



**UNIVERSITY
OF OULU**

FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Muhammad Talha Ali

**REGISTRATION OF DIFFERENTLY STAINED
HISTOPATHOLOGY IMAGES FOR TISSUE
EVALUATION**

Master's Thesis
Degree Programme in Biomedical Engineering
June 2025

Ali M. (2025) Registration of Differently Stained Histopathology Images for Tissue Evaluation. University of Oulu, Degree Programme in Biomedical Engineering, 72 p.

ABSTRACT

Histopathology image registration is crucial for appropriate tissue analysis, especially for the diagnosis of cancer. The presence of staining variation among these whole-slide images leads to several challenges because the stain differs greatly across various tissue types. There are multiple registration techniques, including intensity-based, feature-based, and deep learning-based methods. Although these methods employ different working mechanisms, they all face several challenges, including poor generalization due to staining discrepancies. To address this problem, this study proposed a highly optimized and robust landmark-based registration technique. The technique utilizes anatomical landmarks and employs thin-plate spline interpolation to achieve precise non-rigid alignment, effectively addressing both tissue deformation and staining variability. Moreover, stain normalization was applied using the deep learning-based model StainNet to evaluate how changes in color distribution in images affect the registration approach. Furthermore, the developed approach was comprehensively compared with two well-known intensity-based methods, Elastix and ANTs, under both normalized and raw image conditions. The results from the experiments showed that landmark-based registration was more effective under both conditions in dealing with staining changes than Elastix and ANTs. While Elastix significantly struggled with normalized images due to its reliance on pixel intensities, ANTs showed increased strength after stain normalization; however, they did not perform as well as the method based on landmarks. Although, landmark-based registration is reliable as it is not influenced by varying staining but it is still challenging to annotate the landmarks with significant accuracy that can be improved in future techniques by using automated landmark annotations.

Keywords: Digital pathology, histopathology image registration, stain normalization, thin plate spline, whole slide imaging.

TABLE OF CONTENTS

ABSTRACT	
TABLE OF CONTENTS	
FOREWORD	
LIST OF ABBREVIATIONS AND SYMBOLS	
1. INTRODUCTION.....	8
1.1. Background and Motivation.....	8
1.2. Research Problem.....	8
1.3. Research Objectives.....	9
1.4. Author’s Contributions and the Role of Artificial Intelligence.....	10
2. LITERATURE REVIEW.....	11
2.1. Histopathology Image Registration Overview.....	11
2.1.1. Staining Techniques in Histopathology.....	11
2.1.2. Concept of Digital Pathology.....	12
2.1.3. Characteristics of Whole Slide Images.....	13
2.1.4. Image Registration.....	13
2.1.5. Challenges in Image Registration.....	14
2.2. Mechanism and Classification of Conventional Image Registration.....	15
2.2.1. Transformation Models.....	15
2.2.2. Intensity Based Registration and Similarity Metrics.....	17
2.2.3. Feature Based Registration.....	18
2.2.4. Optimization Strategies in Image Registration.....	20
2.3. Deep Learning in Image Registration.....	21
2.3.1. Applications for Deep Learning in Histopathology.....	21
2.3.2. Deep Learning Based Image Registration.....	22
2.3.3. Transformers and Hybrid Models.....	24
2.4. Stain Normalization.....	24
2.4.1. Histogram Transformation Based Approaches.....	25
2.4.2. Separated Transformation Based Approaches.....	25
2.4.3. Unified Transformation Based Approaches.....	25
2.4.4. Generative and Diffusion Models.....	26
3. METHODOLOGY.....	28
3.1. Dataset.....	29
3.2. Stain Normalization.....	31
3.2.1. StainNet.....	31
3.3. Image Registration Techniques.....	35
3.3.1. Landmark Based Registration.....	35
3.3.2. Intensity Based Technique Using Elastix.....	38
3.3.3. Intensity Based Registration Using Advanced Normalization Tools.....	39
3.3.4. Evaluation Metrics for Intensity Based Registration.....	40
3.3.5. Software Tools.....	41
4. EXPERIMENTS AND RESULTS.....	42
4.1. Dataset.....	42

4.1.1. Normalization Results	43
4.1.2. Registration Results.....	45
4.1.3. Comparative Study	58
5. DISCUSSION	61
5.1. Summary of Key Findings	61
5.2. Comparison with Existing Literature	61
5.3. Future Directions	62
6. SUMMARY	64
7. REFERENCES	65

FOREWORD

I am pleased to present this research study on registration of histopathology images for tissue evaluation. This thesis works on building a reliable method for matching histopathology images stained in various ways, which is essential for accurate multi-modal tissue analysis. The method developed in this work based on landmark-guided registration to enhance the registration process when there are differences in staining and tissue structure.

I am truly grateful to my supervisors, Dr. Anja Keskinarkaus and Dr. Ziaul Hoque for their valuable feedback, insightful comments, and support. The insights from them improved the technical base of the suggested approach. I am also grateful to the Center for Machine Vision and Signal Analysis (CMVS) for giving me the opportunity to conduct my thesis in a state-of-the-art laboratory. The resources and help provided by CMVS were essential in finishing this thesis work appropriately. I would like to express my special thanks to Prof. Tapio Seppänen for his support during my studies.

I am thankful to the Faculty of Information Technology and Electrical Engineering (ITEE) at the University of Oulu, Finland for equipping me technical and academic foundation required to pursue this research. Finally, I am grateful to my parents, siblings, and friends for always facilitating me, giving up a lot, and keeping me motivated in my studies.

Oulu, June 10th, 2025

Muhammad Talha Ali

LIST OF ABBREVIATIONS AND SYMBOLS

DP	Digital Pathology
WSI	Whole Slide Images
H&E	Hematoxylin and Eosin
AI	Artificial Intelligence
DL	Deep Learning
CAD	Computer Aided Diagnoses
IHC	Immunohistochemical
3D	Three-Dimensional
2D	Two-Dimensional
PHH3	Phosphohistone H3
Ki67	Convolutional Neural Network
mpp	micrometers/pixels
GB	Giga-Byte
DICOM	Digital Imaging and Communications in Medicine
DAB	Diaminobenzidine
SSD	Sum of Squared Differences
NCC	Normalized Cross-Correlation
MI	Mutual Information
SIFT	Scale-Invariant Feature Transform
SURF	Speeded-Up Robust Features
FAST	Feature from Accelerated Segment Test
BRIEF	Binary Robust Independent Elementary Features
L-M	Levenberg-Marquardt
GD	Gradient Descent Method
CG	Conjugate Gradient Method
QN	Quasi-Newton Method
GANs	Generative Adversarial Network
IKCG	Iterative Key Point Correspondence Guided
GPU	Graphics Processing Unit
ViTs	Vision Transformers
PViT-AIR	Puzzling Vision Transformer-Based Affine Image Registration
COAD	Colon Adenocarcinoma
CC	Cross Corelation
ANHIR	Automatic Non-rigid Histological Image Registration
TRE	Target Registration Error
rTRE	Relative Target Registration Error
ReLU	Rectified Linear Unit Activation Function
MSE	Mean Squared Error
SSIM	Structural Similarity Index
PSNR	Peak Signal to Noise Ratio
TPS	Thin-Plate Spline
RBF	Radial Basis Function
ANTs	Advanced Normalization Tools

BIRL	Benchmark for Image Registration Libraries
NMF	Non-negative Matrix Factorization
SNMF	Sparse Non-negative Matrix Factorization
SyN	Symmetric Diffeomorphic
LDM	Latent Diffusion Model
CLIP	Contrastive Language Image Pre-training
csv	Comma Separated Values
B-Spline	Basic Spline
N	Number of landmarks
MSE	Mean Squared Error
x, y	Pixel coordinates
C	Constant
μ	Mean intensity
σ	Variance of patch
$\Phi(r)$	Thin Plate Spline (TPS) Radial Basis Function
δ	Displacement field components
*	Convolution operation
remap()	OpenCV function

1. INTRODUCTION

1.1. Background and Motivation

In this era of technological advancement, cancer is still incurable without timely diagnoses. It is one of the rapidly spreading diseases in the world. In the field of digital pathology (DP) and cancer research the high resolution multistained histopathology images play a vital role [1]. These histopathology whole-slide images (WSIs) of tissue sections offer high resolution features required for disease diagnosis and analysis [2]. Before the slide preparation, the tissues are stained differently, mostly including hematoxylin and eosin (H&E) to highlight the tissue morphology more clearly [3]. These WSIs can be used as an essential diagnostic tool for determining diseases based on cellular assessment. To improve and automate this diagnosis process, image registration is performed on these differently stained images to produce a more detailed single image for better visualization [4]. By fusing scans stained in various conditions, medical professionals can extract detailed structural and pathological insights [5].

The goal of image registration is to find a geometric transformation that aligns corresponding objects in two or more images [6, 7]. The resulting registered image combines features of different staining conditions and highlights fine details. This helps in the identification of possible malignancies within histopathology images timely which leads to improved survival outcomes for patients. There are several existing registrations algorithms, but they struggle due to staining variations, non-rigid deformation in tissues and lack of distinctive features in histopathology images [8]. The existing registration algorithms require legit solutions with better generalization capabilities to overcome their current limitations on diverse tissue samples and staining variations. The subfield of Artificial Intelligence (AI), deep learning (DL), is progressing day by day in this domain to overcome this poor generalization issue [2, 9]. DL techniques perform well with large amounts of training data. Implementing a versatile DL tool to these WSIs is challenging due to limited publicly available images and patient privacy concerns.

1.2. Research Problem

The examination of tissue structures in pathology is usually tedious and subjective [9]. The identification and outline of specific structures in histopathological images done manually by pathologists heavily rely on their experience and visual inspection expertise [2]. Numerous research has established that manual methods take too long to perform and have both human errors [10], and inconsistent interpretations between different experts [2]. The introduction of a computer-aided diagnosis (CAD) scheme during the early 1980s provided doctors and clinicians with a better understanding of medical images [9]. Furthermore, the use of emerging technology in the field of DP not only changed the way of diagnosis but also made it effective and error free [2]. The examination of tissue samples through staining methods allows pathologists to identify disease-affected tissues under microscope observation [7]. The medical examination of tissue samples becomes possible through staining techniques, which provide transparent and colored visualizations to pathologists and doctors [11]. The application

of histological stains and Immunohistochemical (IHC) stains produces different visual outcomes in surrounding tissue areas because it highlights changes effectively [7, 12]. The histological staining process most commonly uses hematoxylin, which turns cells and nuclei blue or purple while eosin produces pink coloration in cytoplasm, collagen and muscle fibers [7, 13]. Once the slides are prepared, they are scanned through specialized scanners to produce the virtual slides or WSIs [7]. The color appearance of slide images varies depending on the entire process, including specimen preparation, staining solution, stain coloring, chemistry oxidation, staining time, pH level of solutions, manufacturer reactivity, and use of various scanners [7].

Thus, due to staining inconsistencies, automated analysis faces major drawbacks which require precise image registration techniques for accurate alignment. The WSIs need to be registered for multiple reasons to build 3D models from 2D scanned thin sections [5]. The applications of registration include distortion correction as well as the production of high-resolution composite images from small 2D tiles and the fusion of data from distinct stained sections or different imaging modalities. The importance of histopathology image registration exists, although it faces some implementation hurdles because of several reasons that are mentioned below.

1. Large image sizes usually in gigabytes for WSI [5].
2. Repetitive texture that hinders the identification of globally unique landmarks [5].
3. Non-linear elastic deformation together with sample preparation issues that cause occlusions and missing sections [5].
4. Variations in appearance due to staining procedures [5, 14].
5. Differences in local structure between slices [5, 14].

Traditional registration techniques are divided into two main groups that are feature-based registration and intensity-based registration [7]. Their performances are still struggling with the staining variation, and non-rigid deformations present in the WSI resulting in lack of generalization. To deal with this problem the progress in DL, specifically CNN, has revolutionized histopathological image registration. In comparison with the traditional methods, CNNs learn complex patterns and variations directly from data via convolution, capturing subtle traces in tissue morphology such as color variations, texture differences, and spatial relationships [9]. As a result, CNN-based methods can attain high registration accuracy and consistency. Nonetheless, there is need of huge amount of data for training the CNNs efficiently. Data shortage is also a big concern, mainly for the specialized tissue types.

1.3. Research Objectives

After considering the above details. The key goal of this work is to develop a registration method for WSI stained with different staining protocols. The new method shall overcome major problems such as staining variability, non-rigid tissue distortions, and the limitations of traditional registration methods. Specifically, this study aims to compare multiple existing registration methods and their limitations.

1. Use of novel and adaptive stain normalization technique to introduce the variability for experimentation.

2. Development of an optimized image registration approach that can handle the key issue of stain variation in histopathology images.
3. Improve diagnostic reliability by reducing errors resulting from stain variability and improve automated pathology workflows.
4. Comparison of developed method with pre-existing state-of-the-art techniques.

By fulfilling these goals, this work will facilitate the progress of DP through the improvement of accuracy, speed, and reliability of image registration for histopathology.

The thesis follows the following structure: The entire Chapter 2 is based on a literature review covering all aspects, including a detailed overview of the domain, challenges in image registration, mechanisms and classification of conventional image registration, and DL in image registration. Chapter 3 then covers the methodology, including implementation and architectural details utilized in the thesis. Chapter 4 covers the acquired results and comparisons, while Chapter 5 includes the discussion and comparison with the literature review. Lastly, Chapter 6 concludes the thesis.

1.4. Author's Contributions and the Role of Artificial Intelligence

I have used ChatGPT in my work to understand research papers and also to improve my writing structure. Besides, I have checked the GitHub repositories during development to understand some of the key approaches and for a comparative study, I have utilized the BIRL environment to run the experimentation.

2. LITERATURE REVIEW

2.1. Histopathology Image Registration Overview

Histopathology is the practice of using a microscope to analyze tissue samples and discover signs of disease [14, 15]. It is one of the key methods for the diagnosis of cancer. Pathologists can diagnose numerous diseases, recommend appropriate therapies, and estimate patient outcomes by carefully studying the structure of tissues and the appearance of cells [16]. The process initiates when tissues are processed and stained so that specific structures are made visible [17], then they are examined in detail under a microscope. Within the context of clinical practice, histopathology is the microscopic examination of biopsy samples by a pathologist after the specimen has been prepared, with histological sections stained and mounted on glass slides to bring out fine cellular features [17]. The slices used are extremely thin [18], and principally capture only 2D representations of a certain tissue sample. Further, the use of staining has greatly improved diagnostic reliability by clearing pathologists to find certain biomarkers in tissue sections [15]. The advancements in staining techniques and imaging technologies have further enhanced the ability to identify subtle tissue changes, enabling earlier detection and more accurate diagnoses.

2.1.1. Staining Techniques in Histopathology

Staining is a fundamental technique used in histopathology to visually enhance the structures of tissues. There are many different types of staining techniques but the most commonly used is H&E also considered as a gold standard [12, 18]. The hematoxylin gives the purple effect by binding to nucleic acids, in contrast eosin binds with cytoplasm and extracellular matrix in different variation of pink [7]. Another popular technique known as immunohistochemistry utilizes the antibodies that are combined with specific antigens within tissue [7, 12]. It has been widely used in a large number of pathological evaluation procedures to visualize disease-related molecular biomarkers without losing spatial tissue reference due to its affordability, ease of use, and variety of preparation methods.

Histochemistry uses the cellular and subcellular chemical environment of the sample to bind particular chromophores to a specific tissue component, producing color contrast under a visible light microscope that serves as the foundation for pathologists' and professionals' diagnosis of abnormalities in tissue specimens [12, 16]. For example, the Ki67 biomarker from IHC is frequently used to predict the pathological therapeutic response to neoadjuvant chemotherapy and the recurrence-free survival rate in breast cancer [12]. The Phosphohistone H3 (PHH3) is another common IHC biomarker that the Nottingham Grading System uses to evaluate the mitotic activity useful for breast tumor staging [12]. Figure 1 shows two differently stained slides of lung lobe tissue, one stained with H&E and the other with the Ki67 biomarker from the ANHIR dataset [5].

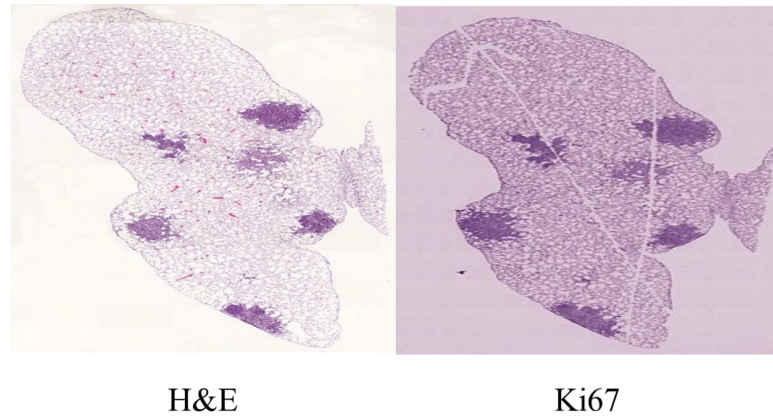


Figure 1. Example different stained images of the lung lobe tissue stained with H&E and Ki67 biomarker from the ANHIR dataset [19], licensed under CC BY-NC-SA.

2.1.2. Concept of Digital Pathology

The introduction of digital pathology (DP) has greatly changed the traditional methods by developing a digital environment [10], focused on methods for acquiring, storing, and interpreting pathology slides through computational approaches. In general, DP refers to an integrated digital infrastructure and imaging tools that support the creation of a digital image-based practice, allowing for the efficient handling, interpretation, and search of WSIs or other digital images. This new methodology is being used more frequently in clinical research, educational institutions, and clinical trials. Because WSIs can now be scanned so quickly, they are able to replace traditional microscopy-based pathology. With this method, it is possible for pathologists to digitally review and compare several stained sections at the same time. This feature becomes very important when combining information from several IHC stains, allowing for a better assessment of biomarker expression.

In this workflow, glass slides are digitized using high resolution scanner and converted into high-resolution images, also known as WSIs [10, 20]. Further, these WSIs are examined by the pathologist on computer screens. The switching to digital platforms from traditional microscopy techniques gives several advantages such as better workflow, simpler remote consultations, and the use of computers to support diagnosis [10]. In addition, the advancement in the domain of image processing leads to the development of major techniques that bring some highly impactful outcomes [10], including image registration, image segmentation, and many others. These are only possible because of the digitization. The sections on glass slides in histopathology are quite thin and only show a two-dimensional image of the tissue. Aligning these sequential two-dimensional slices is highly significant, because it helps pathologists to see the bigger three-dimensional picture. Moreover, WSI makes it easier to use DL algorithms for tasks including virtual staining, precise image segmentation, and tissue classification, which greatly boost diagnostic precision and speed.

