, calculating the exact derivatives for the search direction can be very time-consuming [37]. For this reason, calculating only an approximation of the gradients can be preferable [37]. One example of an approach using an approximation of gradients as search direction is Stochastic Gradient Descent. Several Stochastic Gradient Descent versions exist, but the one presented in this paper is the Robbins Monro and the extended version Adaptive stochastic gradient descent method.

The search direction in a Robbins Monro method can be formulated as  $\tilde{g}_k = g(\mu_k) + \epsilon_k$  [37]. Here  $\tilde{g}_k$  represents the approximation of the gradients at position  $\mu_k$  and  $\epsilon_k$  the approximation error [38]. If the approximation results in a zero error, the method is equivalent to the gradient decent method [38]. The step size in this method is calculated in the same way as for the gradient descent,  $a_k = (a/(k + A))^\alpha$  [38]. As a result, the method require the user to tune the variables  $a$ ,  $A$  and  $\alpha$  for the specific application, which is a large drawback of the method [37]. Even though this approach can be similar to the gradient descent method, the difference in the computation of gradients implies a reduced computational time for the stochastic gradient descent method [37]. This method is also shown to be the best choice in most applications [37].

Another optimisation method is the Adaptive stochastic gradient descent. This method is an extension of the Robbins Monro approach but simplifies the selection of parameters [15]. Still, the method requires the user to set parameters  $A$  and  $a$ , but the method is more robust for larger values of these parameters compared with the Robins Monro method [38]

### 2.3.3 Pyramids and Sampling

The iterative optimisation process of the registration is often implemented with a change in resolution, called pyramid strategy [16]. In this approach the images are firstly being

registered with coarse resolution, then stepping up to finer resolution. The transformation parameters from the previous step is used as the initial for each iteration [16]. This pyramidal approach for blurring the images can also act to prevent aliasing [31].

To lower the computational time for a registration method, a subset of pixels in the image can be used to calculate an approximation of gradients [37]. In every iteration, a new subset with randomly chosen pixels is selected [37]. However, a downside with this approach based on samples is that it might reduce the rate of convergence [37].

## 2.4 Registration process for multiple images

The methods described previously in this chapter have considered registration of only two images, a fixed one  $I_F$  and a moving one  $I_M$ . However, several applications require more than two images to be aligned. When handling a set of images, the registration method can be executed either in a pairwise or groupwise fashion.

In a pairwise registration, a fixed image is manually selected to act as the common coordinate system to which all other images are mapped one by one [39]. In a set of  $n$  images, one is chosen as a fixed image  $I_F$  and the rest as moving images  $I_{Mi}$ , for  $i \in \{1, 2, \dots, n - 1\}$ . The pairwise registration is a widespread technique with high efficiency since all moving images  $I_{Mi}$  are only registered once to the fixed image  $I_F$ . Figure 11 shows an example of a dataset with one selected reference image. One downside of pairwise registration is that it introduces a bias in the process towards the coordinate system of the chosen fixed image [39]. This implies that a choice must be made to determine what image is the most appropriate option for the coordinate system.

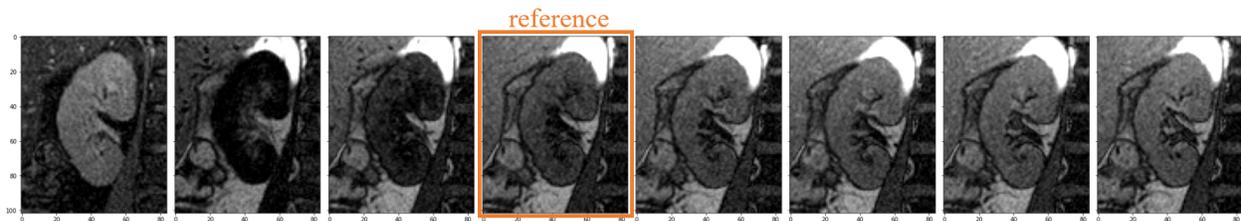


Figure 11: An example of a dataset where the reference image has been selected for the pairwise registration process. The orange square symbolises the reference.

The bias introduced in a pairwise registration can be avoided using a groupwise registration approach. Techniques of groupwise registration have been developed to simultaneously register multiple images [39]. Groupwise registration generates an unbiased and optimal reference to which all images in the set are mapped in a single optimisation procedure [36]. Figure 12 shows an example of a dataset where the reference is generated and not selected from any

existing image. This process involves a high number of parameters and iterations, resulting in a long computational time compared to pairwise registration [39].

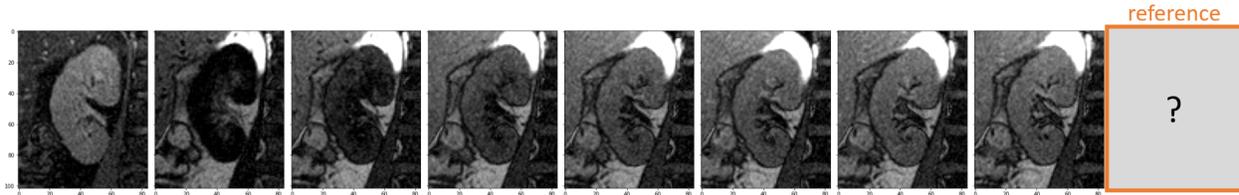


Figure 12: An example of a dataset where the reference image is generated for the groupwise registration process. The orange square symbolises the reference.

## 2.5 Validation of Registration Method

Validating the registration quality is essential for comparing and evaluating methods and their parameters. There are, however, some challenges with the validation of registration methods. One is the lack of gold standards to compare against [40]. For instance, if a gold standard existed it would be intuitive to calculate the difference between this standard and the registered images. The lack of ground truth images is often solved using manually segmented images, performed by experts [40]. This validation method is though time consuming and requires a certain expertise [41].

There are different validation methods, but the two used and described in this thesis is based on segmentations and visual inspection. In segmentation based methods, images are first manually segmented by experts, creating binary masks [41]. Then, the images  $I_M$  and  $I_F$  are registered, and the transform used for this image alignment is saved and applied to the segmented moving  $I_M$  mask [41]. Lastly, the registered segmentations are compared to each other using metrics to see how well they are aligned [41, 42].

It exists several metrics for comparing the registered segmentations [43]. The metrics are in this paper divided into three main categories; spatial overlapping, distance measures and lastly, volumetric approaches [40] [44]. The overlapping spatial measures covered is the Dice coefficient [45]. The distance measures considered are two versions of Hausdorff distances: Absolute Hausdorff Distance and Average Hausdorff Distance. Lastly, the volumetric approach described is the Volume similarity measurement.

### 2.5.1 Spatial Overlapping Metrics

One way to express a comparison between two segmentations is by studying the spatial overlap. The most common overlapping metric for evaluating the segmentation of medical

images is Dice [43]. Dice is a measurement for the intermediate association between two sets [45] and calculated using the formula

$$\text{DICE} = \frac{2|A \cap B|}{|A| + |B|}. \quad (22)$$

Where  $A$  and  $B$  are the masks of the two images to compare, one for the fixed image  $I_F$  and one for the moving image  $I_M$ . The Dice value ranges from zero to one, where one represents a total overlap and zero that there is no overlapping between the two segmentations [40].

A considerable drawback with the overlapping spatial metrics is that they are not considering the positions of the points that are not overlapping [44]. The drawback might imply that other measurements are also needed to compensate for the loss of localisation.

### 2.5.2 Distance Measures metrics

The second category of metrics is the distance measures, considering the spatial distance between the edge points. Compared to the overlapping measurements, the distance measures considers the localisation of the pixels, providing the distance between the edges of two segmentations [46]. One commonly used distance measures to evaluate the degree of overlapping are different variations of Hausdorff distances [46], defined as,

$$H(A, B) = \max(h(A, B), h(B, A)), \quad (23)$$

given two point sets  $A = \{a_1, \dots, a_n\}$  and  $B = \{b_1, \dots, b_m\}$ , including edge points for two respective images. First, the distances between all points in  $A$  and all points in  $B$  are calculated. The minimal distances connecting the points in  $A$  to some point in  $B$  are detected. This gives  $n$  distances. The directed Hausdorff distance  $h(A, B)$  is the maximal out of these values,

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|. \quad (24)$$

A similar procedure, based on  $B$  yield  $h(B, A)$ . The Hausdorff distance is, in turn, the maxima out of these two values, as seen in Equation (23). The equation for Hausdorff distance is in more general term defined with the supremum instead of the maximum in Equation (24). However in the case with a finite number of pixels the equation can be simplified by calculating the maxima instead.

There is a drawback with this Hausdorff metric. Since it only results in the most significant distance, it is very sensitive to outliers, for example, noise [40][43]. Other versions of the Hausdorff distance have evolved to be more robust.

The Average Hausdorff distance is, as the name suggests, the Hausdorff distance averaged on all points [43]. This metric is more stable and less sensitive to outliers than the original Hausdorff distance [43].