2.1.3. Characteristics of Whole Slide Images

WSI is the most widely used form of digitalized histopathological data [21]. These scans are very high dimensional and have large data size [22]. An average sample ranges from 20 mm x 15 mm in dimension, which normally digitalized with a resolution of 0.25 micrometers/pixels (mpp), resulting in an image approximately $80,000 \times 60,000$ pixels in size, equating to a 4.8 giga-pixel image [23, 23]. The following scans are normally captured with 24-bit color, so the resulting image data size is about 15 giga-byte (GB) [23]. The sample sizes can be up to 50 mm x 25 mm from the common 1x3 slides. Moreover, these can be digitized at resolution higher than 0.25 mpp [23]. The oil immersion lenses can capture images up to 100x magnification [23]. Due to the huge amount of information on a microscope slide, pathologists cannot view an entire sample at high resolution [20]. They initially inspect the slide at a relatively low resolution, typically 5 mpp, and gradually zoom into higher resolution for the selected areas of diagnostic significance.

Digital Imaging and Communication in Medicine (DICOM) is an interchange standard that is most commonly used for these scans [23, 24]. The 2D images are organized in a single frame structure, with image data arranged in rows that span the full image [20]. The below Figure 2 shows the simplest organization of WSI [25]. The advanced method of storing 2D picture data is tiled organization, to ensure the image data is arranged in square or rectangular tiles for diagnostic friendly experience [20].

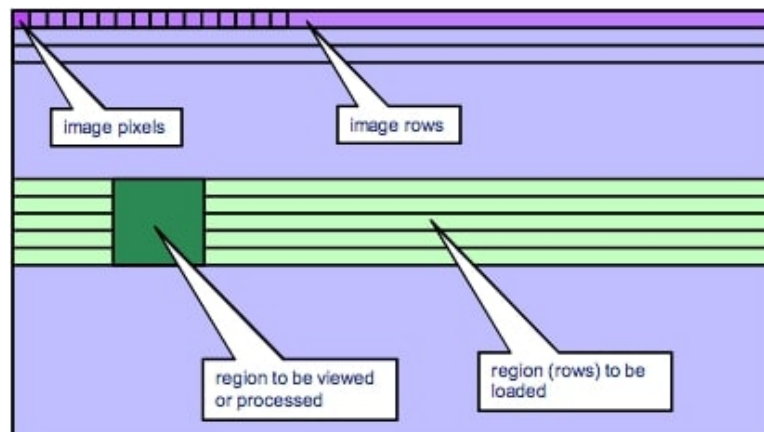


Figure 2. The single frame image organization of an example whole slide image. A picture taken from DICOM website [25].

2.1.4. Image Registration

One of the most significant tasks in biomedical image analysis involves image registration [26], looking for the identification of a geometric transformation between two or more images to have the corresponding projections appear in both images at the same location [5, 7]. It is a very difficult and optimization procedure of using multiple geometric transformations [7]. It is an essential component of numerous detection works, including the image viewpoint detection, fusion, area detection,

similarity monitoring, feature detection, and feature matching [7]. During the registration procedure, the target images must have a few common structures that are projected to be located in similar spots, locations, and orientations [7]. Image registration is used in a variety of analytical studies, such as in studies to determine the extent of tumor tissues, to calculate longitudinal changes, identify cells or tissues impacted by the disease, relate individual anatomy to a standard system, and to carry out guided medical procedures, among other applications [7]. This is crucial for quantitative pathology: once images are aligned, one can correlate spatial patterns of morphology and molecular markers, transfer annotations between stains, or create virtual multiplexed images combining multiple biomarkers [27]. In order to generate a three-dimensional reconstruction from scanned thin slices, digital histopathology images must be registered [7]. Typically, a pathologist examines a lot of slides and then compiles visual information to make a diagnosis [7]. When applying special or immunohistological staining, sequential slides reveal integral information requiring image registration stage to enable an efficient examination [7].

In summary, cross-stain WSI registration is a foundational step for downstream tasks like spatially aligned biomarker co-expression analysis, multi-modal tumor microenvironment studies, and building 3D histopathology atlases [28].

2.1.5. Challenges in Image Registration

The registration of WSIs is commonly considered as a difficult computer vision problem for which no consistent out of the box solution is available [29]. There are several challenges associated with image registration. These challenges are discussed below.

Huge size of WSIs and computational cost: The WSIs are extremely high-resolution, most often on gigapixel-scale, which results in direct processing of them computationally heavy [5, 12]. The classical registration methods might take hours to align two WSIs if performed simply at full resolution. In this regard, DL offers a speed advantage at execution time once a model is trained, such as a CNN that predicts deformation fields. In this regard, DL offers a speed advantage at execution time once a model for predicting deformation field is trained; it can produce an alignment in seconds to minutes [30].

Tissue deformations and missing correspondence: The tissue sections undergo non-linear deformations [29], during preparation cutting, mounting, and staining which can introduce stretching, tearing, or shrinking of tissue, so consecutive sections do not align by simple rigid shifts [12].

Stain variability: In digital pathology, one of the key challenges is stain variability among slides. The staining variations between slides mean the intensity and color distributions can differ considerably for instance, an H&E slide has purple or pink hues showing nuclei and cytoplasm, whereas an IHC slide might have brown Diaminobenzidine (DAB) chromogen highlighting a specific antigen on a subset of

cells [12]. This multi-modal nature violates the assumptions of many registration algorithms that expect similar intensity patterns in both images.

Artifacts and inconsistencies: There are various artifacts and inconsistencies that can occur. The differences in slide orientation [29], that could result in rotation and scaling differences between scanners. In addition, sometimes tissue folds, bubbles, or partially missing regions appear on the slide that affect the image quality. All these issues can confound automatic alignment. In fact, a recent study mentions that the overwhelming dimension of the image, the intricate histological structural change, and the notable variation in tissue appearance in various stains present technical obstacles to the registration of WSI [12].

Despite these challenges, robust registration of differently stained WSIs is highly required after, as it enables more precise comparative diagnostics and research analyses. The next sections review the types of image registration methods and their suitability to this domain, as well as the typical workflows and recent advances addressing the above challenges.

2.2. Mechanism and Classification of Conventional Image Registration

Over the years, traditional image registration methods have been used to align tissue images, for instance Elastix and ANTs. The traditional registration algorithm is either constructed as continuous optimization problem or discrete optimization problem [31]. Methods usually consist of three main components: the transformation method, similarity metrics, and the optimization strategy. These methods are generally grouped on how similarity between the images is calculated into the following two general categories: intensity based, and feature based [7, 32].

2.2.1. Transformation Models

The fundamental component of any registration method is the selection of a geometric transformation model that maps the point from the moving image to the fixed image. The transformation defines the allowed motion or deformation of the tissue required to align the images. These types include Rigid Transformation, Affine Transformation and Non-Rigid Transformation [33]. The details of these transformations are given below,

Rigid Transformation: This is the simplest transformation model compared to others. This transformation is based on the rotation and translation [34]. During the transformation it preserves distance and angles [34], and there is no scaling and shearing. Rigid models are computationally efficient, require only a few parameters and are often used for initial alignment [34]. However, they cannot accommodate

any tissue deformations. For differently stained serial sections, a pure rigid alignment usually leaves many local misalignments because tissue undergoes non-uniform distortion. It is mostly suitable for registering rigid body scans like bones.

Affine Transformation: An affine model extends rigid transforms by including scaling and shear [33]. It allows for linear distortions, making it suitable for slightly deformed tissues or images captured from different angles. This added flexibility can be correct for global stretching or compression of tissue. It remains relatively fast to compute and is a common choice for an initial alignment stage [35]. In histopathology, the affine alignment can account for differences in slide rotation, magnification, and aspect ratio between scans. Nevertheless, like rigid transforms, an affine mapping will typically not perfectly align fine structures if the tissue has bent or non-linearly distorted. Further finer adjustments are often needed after an affine step.

Non-Rigid Transformation: This transformation is also known as a deformable transformation. Deformable models allow local, non-linear warping of the moving image to align it with the fixed image. Rather than a single global formula, non-rigid transformations are often defined by interpolating a dense displacement field or using basis functions. One popular parametric approach is free-form deformation with Basis spline (B-splines) [34]. In this, the image is overlaid with a coarse control point grid, and each grid node can move, and a smooth spline interpolation defines the pixel-wise displacement from those node shifts [35]. Another approach is thin-plate spline (TPS) warping, which fits a smooth elastic sheet through the corresponding landmark points. The downside is computational cost and the risk of misregistration if not properly regularized. With so much freedom, a deformable algorithm might align incorrect features, especially in multi-stain scenarios where some structures exist only in one stain or introduce biologically implausible shape changes. Regularization and physics-inspired constraints like diffeomorphic mappings are used to keep the deformation smooth and realistic. Despite these challenges, deformable registration is considered essential for high-accuracy alignment of serial WSIs, since rigid or affine alone cannot handle complex non-linear differences between sections [35]. Indeed, in a large study consisting of 481-pair study for registration from the ANHIR challenge, the top-performing methods all employed non-rigid refinement, with the diffeomorphic ANTs algorithm achieving the best accuracy at the cost of long runtime [5].

Each transformation model has trade-offs in accuracy and complexity comparison. The simpler transformation models are fast and robust to overfitting but often insufficient for precise alignment of differently stained tissue sections. In contrast, complex models can achieve pixel-level alignment of fine structures at the expense of greater computational load and the need for careful validation. In practice, many WSI registration pipelines use a multi-stage strategy that starts with a rigid or affine alignment to achieve the positioning correctly, then apply a deformable refinement on top [35]. This leverages the strengths of each model.

2.2.2. *Intensity Based Registration and Similarity Metrics*

Intensity-based methods, also known as voxel-based methods consider image registration as an optimization problem. They find the transformation parameters that maximize a statistical similarity between the fixed and moving image intensities [7]. These approaches do not explicitly identify correspondences between specific points instead the entire intensity pattern contributes to the alignment score. Similarity metric that are most used in traditional image registration are Sum of Squared Differences (SSD), Normalized Cross-Correlation (NCC), and Mutual Information (MI).

Sum of Squared Differences: The SSD is considered as simplest and efficient [36]. It is one of the measures of match that is based on pixel-by-pixel intensity differences between the two images [37]. This metric assumes the SSD is minimum when two images become identical and perfectly aligned. SSD is effective for mono-modal registration under consistent lighting. Mono modal are the images that come from the same imaging modality. However, for different stains, sometimes SSD is problematic, for instance a dark pixel in an H&E image might correspond to a pixel that is blank in an IHC image, yielding a large intensity difference even when correctly aligned. Thus, SSD is very sensitive to intensity and color differences, so it generally fails for multi-stain WSI alignment unless the images are first normalized to have similar intensity ranges, or one extracts a common intensity channel present in both. SSD also treats all intensity differences as noise to eliminate, which may not be appropriate if each stain has a unique contrast for different structures.

Normalized Cross-Correlation: The NCC measures the linear correlation between intensities of the two images. NCC is invariant to uniform brightness and contrast changes between images [38] for instance if one image is globally darker, it can still yield a high NCC when aligned, unlike SSD. This makes NCC somewhat more robust than SSD to staining differences [37], especially if the stain variation is mainly an intensity scaling. But NCC still assumes a linear relationship between the two images intensities. In histology, different stains may highlight completely different structures, violating this assumption. For example, nuclei might be dark in both H&E and an IHC stain with hematoxylin counterstain, but an IHC's DAB precipitate marks only certain cells that have no direct counterpart intensity in H&E. In such cases, NCC may not find a meaningful maximum, as large portions of the images are uncorrelated. NCC can be effective if applied to images that have been processed to enhance commonly visible structures or if focusing on a single channel common to both like aligning based on the hematoxylin channel present in both H&E and some IHC slides. In practice, normalized cross-correlation is often used for registering re-stained slides or images where one channel is consistent across stains. It is also used within pyramidal coarse alignment since it's fast to compute via convolution methods.

Mutual Information: The most widely used similarity metric for multi-modal image registration is mutual information. It is a statistical metric derived from the information theory [39]. It shows the amount of information that one random variable provides about another random variable and calculates the statistical dependence of two random

variables [39]. It is one of the very successful techniques for multimodal medical image registration. It is not necessary to make any assumptions on the relationship between the pixel cells in images [39]. Mutual information has been demonstrated to be the most reliable and accurate way of multimodal medical image registration without the need for any previous feature extraction, segmentation or pre-processing of the images [39]. MI is also considered as one of the requirements for registration [40]. MI is a mathematical metric that measures how much information about one dataset can be acquired from understanding the distribution of another dataset [40]. One practical challenge with MI is computational cost. Computing MI involves estimating joint intensity histograms or probability distributions at each iteration of the optimizer [39]. Overall, MI provides a robust criterion for cross-stain WSI registration, especially for global alignment stages.

Intensity-based registration typically involves choosing one of the above metrics and then using an optimization algorithm to hunt for the transformation that maximizes the similarity or minimizes the cost derived from it. Common optimizers include gradient descent, simplex methods like Powell’s method. Powell’s method performs a series of 1-D bracket searches along multiple directions and doesn’t require gradient computation [39]. Because WSIs are huge, intensity methods are almost always implemented in a multi-resolution pyramid: the images are first heavily downsampled to find an initial alignment, then progressively higher resolutions are registered, using the result of the previous level as initialization [27]. This coarse-to-fine approach greatly speeds up convergence and improves robustness. Even so, intensity-based WSI registration can be slow for instance, the ANTs deformable algorithm using Mattes’ mutual information achieved top accuracy but with very long runtimes on WSIs [35]. Another consideration is masking out irrelevant regions. Most often the background of slides is excluded from the metric computation using a tissue mask, to focus the metric on overlapping tissue only [27].

The intensity-based methods generally require the images to be roughly pre-aligned; if the initial offset or rotation is too large, the optimizer can get trapped. That is why intensity-based methods are often combined with feature-based methods for a two-stage process. In summary, intensity-based registration leverages all pixel data and can achieve very accurate alignments when the similarity metric is appropriate for the modality. For differently stained slides, mutual information is the preferred metric due to its ability to handle different intensity mappings. Nonetheless, pure intensity-based approaches can struggle if the images have very little overlapping information.

2.2.3. Feature Based Registration

Feature-based registration focuses on identifying and matching distinct features within images to determine the spatial transformation required for alignment. This approach is particularly effective when dealing with images that have significant differences in intensity values or originate from different modalities. The key steps involved in this area are feature detection, feature description, feature matching, transformation estimation and image resampling and transformation.

Feature detection involves identifying salient features such as corners, edges, blobs, or other distinctive structures in both reference and target images [7, 41]. An ideal

feature detection technique should be robust to image transformations such as rotation, scale, illumination, noise and affine transformations [41]. Common detectors include the following.

Scale-Invariant Feature Transform (SIFT): SIFT recognizes features invariant to scale, rotation, and illumination changes. Notably, SIFT features have been widely used for WSI registration [7]. SIFT can often detect corresponding structures in H&E and IHC slides, for instance the outline of a group of nuclei or a tissue boundary will generate similar gradient patterns in both images regardless of color. Hoque et al. (2022) specifically developed a multi-stained feature matching algorithm based on SIFT for WSIs, with adaptive smoothing to account for stain differences [7].

Speeded-Up Robust Features (SURF): SURF is an accelerated version of SIFT, offering faster computation while maintaining robustness [42]. It is up to some point similar in concept as SIFT, in that they both focus on the spatial distribution of gradient information [42]. It integrates the gradient information within a sub patch, on the other hand SIFT depends on the orientations of the individual gradients [42]. This trait makes the SURF less sensitive to noise [42].

Oriented FAST and Rotated BRIEF (ORB): ORB combines the Feature from Accelerated Segment Test (FAST) key point detector and Binary Robust Independent Elementary Features (BRIEF) descriptor, optimized for performance and efficiency [41]. In the beginning to find out the key points it utilizes the FAST [41]. FAST does not compute the orientation and is rotation variant [41]. It measures the intensity weighted centroid of the patch with located corner at the center [41].

Further step is feature description. Once features are detected, they are described using vectors that capture the local image appearance around each feature point. These descriptors must be robust to variations in imaging conditions. Then feature matching establishes correspondences between feature descriptors from the reference and target images [43]. Techniques like the Nearest Neighbor Search or Hamming distance are commonly employed. Transformation estimation computes the spatial transformation that best aligns the matched features. Robust estimation methods like Random Sample Consensus (RANSAC) help in mitigating the influence of outliers. Then, image resampling and transformation apply the estimated transformation to warp the target image, aligning it with the reference image [43].

Feature-based methods work in both manual extractions and DL based registration approaches [44]. Traditional methods such as SIFT and ORB use gradient, grayscale, and other information to manually design feature points for matching. Nevertheless, this technique concentrates on local features and yields sparse feature points. In contrast, deep learning-based feature extraction algorithms consider characteristics of diverse sizes and scales, contributing to more comprehensive features [44].

There are two types of DL techniques: detector-based and detector-free approaches [44]. Whether a feature extraction network is required determines how these two categories vary from one another. The detector-based approach uses the response values of the network to detect interest points. It constructs descriptors using features extracted by CNN and U-Net, such as SuperPoint [45]. These interest

points and descriptors then match and get the corresponding feature pairs. Detector-free techniques, such as correspondence transformer-based image matching network, employ the designed network for end-to-end matching without requiring a dedicated detector to identify interest sites [45].

Feature-based registration is usually faster than intensity-based for large images because it reduces the data to a sparse set of points and descriptors. Instead of comparing millions of pixels, one might match a few thousand features. Additionally, key point methods are more forgiving to large initial misalignments. They can detect features all over the image and match even if one image is rotated or translated far relative to the other, whereas intensity methods might fail to converge if starting too far off. The limitation of feature methods is that they may yield only a coarse alignment if features are too sparse. Fine details not captured by the features will remain misaligned unless further refinement is done. Also, the accuracy of alignment depends on feature localization accuracy, which for blob features might be a few pixels. For WSI, a few-pixel error at low resolution could translate to dozens of microns at full resolution. Therefore, pure feature-based alignment might not suffice for precise cell-level mapping, often a downstream deformable adjustment or an intensity-based fine registration is applied on top of the feature-based result. Both feature-based and intensity-based registration methods have their unique strengths and are chosen based on the particular requirements of the imaging scenario.

Recent advancements, particularly the integration of machine learning techniques and the development of hybrid methods, have further enhanced the capabilities and applicability of these registration approaches. As the field progresses, the combination of traditional methods and modern computational techniques continues to bring improvements in accuracy, efficiency, and robustness of image registration processes.

2.2.4. Optimization Strategies in Image Registration

The correlation of spatial transformation between the images is the basis for image registration, which can be transformed from a registration problem into a mathematical parameter optimization problem [46]. The objective of the optimization method is to improve the similarity measure degree and speed up to reach the ideal value [46]. It is also considered as a core stage in the image registration workflow [47]. Numerous optimization techniques employ the gradient direction to find out the update step that might be not optimal [47]. An imperfect update may lead to a lot of iterations, which would greatly lengthen the computation time or reduce the precision of the registration results. A better algorithm that offers more accurate iteration steps is Levenberg–Marquardt (L-M) optimization [47].