### 2.5.3 Volumetric metrics

Another way to validate the performance of the registration is by comparing the volume between the two segmented masks. One example of a volumetric approach divides the absolute difference in volume by the sum of the compared volumes [43]. One of the properties of this method is its independence of grade of overlapping. Accordingly, the volumes can be similar even if there is no overlapping between the segmentations [43]. The formula for the volumetric measure is

$$VS = 1 - \frac{|A - B|}{|A| + |B|}. \quad (25)$$

Here  $A$  is the mask of the fixed image  $I_F$  and  $B$  is the moving image  $I_M$ . A value close to zero implies a similar volume between the masks, while a result close to one instead implies a significant difference in volume between the two segmentations.

## 3 Methodology

The first part of the project consisted of a literature study. This study was performed by searching for articles through databases using various keywords. Examples of databases used are EDS provided by Chalmers library and Google scholar. The purpose of the literature study was first to find and gather information to understand different methods and techniques for image registration. The objective was also to identify promising methods specifically for registration of MRI images of kidneys.

The second part of the project was the performance study. Here, an algorithm was developed to test the state of the art methods and promising techniques on given MRI kidney data. The objective was to see how well the methods perform and better understand how settings could be selected and parameters tuned to yield adequate results. The algorithm was developed in Python, using Visual Studio Code and Jupyter Notebooks. The library Insight Toolkit

(ITK) [47] was used for its software tools for image analysis. The toolbox Elastix [48] was used for the image registration.

The following sections present the method used for this project. Firstly the data used in the study is described, followed by how the images are processed before the registration. Then the setup of the performance study is described as well as an in-depth explanation of what settings were investigated and why.

### 3.1 Data used in study

The data used for the implementation was kidney MRI images provided by Antaros Medical. Figure 13 illustrates the structure of the data. The data consisted of scans collected from 6 subjects A-F, including the MRI sequences MOLLI and DWI. Each scan were in 2D and include three coronal cross sections of MOLLI images: two at random frontal positions and one in the centre of the kidney, and one centred cross section for the DWI images. The original MOLLI images had a size of  $288 \times 288$  pixels and DWI  $210 \times 210$ . The left and right kidneys were registered individually, consequently the total number of images of kidneys to register for MOLLI was 36 and 12 for DWI. The MOLLI images were acquired at different inversion times, in this case, it varied between 8 and 11 instances, identified as N in Figure 13. The DWI images are, on the other hand, acquired for different b-values, resulting in 4 instances.

The datasets were selected to get a variety of misalignments between the sets. There is also a difference in the degree of movement between the DWI and MOLLI images. The first image in all MOLLI datasets have a different intensity than the rest, where the kidney appear white. However, in the MOLLI dataset B, the first two images have this characteristics.

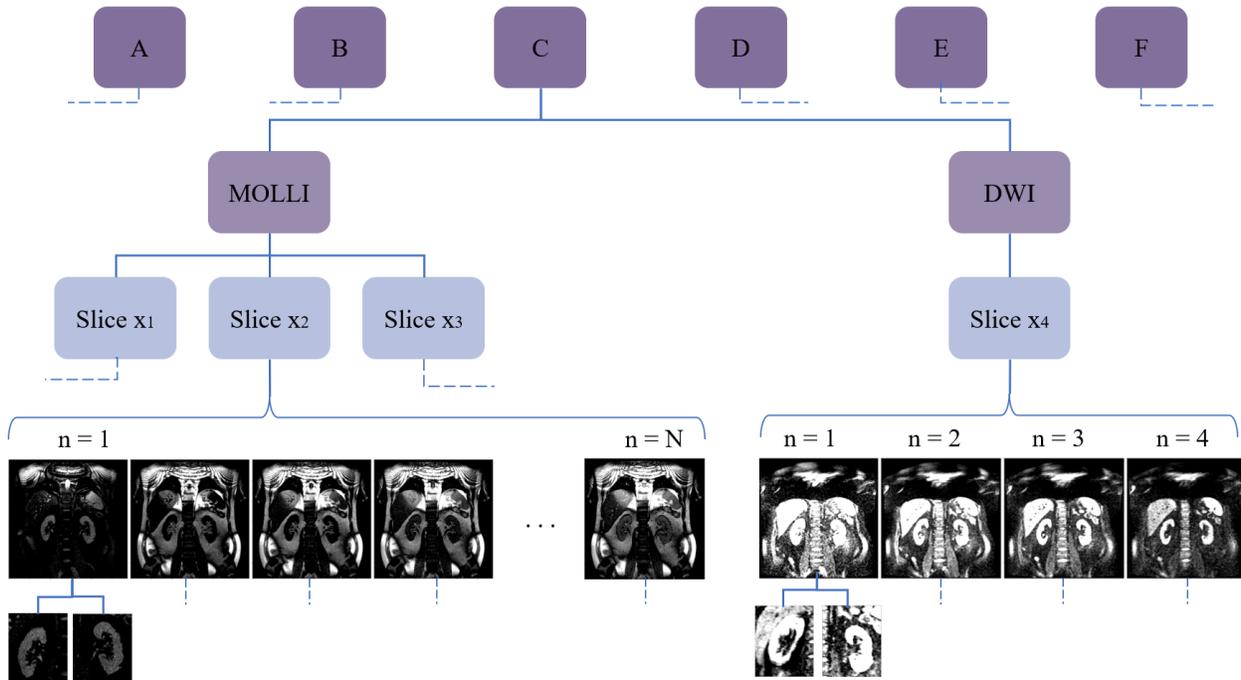


Figure 13: Visualisation of the data structure used in this project. Here A-F are the datasets, each from a different scan occasion and subject. Each dataset contain both MOLLl and DWI images. There are three MOLLl slices  $x_1$ ,  $x_2$  and  $x_3$ . Each MOLLl slice contain a set of images  $n = 1, 2, \dots, N$  acquired with different inversion times. The number off different inversion times differ where  $N$  reaches from 8 for some datasets, to 11 for others. Each DWI image contain one slice,  $x_4$ , containing four images captured with different b-values. The DWI slice  $x_4$  has the same frontal position as the MOLLl slice  $x_2$ . Each image is then divided into left and right kidney before the registration.

### 3.2 Image Pre-processing

The images were pre-processed in order to support the registration. Pre-processing in this project includes limiting the region of interest (ROI) and upsampling the images with different sizes. First, the images were cropped into images containing left and right kidneys separately. The benefit of this is the reduced amount of background information influencing the registration [7]. Furthermore, the kidneys move independently, and therefore, the performance can be improved when registrating each kidney separately [7]. The ROIs used were created in the image pre-processing program ImageJ. The ROIs used to crop the left kidney had the same size for all data sets, and the ones used to crop the right kidney had all the same size. Although, the positions of the ROIs might differ between datasets. The reason for this is the various positions of the kidneys due to anatomical differences and various positions between subjects during scanning.

The second step is upsampling. The pixel sizes between the MOLLI and DWI images used in this project are different. For this reason, a good registration between these two types of images is not likely. Therefore, an upsampling of the DWI images is carried out. The upsampling process uses linear interpolation to approximate intensities.

### 3.3 Performance study

In order to analyse promising methods from the literature study on kidney MRI, specifically MOLLI and DWI, these methods were applied in a performance study. The settings investigated for this study includes: i) the registration process, ii) choice of transform, iii) choice of similarity measure. For some of the settings parameters were tuned or investigated to get the best performance of the settings. These included comparing reference images for pairwise registration, selection of grid spacings for the B-spline transform, and the effect from different histogram bins for MI. See Figure 14 for an overview of the different settings and parameters used in the study. These options are described in more detail in Chapter 2.

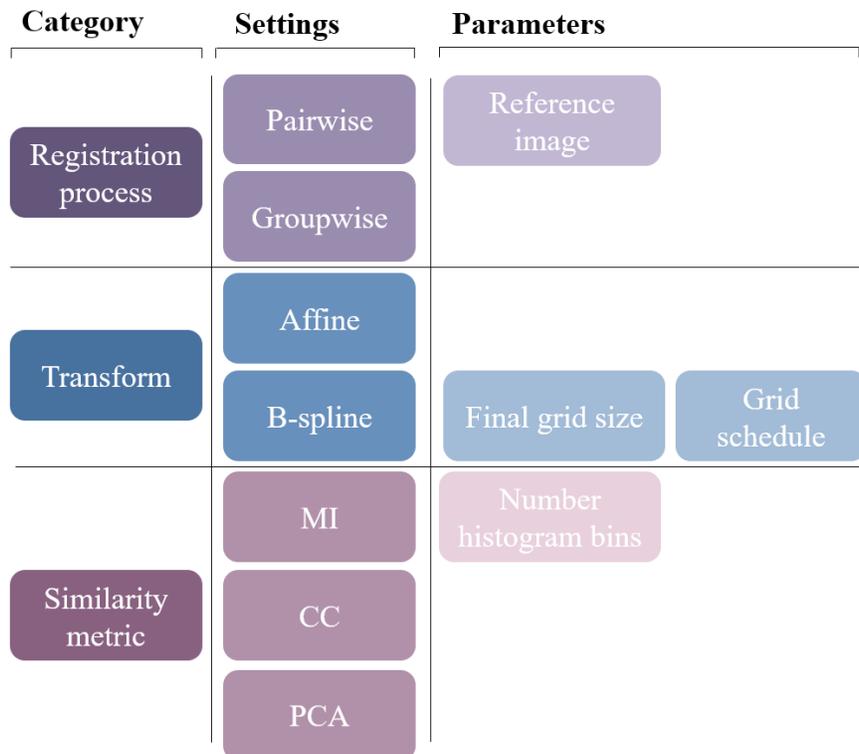


Figure 14: Illustration of the different settings investigated in the performance study. Parameters have been tuned for some of these settings: for pairwise registration the reference of choice is tested, for the B-spline transform different grid spacings have been investigated, and for MI the effect from different number of histogram bins have been tested.

In the performance study, all tests were first performed on MOLLI and DWI individually to understand how different settings affected the intra-modality registrations. Then MOLLI and DWI were registered together to investigate the effects on inter-modality registration. Some tests were performed solely on pairwise registration, for example, choice of reference images.

### 3.3.1 Selection of reference image

In pairwise registration the selection of reference image can have an effect of the registration result, like explained in section 2.4. A test was performed to see what the effect of different reference images was and if any conclusions could be drawn. This test was performed on the data sets A-F for MOLLI and DWI images. The settings used was a B-spline transform with MI. The images tested as references were image 1-8 in the order of inversion time for MOLLI, and image 1-4 in the order of b-value for DWI.

### 3.3.2 Similarity metrics to investigate

The similarity metrics investigated are CC and MI for pairwise registration and CC and PCA for groupwise registration. As described in section 2.2, MI is considered a promising technique, especially for medical images with differing intensities, and was therefore chosen for testing. PCA was selected since it, according to the literature, is regarded as a commonly used similarity metric for groupwise registration. However, studies also imply that CC is not suitable for images with a big variation in intensities, which is the case for the images used in this study. Therefore the study only tested CC on a few data sets to confirm or deny these findings from the literature study.

Furthermore, a test was explicitly conducted with MI to investigate the effect of the selected number of histogram bins. This test was done with both Affine and B-spline transformation models for datasets A-F. The number of bins tested was 8, 16, 32, 64, and 128. The image used as a reference for registration of MOLLI images and MOLLI and DWI images together was the fifth inversion time. The second b-value was used as a reference for the registration of DWI images.

### 3.3.3 Geometric transformation models to investigate

The two different geometric transformation models used for testing in this thesis were an Affine transformation model and a B-spline one. The Affine transform was selected due to its properties of rotating, scaling, and shearing to align the images but still preserving the

internal structures of the original images. The B-spline transformation was selected due to its interesting property of aligning both the contours and the internal structures. Before the B-spline transformation was applied to the images, they were pre-aligned using a rigid transform.

Previous studies have indicated an increased flexibility of the B-spline transform using a smaller grid spacing. It is also stated that an hierarchical change of the grid spacing can improve the precision and computational time. Therefore, it was interesting to investigate the effect of these approaches on the images used in the current project.

The final grid spacing's tested were: 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 pixels. Additionally, four different grid schedules were tested, schedule 1: {4, 2, 1}, schedule 2: {6, 3, 1}, schedule 3: {8, 4, 1} and schedule 4: {10, 5, 1}. Where a schedule determine by how much the final grid spacing will be multiplied with for each iteration of the pyramid strategy. The schedules are built to start with a bigger spacing for the more coarse resolution and finish at one for the finest resolution to reach the desired final grid spacing. The datasets used for this test were A-D. The last two sets were excluded due to the large computational time needed for this test. Furthermore, tests with all grid schedules were only performed using a pairwise registration. For a groupwise registration, only grid schedule 3 was tested and a few random samples were chosen for investigation of the effect of changing final grid spacings.

### 3.3.4 Optimisation method to investigate

An adaptive stochastic gradient descent method was used as the optimisation method throughout this project. The adaptive approach was chosen since fewer parameters need to be selected and tuned than for the other stochastic gradient methods. The adaptive stochastic gradient descent methods parameters were set automatically during the registration, and nothing was manually changed.

### 3.3.5 Selection of registration process and transformation model

In order to identify the methods that performs best on kidney MRI images, four different promising combination of settings were tested. All tests that were performed using a B-spline transformation model used a final grid spacing of 60 pixels and the third grid spacing schedule {8, 4, 1}. Furthermore, the images registrated using B-splines were also pre-aligned using a rigid transformation. The test using MI as similarity metric used a number of 16 histogram bins. Lastly, the pairwise registrations used the fifth or eight image as the fixed one. These choices were in most cases based on their performance from previous individual test presented above. The four different combination tested can be seen in Figure 15.

	Combination 1	Combination 2	Combination 3	Combination 4
Registration process	Pairwise	Pairwise	Groupwise	Groupwise
Transform	B-spline	Affine	B-spline	Affine
Similarity metric	MI	MI	PCA	PCA

Figure 15: An image showing the four different combinations tested. The first column represent the settings and the four following columns each represent a combination. The first row show what registration process is used in the corresponding combination, the second row show the transform used and the last row present what similarity metric used.

### 3.3.6 Validation of methods

The quality of the registration method was determined by using segmentation and visual inspection. First, segmentations of the original images were manually created by experts. The segmentations are binary masks, where the background is assigned a value of zero and the pixels containing kidney tissue a value of one. Then, the same transformations found registering the images were applied to the corresponding masks. As a result, the masks were transferred to the same coordinate systems as the registered images. Example of a set of segmented and transformed left kidneys can be seen in Figure 16.

Once the masks were transformed, they could be compared using metrics. The metrics used to evaluate the performance of the methods in this paper were the Dice similarity measure. Other metrics, for example, Hausdorff distance and Average Hausdorff distance, were during the test used to confirm the observations and identify outliers but not used for the result and analysis of the registration performance. The reasons for excluding these metrics were because Hausdorff and Average Hausdorff, in most cases, followed the same pattern as the Dice metric. For example, when there is a low Dice value, Hausdorff and Average Hausdorff result in high values and vice versa. The validation based on the volumetric approach was totally excluded since the alignment of the images is dependent on the overlap of the kidneys, and this metric did not concern that.

The metrics only describe the overlap between two masks. Therefore, an evaluation of multiple masks was performed comparing all transformed masks in the data set to each other, left and right kidneys separately. Lastly, a mean value including both kidneys were calculated

and used as the final value representing the performance of the registration.

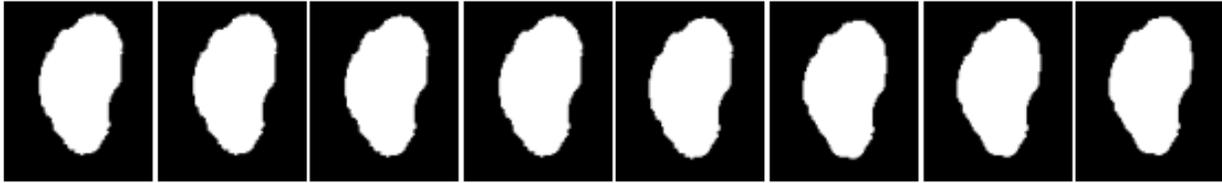


Figure 16: Set of expert segmented masks that are transformed and compared using metrics.

## 4 Results

In this section, the results from the different performance studies are presented. Some studies are presented alone, while others are compared and presented together.

### 4.1 Selection of reference image

Here the results are presented for registration using pairwise registration with different images as reference. For MOLLI images the results from having the first image as a reference was drastically lower compared with the rest, which stayed quite consistent regardless of the choice of reference. This can be seen in Figure 17. The resulting Dice value for dataset F with reference image 1 is 0.871 which has a great effect of the mean Dice value. For dataset B, registration using both the first and the second image as reference results in lower values than the rest. For further tests a reference image between 3-8 is used for MOLLI images.

Figure 18, show the characteristics of choice of reference image for DWI images. The Dice values stay similar for different reference images.

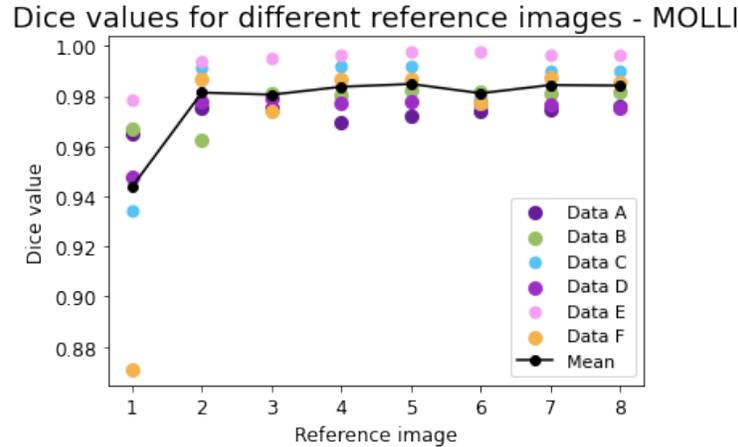


Figure 17: Dice values for pairwise registration of MOLLI images using different images as a reference. The Dice value is the mean over all three slices in each dataset. Note how the registration with the first image as a reference result in lower values for all datasets except dataset B, which have low values for both the first and the second image.