Nevertheless, the L-M technique has a significant computational complexity due to the huge number of factors in nonrigid image registration. The steepest descent method is a popular optimization technique for image registration; it converges quickly at the beginning but slows down when the parameter vectors are approaching the solution set. The Gauss-Newton approach provides a more accurate search direction by determining the update step by utilizing both the first and second order differential of the desired function. The geometric transformation in image registration corresponds to the set of transformation parameters that are optimized during the registration process [46]. They

are further categorized into two categories, that is linear transformation and non-linear transformation [46].

Linear transformation: When the internal structure of an image does not exhibit visible distortion or deformation, linear transformation is utilized for the registration of bone structures and tissues. It is carried out by computing translational and rotational vectors [46].

Non-linear transformation: The primary advantage of non-linear transformation over linear transformation is its ability to register local deformations, which results in more accurate registration [46]. Most commonly, optimization methods iteratively modify the transformation parameters until a predefined similarity measure reaches its optimal value or convergence criteria are met. Depending on the type of variable, optimization methods can be divided into two groups: continuous variables and discrete variables [46].

In a nutshell, the continuous variable optimization method solves the differential in an objective function, with parameters that can be set in a variety of ways. The gradient descent method (GD), conjugate gradient method (CG), Powell optimization method, Quasi-Newton method (QN), Levenberg-Marquardt method (LM), and stochastic gradient descent method are the often-employed techniques in the optimization of continuous variables [46]. On the other hand, only discrete values can be used for the discrete optimization of variables. Discrete variable optimization basically use the Markov random field to carry out the optimization in image registration applications [46].

2.3. Deep Learning in Image Registration

Deep learning is an area of machine learning that concentrates on algorithms inspired by the structure and function of the human brain, known as artificial neural networks. It involves training deep neural architectures composed of multiple layers to automatically learn hierarchical feature representations from data [48]. By leveraging large datasets and the backpropagation algorithm, deep learning models have achieved dramatic improvement in tasks like image recognition [48]. In DP, DL has become a transformative tool, enabling automated analysis of WSIs for diagnosis, prognosis, and research. Indeed, pathology is considered one of the medical fields broadly transformed by DL, known its ability in image analysis [49].

2.3.1. Applications for Deep Learning in Histopathology

The integration of DL into DP has continued to drive the growth in the field. Staining tissue samples manually in histopathology is a labor-intensive, expensive, and time-consuming procedure [16, 50]. Also, creating staining variability among different

labs which may result in misdiagnoses [16]. Further, the manual staining technique is prone to degrading [16, 50]. DL supports virtual staining and provides real histological imaging outcomes without needing the traditional physical staining procedures [16]. In addition, there are several image processing techniques that rely on DL. Recent advances in computational models for robust registration and localized co-expression analysis of several whole-slide images have enhanced the analytical abilities in terms of the serial histology sections stained differently. These frameworks allow for precise and accurate alignment of tissue sections, which makes accurate co-localization of the biomarkers, and more detailed and comprehensive tissue analysis possible. Stain normalization is another technique used to counter the staining variability mostly based on DL that produces good results [1, 3, 51].

Overall, the transition to DP, facilitated by the progress in WSI and DL technologies, has significantly improved the possibilities of the analysis and interpretation of the histopathological images. These advanced methodologies make it easy to integrate and register differently stained images to generate a 3D structure from 2D slides hence allowing for a better understanding of tissue pathology and increased diagnostic accuracy and future treatment [7].

2.3.2. Deep Learning Based Image Registration

Unlike traditional registration methods that often struggle with large appearance differences between stains, deep models can learn complex, stain-invariant features or even translate images to a common stain domain, thereby improving cross-stain alignment accuracy [12, 52]. These CNNs, Generative Adversarial Networks (GANs), and more advanced architectures that include the transformers and diffusion models have been applied to better handle multi-stain variability in WSIs. When compared to traditional methods, data-driven methods dramatically benefit from many paired and unpaired images, which is why DL-based deformable registration methods have recently drawn a lot of interest from researchers [31]. DL-based models can improve deformable registration functionality. First of all, deep neural networks are capable of accelerating the training phase of the iterative optimization process and achieving rapid inference during the test phase [31]. Second, neural networks can assist with registration by approximating the similarity between image pairings.

Furthermore, without first constructing a transformation model, the deformation field can be predicted directly using an end-to-end model [31]. These methods can be split into three main categories that are supervised, unsupervised and adversarial registration [53]. Ground-truth deformation fields are needed for supervised registration. The generator and discriminator networks that form the basis of the adversarial technique have similar drawbacks [53]. Furthermore, training adversarial networks is not simple. Registered image pairings that represent the ground-truth alignment are required in order to train the discriminator. For some tasks, it can be achieved [53]. However, it is typically expensive and time-consuming. Conversely, unsupervised techniques don't need any ground truth. They work by minimizing a specific cost function, and the accuracy of registration primarily depends on the following factors that are an appropriate selection of the similarity measure, enforcing plausible deformations with a regularization term, the capacity to converge during

training, and a generalization ability. The unsupervised registration can be seen to speed up the classical, iterative image registration [53]. Many DL approaches use CNN architectures to directly predict the spatial transformation, aligning two WSIs of the same tissue. These models often adopt an unsupervised strategy, so instead of needing ground-truth for deformation fields, they optimize a similarity metric between the moving and fixed image to learn the alignment. For instance, Wodzinski et al.'s DeepHistReg framework employs a multi-stage CNN pipeline which includes tissue segmentation, initial coarse alignment, followed by trained affine and non-rigid registration networks to iteratively register differently stained slide [53]. Such end-to-end CNN methods can operate at multiple scales, capturing both global alignment and local deformations. Recent work has also integrated deep feature correspondences to enhance CNN-based registration.

Wei et al. propose an Iterative key point correspondence guided (IKCG) unsupervised registration network guided by learned key point correspondences, using deep feature descriptors to match points across stains [54]. By leveraging features from pre-trained CNNs alongside features learned on histology data, their approach robustly identifies corresponding anatomical structures despite stain differences [54]. This deep feature guidance enables the model to handle large local distortions that conventional algorithms often fail to align [54]. Overall, CNN-based methods provide a fast inference once trained, often registering images in seconds or minutes, much faster than iterative classical methods and can learn to be resilient to staining variability through data-driven feature learning. The main challenge with the learning-based registration of histology images relates to the high resolution of these images, coupled with large and complex deformations [53]. The deep learning methods suffer from large Graphics Processing Unit (GPU) memory utilization. The higher the image resolution, the larger the necessary receptive field and the required GPU memory. The simplest solution is to downsample the images. This approach was used by the TUB team to apply the volume tweening network during the ANHIR challenge [5]. However, downsampling the images reduces the registration quality and makes it harder to register fine details [53].

Additionally, sensitivity of CNN to domain changes may lead to inadequate generalization [55]. In medical imaging, different institutions have different data acquisition conditions, which causes domain shift and changes in image attributes. One well-known example is stain variation in histopathology slides [55]. Domain shifts can happen in medical imaging when testing data from various facilities, all having its own acquisition techniques. The tissue is stained with several dyes to create WSI, with H&E being the most popular. Although the staining procedure makes the underlying structure of tissues visible under a microscope, it can also cause variations in image contrast and color because different chemicals and staining techniques are used. Second, slides are scanned using a scanner after staining; the optical properties of the scanner and post-processing filters vary depending on the brand. The image properties variation among different centers can heavily degrade the performance of machine learning algorithms that were trained on the data from a single institution when testing slides from external centers. Developing frameworks robust to changes in image acquisition parameters is essential to achieve generalization. Data augmentation is a commonly used blend to establish some degree of invariance to a CNN [56, 57], it also can be used to tackle class imbalance [58]. By adding distortion to the data and

avoiding overfitting, augmentation can maintain training and can also add previous details to the model [56].

2.3.3. Transformers and Hybrid Models

Building on the success of CNNs, researchers have started exploring Vision Transformers (ViTs) for registration tasks, given their strength in modeling long-range dependencies. Transformers can capture global context across a WSI, complementing CNNs local feature extraction. Golestani et al. introduced a hybrid CNN-transformer model for aligning multi-modal breast tissue images [59]. In their Puzzling Vision Transformer-Based Affine Image Registration (PViT-AIR) approach, a Vision transformer processes the entire tissue microsection to understand global spatial relationships, while CNN-based modules focus on fine-grained alignment of image segments [59]. This combination allowed simultaneous registration and stitching of multiple histology image patches to a large field of view radiology image through a puzzling mechanism [59]. The transformer’s global awareness improved the model’s ability to match corresponding regions even when tissue context was very different between modalities.

Notably, because obtaining ground-truth for such cross-modality registration is extremely difficult, they trained the model in a weakly-supervised manner using synthetic data, essentially creating pseudo-paired examples by overlaying and transforming images within the same modality [59]. This data augmentation strategy helped overcome the scarcity of aligned multi-stain training data and yielded state of the art accuracy in specifically radiology pathology registration. The use of transformers in histology image alignment is still emerging, but early results indicate they can enhance performance by providing a more holistic understanding of tissue layout.

2.4. Stain Normalization

In most cases, stain normalization procedures preserve the other information in the processed image while converting the source image to the target image color style [60]. This can be broadly categorized into two classes: conventional methods and deep learning-based methods [11, 61]. Staining processes are often subject to variability, which complicates the automation of tissue area analysis across different staining techniques [62]. Variations in biological factors across different tissue types and inconsistencies in staining protocols contribute to high levels of visual appearance variability [62]. One way to reduce color variations present in whole slide images is to convert RGB into grayscale. Grayscale conversion can lead to a loss of several image information, as grayscale images lack the color depth and important details of RGB images [62].

To match the features of images from source and target domains, the traditional ones, which are frequently based on physics [63], use statistical techniques [64]. The color matching and stain-separation methods are two kinds of traditional methods that are mainly obtained by analyzing, converting, and matching the components of the

color. The color matching methods calculate the mean and standard deviation of source images and standardize them to match a reference image in Lab color space [65]. The stain-separation procedures try to distinguish and normalize every staining channel separately.

The target of DL-based techniques is to match the characteristic distributions of source and target slides by first classifying pixels into distinct stain components [66]. Additionally, stain normalization could be seen as a style transfer issue, transforming data from a different domain into the target center's domain [67]. Normalization has the benefit of decoupling the domain invariance and classification tasks [55], but it also has the drawback of requiring an additional step to normalize each image before it can be classified. The recent DL-based methods allow the learning of stain transfer mappings in a data-driven manner where fine-grained structural features are retained [65]. According to Hoque et al. [62], these techniques are further grouped into four categories based on their methodologies [68] as listed below,

2.4.1. Histogram Transformation Based Approaches

These approaches standardize the color appearance of histopathological images by aligning the color distributions of source images to those of a reference image using statistical transformations [62]. Reinhard et al. introduced a landmark based method that operates in the LAB color space [64], transferring the mean and standard deviation of color channels from the reference image to the source image. Zheng et al. extended this idea using HSV space to control illumination and saturation components for improved normalization [68]. While histogram transformation-based methods are computationally efficient and simple to implement, they often ignore the underlying staining characteristics of histological structures, leading to color artifacts and inconsistent normalization, especially in regions of interest.

2.4.2. Separated Transformation Based Approaches

These methods treat each stain as an independent component and apply separate transformations to normalize them. The foundational technique here is color deconvolution, which separates stains based on their absorbance spectra using the Beer-Lambert law [62]. Khan et al. [60] used nonlinear mapping guided by staining vectors. Bejnordi et al. [66] applied hue-saturation-density transformations to align chromatic and density distributions. Although separated transformation methods offer greater control over individual stain normalization, they require additional computation and may cause color discontinuities and structural distortion due to per-pixel transformation inconsistencies.

2.4.3. Unified Transformation Based Approaches

The drawback of the above-mentioned approaches is the color discontinuity and the high computational time. In this context, several approaches have been

proposed to establish a unified transformation for pixels in histopathology images [62]. Instead of pixel-based stain classification, these techniques offer pixel-based extraction of independent stain components. To overcome the limitations of per-stain transformations, unified transformation methods estimate a global mapping that simultaneously accounts for all stain components. These methods often rely on non-negative matrix factorization (NMF) or sparse NMF (SNMF) to extract stain vectors and weights [62]. Macenko et al. [63] proposed a method based on singular value decomposition in the optical density space. Zhou et al. [41] optimized the deconvolution matrix to improve hematoxylin component decomposition. Adaptive color deconvolution methods enhance stain separation by embedding stain intensity and proportion constraints in the optimization process [62]. While unified transformation approaches improve consistency and structure preservation, they remain computationally complex and require prior knowledge for optimization. Moreover, challenges persist when one stain is weakly represented, or immunohistochemical staining does not conform to linear assumptions.

2.4.4. Generative and Diffusion Models

Recent advances in deep learning have introduced Generative Adversarial Networks (GANs) and diffusion models for stain normalization.

Generative Adversarial Network: GANs aim to bridge the appearance gap between different stains by learning unpaired image-to-image translation mappings. GANs have become a powerful tool to bridge the appearance gap between different stains. Rather than directly computing correspondences in two vastly different images, the idea is to translate one staining to another, effectively normalizing the appearance before registration [30]. Mostly, images are differently stained to highlight the key feature [30]. Cycle-consistent GANs (CycleGANs) have been widely adopted for such unpaired image-to-image translation[30]. For instance, Xu et al. utilized a CycleGAN-based approach to generate synthetic IHC-stained images from H&E slides [30]. By doing so, regions of tumor that are difficult to discern on H&E alone could be virtually stained for Ki-67 or any other stain, improving downstream alignment and analysis [30]. In their framework, adversarial training was augmented with class-specific feature vectors to better preserve diagnostically relevant structures during translation [30]. Likewise, Roy et al. developed a translation then registration approach. A modified CycleGAN first produces a synthetic IHC image from an H&E WSI, after which a deformable CNN registers the synthetic IHC to the actual IHC image [12]. To ensure that generative translation does not alter tissue morphology, they imposed a feature consistency loss using a pre-trained VGG network, so that the synthetic image shares structural features with the real target [30]. This GAN-assisted pipeline termed CGNReg was able to align H&E/IHC serial sections at full WSI resolution and showed significantly improved alignment over traditional methods in cross-stain registration tests [12].

Diffusion Models: are the class of generative models that learn to generate images by reversing a gradual noise-adding process with iterative denoising [22]. They are

trained to model the distribution of the data by predicting the noise that is added at each timestep in a Markov chain, which allows high-fidelity sample generation from pure gaussian noise [69]. A neural network is then trained in reverse, step by step to denoise the sample and recover the original structure [70]. Such an approach enables diffusion models to learn highly complex data distributions and produce realistic, high-fidelity images with strong diversity and stability outperforming traditional GANs on many visual synthesis tasks [51]. To reduce the high computational cost of working directly in the pixel space, Latent Diffusion Models (LDMs) were suggested, and these compress the image down to a lower-dimensional latent space prior to applying the diffusion process [71]. Stable diffusion is a popular realization of LDMs utilizing Variational Autoencoder (VAE) for image encoding and decoding, U-Net backbone is used for denoising in the latent space, and conditioning mechanism based on transformer-based encoders such as Contrastive Language image pre-training (CLIP) [72]. Diffusion architecture provides efficient generation of high-resolution images with semantic guidance, and this makes stable diffusion a scalable and flexible framework for natural image synthesis [72].

Diffusion Probabilistic Models: the latest generation of deep generative models, diffusion probabilistic models, are also being explored for stain normalization and augmentation tasks. Diffusion models offer a new way to perform image to image translation by iteratively refining images through learned denoising processes, which can produce highly realistic outputs without some of the pitfalls of GANs [73]. StainDiff is a recently proposed diffusion-based method for histology stain style transfer [73]. It introduces a cycle-consistent diffusion process that learns to translate an image from one stain domain to another while preserving tissue structure. This approach was shown to avoid issues like mode collapse that GANs may suffer from, and it can generate multiple variant outputs to capture the variability in staining [73]. Although diffusion models are computationally intensive, their ability to produce reliable and diverse stain-normalized images can be very valuable as a preprocessing step for registration. By using such models to standardize color and intensity across slides, registration algorithms, whether traditional or CNN-based can achieve more robust performance.

The incorporation of diffusion models into the registration pipeline is still in emerging stages, but it represents a promising edge for handling stain variability through data augmentation and normalization. The diverse methods reviewed above highlight the trade-offs between computational cost, stain specificity, structural preservation, and generalizability. Classical methods like Reinhard and Macenko remain useful as baselines, while matrix factorization-based and adaptive deconvolution techniques offer more robust performance. GAN and diffusion-based approaches represent the cutting edge, although their reliance on large datasets and training stability remains a concern.

3. METHODOLOGY

This section outlines the complete approach designed to perform the image registration of histopathological tissue samples prepared with various chemical staining techniques. The proposed pipeline utilizes DL-based stain normalization techniques to introduce the variability and non-rigid image registration methods to achieve high spatial precision. A histopathology dataset discussed in the Section 3.1 serves as the initial input to the pipeline which undergoes stain normalization to increase the dataset variability. The next stage employs a DL model to normalize images of other stains into an H&E-stained appearance discussed in Section 3.2. The harmonized output from the previous step moves to the registration module, where it undergoes spatial alignment with the target image. This study employs three registration techniques, which are discussed in Section 3.3 of this chapter, along with their respective evaluation metrics. The Target Registration Error (TRE) is used as the measure of registration accuracy across all three methods for the comparative study in Section 4.1.3. The below Figure 3, illustrates the methodology followed in the study.