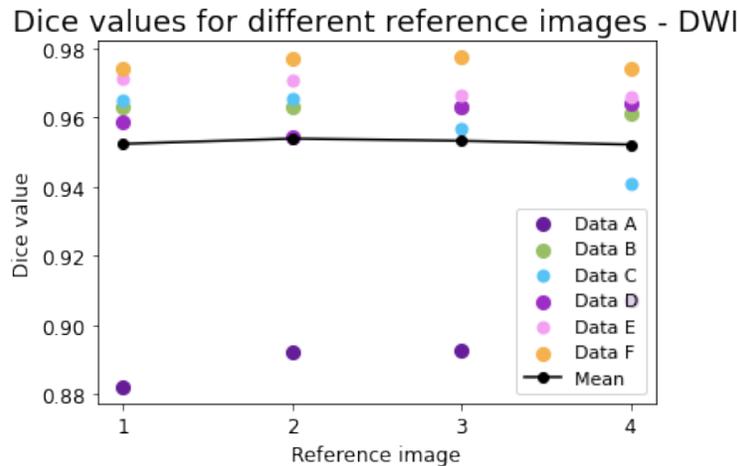


Figure 18: Dice values for pairwise registration of DWI images using different images as a reference. Note how the trend is consistent over all reference images.

## 4.2 Investigation of different similarity metrics

Sample tests were performed using a CC similarity metric. An example of MOLLI images in dataset B registered using CC as similarity metric is showed in Figure 19. From visual validation, it was observed that the inverted images in the MOLLI data sets were distorted and therefore badly registered, while images with similar intensities were well aligned.

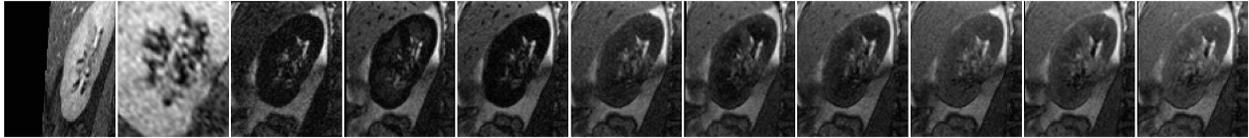


Figure 19: Result from a test using cross correlation as similarity metric together with an Affine transformation model using a pairwise approach. It can be identified that the two images with inverted intensities results in bad alignment.

In Figures 20 - 22 the results of using MI with different number of histogram bins are presented. From the figures, it can be observed that the patterns are similar for both B-spline and Affine transformation models. Registration of MOLLI images with different number of histogram bins is seen in Figure 20, note how the lowest Dice value represent the lowest number of histogram bins. For DWI images, Figure 21, the result show that the Dice values is almost constant for all number of histogram bins tested. The last combination, registration of MOLLI and DWI images together, shows its lowest Dice value using a low number of histogram bins, then a peak can be identified followed by a slight decrease for higher number of bins. For Affine transformation the peak occurs at 32 histogram bins and for B-spline transformations at 16 bins.

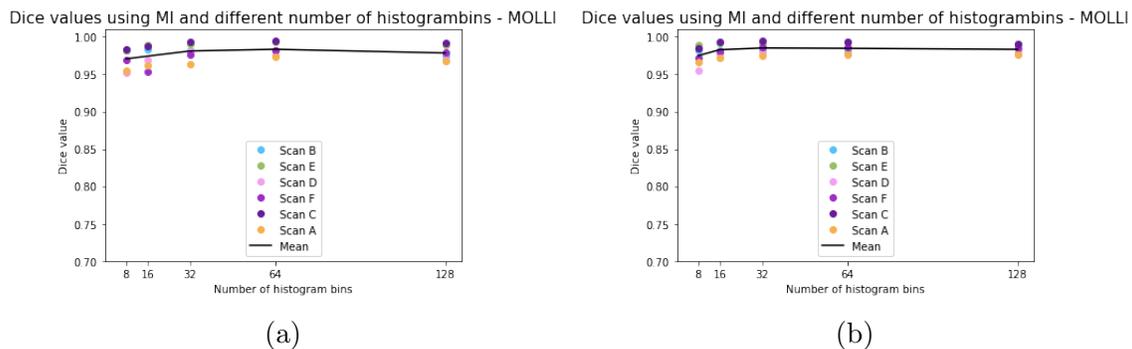


Figure 20: Registration of MOLLI images with different number of histogram bins. (a) Registration using an Affine transformation and (b) a B-spline model. The test is performed using both Affine transformation model and a B-spline model. It can be observed that the lowest number of bins result in the lowest Dice value.

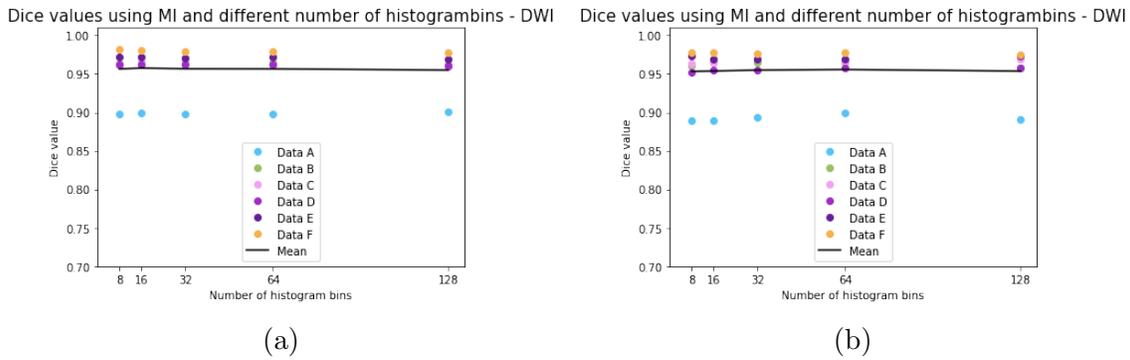


Figure 21: Registration of DWI images with different number of histogram bins. (a) Registration using an Affine transformation and (b) a B-spline model. The test is performed using both Affine transformation model and a B-spline model. It can be observed that the Dice value is constant and independent on the number of histogram bins.

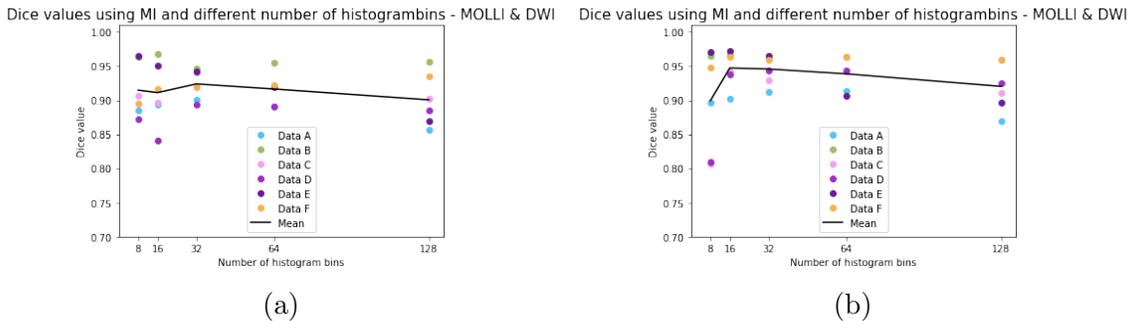


Figure 22: Registration of MOLLII and DWI images together with different number of histogram bins. (a) Registration using an Affine transformation and (b) a B-spline model. The test is performed using both Affine transformation model and a B-spline model. Both transformation models result in a peak and then a decrease.

The images registered with different number of histogram bins are also visually inspected. The result is presented in Figures 23 and 24. It can be observed that the images with odd intensities are distorted, while the ones with similar intensities are still well aligned.

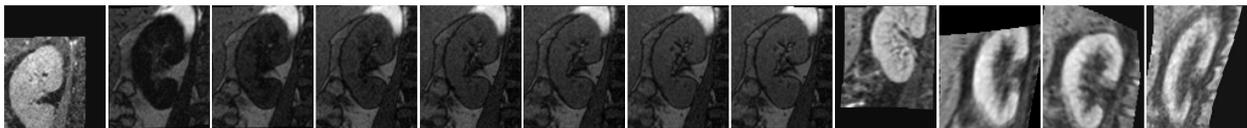


Figure 23: In this image, pairwise registration of MOLLII and DWI images using 8 histogram bins for an Affine transformation. The fifth image from the left is the reference image, and dataset D is used.

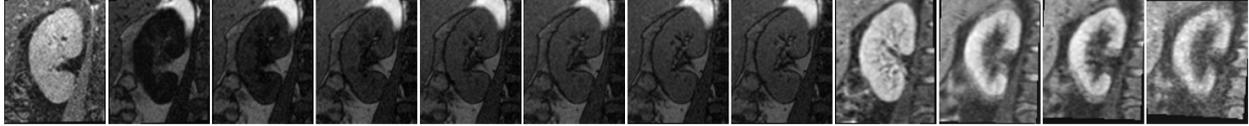


Figure 24: In this image, pairwise registration of MOLLI and DWI images using MI and 32 bins is shown. The fifth image from the left is the reference image, and dataset D is used. It can be observed that there are less distortions and the images are well aligned.

### 4.3 Investigation of B-spline grid spacing

The effect different spacings of the grid in the B-spline transformation have on the registration is presented in this section. First the observations of pairwise registration is presented, then moving on to groupwise.