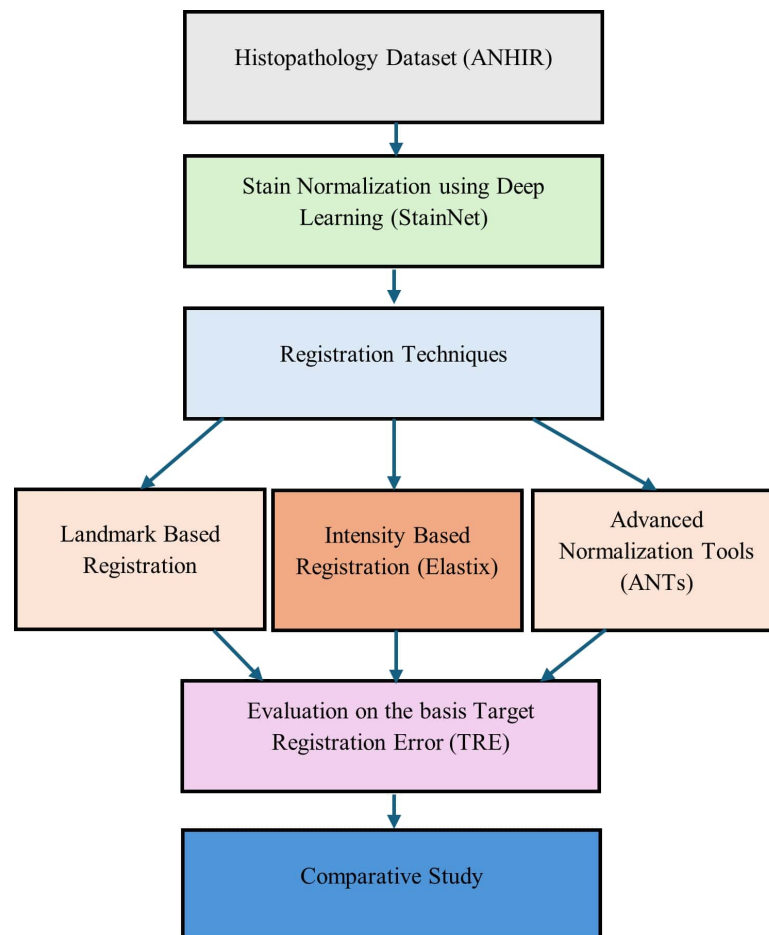


Figure 3. Methodology for the histopathology image registration study, illustrating the process from dataset preparation to comparative analysis based on TRE.

3.1. Dataset

The dataset utilized in this study is the Automatic Non-rigid Histological Image Registration (ANHIR) dataset. The ANHIR dataset was earlier released for the ANHIR grand challenge at IEEE International Symposium on Biomedical Imaging (ISBI) 2019 [74]. The ANHIR dataset serves for benchmarking and evaluating purposes of modern non-rigid image registration methods in digital pathology because it addresses common issues like multi-stain variability and tissue deformation along with scanning artifacts that traditional techniques fail to handle. The dataset consists of whole slide images, high resolution up to 40x magnification of tissues including lesion, lung lobes, breast, mice kidney and mammary glands [5]. The initial dimension of images is 100 000 x 200 000 pixels, stained with multiple dyes. The images are arranged in sets of successive slices [74]. The data set consist of whole-slide histological images that were obtained from eight distinct tissue types. Figure 4 shows the sample images from the dataset of different tissues.

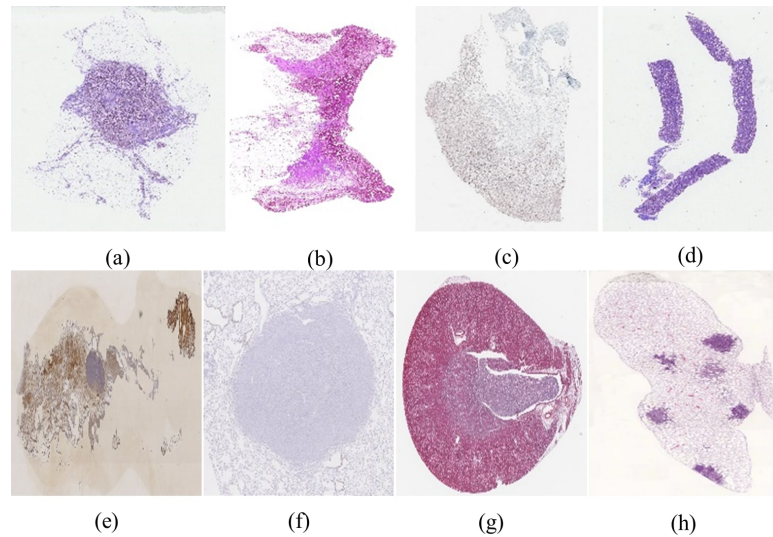


Figure 4. Sample images from the dataset: (a) breast, (b) COAD (colon adenocarcinoma), (c) gastric, (d) kidney, (e) mammary gland (f) lung lesion (g) mice kidney and (h) lung lobes.

Each of the sample images in Figure 4 has a corresponding set of differently stained variations. The presence of multiple staining variations creates substantial challenges because of color variations, contrast and texture differences that make this dataset suitable for evaluating stain invariant image registration method. Below Table 1 shows the list of biological marker codes and their corresponding names used for staining histology tissue samples in the ANHIR dataset and Table 2 shows the dimensional characteristics of images.

Table 1. List of biological marker codes and their corresponding names used for staining histology tissue samples in the ANHIR dataset [19]

Code	Name
Ki67	Antigen Ki-67
PR	Progesterone receptor
proSPC	Prosurfactant protein C
HER2	Human epidermal growth factor receptor 2
H&E	Hematoxylin and Eosin
ER	Estrogen receptor
CD31	Platelet endothelial cell adhesion molecule
Cc10	Clara cell 10 protein

Table 2. The dimensional characteristics of images in the ANHIR dataset [19]

Name	Tissue	Magnification	Resolution [$\mu\text{m}/\text{pixel}$]	Average size [pixels]
Breast	Human breast	40x	0.2528	65k \times 60k
COAD	Colon adenocarcinoma	10x	0.468	60k \times 50k
Gastric	Gastric mucosa	40x	0.2528	60k \times 75k
Kidney	Human kidney	40x	0.2528	18k \times 55k
Mammary gland	Mammary glands	10x	2.294	12k \times 4k
Lung lesion	Lung lesion	40x	0.174	18k \times 15k
Mice kidney	Mice kidney	20x	0.227	37k \times 30k
Lung lobes	Whole mice lung lobes	10x	1.254	11k \times 6k

Quantitative assessment becomes possible through the manual annotation of landmark points in ANHIR dataset for the specific image pairs. The manually placed points serve as reference markers that link specific areas in both moving and target images for measuring TRE, which represents a common metric in medical image registration [74]. The landmarks are stored in a comma separated values (csv) file format, which is quite easy to use and includes a standard ImageJ structure and coordinate frame. The top left corner of the image is assigned to coordinate the origin the system (0,0). Regarding annotation quality, two experts completed each landmark annotation, and one additional expert independently verified each annotation [74]. This multi-expert approach ensures high annotation accuracy and reliability. The consistent verification process minimizes human error and enhances the robustness of the landmark data. Landmarks serve as valuable tools for assessing non-rigid registration performance when images undergo severe tissue deformations. Landmarks serve as valuable tools for assessing non-rigid registration performance when images undergo severe tissue deformations [5]. Figure 5 illustrates the images of the lung lesion and rat kidney, warped with the landmarks.

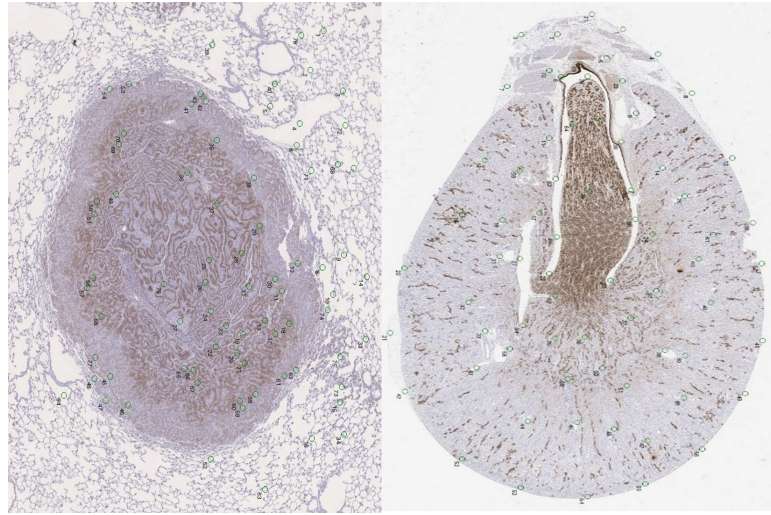


Figure 5. Images warped with their respective landmarks.

3.2. Stain Normalization

The next step is stain normalization. As stated earlier, staining variability is one of the key issues that results in poor performance for the registration algorithm. This is a critical gap that can be optimized. To determine the impact of staining variability on registration performance, the stain normalization was applied to the images using a DL-based method, StainNet. The primary goal of this step is to introduce variability and make it possible to assess the robustness of different registration approaches under varying conditions. This allowed to systematically evaluate how the intensity-based registration algorithms and the landmark-based registration method respond to stain normalization. Such an approach is essential given that stain normalization can sometimes introduce synthetic alterations in intensity distributions, raising questions about its real-world applicability. Therefore, by applying normalization aim was to gain deeper insights into their stain-invariance capabilities and practical usability in diverse histopathology settings.

In this work, two different methods were tried that include the efficient pretrained StainNet model for stain normalization and Reinhard stain normalization coupled with stable diffusion model for denoising. For stain normalization, these two methods were experimented to select the best performing, which was StainNet. DL based techniques are efficient at stain normalization, but they are still lacking computational efficiency and robustness [65].

3.2.1. *StainNet*

Most of the stain normalization methods, such as Reinhard normalization or Macenko, have been utilizing statistical models based on color space transformations [75]. However, classical approaches are not well suited to generalizing other tissue

structures, but also have difficulty managing complex, nonlinear stain variations. To overcome this StainNet is one of the deep learning-based methods proposed [65].

StainNet Architecture:

StainNet is a convolutional neural network designed for stain normalization via channel-wise color transformation using only 1×1 convolutional filters originally developed by Kang et al. [65]. This study applies optimization and utilizes architecture in the workflow for the normalization task. The original design goal of StainNet is to preserve spatial structure while learning an end-to-end mapping of the input stain style to a reference stain in the RGB color space [65, 75]. To learn color mapping using distillation learning, the model employed StainNet as the student network and StainGAN as the teacher network [65].

The implementation consists of two stages: The first stage involves the training of StainGAN, a generative adversarial network with the couple of generators and discriminators. The second stage involves the development of StainNet that is built on CNN. For StainNet to learn the transition from the source color space to the target color space, both the source and target images must be matched. In this way StainNet is utilized as the student network and StainGAN as the instructor network because it is challenging to obtain matched images and align them precisely in practice. In other words, StainNet learns the StainGAN output using the L1 loss [65]. The following Figure 6 illustrates the complete architecture of StainNet.

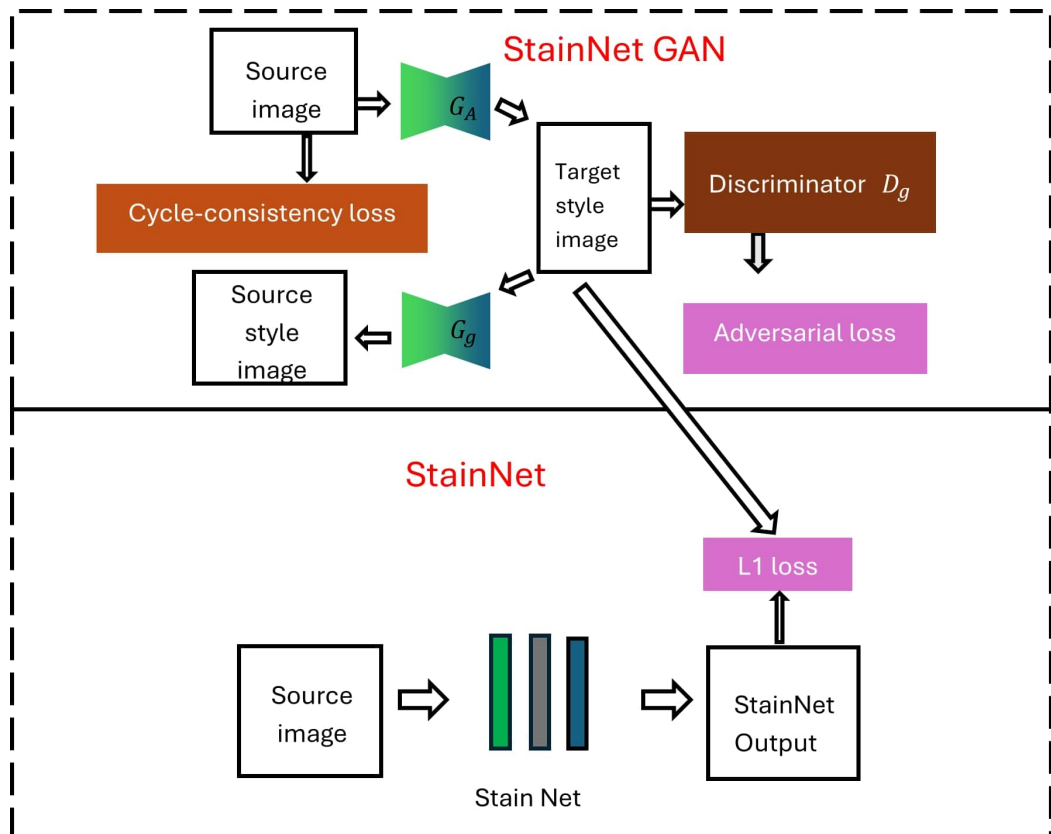


Figure 6. The architecture of the StainNet model.

StainGAN:

As discussed earlier, there are couple of generators (G_A and G_B) and discriminators (D_A and D_B) in the framework of StainGAN. The image is transferred from the source domain to the target domain using G_A , and from the target domain back to the source domain using G_B . The generator G_A transforming the source image x_s to the target domain is expressed in Equation (1)

$$x_t = G_A(x_s) \quad (1)$$

Further, the generator G_B transforming the target image x_t back to source domain is expressed in Equation (2)

$$x_s = G_B(x_t) \quad (2)$$

where x_s is the source image and x_t is the target image in Equations (1) and (2).

Images created by G_A and a real target image are distinguished by D_A , while images created by G_B or a real source image are distinguished by D_B . StainGAN suffers from two types of losses: adversarial loss and cycle-consistency loss. The cycle-consistency loss ensures that G_B can reconstruct the images produced by G_A to the source image and G_A can rebuild the images produced by G_B to the target image [76]. Equation (3) describes the cycle consistency of cycle GAN [76]

$$\mathcal{L}_{cyc} = \sum_x \|G_B(G_A(x)) - x\|_1 \quad (3)$$

where x is an image from the source domain, G_A and G_B are the generators.

Moreover, the goal of adversarial loss is to make sure that the generated images stain distribution matches the actual distribution [61]. Equation (4) describes the calculation of adversarial loss for the GAN model [76]

$$\mathcal{L}_{GAN} = \sum_x [\log D(x)] + \sum_x [\log (1 - D(G(x)))] \quad (4)$$

where $G(x)$ is the generated image created by the generator G , $D(x)$ is the discriminator's output for a real image, and $D(G(x))$ is the discriminator's output for the generated image.

Equation (5) describes the total loss for StainGAN, where λ is a weight to balance the loss terms [76]. This collectively shows the adversarial losses for both domains and the cycle-consistency loss to guide the training of the StainGAN framework

$$\mathcal{L}_{\text{StainGAN}} = \mathcal{L}_{A_{GAN}} + \mathcal{L}_{B_{GAN}} + \lambda \cdot \mathcal{L}_{cyc} \quad (5)$$

where $\mathcal{L}_{A_{GAN}}$ is the adversarial loss for domain A (source to target), $\mathcal{L}_{B_{GAN}}$ is the adversarial loss for domain B (target to source), \mathcal{L}_{cyc} is the cycle-consistency loss, which ensures that a source image translated to the target domain and back reconstructs

the original image, and λ is a weighting parameter that balances the importance of the cycle-consistency loss relative to the adversarial losses.

StainNet:

Convolution operations in the present convolutional neural network use a kernel size of at least 3×3 . A 3×3 or larger convolution, on the other hand, conducts a weighted summation in the nearest neighbors of the input image [65]. As a result, the nearest area of the input image invariably influences the pixel value in the output image. The 1×1 convolution has nothing to do with neighboring values, in contrast to the 3×3 convolutions, and it only maps one pixel. In other words, it can maintain the source information and would not be impacted by texture [65]. The mapping relationship from StainGAN is then extracted using a completely 1×1 convolutional neural network called StainNet. ReLU is utilized as a convolutional layer to improve the non-linear mapping capability, except for the final convolutional layer. By default, convolutional layers with 32 channels were employed, considering the trade-off between computational efficiency and performance. The model Equation (6) for 1×1 CNN with ReLU activations

$$f(x) = W_3 \cdot \text{ReLU}(W_2 \cdot \text{ReLU}(W_1 x + b_1) + b_2) + b_3 \quad (6)$$

where $f(x)$ is the final output of the neural network for input x , W_1 , W_2 , W_3 are weight matrices for the first, second, and third layers, respectively. b_1 , b_2 , and b_3 are the corresponding bias vectors. ReLU is the rectified linear unit activation function.

Further, the loss function for the StainNet can be calculated as shown in Equation (7)

$$\mathcal{L}_{\text{StainNet}} = \sum_{x_s} \|\text{StainNet}(x_s) - G_A(x_s)\|_1 \quad (7)$$

where $\mathcal{L}_{\text{StainNet}}$ represents the training loss used for the StainNet model. x_s is an image from the source domain. $\text{StainNet}(x_s)$ is the output of the StainNet model, which performs stain normalization on the source image. $G_A(x_s)$ is the target-styled image generated by the teacher network G_A (from StainGAN).

Utilization of StainNet Architecture:

This study utilizes StainNet by implementing architecture in a highly optimized way and loading the pretrained weights. The model consists of three consecutive 1×1 convolutional layers with 32 intermediate channels, where ReLU activation functions are applied after the first and second layers to introduce non-linearity, while the final layer produces the output without activation. The model performs pixel-wise stain normalization on histopathology images. The pipeline includes preprocessing input images to match the expected range of the model, performing GPU-accelerated inference, and post-processing the outputs to standard RGB format. This integration enabled the effective deployment of StainNet for scalable and reproducible stain normalization in this study.

Evaluation Metrics:

Digital images are subject to a wide variety of distortions during acquisition, processing, compression, storage, transmission and reproduction, any of which may result in a degradation of visual quality. To evaluate the performance, the similarity between the normalized image and the target image, and the consistency between the normalized image and the source image was measured with Structural Similarity Index Measure (SSIM) and Peak signal to Noise ratio (PSNR).

Structural Similarity Index: SSIM is a perceptual metric that measures the similarity between two images [77]. The extent of source information preservation is weighed by the SSIM of the source image (SSIM Source), which also was used to measure the similarity between the normalized image and the source image [65]. This method is particularly useful for stain normalization. It estimates the perceived quality of the images [77]. It can be calculated as shown in Equation (8)

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (8)$$

where x and y are the patches from the reference and test images respectively, μ_x , μ_y are the mean intensity values of x and y , σ_x^2 , σ_y^2 are the variances of x and y , σ_{xy} is the covariance between x and y , and C_1 , C_2 are small constants to stabilize the division.