No considerable difference or trend was found for a pairwise approach for registration using different grid schedules. This result applies to all three combinations of images. Figure 27 show the result for the registration of DWI and MOLLI images together. This graph shows that the Dice value improves as the final grid spacing increases. The same applies to the result for MOLLI images seen in Figure 28, where the lowest Dice value is found for the lowest final grid spacing. Turning now to the result of registering DWI images. It can be observed in Figure 29 that the Dice value is constant and independent of the different final grid spacings tested in this report. For a groupwise registration, it could be observed that the values were consistent for all grid spacings, and no large changes could be observed.

An example of pairwise registered MOLLI and DWI images together using a B-spline grid of 10 voxels is seen in Figure 25. The images were registered using the fifth image as a fixed image. It can be observed from the figure that the images with inverted intensities compared to the fixed image are distorted. In Figure 26, the same images are registered but this time using 60 voxels as final grid spacing. Note how no distortion can be seen.

The result from similar tests using groupwise registration for a fixed grid schedule and for final grid sizes ranging from 10-100 showed very similar Dice values regardless of the final grid size.

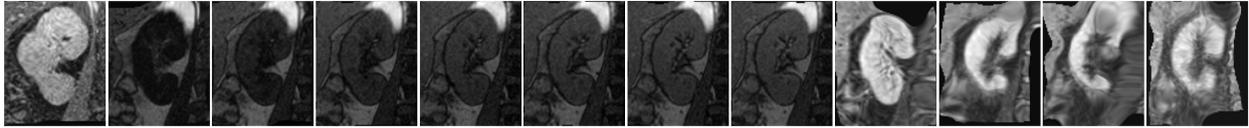


Figure 25: In this image, pairwise registration of MOLLI and DWI images using 10 as a final grid spacing and schedule 3 is shown. The fifth image from the left is the reference image, and dataset D is used. It can be observed that the images with different intensities compared to the fixed image is distorted.

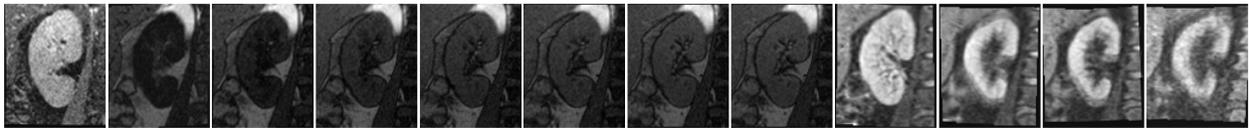


Figure 26: In this image, pairwise registration of MOLLI and DWI images using 60 as a final grid spacing and schedule 3 is shown. The fifth image from the left is the reference image, and dataset D is used. It can be observed that there are no distortions and the images are well aligned.

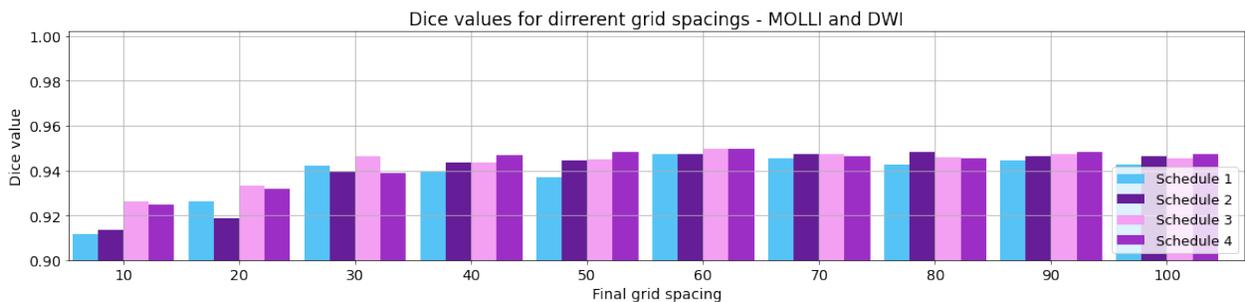


Figure 27: Graph showing the Dice values for registration using ten different final grid spacings and four different schedules. Each bar is the mean value of all datasets and slices.

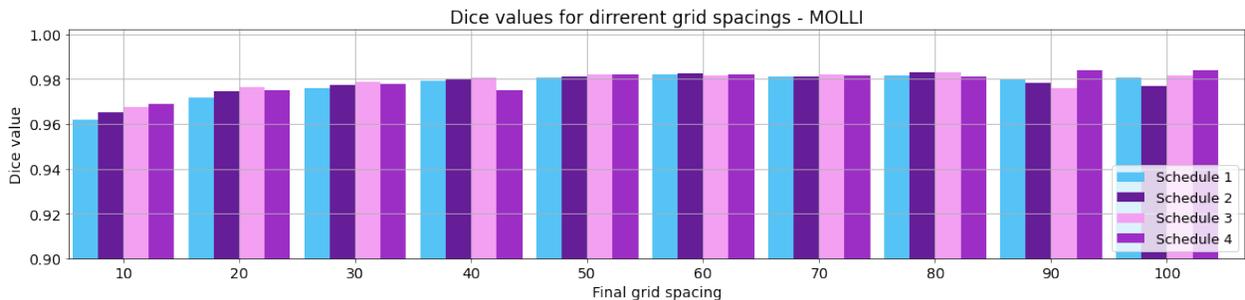


Figure 28: Graph showing the Dice values for registration using ten different final grid spacings and four different schedules. Each bar is the mean value of all datasets.

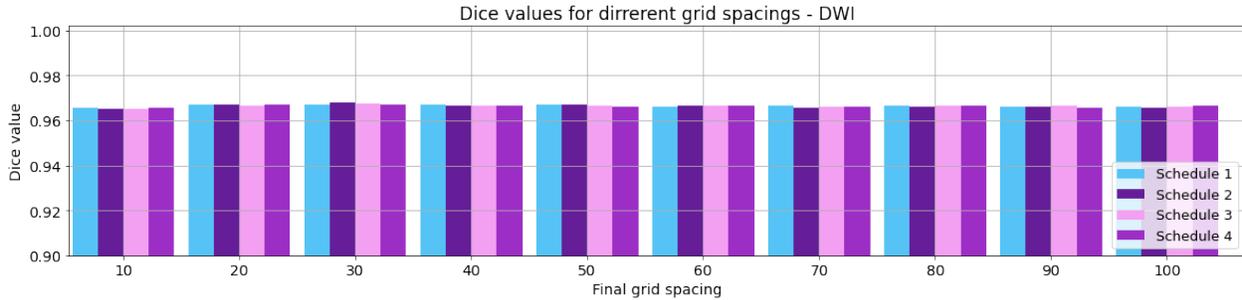


Figure 29: Graph showing the Dice values for registration using ten different final grid spacings and four different schedules. Each bar is the mean value of all datasets.

#### 4.4 Selection of registration process and transformation model

Here the result for groupwise and pairwise registration using the two geometric transformation models B-spline and Affine are presented. The graphs in Figures 30-32 show the resulting Dice values for each dataset A-F for respective registration method. Note that the graphs have different minimum limits on y axis. The blue line in the images represent the original Dice values before the images have been registered, based on the manual segmentation. For the MOLLI images, the presented Dice value is the mean over the Dice values for all three slices. The observations are in this subsection divided into MOLLI, DWI and MOLLI and DWI registered together.

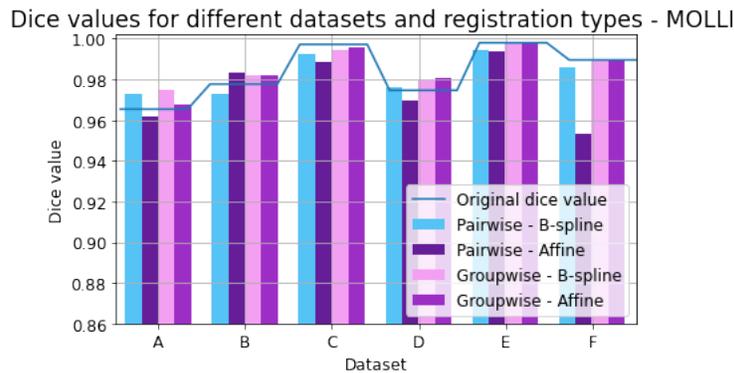


Figure 30: Barplot for Dice values for different combinations of pairwise and groupwise registration together with Affine and B-spline transformation. The result is shown for datasets A-F for MOLLI images. The blue line shows the Dice value before registration.

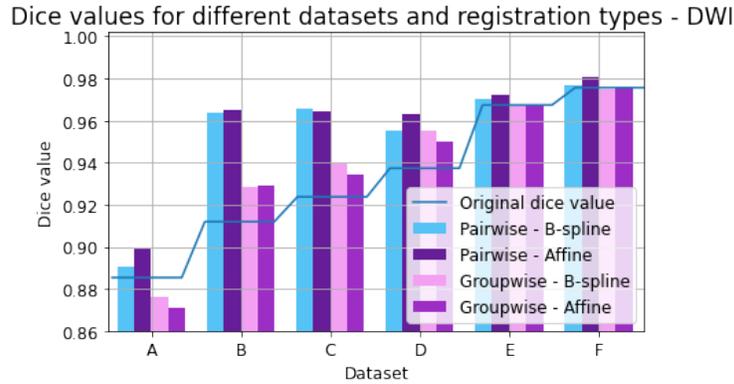


Figure 31: Barplot for Dice values for different combinations of pairwise and groupwise registration together with Affine and B-spline transformation. The result is shown for datasets A-F for DWI images. The blue line shows the Dice value before registration.