Peak Signal to Noise Ratio: PSNR is used to calculate the ratio between the maximum possible signal power and the power of the distorting noise which affects the quality of its representation [77]. This ratio between two images is computed in decibel form [77]. PSNR is a traditional metric that quantifies image fidelity based on the Mean Squared Error (MSE) between a target and test image. It can be calculated as shown in Equation (9)

$$PSNR(x, y) = 10 \cdot \log_{10} \left(\frac{L^2}{MSE(x, y)} \right) \quad (9)$$

where L is the maximum pixel value, x and y are the pixel value from the two images after the mean squared error.

3.3. Image Registration Techniques

Three different registration methods were applied. A landmark-based registration method was developed in this thesis work; the other two are prebuilt Elastix and ANTs methods.

3.3.1. Landmark Based Registration

The registration approach employed is a non-rigid, landmark-based geometric method, designed to align multi-stained histopathological images. The following approach

aligns two images by allowing non-rigid deformations like stretching, bending, or warping so that corresponding structures match between them. The method is based on the classical TPS interpolation technique, which has been extensively used in medical image registration due to its smooth deformation modeling and minimal bending energy properties. The pipeline consists of five key stages: displacement vector estimation from anatomical landmarks, TPS-based interpolation using RBF, gaussian smoothing of the displacement field, image warping using remapping functions, and registration accuracy evaluation using TRE. The combination of TPS-RBF interpolation, chunk-wise processing, Gaussian regularization, and manual landmark transformation, coupled with automated TRE-based validation and overlay visualization, provides a flexible and interpretable framework suitable for computational pathology research.

Landmark-Based Displacement Estimation:

For each pair of fixed (H&E) and moving (immunohistochemically stained) images, corresponding anatomical landmarks were used to compute displacement vectors. These coordinates were stored in csv file format. The dataset for each image pair was split into training (70%) and validation (30%) subsets. The training landmarks guided the construction of the registration model, while the validation subset was used for accuracy evaluation.

Displacement Vector Estimation:

We have two images the moving image and the fixed image. By using the image coordinates the displacement estimation can be computed as shown in Equation (10). The following are the coordinates of the images

Coordinates of the moving image = (x_i^m, y_i^m)

Coordinates of the fixed image = (x_i^f, y_i^f)

$$\Delta x_i = x_i^f - x_i^m, \quad \Delta y_i = y_i^f - y_i^m \quad (10)$$

where Δx_i and Δy_i vectors represent how much each point in the moving image must shift to align with the fixed reference image.

Thin Plate Spline Interpolation with Radial Basis Function:

To extend the displacement information from sparse landmark locations to every pixel in the image, Radial Basis Function (RBF) interpolation using a Thin-Plate Spline (TPS) kernel was applied. The TPS-RBF formulation minimizes bending energy while ensuring smooth warping of the image. The displacement at any arbitrary pixel (x, y) is computed as shown in Equation (11)

$$D(x, y) = \sum_{i=1}^N \omega_i \cdot \Phi(\|(x, y) - (x_i, y_i)\|) \quad (11)$$

where $\Phi(r) = r^2 \log r$ is the TPS kernel function, (x_i, y_i) are training landmark coordinates, ω_i are the interpolation weights run from training displacements. The interpolation was applied in memory-efficient chunks to accommodate high-resolution histopathology images, ensuring scalability of the method without significant computational overhead.

Gaussian Smoothing of Displacement Field:

The interpolated displacement field from Equation (11), further smoothed using a Gaussian kernel $G_\sigma(x, y)$ to remove localized interpolation noise and ensure smooth, anatomically plausible deformations can be seen in Equation (12)

$$D_{\text{smooth}}(x, y) = G_\sigma(x, y) * D(x, y) \quad (12)$$

where $*$ denotes convolution. This step helps in reducing abrupt displacements that may arise due to slight inconsistencies in landmark placement.

Image Warping and Registration Output:

The smoothed displacement field $(dx(x, y), dy(x, y))$ was applied to the moving image using OpenCV `remap()` function, producing the registered image shown in Equation (13)

$$(x', y') = (x + dx(x, y), y + dy(x, y)) \quad (13)$$

where each pixel in the moving image is shifted according to the computed field, resulting in a warped image aligned to the fixed image.

Evaluation Metrics:

The accuracy of the registration was evaluated using TRE, calculated on the validation landmarks as shown in Equation (14)

$$TRE = \frac{1}{N} \sum_{i=1}^N \sqrt{(x_i^f - x_i^r)^2 + (y_i^f - y_i^r)^2} \quad (14)$$

where (x_i^f, y_i^f) are the fixed validation landmark coordinates and (x_i^r, y_i^r) are the corresponding transformed coordinates from the moving image. TRE is the average spatial error and was reported in pixels.

3.3.2. Intensity Based Technique Using Elastix

To perform baseline image registration between differently stained histopathology slides, Elastix, an open-source medical image registration toolkit built upon the ITK library was used [78]. Elastix provides modular control over each component of the registration pipeline through parameter files, allowing for customized workflows suited to multi-modal, high-resolution image data.

In this work, affine transformation was employed using Elastix, which models global spatial alignment through translation, rotation, scaling, and shearing. Although affine transformation is limited in correcting local non-linear deformations, it is computationally efficient and effective for tissue sections with consistent orientation and scale. The configuration details are summarized in below Table 3.

Table 3. Configuration settings used in the registration process

Component	Setting
Transformation Model	Affine Transform
Similarity Metric	Mutual Information (MI)
Optimizer	Adaptive Stochastic Gradient Descent
Multiresolution Strategy	6 pyramid levels with recursive down sampling
Sampling Strategy	4096 samples per iteration
Interpolation	B-spline

Transformation Model: An affine transformation is used that enables linear transformation in 2D space, suitable for correcting misalignments between stained images.

Similarity Metric: Mutual information is used that captures statistical dependence between intensities of the fixed and moving images. The 64 histogram bins were used to compute mutual information with enhanced sensitivity.

Optimizer: Adaptive Stochastic Gradient Descent is used that is capable of handling noisy gradients and converging effectively in high-dimensional spaces. The optimizer was configured to perform up to 999 iterations per resolution level.

Multi-Resolution Strategy: Registration was executed across six resolution levels, progressively refining alignment from coarse to fine scales. This hierarchical pyramid strategy improves robustness and reduces the risk of local minimum.

Interpolator and Resampler: B-spline interpolation of order 2 was used during registration to compute gradients smoothly, and order 3 was used in the final resampling step to maintain spatial accuracy in the output image.

Sampling Strategy: A total of 4096 spatial samples were randomly selected per iteration to compute the similarity metric, balancing accuracy with computational efficiency.

This affine configuration is well-suited for aligning histopathology images that share similar global structures but differ in staining. However, due to its linear nature, it may struggle in regions exhibiting local deformation or heterogeneous tissue appearance, which are common in WSIs. Nonetheless, the affine model served as a reliable baseline for comparative evaluation against more complex or learning-based registration strategies. This experiment was conducted within the BIRL (Benchmark for Image Registration Libraries) Borda environment, which facilitated reproducible benchmarking and evaluation.

3.3.3. Intensity Based Registration Using Advanced Normalization Tools

To handle complex tissue deformations and achieve fine-grained alignment between differently stained histopathological images, ANTs a widely recognized, high-performance software suite for medical image registration was employed. Developed originally for neuroimaging applications, ANTs have evolved into a general-purpose image registration toolkit and is consistently ranked among the most accurate in benchmark studies. Its flexible design allows users to select from a variety of transformation models, similarity metrics, regularization strategies, and optimization procedures, making it well-suited for histology applications involving multi-stained slides with structural variations [79].

One of the key strengths of ANTs lies in its ability to model non-linear and large-deformation transformations while preserving anatomical topology. This is especially critical in histopathology, where tissue sections may undergo warping, folding, tearing, or variability due to staining protocols. Among its various models, we chose the Symmetric Normalization (SyN) transformation, which is a diffeomorphic registration method that computes a smooth, invertible deformation field to align the moving image to the fixed image. Unlike traditional forward-only methods, SyN simultaneously optimizes both forward and inverse transformations, producing a symmetric and topology-preserving mapping. The configuration details are summarized in below Table 4.

Table 4. Configuration details for ANTs SyN registration

Component	Setting
Transformation Model	Symmetric Diffeomorphic (SyN)
Similarity Metric	Cross-Correlation (CC)
Optimizer	Gradient Descent
Multiresolution Strategy	4 resolution levels
Sampling Strategy	Dense
Interpolation	Linear interpolation

Transformation Model: SyN is used with Gaussian smoothing of the velocity field to enforce regularization and prevent unrealistic deformations.

Similarity Metric: The cross-correlation (CC) is used as the similarity metric, capturing local structural correspondence between images while being robust to global

intensity differences. This is an essential property when working with images from different staining modalities.

Optimization Strategy: Registration was carried out using a multi-resolution approach, where image pyramids were created by progressively downsampling the input images. Optimization was performed independently at each level using gradient descent until convergence thresholds on metric change and parameter norm were satisfied.

Interpolation and Resampling: Linear interpolation was used to ensure smooth estimation of intensities during transformation. Final warped images and landmark coordinates were resampled at full resolution for evaluation.

During execution, each moving image was registered to its corresponding fixed image, and the computed transformation was applied to both the image and its associated landmark annotations. ANTs exhibited strong performance in cases with high degrees of local deformation and tissue complexity, especially in regions like glandular structures or loosely packed tumor zones. However, due to its intensity-driven nature, registration performance was sometimes affected when stain normalization altered the texture or contrast patterns of the input images. Additionally, the computational cost of ANTs was significantly higher compared to affine or linear methods such as those used in elastix.

3.3.4. Evaluation Metrics for Intensity Based Registration

To evaluate the performance of the registration methods we used the following geometric accuracy metrics.

Mean Target Registration Error:

The TRE measures the average Euclidean distance between corresponding landmarks in the fixed (target) image and the registered (warped) moving image. The mean TRE was computed as shown in Equation (15)

$$\text{Mean } TRE = \frac{1}{N} \sum_{i=1}^N \left\| P_i^{\text{fixed}} - P_i^{\text{warped}} \right\|_2 \quad (15)$$

where N is the total number of landmarks, P denotes the landmark, and i^{th} denotes the index.

Relative Mean Target Registration Error:

To track the variation in image resolution and size, mean rTRE is also computed. It normalizes the TRE by the diagonal length of the image. It was computed as shown in Equation (16)

$$\text{Mean } rTRE = \frac{1}{N} \sum_{i=1}^N \frac{\|P_i^{\text{fixed}} - P_i^{\text{warped}}\|_2}{D} \quad (16)$$

where D represents the diagonal length of the image.

3.3.5. *Software Tools*

For the implementation in this study, Python version 3.10 was used within the Google Colab environment. The key Python libraries included PyTorch for the implementation of StainNet, along with Torchvision. Furthermore, the Elastix and ANTs methods were executed within the BIRL (Benchmark for Image Registration Libraries) [74] framework.

4. EXPERIMENTS AND RESULTS

This section presents the experiments conducted during the research and the results obtained. Initially, six different tissue types were selected from the ANHIR dataset. Following the application of stain normalization, three different methods for image registration were evaluated. These experiments were performed on both normalized and non-normalized images to evaluate the impact of staining variability. The aim was to assess how different registration methods, particularly intensity-based approaches, respond to such variability, and to examine the robustness of the proposed landmark-based method.

4.1. Dataset

In the first step, the dataset was organized according to corresponding image pairs for registration, including the H&E-stained reference image for each specific tissue type and the associated csv file containing landmark annotations. For stain normalization, the csv file was not required; only the H&E-stained image of the respective tissue was used for comparison. Table 5 outlines the selected tissues and respective images for normalization and Figure 7 showing the proportion of each tissue category. Images were chosen for stain normalization based on the availability of H&E-stained image.

Table 5. Summary of tissue types and corresponding image samples used in the dataset

Tissue Types	Number of Samples	Total Images
Lung lobes	4	16
Mammary gland	2	8
Lung lesion	3	12
Breast	5	10
Mice kidney	2	10
Kidney	5	15
Total	21	71

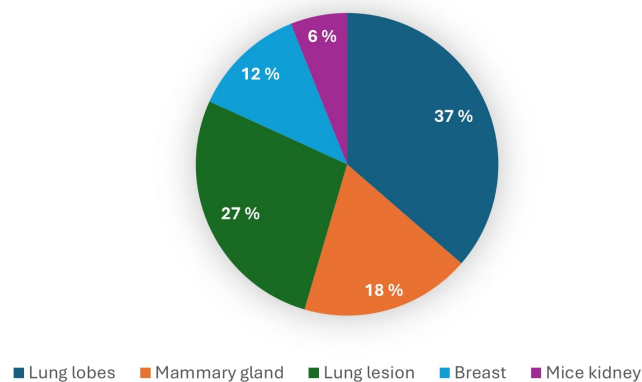


Figure 7. Proportion of image samples per tissue type used in the dataset.

4.1.1. Normalization Results

After organizing the dataset by tissue type, stain normalization was performed using the StainNet architecture, whose implementation details have been previously outlined in the methodology Section 3.2. For each tissue group, the H&E stained image was used as the reference template during normalization. The objective was to transform the differently stained images (e.g., HER2, PR, Ki67) into H&E-like representations to introduce controlled variability in staining appearance. Figure 8 presents an image grid illustrating representative results for various tissue types, including both the original target (H&E) image and the corresponding stain-normalized outputs.

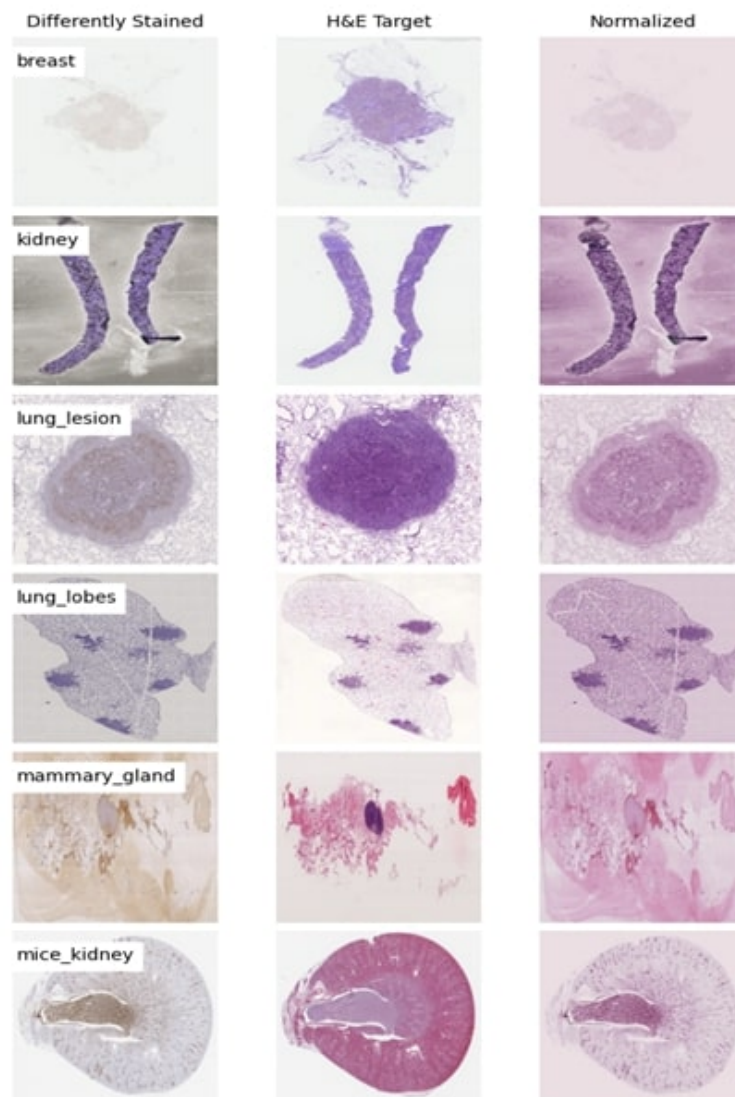


Figure 8. Stain Normalization Results Across Tissue Types Using StainNet: Comparison of Differently Stained Inputs, H&E Targets, and Normalized Outputs.

To quantitatively evaluate the visual similarity between the stain-normalized images and their corresponding H&E reference images, we computed two of the most commonly used image quality metrics: Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR). These metrics help assess how

well the stain normalization preserved structural information while introducing visual differences. Importantly, controlled variability was deliberately introduced during the normalization process to mimic realistic variations in staining, ensuring that the resulting images remain natural in appearance while challenging the robustness of registration algorithms.

The computed SSIM and PSNR values were averaged across each tissue group to capture the overall impact of normalization. These results are presented in Table 6, accompanied by a Figure 9 for intuitive visualization of similarity metrics and variability across tissue types.

Table 6. Average SSIM and PSNR values computed between stain-normalized and H&E images across different tissue groups

Tissue Group	Average SSIM	Average PSNR (dB)
Breast	0.9438	25.43
Kidney	0.9736	24.98
Lung lesion	0.9526	26.58
Lung lobes	0.9908	30.27
Mammary gland	0.9797	25.31
Mice kidney	0.9888	29.61

Figure 9 shows the average SSIM and PSNR values for each tissue group after applying stain normalization with the StainNet architecture. These metrics were calculated by comparing the stain-normalized images with their corresponding H&E reference images to check how well the normalized images look and keep the same structure. The PSNR values for lung-lobes and mice-kidney tissues turned out to be the highest, at 30.27 dB and 29.61 dB. This indicates that stain-normalized images of these tissue types are still quite similar in intensity to their original H&E images. Both groups scored an SSIM of 0.99, meaning all the details of the tissue structure, edges and fine features were kept almost perfectly.

Unlike the others, the kidney, breast and mammary gland tissues had slightly lower PSNR values (between 24.98 dB and 25.43 dB), suggesting that the intensity variation was greater during normalization. This variation fits with the goal of StainNet, which is to allow for realistic changes in staining without affecting the anatomy. Though PSNR was not very high, the SSIM values for these tissues stayed at 0.94 or higher, proving that the main features were still preserved in all cases. Overall, using StainNet for normalization shows that it can deliver a wide variety of realistic stains without affecting the structure. It allows us to test the main goal of the performance study, which is to compare how different registration algorithms handle the same, but naturally changing, data structures.