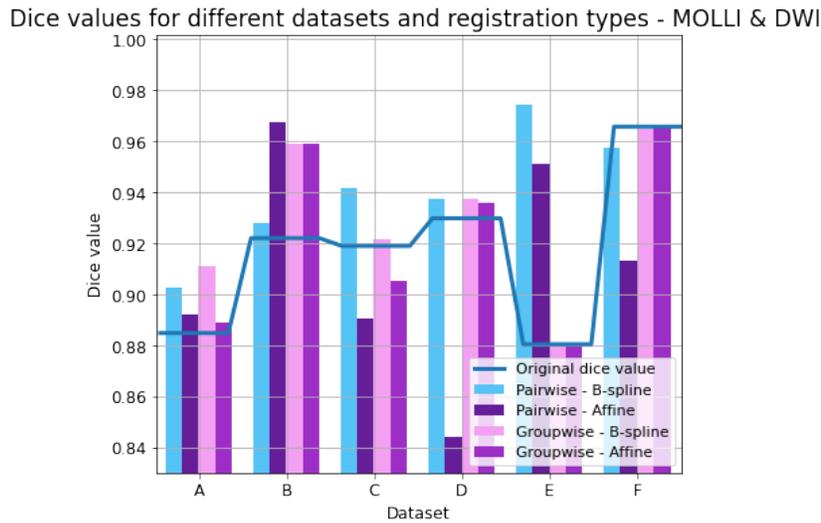


Figure 32: Barplot for Dice values for different combinations of pairwise and groupwise registration together with Affine and B-spline transformation. The result is shown for datasets A-F for MOLLI together with DWI images. Note how the results vary and lack a distinct pattern. The blue line shows the Dice value before registration.

Table 1: The mean differences between Dice value after registration and Dice value before registration over all datasets A-F. Note how the values for pairwise registration are negative for both Affine and B-spline transformation, while staying positive for groupwise registration.

MOLLI	
Option	Mean difference
Pairwise B-spline	-0.00138
Pairwise Affine	-0.00876
Groupwise B-spline	0.00259
Groupwise Affine	0.00179

Table 2: The mean differences between Dice value after registration and Dice value before registration over all datasets A-F. For DWI images. Note how all values are positive, but higher for the pairwise options.

DWI	
Option	Mean difference
Pairwise B-spline	0.02020
Pairwise Affine	0.02394
Groupwise B-spline	0.00677
Groupwise Affine	0.00443

Table 3: The mean differences between Dice value after registration and Dice value before registration over all datasets A-F. For MOLLI images registered together with DWI images.

MOLLI and DWI	
Option	Mean difference
Pairwise B-spline	0.02332
Pairwise Affine	-0.00716
Groupwise B-spline	0.01203
Groupwise Affine	0.00570

The graphs in Figures 30 - 32 show that there is a variation of the results between the different datasets. In Figure 30, showing the results for registration of MOLLI images, the mean value is represented for the three slices in each dataset A-F. The bar plots for the MOLLI images show rather even results for the different options with only one deviating result of the pairwise-affine registration on dataset F reaching a Dice value of 0.953. Table 1 show the the mean differences from the original Dice values over all datasets for MOLLI images, rounded to five decimal places. This differences show that groupwise registration perform better than the pairwise registration, for both transformation models tested. In addition, it can also be seen that the mean difference for groupwise registration using B-splines are slightly higher than groupwise registration using an Affine transformation model.

In Figure 31, the Dice values for different registration options of DWI data is presented. Dataset A stands out for having Dice values below the original Dice values for both groupwise options. The rest of the bars however are above the original Dice values. The pairwise options results here in higher Dice values than the groupwise options for all sets, where pairwise Affine performs slightly better than pairwise B-spline. The mean difference between the resulting Dice values and the original Dice values over all datasets can be seen in Table 2.

It can be seen from the barplot in Figure 32 that the results for the different options are varying when MOLLI and DWI images is registered together. For example is pairwise Affine showing great results for dataset E, but a very low Dice value for dataset D. Pairwise B-spline show higher Dice values than original Dice in all cases except for dataset F. Groupwise Affine show higher or equal Dice values than original Dice in all cases except for dataset C. Groupwise B-spline results in higher or equal Dice values for all datasets. When comparing over all datasets together, pairwise B-spline show the highest value, see Table 3.

## 5 Discussion

This chapter presents a discussion about the results from the previous chapter. The discussions cover the findings, limitations and recommendations for further research.

### 5.1 Selection of reference image

One of the questions raised while implementing the pairwise approach was if there existed an optimal choice of the reference image and, if so, which one. The literature indicates that a bias is introduced toward the coordinate system of the chosen reference image using a pairwise approach.

The result from this study shows that the choice of reference image can have a large effect on the registration result. Even though there can be small changes between different references within each dataset, it can be observed that images 3-8 for MOLLI and all images for DWI work as good references. All these images have in common that non have a deviating intensity. From the test with MOLLI images, see Figure 17 in the result, it can be seen that when an image with inverted intensity was used as a reference, the results are lower than for the rest. This was image 1 for all datasets, but also image 2 for dataset B. These tests were performed with MI as a similarity metric, known for performing well on images with different intensities. Due to the bad performance of CC as a similarity measure early in the study, it was not investigated for this test. However, it would be interesting to compare the results if the same test had been done with CC as a metric instead of MI to confirm the hypothesis that the use of MI reduces the variety of the results for the different reference images.

The outcome from the test on MOLLI images correlates with what is observed on the test for the DWI images. Given the result from the MOLLI images, it could be expected that the effect of changing the reference image for DWI would not have a big impact on the result since all DWI images have similar intensity.

These tests were only performed on MOLLI and DWI images separately. Thus it would be interesting to do a similar investigation for MOLLI registered together with DWI to see if the results would be any different.

## 5.2 Investigation of different similarity metrics

Several reports have shown that MI is the most common similarity metric for medical images with varying intensities. Studies also indicate that CC is a similarity metric sensitive to intensity changes. The results from this report is in line with these studies and shows that MI outperforms CC on the types of images used in this study. The tests for CC are only performed on a few datasets. However, it may be the case that CC works well for registration of DWI images only, since these images all have more similar intensities. Hence, there is room for further investigations.

The results from testing MI with different histogram bins show that a low number of bins is not optimal for registration of MOLLI images nor MOLLI and DWI images together. The observations found using visual inspection indicate that the images with deviating intensities from the rest are the ones not aligned. A possible explanation for this is that when using a low number of histogram bins for images with different intensities, the pixels might be divided differently. Meaning that the pixel intensities for some detail might be grouped in one histogram for one of the images but split into different histograms for another. This means that the metric will handle these images as if they contain different details. This theory, that a set of images with varying intensities would be more sensitive to the number of histogram bins, is supported when analysing the result on the DWI datasets. The DWI images are less dependent on the number of histogram bins, which could be explained by the fact that the images are more similar in general.

## 5.3 Investigation of B-spline grid spacing

As mentioned in the literature review, a finer grid spacing can improve registration accuracy. On the contrary, the results from this study indicate improved and more consistent results for a higher final grid spacing, especially for MOLLI images and MOLLI images registered together with DWI images. This finding was unexpected considering the benefits of a finer grid stated in the literature. One explanation for this rather contradictory result is that the

flexible characteristics of a B-spline transformation might be redundant for the misalignment of kidneys present between the images. The bigger the spacing between the control points of the grid, the more the transform starts to resemble an Affine transform. Consequently, an Affine transformation might be enough for this kind of registration.

The result for different final grid spacings when registering DWI images stayed consistent. The reason for this is not clear and would need further investigation. The same applies to groupwise registration where it also would be interesting to investigate the effect of a different number of resolutions instead of different schedules. One hypothesis is that fewer resolutions will result in a worse registration performance.

The grid schedule showed no trend or large difference on the tests. Therefore, it is likely that the most important property is the final grid size, which is used for the final resolution. For some of the tested schedules, the grid spacing in the first resolutions is much larger than the image, which might imply that grid spacing is too sparse to have an effect on the image. However, in that case, the outcome should be different for the first two schedules registered with low final grid spacing.

## 5.4 Selection of registration process and transformation model

The next focus of this project was to see how well methods performed and understand how settings and parameters could be tuned. These tests used the findings for different settings from previously presented studies to compare different combinations. The results from these tests further strengthen the sense of complexity of registration of kidney MRI images.

It is hard to determine a trend between the different options presented in this study, which was a combination of pairwise or groupwise registration with Affine or B-spline transformation. A small indication is that a groupwise approach is better for registering MOLLI images. On the other hand, a pairwise approach shows the best and most robust result for the registration of DWI images. As shown in previous tests, a pairwise approach depends on the choice of a reference image. For this test, the reference image is chosen arbitrarily from the span presented in this previous test. Since the reference image is chosen arbitrary and not dependent on the best choice for each data set, there is a possibility that another choice can improve the result. However, in a real-world scenario, there will be no test defining the optimal reference image before performing a registration.

Several prior studies pointed out the benefits of using a non-rigid transformation model as B-splines, for example, the possibility of deforming both local structures and global structures. Considering the benefits presented in the literature, the results of this test indicate that the difference between a B-spline transformation and an Affine transformation is not that significant. Therefore, a trend for an optimal transform is hard to identify for the registration

of MOLLI images and MOLLI and DWI images together. There is a slight trend toward using an Affine model for the registration of DWI images. However, the difference between the two models is minimal.

The registration of MOLLI and DWI images together shows a very diffuse result for both different options and datasets, and it is hard to determine a tendency toward one of the options. Consequently, Table 3 showing the mean difference between the resulting Dice value and the original Dice value is misleading since the results vary and are not robust enough to result in findings. A possible explanation for this outcome might be the complexity of the methods. Several parameters can be tuned, both independently and in relation to each other, to improve the outcome of a transformation. In this study, only a few of them are tested and tuned. Therefore, these results must be interpreted with caution since there is a possibility that other parameters can improve the results. Furthermore, the previous tests only consider the best choice for registration of MOLLI images and DWI images separately. Therefore, future research could take a step back to analyse the optimal settings for reference images and grid spacing for registration of MOLLI and DWI images together.

Some of the datasets in this project were initially very well aligned, especially MOLLI images. The results after registration of these sets are, in most cases, either a deterioration, especially applying an Affine pairwise registration, or the same as the original ones, mainly for the two groupwise approaches. The Dice values after registration are still high for those resulting in deterioration. However, they are not as high as they were originally. When visually observing the resulting images, the changes before and after registration are, in these cases, minimal and barely visible to the human eye.

In conclusion, there is no definite answer to what settings are optimal for registration of the images in this study. However, registration of the DWI images shows promising results, even though the optimal settings are hard to conclude. For MOLLI the results are not quite as good but still show potential. Nevertheless, there is room for further investigations to find the best parameters for the registration of MOLLI and DWI images together.

## 5.5 Validation of methods

The two methods used for validation in this project were Dice and visual inspection. In this project, the validation of multiple images has been performed by comparing all masks with each other. Other ways of doing it could instead have been to compare all registered masks to the fixed image  $I_F$  for pairwise registration and, in a groupwise approach, compare all masks to a guessed "gold standard". This would though introduce a bias toward the mask chosen as the ground truth image in the groupwise registrations.

The masks used for validation were created by manually segmenting kidneys by experts.

This implies that a Dice value of 1 equals a registration exactly the same as what the manual would result in. That means that the validation using Dice compares the registration with the manual performance and is not necessarily the best possible. Some images are of lower quality where the edges can be blurry and hard to segment, which means that observer variation is possible. If a registration ends up more correctly aligned than the manual ones, these registrations will be shown as lower Dice values for the registered image than the original value. By visual observations however, it can be seen that this is not true for all cases. During the analyses of the results, it is important to keep in mind that the Dice value is a good method for validation but not incontestable.

Comparing all masks result in several Dice values, and therefore, a mean value is calculated to represent the registration result for all images. A mean value is used due to the insight that a median would result in a total overlap for datasets where more than 50% of the images were totally aligned before registration. A mean value, however, is more sensitive to outliers. Therefore, visual inspection and sanity checks are of importance, and an outlier detector was used to inform if any values would differ significantly from other observations. If outliers were detected, a thorough visual inspection was conducted. Still, the Dice values were all given as a mean for all images registered in a set, which means that the overall score is affected even if only one registered image is off. This is particularly interesting for the first image in the MOLLI datasets, which has an inverted intensity and is observed to have more difficulties to register than the rest. From visual validation, it was observed multiple times how all time instances were well registered except for the first. Example of this can be seen in Figure 19. In a real-world application, it is still possible to manually segment the inverted image in a dataset and then register the rest automatically and still save time and resources. Therefore, it would be interesting to test these methods on the images with similar intensities only. It is nevertheless good to keep in mind when analysing the Dice values that these are based on all images in the dataset and one bad registered image will have a large impact on the resulting Dice value.

## 5.6 Suggestions for future work

There are several things to investigate further based on this study. For instance, it would be interesting to try different sizes and positions of the ROIs to see if that would impact the result. The size of the ROIs used in this project was constant for all data sets. However, they do not have the same position in all scans, since the position of the kidney differ. This implies that different data sets might include varying details of background, for instance, the spine or liver. As a result, the algorithm may try to align the images with focus on the spine instead of the kidneys. Therefore, further tests would be needed where smaller ROIs that include less surrounding details are used.

The tests are performed on fewer number of DWI images and DWI and MOLLI images to-

gether than on only MOLLI images. Therefore, there might be a bias toward the registration of MOLLI images alone. Further research should be undertaken to investigate inter-modality registration of MOLLI images together with DWI images, and not just the two individually. Also, more DWI data is needed to avoid possible bias towards the MOLLI images.

A greater focus on the similarity metrics could produce interesting findings that account more for more accurate registration. As mentioned in Chapter 2, some research suggest a combination of MI and gradient information, something that could be achieved by using the image gradients instead of the images them selves for the registration process. This may be a promising technique and would be interesting to explore.

Despite the promising possibilities with Elastix, different methods and toolboxes could be investigated. Even Elastix itself have more options to explore and more parameters to adjust. Some parts of the registration process have been left out in this study. For example, the optimisation process which is also very complex and has great opportunities to affect the result.

## 6 Conclusion

This study set out to investigate the current promising registration techniques used today for kidney MR images as well as how these affect the registration of MOLLI and DWI images. The elements prioritised to investigate in this study were i) the registration process (groupwise and pairwise registration) ii) transformation models (B-spline and Affine), iii) choice of similarity measure (MI and CC) and iii) tuning of parameters such as the number of histogram bins for MI and the size of grid spacing for the B-spline transform.

This study has identified great potential in image registration of kidney MR images, however further research is still needed. The experiments confirmed that MI is a promising similarity metric for this area while CC is not reaching the same good results. Even when using MI, the registration is sensitive to using images with large differences in intensities. There are naturally intensity differences between MRI images taken at different time instances. However, when one or two images have almost an inverted intensity, it may be good to totally separate these images from the datasets and manually register them afterwards.

Overall, this study urges the complexity of selecting registration settings and parameter tuning. There were difficulties in defining optimal settings for combinations of pairwise and groupwise registration together with Affine and B-spline transformation. There is no definite answer to the impact of the results, therefore further work needs to be done to establish more knowledge behind the effects of these settings.

Considerably more work will need to be done to explore different settings and parameters

to tune. An investigation of the optimisation process could produce interesting findings to improve the overall registration. Future work need more data to test the methods on and also more DWI data to match the amount of MOLLI data in order to avoid bias. Additionally, more work has to be accomplished on inter-modality registration on MOLLI and DWI together, which is not only based on the findings from intra-modality registration to analyse the possibilities fully.

Finally, there are several choices to take into consideration designing a proper registration method for kidney MRI images. Several questions still remain unanswered, and further research is needed to determine the most robust image registration method for reducing the motions in MRI images of kidneys. The findings in this project contribute in several ways to the understanding of image registration of MRI images of the kidneys and provide a base for future research in this field.

## References

- [1] “Glass kidney.” [Online]. Available: [https://quest-eb-com.eu1.proxy.openathens.net/search/139\\_3915085/1/139\\_3915085/cite](https://quest-eb-com.eu1.proxy.openathens.net/search/139_3915085/1/139_3915085/cite) (Accessed: 2022-05-16).
- [2] National Institute of Diabetes and Digestive and Kidney Diseases. Your kidneys and how they work. [Online]. Available: <https://www.niddk.nih.gov/health-information/kidney-disease/kidneys-how-they-work#why> (Accessed: 2022-01-20).
- [3] A. Levin, M. Tonelli, J. Bonventre, J. Coresh, J.-A. Donner, A. B. Fogo, C. S. Fox, R. T. Gansevoort, H. J. Heerspink, M. Jardine *et al.*, “Global kidney health 2017 and beyond: a roadmap for closing gaps in care, research, and policy,” *The Lancet*, vol. 390, nr. 10105, ss. 1888–1917, 2017.
- [4] N. M. Selby, P. J. Blankestijn, P. Boor, C. Combe, K.-U. Eckardt, E. Eikefjord, N. Garcia-Fernandez, X. Golay, I. Gordon, N. Grenier *et al.*, “Magnetic resonance imaging biomarkers for chronic kidney disease: a position paper from the european cooperation in science and technology action parenchima,” *Nephrology Dialysis Transplantation*, vol. 33, 2018.
- [5] L. Xing and Y. Wu, “Qualitative and quantitative mri using deep learning,” 2021. [Online]. Available: <https://appft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fmetahtml%2FPTO%2Fsearch-bool.html&r=1&f=G&l=50&co1=AND&d=PG01&ts1=20210313046.PGNR.&OS=DN/20210313046&RS=DN/20210313046>
- [6] K. E. Keenan, Z. Gimbutas, A. Dienstfrey, K. F. Stupic, M. A. Boss, S. E. Russek, T. L. Chenevert, P. Prasad, J. Guo, W. E. Reddick *et al.*, “Multi-site, multi-platform comparison of mri t1 measurement using the system phantom,” *PloS one*, vol. 16, nr. 6, 2021. [Online]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0252966>
- [7] F. Zöllner, A. Šerifović Trbalić, G. Kabelitz, M. Kociński, A. Materka, and P. Rogelj, “Image registration in dynamic renal mri—current status and prospects,” *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 33, 2020.
- [8] G. Delso, L. Farré, J. T. Ortiz-Pérez, and et al., “Improving the robustness of molli t1 maps with a dedicated motion correction algorithm,” *Scientific Reports*, vol. 11, nr. 18546, 2021.
- [9] K. Chow and R. Thompson, “Myocardial t1 mapping—comparison of techniques,” *MAGNETOM Flash*, vol. 49, s. 96, 2015.
- [10] A. J. Taylor, M. Salerno, R. Dharmakumar, and M. Jerosch-Herold, “T1 mapping: Basic techniques and clinical applications,” *JACC: Cardiovascular Imaging*, vol. 9, nr. 1, ss. 67–81, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1936878X15008670>