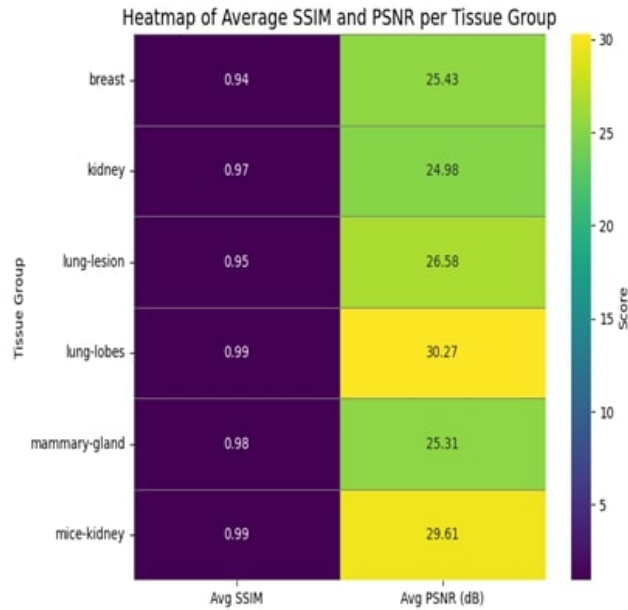


Figure 9. Stain normalization results across tissue types using StainNet: comparison of differently stained inputs, H&E targets, and normalized outputs.

4.1.2. Registration Results

Following stain normalization, a number of comprehensive sets of image registration experiments were conducted to evaluate the alignment performance under controlled stain variability. The first approach involved a landmark-based registration method, which was designed using pre-annotated anatomical landmarks provided by the ANHIR dataset. This method served as a structurally guided baseline, relying solely on spatial correspondence rather than pixel intensities.

After establishing the baseline with the landmark-based method, the evaluation was extended to include two widely used intensity-based registration techniques, Elastix and ANTs. These methods were chosen to assess the performance of automated registration algorithms that operate directly on image intensity patterns, which are more likely to be affected by staining variability. By applying all three techniques, landmark-based, Elastix, and ANTs on both original stain pairs and stain-normalized pairs, the aim was to systematically compare their robustness and sensitivity to color variation in histopathological image registration.

Landmark Based Registration Results:

For the first phase of registration evaluation, landmark-based registration technique was employed. This method leverages spatial alignment between predefined landmark points on the moving and fixed images to compute a non-rigid transformation using TPS interpolation. The goal was to establish a structurally informed baseline that is independent of intensity information and thus potentially more robust to staining variability.

Representative subset of image pairs was selected from the ANHIR dataset presented in Table 7, which provides a diverse collection of multi-stained histopathological images across various tissue types. The dataset includes pre-annotated landmark pairs for each image, allowing precise and reproducible evaluation of geometric registration performance. For the experiments, six tissue types were used, each comprising several image pairs with associated landmark annotations. Registration was performed under two conditions: using the original stained moving images and the stain-normalized versions, both aligned to their corresponding H&E-stained fixed images.

To quantitatively assess the effectiveness of the landmark-based registration, the mean TRE was computed for each image pair. This metric calculates the average Euclidean distance between each corresponding landmark in the fixed and registered moving images. A lower TRE indicates a higher degree of alignment and better registration accuracy. Comparing the mean TRE values obtained from the original and stain-normalized conditions across all tissue types, enabled to directly quantify the impact of stain normalization on landmark-based registration performance. These results from the baseline were used for the subsequent comparison with fully automated intensity-based registration techniques, such as Elastix and ANTs.

The dataset includes tissue samples from lung lobes, where markers such as ProSpC, Ki67, Cc10, and Cd31 were used to stain different cell types. These are registered to HE-stained references to assess structural consistency. The mammary gland samples include stains like CNEU, ER (Estrogen Receptor), and PR (Progesterone Receptor), commonly used in breast cancer diagnostics, registered to HE variants (HE37 and HE39). Similarly, lung lesion tissues (pathological lung samples) include diverse staining such as ProSpC, Ki67, Cc10, and Cd31 to capture inflammation or abnormal growth, challenging registration accuracy due to structural distortions. The full form of stains can be found in the following Table 1.

Additionally, the dataset includes breast tissue samples stained with HER2, a key biomarker in breast cancer, all aligned to HE. Mice kidney samples stained with Pycytokeratin and CD31 and registered to either HE or PAS serve to test the generalization of the registration approach across samples. Finally, human kidney samples stained with MAS and registered to PAS offer complex internal structures like glomeruli and tubules that pose challenges for non-rigid alignment.

Table 7. Mapping of tissue types, sample identifiers, and corresponding stain pairs used for registration

Tissue Type	Sample Number	Moving Image	Fixed Image
Lung lobes	lung lobes_1	ProSpc	HE
		Ki67	HE
		Cc10	HE
	lung lobes_2	Cd31	HE
		ProSpc	HE
		Cc10	HE
	lung_lobes_3	Cd31	HE
		ProSpc	HE
		Ki67	HE
	lung_lobes_4	Cd31	HE
		Cc10	HE
		Ki67	HE
Mammary gland	Mammary gland_1	CNEU	HE39
		ER	HE39
		PR	HE39
		PR	HE37
		ER	HE37
		CNEU	HE37
Lung lesion	lung_lesion_1	Cc10	HE
		Ki67	HE
		ProSpc	HE
	lung_lesion_2	Cd31	HE
		Cc10	HE
		ProSpc	HE
lung_lesion_3	Cc10	HE	
	Ki67	HE	
	ProSpc	HE	
Breast	Breast_1	HER2	HE
	Breast_2	HER2	HE
	Breast_3	HER2	HE
	Breast_4	HER2	HE
Mice kidney	Mice kidney_1	Pycytokeratin	HE
	Mice kidney_2	CD31	PAS
Kidney	Kidney_1	MAS	PAS
		MAS	PAS

The results in Table 8 and Table 9 demonstrate that the mean TRE values remain consistent across both the original and stain-normalized image pairs for all tissue types. This indicates that stain normalization had no significant impact on the registration performance when using the landmark-based method. The reason for this invariance lies in the fundamental nature of the technique itself: Landmark-based registration relies solely on spatial correspondences between anatomical structures, rather than on pixel intensity values or color distributions. Since the landmarks used were pre-defined and anatomically grounded, changes in stain appearance introduced by normalization did not affect the geometric alignment process. These findings confirm that landmark-

based registration provides a robust structure baseline that is inherently resistant to staining variability, making it a valuable reference for comparison with intensity-driven registration methods in the subsequent sections.

Table 8. Mean(TRE) before and after stain normalization across different sample pairs

Sample Number	Moving Image	Fixed Image	(mean)TRE (px)	(TRE)After Normalization (px)
Lung lobes_1	ProSpc	HE	1.73	1.73
	Ki67	HE	3.50	3.50
	Cc10	HE	2.50	2.50
Lung lobes_2	Cd31	HE	1.95	1.95
	ProSpc	HE	2.39	2.39
	Cc10	HE	2.80	2.80
Lung_lobes_3	Cd31	HE	1.03	1.03
	ProSpc	HE	1.42	1.42
	Ki67	HE	2.36	2.36
Lung_lobes_4	Cd31	HE	1.67	1.67
	Cc10	HE	1.82	1.82
	Ki67	HE	2.60	2.60
Mammary Gland_1	CNEU	HE39	3.77	3.77
	ER	HE39	1.36	1.36
	PR	HE39	4.02	4.02
	PR	HE37	4.95	4.95
	ER	HE37	2.35	2.35
	CNEU	HE37	1.49	1.49
Lung_lesion_1	Cc10	HE	8.14	8.14
	Ki67	HE	9.47	9.47
	ProSpc	HE	9.06	9.06
Lung_lesion_2	Cd31	HE	9.35	9.35
	Cc10	HE	10.75	10.75
	ProSpc	HE	12.10	12.10
Lung_lesion_3	Cc10	HE	9.56	9.56
	Ki67	HE	11.50	11.50
	ProSpc	HE	9.42	9.42
Breast_1	HER2	HE	6.06	6.04
Breast_2	HER2	HE	14.94	14.94
Breast_3	HER2	HE	18.82	18.82
Breast_4	HER2	HE	15.64	15.64
Mice kidney_1	Pycytokeratin	HE	5.87	5.87
Mice kidney_2	PAS	Cd31	4.34	4.34
Kidney_1	MAS	PAS	15.95	15.95
	MAS	PAS	9.20	9.20

Table 9. Average mean TRE before and after stain normalization across tissue groups

Tissue Group	(mean)TRE (px)	(mean)TRE after Normalization (px)
Lung lobes	2.145	2.145
Mammary gland	2.99	2.99
Lung lesion	9.92	9.92
Breast	13.86	13.86
Mice kidney	5.10	5.10
Kidney	12.57	12.57

The qualitative assesment of the registration method is done by comparing visuals and evaluating landmarks. Figure 10 displays the fixed H&E image, the corresponding moving stained image, and the warped image that is obtained after registration for two types of tissue. They demonstrate how similar the structures are after the registration process. To further support the analysis, Figure 11 demonstrates a landmark-based evaluation by overlaying annotated landmarks on source, target, and warped images. The true shift, estimated (warped) shift, and TRE vectors are visualized to show the geometric accuracy of registration.

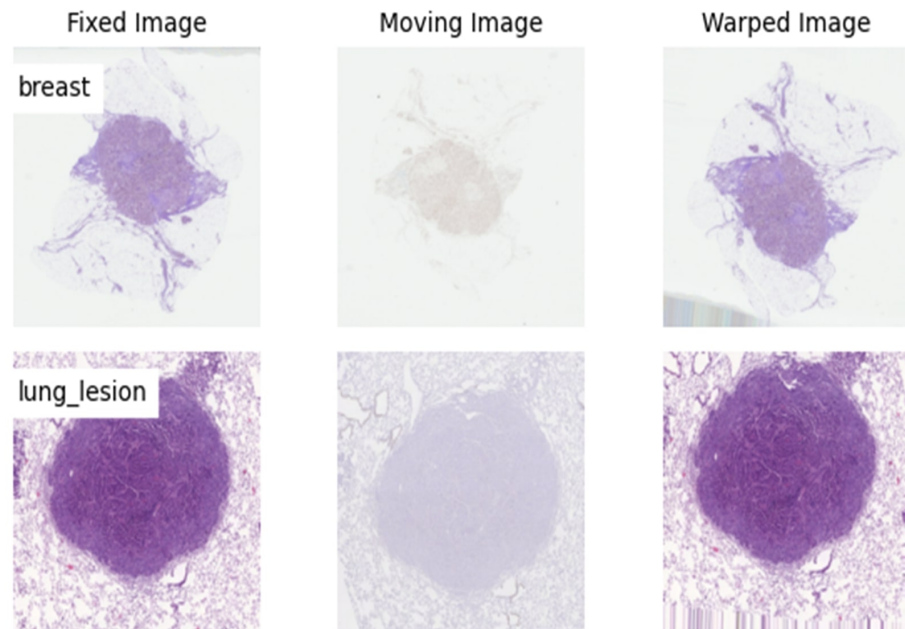


Figure 10. Visual comparison of registration results. Each row shows a tissue case (breast and lung lesion) with its fixed H&E image, the moving immunostained image, and the warped output after registration.

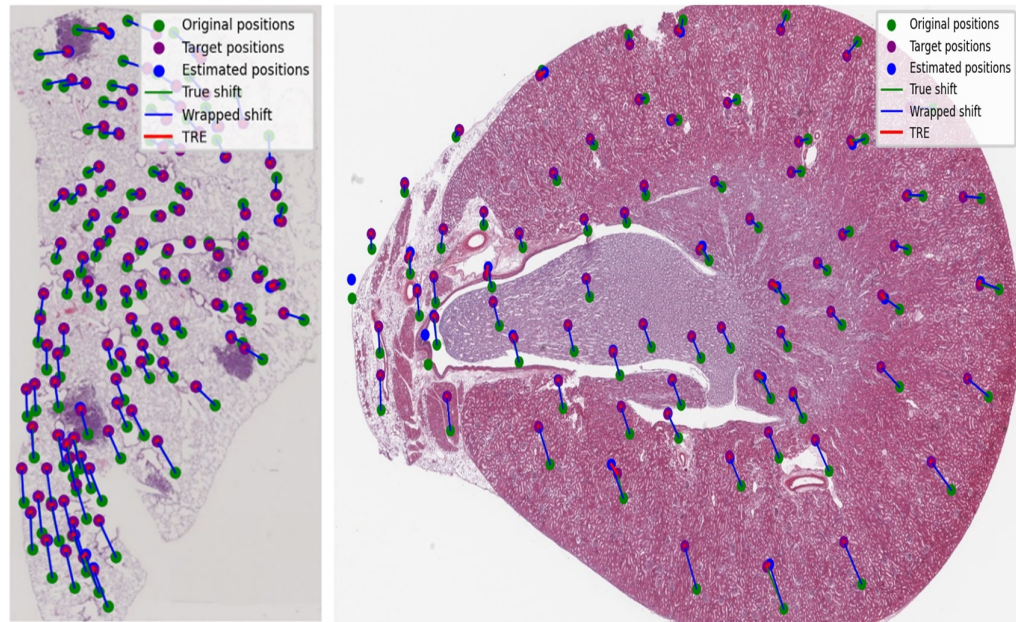


Figure 11. Landmark-based visualization of registration performance. The figure shows original (green), target (magenta), and estimated (blue) landmark positions for two tissue samples. The lines represent true shift (green), warped shift (blue), and TRE (red) vectors to visually compare registration accuracy.

Performance Results of Elastix Method:

After completing the landmark-based registration experiments, the intensity-based registration method was evaluated to further assess the impact of stain normalization. Unlike the landmark-based approach, which relies on spatial landmark correspondences, intensity-based methods align images by comparing pixel intensities. These methods are more susceptible to staining variations, making them ideal for testing how stain normalization affects registration accuracy.

For the intensity-based registration experiments, a subset of diverse image pairs was selected from the dataset shown in Table 10, ensuring a balanced representation of tissue types. These pairs were tested under two conditions: original stain and stain-normalized versions. The inclusion of both original and normalized images allowed us to directly compare the robustness of Elastix and ANTs in handling staining variability. By drawing these diverse pairs, the aim was to evaluate the generalizability and performance of both methods across a diverse range of tissue types and staining conditions.

Table 10. List of tissue group samples with corresponding moving and fixed stain image pairs used in the registration process

Tissue Group	Moving Image	Fixed Image
Lung_lobes_1	ProSPC	H&E
Lung_lobes_1	Cc10	H&E
Lung_lobes_2	cd31	H&E
Lung_lobes_2	ProSPC	H&E
Lung_lobes_3	ProSPC	H&E
Lung_lobes_4	Cc10	H&E
Lung_lobes_4	ki67	H&E
Mammary_gland_1	PR	H&E
Lung_lesion_1	Cc10	H&E
Lung_lesion_3	Cc10	H&E
Lung_lesion_3	Ki67	H&E
Breast_2	HER2	H&E
Mice_kidney_1	PanCytokeratin	H&E
Mice_kidney_2	CD31	PAS
Kidney_1	MAS	PAS

The Elastix registration toolkit is used first to perform the registration and is evaluated with several metrics such as mean TRE, maximum TRE, median TRE and mean rTRE. To evaluate the effect of stain normalization on registration accuracy, these metrics were calculated separately for raw (non-normalized) images and for stain normalized images. The configuration setting that was used for the procedure is described in Table 3 in the methodology Section 3.3.2. We find that registration accuracy varies widely across tissue types and stain combinations.

In general, samples with moderate tissue deformation and consistent staining, such as mice_kidney_2 and mice_kidney_1—had low TRE and rTRE values, with rTRE mean as low as 0.001 and 0.006, respectively, implying accurate alignment. On the other hand, difficult cases such as kidney_1 and lung_lesion_3 resulted in much greater errors, with TRE mean values of 31.145 and 31.515, respectively, and a TRE max of 86.73, indicating that Elastix was not successful in aligning images with substantial stain variability or significant morphological differences.

Furthermore, lung_lobes_2 showed the highest registration error among the lung lobe samples, with a TRE mean of 27.254 and rTRE mean of 0.039. This is likely due to the use of Cd31 staining, which targets endothelial structures, resulting in sparse or non-overlapping features relative to its H&E counterpart. In addition, mammary_gland_1 had a high TRE max of 44.256, suggesting the presence of localized mismatches, even though the overall alignment appeared acceptable with a TRE mean of 15.934. The results can be seen in Table 11.

The normalized images of several tissue samples, particularly mice_kidney_1, lung_lobes_2, and lung_lobes_3, resulted in notably high TRE mean and rTRE mean values. For instance, lung_lobes_3 exhibited a TRE mean of 300.950 and an rTRE mean of 0.487. In contrast, some tissues such as breast_2 and mice_kidney_2 demonstrated relatively small registration errors even after stain normalization. This suggests that normalization may still be beneficial in cases where stain differences are minimal or where anatomical structures are well-preserved. Nonetheless, the

overall trend underscores that staining variability significantly affects registration performance. The post normalization results can be seen in the Table 12.

Table 11. Elastix registration performance without stain normalization. TRE and rTRE statistics are computed per tissue case

Tissue	Moving Image	Fixed Image	TRE Mean	TRE Max	TRE Median	rTRE Mean
Lung_lobes_1	ProSPC	H&E	13.898	20.65	14.616	0.024
Lung_lobes_1	cc10	H&E	15.024	21.347	15.170	0.026
Lung_lobes_2	cd31	H&E	27.254	39.416	26.826	0.039
Lung_lobes_2	ProSPC	H&E	15.236	33.168	13.227	0.022
Lung_lobes_3	ProSPC	H&E	9.627	15.530	9.169	0.015
Lung_lobes_4	cc10	H&E	14.701	18.075	15.106	0.023
Lung_lobes_4	ki67	H&E	19.503	28.250	19.283	0.030
Mammary_gland_1	PR	H&E	15.934	44.256	15.832	0.008
Lung_lesion_1	Cc10-5	H&E	11.567	29.100	10.820	0.010
Lung_lesion_3	Cc10-5	H&E	18.490	54.575	16.987	0.016
Lung_lesion_3	Ki67	H&E	31.515	86.730	29.128	0.028
Breast_2	HER2	H&E	20.763	64.179	17.973	0.012
Mice_kidney_1	PanCytokeratin	H&E	8.563	22.866	8.278	0.006
Mice_kidney_2	CD31	PAS	3.502	18.512	2.641	0.001
Kidney_1	MAS	PAS	31.145	53.004	32.896	0.010

Table 12. Registration results using Elastix with stain-normalized moving images

Tissue	Moving Image	Fixed Image	TRE Mean	TRE Max	TRE Median	rTRE Mean
Lung_lobes_1	ProSPC	H&E	41,670	125,270	24,111	0,011
Lung_lobes_1	cc10	H&E	40,027	129,502	26,816	0,072
Lung_lobes_2	cd31	H&E	102,179	141,223	102,580	0,149
Lung_lobes_2	ProSPC	H&E	102,027	138,245	101,631	0,149
Lung_lobes_3	ProSPC	H&E	300,950	441,688	294,429	0,487
Lung_lobes_4	cc10	H&E	20,174	254,369	88,870	0,540
Lung_lobes_4	ki67	H&E	100,495	133,078	99,955	0,158
Mammary_gland_1	PR	H&E	57,419	89,222	55,283	0,031
Lung_lesion_1	Cc10-5	H&E	60,650	148,049	56,543	0,053
Lung_lesion_3	Cc10-5	H&E	34,798	71,902	31,555	0,031
Lung_lesion_3	Ki67	H&E	103,237	226,427	95,913	0,093
Breast_2	HER2	H&E	19,411	71,119	18,018	0,011
Mice_kidney_1	PanCytokeratin	H&E	142,887	163,634	144,053	0,102
Mice_kidney_2	CD31	PAS	3,545	19,795	2,703	0,002
Kidney_1	MAS	PAS	128,009	191,228	124,201	0,045

It was shown that staining variability can have a profound effect on intensity-based registration accuracy, as demonstrated by comparing the registration performance of

raw versus stain-normalized images using Elastix can be seen in Figure 12. In the experiments involving raw images, several tissue pairs such as mice_kidney_2 and mice_kidney_1 exhibited low TRE mean values of 3.502 and 8.563, respectively, along with low rTRE mean values, indicating good structural alignment. Conversely, other raw samples such as kidney_1 and lung_lesion_3 showed relatively higher error values, though still within an acceptable range for medical image registration.

However, the registration accuracy degraded significantly after applying stain normalization in most cases. Tissue pairs that initially demonstrated good alignment, such as lung_lobes_3 and mice_kidney_1, showed dramatic increases in TRE mean (300.950 and 142.887, respectively) and elevated rTRE mean values, reflecting poor spatial correspondence post-normalization. Only a few tissue pairs, notably breast_2 and mice_kidney_2, maintained or slightly improved their registration performance following normalization, suggesting that its effectiveness may depend on stain similarity and preservation of anatomical features.

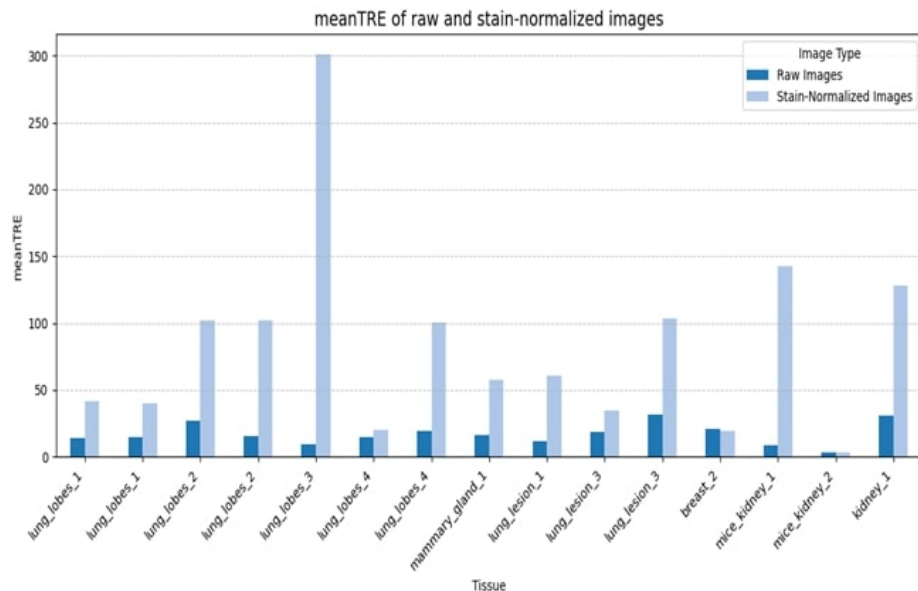


Figure 12. Comparison of mean TRE values between raw and stain-normalized images across different tissue samples. Stain normalization led to an increase in mean TRE for several cases such as lung_lobes_3 and mice_kidney_1, indicating potential challenges in handling significant stain variability.

For the qualitative assessment of Elastix. Several important features are included in the landmark overlays can be seen in Figure 13: green marks the original position in the source image, magenta shows the corresponding position in the fixed image, and blue marks the position found after registration. The green lines mark the true difference between the old and new points, and the blue lines mark the differences caused by the transformation. Red lines point out the TRE, which show the difference between the warped and real target locations. With no stain normalization, the top image suggests that the landmarks are well-aligned and the TRE vectors are not very long, which means the registration is accurate. On the other hand, the bottom image with stain normalization reveals that the red TRE vectors are longer and more frequent, which indicates that the registration did not perform well.

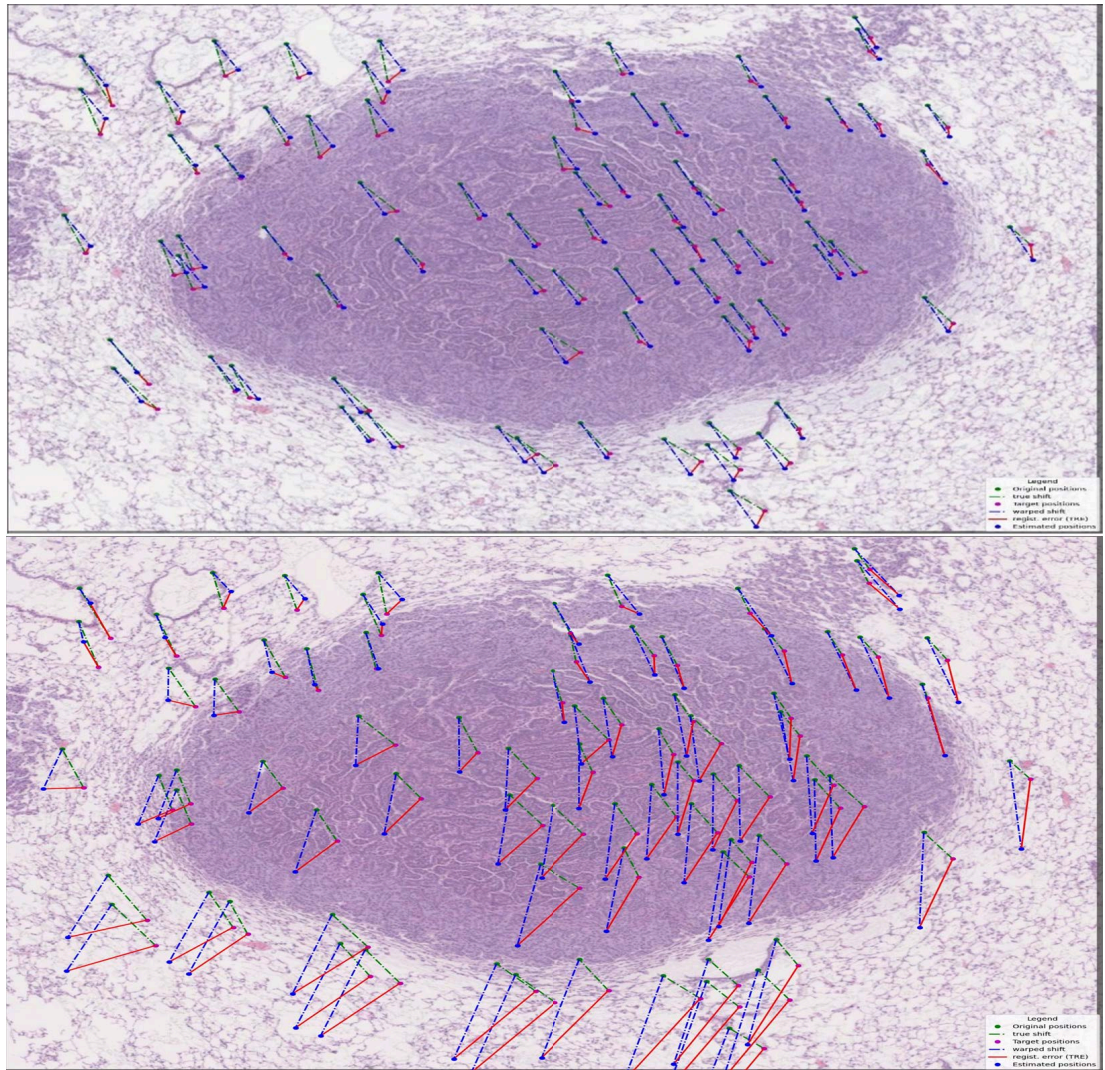


Figure 13. Visualization of registration accuracy for lung lesion tissue. The top image shows the warped landmark alignments without stain normalization, whereas the bottom image displays the alignment after applying stain normalization. Note that the red lines in the bottom image indicate increased TRE values, reflecting degraded spatial correspondence following normalization.

Performance Results of ANTs Method:

After evaluating the performance of Elastix, the analysis was further extended using ANTs, a widely used intensity-based image registration framework known for its high accuracy in medical image alignment. ANTs utilize a Symmetric Normalization (SyN) transformation model combined with mutual information or cross-correlation as similarity metrics, making it particularly suitable for non-linear and deformable registration tasks. The configuration setting that was used is described in the Table 4. In this study, ANTs was applied to the same subset of tissue image pairs used in the Elastix experiments can be seen in the Table 10, both with original stained images and their stain-normalized counterparts. This allowed for a direct comparison between the two intensity-based methods and provided further insight into how staining variability

affects the performance of automated registration pipelines. The goal was to assess whether ANTs can offer improved robustness and alignment accuracy under variable staining conditions introduced through normalization.

The registration performance for ANTs was robust and consistent across most tissue types. Table 13 presents the results for ANTs registration method without stain normalization. Especially good results were obtained for the lung_lobes_2 and lung_lobes_3 achieved low TRE mean of 3.908 and 3.347, respectively and rTRE below 0.006. In mice_kidney_2 and mice_kidney_1, the TREs were also good, with values of 4.852 and 5.625 and rTRE as low as 0.002 and 0.004. A particular exception was a mammary_gland_1 with a very high TRE mean of 182.709 and rTRE mean of 0.098. This indicates that ANTs had a very hard time with this specific stain combination which may be due to poor structural overlap or large stain dependent contrast differences. Lung_lesion_3 exhibited relatively high registration errors with the TRE mean of 27.905, probably because of complex pathological structures and different staining characteristics. To further investigate the influence of stain normalization on registration accuracy, the ANTs framework was applied to the stain-normalized moving images, using the same parameter configuration as in the raw image experiments.

Table 13. Registration results using ANTs without stain normalization

Tissue	Moving Image	Fixed Image	TRE Mean	TRE Max	TRE Median	rTRE Mean
Lung_lobes_1	ProSPC	H&E	4,315	16,838	2,864	0,008
Lung_lobes_1	cc10	H&E	7,883	34,748	5,166	0,014
Lung_lobes_2	cd31	H&E	3,694	19,720	2,253	0,005
Lung_lobes_2	ProSPC	H&E	3,908	23,798	1,607	0,006
Lung_lobes_3	ProSPC	H&E	3,347	13,153	2,137	0,005
Lung_lobes_4	cc10	H&E	3,404	10,363	2,775	0,005
Lung_lobes_4	ki67	H&E	4,092	12,338	3,252	0,006
Mammary_gland_1	PR	H&E	182,709	240,223	182,879	0,098
Lung_lesion_1	Cc10-5	H&E	8,355	25,415	6,767	0,007
Lung_lesion_3	Cc10-5	H&E	12,437	50,005	8,552	0,011
Lung_lesion_3	Ki67	H&E	27,905	78,367	24,967	0,025
Breast_2	HER2	H&E	18,566	65,489	15,412	0,011
Mice_kidney_1	PanCytokeratin	H&E	5,625	22,927	4,817	0,004
Mice_kidney_2	CD31	PAS	4,852	25,843	3,556	0,002
Kidney_1	MAS	PAS	16,082	34,348	15,152	0,006

In general, ANTs showed stable and often improved performance Table 14, presents the results after stain normalization, especially for tissue samples that had moderate errors in the raw image condition. For example, the previously challenging case mammary_gland_1 had an extremely large TRE mean of 182.709 on raw images and after normalization this error was dramatically reduced to a TRE mean of 5.181 and rTRE mean of 0.003, suggesting that normalization effectively synchronized appearance and structural correspondence in this case. Likewise, lung lobe samples like lung_lobes_3 and lung_lobes_2 also gave further improvements or maintained already low TRE values of 2.535 and 3.540, respectively with rTRE values below

0.006 which indicates good and consistent alignment. The results of this study indicate that ANTs, because of their robust transformation model and its metric based on mutual information, can utilize the advantages of stain normalization in many circumstances. Some exceptions however remained, mainly in lung_lesion_3 and kidney_1. The former had a TRE mean of 42.507 and the latter had an extremely high TRE mean of 219.149 with an rTRE of 0.077, indicating that in some stain tissue combinations normalization may have introduced artifacts or distortions which deceived the registration algorithm.

Table 14. Registration results using ANTs with stain-normalized moving images

Tissue	Moving Image	Fixed Image	TRE Mean	TRE Max	TRE Median	rTRE Mean
Lung_lobes_1	ProSPC	H&E	4,676	21,713	2,848	0,008
Lung_lobes_1	Cc10	H&E	8,263	33,509	6,128	0,015
Lung_lobes_2	Cd31	H&E	3,540	21,236	2,165	0,005
Lung_lobes_2	ProSPC	H&E	3,997	21,696	1,535	0,006
Lung_lobes_3	ProSPC	H&E	2,535	10,273	1,417	0,004
Lung_lobes_4	Cc10	H&E	3,240	10,965	2,977	0,005
Lung_lobes_4	Ki67	H&E	4,873	14,224	3,764	0,008
Mammary_gland_1	PR	H&E	5,181	40,563	3,362	0,003
Lung_lesion_1	Cc10-5	H&E	8,042	29,123	6,708	0,007
Lung_lesion_3	Cc10-5	H&E	12,833	47,828	9,519	0,012
Lung_lesion_3	Ki67	H&E	42,507	78,612	41,702	0,038
Breast_2	HER2	H&E	18,712	67,312	15,779	0,011
Mice_kidney_1	PanCytokeratin	H&E	6,790	25,499	6,124	0,005
Mice_kidney_2	CD31	PAS	4,847	20,079	3,558	0,002
Kidney_1	MAS	PAS	219,149	270,360	242,952	0,077

Figure 14 illustrates the mean TRE for each tissue sample using the ANTs registration pipeline on both raw and stain-normalized images. Overall, the comparison shows that ANTs maintained robust performance across both conditions, with stain normalization leading to mixed results depending on the tissue type.

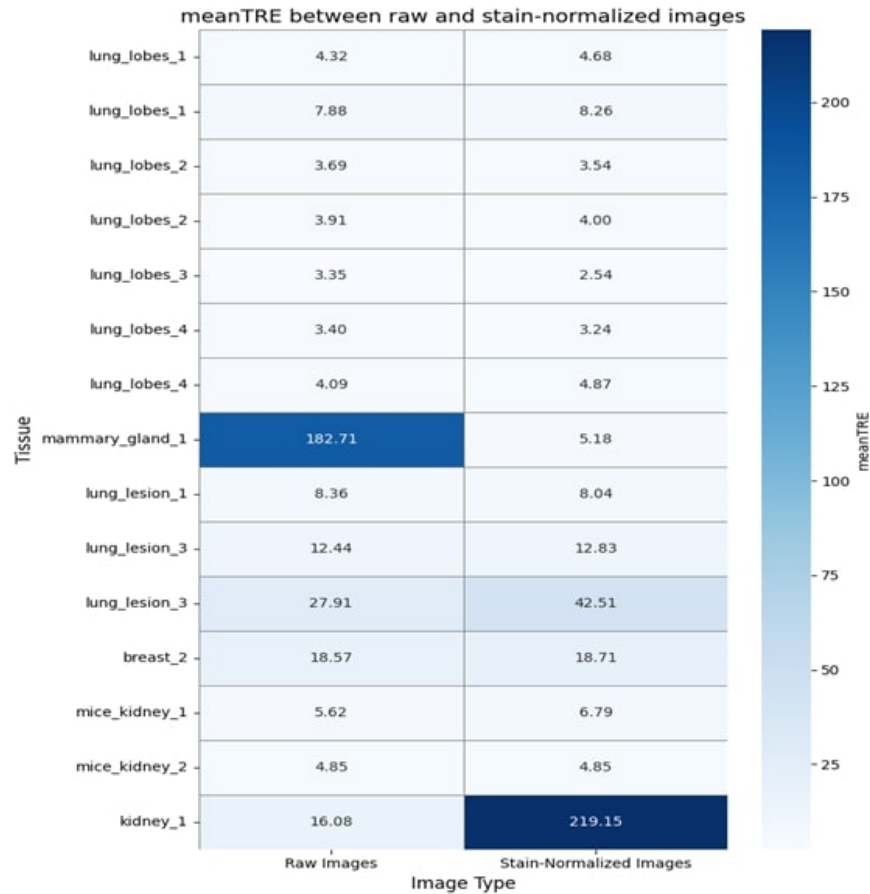


Figure 14. Heatmap comparison of mean Target Registration Error (TRE) between raw and stain-normalized images across various tissue samples using the ANTs registration framework. Darker shades indicate higher TRE values, highlighting significant error reduction in mammary_gland_1 post-normalization, and error increase in kidney_1.

Figure 15 and Figure 16 give examples of the warped mammary_gland_1 sample tissue before and after stain normalization. These findings highlight that while ANTs generally perform well even without normalization, their performance can benefit from stain normalization when there is severe staining inconsistency, as seen in mammary_gland_1.

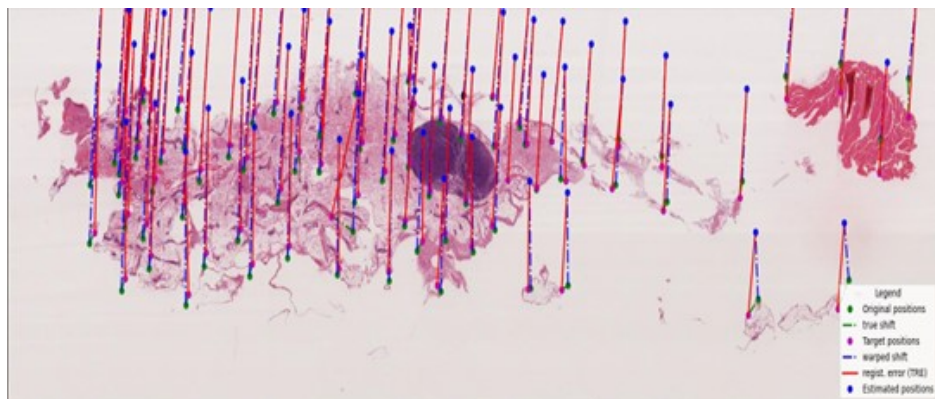


Figure 15. The warped tissue (mammary_gland_1) with transformation landmarks before normalization.