- [11] V. Baliyan, C. Das, R. Sharma, and A. Kumar, “Diffusion weighted imaging: Technique and applications,” *world journal of radiology*, vol. 8, ss. 785–798, 09 2016.
- [12] A. D. Elster. Dwi b-value. [Online]. Available: <https://mriquestions.com/what-is-the-b-value.html>
- [13] B. Zitová and J. Flusser, “Image registration methods: a survey,” *Image and Vision Computing*, vol. 21, nr. 11, ss. 977–1000, 2003. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0262885603001379>
- [14] J. Little, D. Hill, and D. Hawkes, “Deformations incorporating rigid structures [medical imaging],” in *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*, 1996. doi: 10.1109/MMBIA.1996.534062 ss. 104–113.
- [15] S. Klein\*, M. Staring\*, K. Murphy, M. A. Viergever, and J. P. Pluim, “elastix: a toolbox for intensity-based medical image registration,” *IEEE Transactions on Medical Imaging*, vol. 29, nr. 1, ss. 196 – 205, January 2010.
- [16] F. P. Oliveira and J. M. R. Tavares, “Medical image registration: a review,” *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 17, nr. 2, ss. 73–93, 2014. [Online]. Available: <https://doi.org/10.1080/10255842.2012.670855>
- [17] P. Devadas, G. Kalaiarasi, and M. Selvi, “Intensity based image registration on brain mri images,” in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020. doi: 10.1109/ICIRCA48905.2020.9183191 ss. 257–262.
- [18] D. J. Eck. Introduction to computer graphics. [Online]. Available: <https://math.hws.edu/graphicsbook>
- [19] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes, “Nonrigid registration using free-form deformations: application to breast mr images,” *IEEE Transactions on Medical Imaging*, vol. 18, nr. 8, ss. 712–721, 1999.
- [20] R. H. Bartels, J. C. Beatty, and B. A. Barsky, “An introduction to the use of splines in computer graphics and geometric modeling.” [Online]. Available: <https://www.osti.gov/biblio/5545263>
- [21] F. G. Zöllner, A. Šerifović-Trbalić, G. Kabelitz, M. Kociński, A. Materka, and P. Rogelj, “Image registration in dynamic renal mri—current status and prospects,” *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 33, nr. 1, ss. 33–48, 2020. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/31598799/>
- [22] J. V. Hajnal and D. L. Hill, *Medical image registration*. CRC press, 2001.
- [23] J. H. Song, “Methods for evaluating image registration,” 2017.
- [24] J. Modersitzki, *Fair: Flexible Algorithms for Image Registration*. Society for Industrial and Applied Mathematics, 2009. ISBN 089871690X

- [25] T. M. Cover, J. A. Thomas *et al.*, “Entropy, relative entropy and mutual information,” *Elements of information theory*, vol. 2, nr. 1, ss. 12–13, 1991.
- [26] D. L. Hill and D. J. Hawkes, “Across-modality registration using intensity-based cost functions,” *Handbook of Medical Imaging: Processing and Analysis*, ss. 537–553, 2000.
- [27] A. Rajwade, A. Banerjee, and A. Rangarajan, “Continuous image representations avoid the histogram binning problem in mutual information based image registration,” in *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006*. IEEE, 2006, ss. 840–843.
- [28] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, “Image registration by maximization of combined mutual information and gradient information,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, S. L. Delp, A. M. DiGoia, and B. Jaramaz, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000. ISBN 978-3-540-40899-4 ss. 452–461.
- [29] A. Andronache, M. von Siebenthal, G. Székely, and P. Cattin, “Non-rigid registration of multi-modal images using both mutual information and cross-correlation,” *Medical Image Analysis*, vol. 12, nr. 1, ss. 3–15, 2008, special Issue on The Third International Workshop on Biomedical Image Registration – WBIR 2006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1361841507000588>
- [30] P. Bourke, “Cross correlation,” *Cross Correlation”, Auto Correlation—2D Pattern Identification*, 1996.
- [31] M. Sonka and J. M. Fitzpatrick, “8. image registration,” ss. 449–506, 2009. [Online]. Available: <https://app.knovel.com/hotlink/khtml/id:kt0087DEA5/handbook-medical-imaging/image-registration>
- [32] P. Artal, “18.6.1.4 grating acuity,” ss. 288–290, 2017. [Online]. Available: <https://app.knovel.com/hotlink/khtml/id:kt011MNAH4/handbook-visual-optics/grating-acuity>
- [33] L. I. Smith, “A tutorial on principal components analysis,” 2002. [Online]. Available: <https://ourarchive.otago.ac.nz/bitstream/handle/10523/7534/OUCS-2002-12.pdf>
- [34] M. Polfiet, S. Klein, W. Huizinga, M. M. Paulides, W. J. Niessen, and J. Vandemeulebroucke, “Intrasubject multimodal groupwise registration with the conditional template entropy,” *Medical Image Analysis*, vol. 46, ss. 15–25, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1361841518300288>
- [35] W. Huizinga, D. H. Poot, J.-M. Guyader, R. Klaassen, B. F. Coolen, M. van Kranenburg, R. Van Geuns, A. Uitterdijk, M. Polfiet, J. Vandemeulebroucke *et al.*, “Pca-based groupwise image registration for quantitative mri,” *Medical image analysis*, vol. 29, ss. 65–78, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1361841515001851>

- [36] J.-M. Guyader, W. Huizinga, D. H. Poot, M. van Kranenburg, A. Uitterdijk, W. J. Niessen, and S. Klein, “Groupwise image registration based on a total correlation dissimilarity measure for quantitative mri and dynamic imaging data,” *Scientific reports*, vol. 8, nr. 1, ss. 1–14, 2018. [Online]. Available: <https://www.nature.com/articles/s41598-018-31474-7>
- [37] S. Klein, M. Staring, and J. P. W. Pluim, “Evaluation of optimization methods for non-rigid medical image registration using mutual information and b-splines,” *IEEE Transactions on Image Processing*, vol. 16, nr. 12, ss. 2879–2890, 2007.
- [38] S. Klein, J. Pluim, M. Staring, and M. Viergever, “Adaptive stochastic gradient descent optimisation for image registration,” *International Journal of Computer Vision*, vol. 81, ss. 227–239, 03 2009.
- [39] M. Pilia, “Groupwise whole-body mr image registration guided by zero-average volume changes,” 2018. [Online]. Available: <https://www.diva-portal.org/smash/get/diva2:1273517/FULLTEXT01.pdf>
- [40] K. O. Babalola, B. Patenaude, P. Aljabar, J. Schnabel, D. Kennedy, W. Crum, S. Smith, T. F. Cootes, M. Jenkinson, and D. Rueckert, “Comparison and evaluation of segmentation techniques for subcortical structures in brain mri,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008*, D. Metaxas, L. Axel, G. Fichtinger, and G. Székely, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. ISBN 978-3-540-85988-8 ss. 409–416.
- [41] W. Crum, O. Camara, and D. Hill, “Generalized overlap measures for evaluation and validation in medical image analysis,” *IEEE Transactions on Medical Imaging*, vol. 25, nr. 11, ss. 1451–1461, 2006.
- [42] M. Bach Cuadra, B. Platel, E. Solanas, T. Butz, and J.-P. Thiran, “Validation of tissue modelization and classification techniques in t1-weighted mr brain images,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2002*, T. Dohi and R. Kikinis, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. ISBN 978-3-540-45786-2 ss. 290–297.
- [43] A. A. Taha and A. Hanbury, “Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool,” *BMC Medical Imaging*, vol. 15, 2015.
- [44] S. K. Warfield, K. H. Zou, and W. M. Wells, “Validation of image segmentation and expert quality with an expectation-maximization algorithm,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2002*, T. Dohi and R. Kikinis, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. ISBN 978-3-540-45786-2 ss. 298–306.
- [45] L. R. Dice, “Measures of the amount of ecologic association between species,” *Ecology*, vol. 26, ss. 297–302, 1945.

- [46] D. Huttenlocher, G. Klanderman, and W. Rucklidge, “Comparing images using the hausdorff distance,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, nr. 9, ss. 850–863, 1993.
- [47] Itk. [Online]. Available: <https://itk.org/>
- [48] elastix. [Online]. Available: <https://elastix.lumc.nl/>