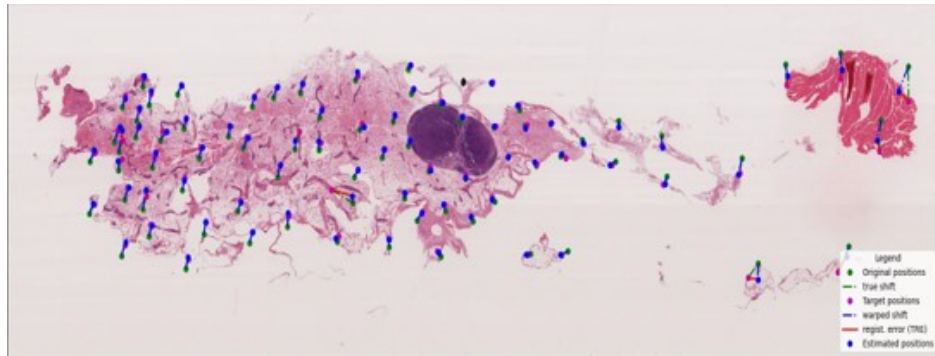


Figure 16. The warped tissue (mammary_gland_1) with transformation landmarks after normalization.

4.1.3. Comparative Study

This study further comprehensively analyzed histopathological image registration algorithms to check how well they align raw histological images from various tissue types and how stain normalization impacts their performance. In previous Section 4.1.2 all three methods result well-presented and discussed separately for both conditions i.e. raw images as well as normalized images. The registration methods were Elastix, ANTs and a developed Landmark-based technique. Using all these techniques allowed us to see how various methods behave in different situations. This study explored how changes in stain color can be a problem for registration, since there is often a big difference in the intensity of staining in histopathological images. So, the tissue samples were also processed by stain normalization to see how this step changes the accuracy of each registration algorithm. The main goal was to check if each method deals with issues caused by different stains and the complexity of tissues. This work finds out how various registration methods behave with staining changes, and which one works the best in such conditions. The outcomes of this comparison will show the best registration approaches for analyzing histological images, especially when stain normalization is applied beforehand. The next part reports on the results of the comparison by studying the mean TRE, the amount of variability between methods and how well each method handles staining differences after normalization.

It can be seen in Figure 17 that on raw images, landmark-based methods performed best in terms of mean TRE, having a median value below 5 and a narrow range between the highest and lowest values, showing they were both accurate and consistent for all tissue types. ANTs performed well, with a median TRE almost the same as Landmark's based, but its results were less consistent overall and often showed a higher upper quartile, indicating that ANTs results can vary with the type and complexity of the data used. When compared to Elastix, it was consistently better at performing its tasks. Unlike the others, Elastix had the highest overall TRE values, with a wider interquartile range and several outliers extending beyond 30 TRE units. Thus, it becomes clear that Elastix may not perform as well in registration when the data has different stains or structures and this is made worse if the data is not preprocessed or the stains not normalized. All in all, landmark-based registration delivers the best alignment and ANTs comes next, since it is both flexible and strong. Elastix, despite

being efficient in some cases, does not handle the natural variation in histological images well, which makes it important to choose a method that fits the data and the goal.

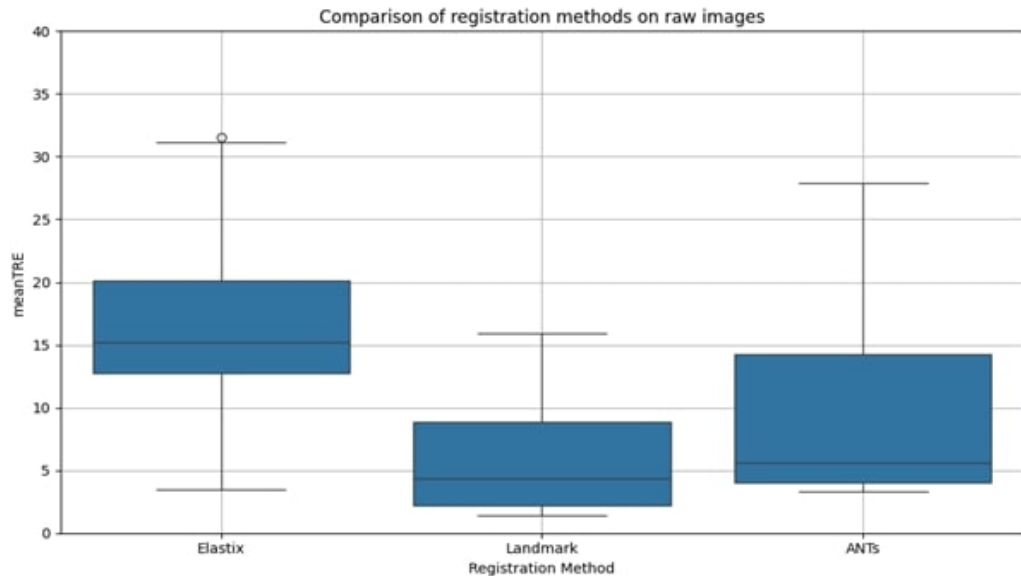


Figure 17. Comparison of registration methods on raw images using box plots of mean TRE values. Landmark-based registration achieved the lowest variation and mean error, while Elastix exhibited higher variance and ANTs showed balanced performance.

Further, the comparison aims to find out how much stain normalization influence the ability of each registration method to handle variations in staining and the alignment precision. After normalizing data with staining, the performance of each group is quite different. The ANTs method has the lowest mean TRE, which is close to zero, with a narrow interquartile range and little variability, suggesting it was very effective at aligning the images. The landmark-based method also gave satisfying results, having close to the same mean TRE as ANTs, but with a bit more variation, which shows it is a dependable method, although differences in tissue structures and landmark spotting can make it less precise. Elastix had considerably higher average TRE values than ANTs and the Landmark-based approach. The wider interquartile range and presence of outliers suggest that Elastix is more sensitive to staining differences, which can affect its performance when applied to tissues with higher collection in staining patterns.

In addition, the top quarter of Elastix scores is much higher and a few extreme values are present, pointing to situations where the approach struggled to normalize stains, probably due to its sensitivity to stain intensity differences. After stain normalization, ANTs appear to be both the most reliable and accurate method, dealing with different staining levels and maintaining accurate alignment of structures. Landmark-based methods are similarly successful in competitions, but their results depend on the quality of the landmarks, while Elastix is less affected by stain normalization and might require more advanced preparations to work well with different staining methods.

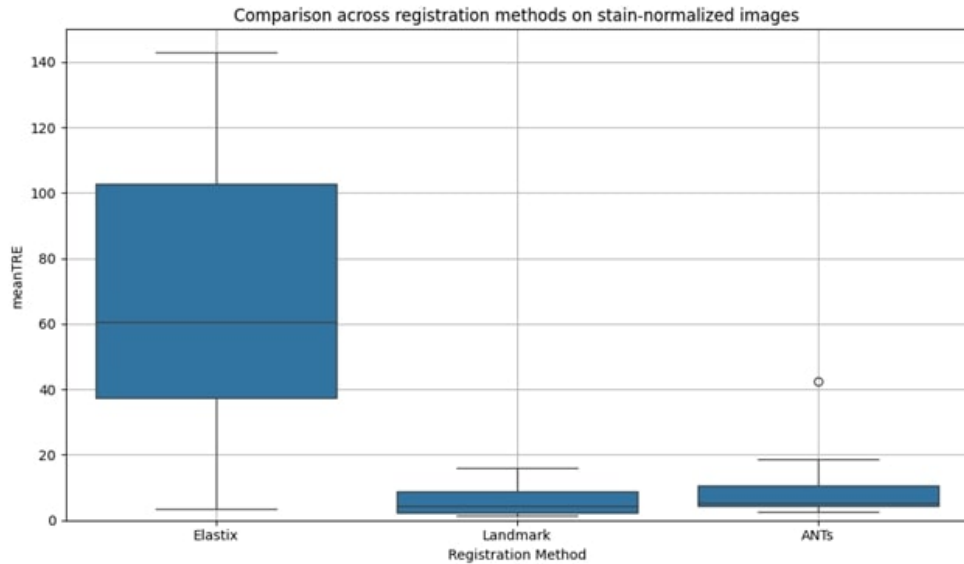


Figure 18. Comparison across registration methods on stain-normalized images. The box plot shows that Landmark-based registration maintains the lowest mean TRE and smallest variance, while Elastix exhibits the highest error spread. ANTs provides a balance with moderate error and variation.

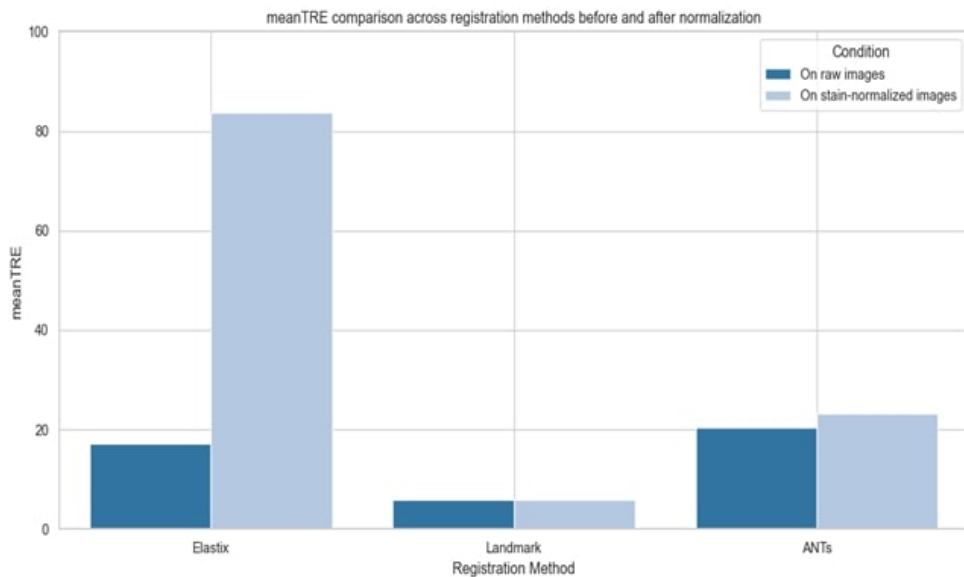


Figure 19. Mean TRE comparison across registration methods before and after stain normalization. While Landmark-based registration shows consistently low error under both conditions, Elastix demonstrates a significant increase in mean TRE after normalization. ANTs performance remains relatively stable.

Figure 19 shows collectively the outcomes of all three registration methods (Elastix, Landmark-based and ANTs) for both kinds of images: raw and stain-normalized. It shows how the methods perform differently, so we can clearly see how each one handles different types of stains. Using the mean TRE for all methods in both conditions, the plot clearly illustrates how stain normalization affects registration accuracy.

5. DISCUSSION

5.1. Summary of Key Findings

The goal of this study was to evaluate the accuracy of a newly developed landmark-based registration method, as well as two intensity-based methods, Elastix and ANTs, in registering histopathology images that have been stained differently. The primary objective was to determine how stain normalization affects the accuracy and robustness of these registration approaches.

The study found that the developed landmark-based registration method, utilizing pre-set anatomical landmarks, was not significantly influenced by the different stains used. The TRE value showed negligible little change before and after stain normalization, indicating that this method is not sensitive to variations in staining. In contrast, Elastix and ANTs, both of which rely on intensity, were affected differently by stain normalization. Specifically, Elastix performed less effectively than ANTs after normalization, particularly in tissues with significant staining differences.

Overall, the landmark-based registration method demonstrated superior accuracy, especially in cases with considerable variation in staining. This robustness to staining differences makes the landmark-based approach a strong option for histopathology image registration, which is particularly important given the common occurrence of staining variability. By comparing the landmark-based method with existing intensity-based methods, the landmark-based approach is better suited to handle these staining differences, making it an excellent choice for histopathology image registration.

5.2. Comparison with Existing Literature

The findings of this study align with existing literature that emphasizes the impact of stain variability on image registration. Previous studies, such as those by Kang et al. (2019) and Xu et al. (2021), have explored the challenges of cross-stain image registration, highlighting how staining differences complicate the alignment process for intensity-based methods [65, 68], like Elastix and ANTs. In contrast, landmark-based methods, which focus on aligning anatomical structures rather than relying on pixel intensities, have shown greater resilience to staining variations. Further, normalizing stains with StainNet highly impacted on these intensity-based methods, though the improvement is more obvious in ANTs and Elastix struggles. This means that ANTs which uses mutual information and can handle various deformations, is more flexible in processing-stained images than Elastix which depends on pixel intensity.

By applying stain normalization, the study was able to assess how the registration methods handle differences in staining. It was expected that landmark-based registration would not change its accuracy when using stained or unstained images. Because it depends on set anatomical points, the method is not as influenced by changes in staining. On the other hand, both Elastix and ANTs saw a rise in registration errors after stain normalization, mainly in areas with big structural and staining differences (e.g., mammary gland, lung lesion). This points to the fact that stain normalization can change the way intensities are seen in images, which may

disrupt registration methods that depend on matching pixels. Even though ANTs performed better than Elastix, it experienced performance issues when handling tough tissue transformations or major inconsistencies in staining. This result agrees with previous studies by Bejnordi et al. (2017) [80], and Zheng et al. (2019) [68], which stated that stain normalization can improve registration but should be used cautiously to avoid artifacts and ensure accuracy.

Landmark-based registration is especially strong because it handles variations in staining well. It makes use of TPS interpolation to match images according to anatomical landmarks, which are not affected by the intensity changes caused by various stains. According to the study, using landmarks for registration gives predictable and accurate results, with very little difference in mean TRE values for various tissues. Even so, this method is limited because it depends on predetermined anatomical points that need to be figured out in the images. Even though the ANHIR dataset includes landmark annotations, this process may take a lot of time and can have multiple errors. This type of registration also assumes that tissue deformations are smooth, which might not fit images with very complex or distorted areas. For such cases, ANTs and Elastix can be more flexible in dealing with local changes in tissue.

5.3. Future Directions

The results of this study suggest that image registration is valuable in DP. Landmark-based registration is both very accurate and computationally efficient, making it suitable for uses such as tumor detection and the study of biomarker expression together. Since landmark-based registration is not affected much by staining differences, it is especially valuable in cases where images from several stains are studied together, such as in multi-modal tumor microenvironment analysis. Meanwhile, intensity-based methods are sensitive to different staining patterns and their performance may decline a lot when dealing with challenging stains. Stain normalization reduces some of the sensitivity, yet it cannot fully correct the misalignment caused by big differences in staining between samples. Thus, clinicians ought to pick the best registration method, depending on what they are analyzing in the tissue samples.

Future studies could work on improving the landmark-based registration method in a way to automate the process of annotating landmarks with DL. The spotting landmarks in images is slow and requires a lot of effort which makes it hard to use landmark-based methods on large sets of data. If CNNs or transformer-based architectures are trained on large, labeled datasets, they can be used to automate the identification of landmarks which would save a lot of time and effort. Adding DL-based landmark detection would allow the method to be used on many tissue types, staining methods and image types. This may result in a more automated, reliable and scalable process in hospitals, mainly for tasks that require multi-modal image registration and annotations are not abundant.

As well as automating landmark annotation, looking into combining stain normalization with hybrid registration techniques opens a promising direction for future research. Using both landmark-based and intensity-based methods (such as ANTs or Elastix) might create hybrid models that make use of the best features

of each approach. The landmark-based technique is strong because it focuses on anatomical points and is not affected by staining differences, but intensity-based methods are flexible and can handle small changes in tissue images. Combining these techniques can improve how accurately registration is performed, mainly in challenging conditions with both uneven staining and flexible tissue movements. Using hybrid methods, first the landmarks are used to roughly align the images and then more details are added with intensity-based methods. Using these hybrid models may be very helpful for analyzing complex medical data that uses several stains and needs precise alignment.

Also, adding generative models like CycleGANs or Stable Diffusion to the registration process can greatly strengthen the method when there are big differences in staining. It has been shown in recent studies that CycleGANs are very good at changing images from one form to another and they can create images that look like the target stain even if the images were stained differently. Stain normalization is one of the complex tasks that Stable Diffusion can handle by repeatedly improving the image during a denoising process. The use of generative models in registration would lead to improved stain normalization which would result in more consistent and accurate cross-stain registration. Because these models are data-driven, they can figure out complex links between stains, so staining differences have less effect on registration. But since generative models take a lot of computing resources, more work could be done to enhance their performance and make them more real-time applicable, especially in medical settings.

Lastly, future studies should concentrate on using deep learning for landmark annotation, using different registration methods together to enhance results with different stains, including generative models to make stain normalization more reliable and developing more advanced multi-resolution techniques for working with high-resolution images. As a result, registration will be more automated, more accurate and scalable which will help digital pathology workflows in both research and clinical areas.

6. SUMMARY

This thesis focused on solving problems in histopathology image registration by designing a landmark-based method and comparing it with two state-of-the-art intensity-based techniques, Elastix and ANTs. In medical diagnostics, particularly for cancer, histopathological images, which are often stained in various ways, are crucial. Still, it is not easy to align these images because different stains and techniques can result in significant variation in tissue samples. The study aimed to investigate the performance of different registration methods when the images being aligned have been stained differently. The purpose of this research was to examine if the landmark-based method, developed in this study, could provide a better and more efficient approach to image registration than existing intensity-based methods, mainly when dealing with complicated and different staining standards. It was found that the landmark-based method provided better registration accuracy than Elastix and ANTs, mainly when the staining was highly variable.

Additionally, it was examined whether stain normalization affected the accuracy of registration. Before registering the histopathology images, StainNet was used for stain normalization. The tests showed that the landmark-based registration did not depend on stain normalization since it performed equally well on both types of images. Due to its strength, landmark-based registration is a suitable choice for cases where high accuracy is required for various types of tissue and staining methods. The findings of this study are significant for the use of image registration in digital pathology (DP). Landmark-based registration is highly accurate, efficient, and reliable, making it a suitable choice for clinical work that requires precise image alignment, such as tumor detection, studying different biomarkers, and assessing tumor microenvironments. The landmark-based method is well-suited for situations that involve comparing stains, as it is not affected by differences in staining and can process data from multiple stains. This is particularly important in clinics, where various staining methods are used to examine tissue samples and multiple biomarkers must be checked simultaneously.

In summary, the study demonstrates that the landmark-based registration method is highly effective in histopathology image registration because it is not easily affected by different stains and dyes. This method is a good choice for comparing slides stained differently in digital pathology. Although ANTs and Elastix provide important alternatives, their sensitivity to different staining methods shows that it is essential to select the correct registration method for each type of tissue. It provides a starting point for further research, such as automating landmark annotation, integrating different registration strategies, and utilizing generative models to enhance the accuracy of stain normalization and registration. These advances enable the registration process to be more scalable, efficient, and robust, thus improving digital pathology and improving the way diagnoses are made in clinics.

7. REFERENCES

- [1] Yang H. et al. (2025) SAStainDiff: Self-supervised stain normalization by stain augmentation using denoising diffusion probabilistic models. *Biomedical Signal Processing and Control* 107, p. 107861. DOI: <https://doi.org/10.1016/j.bspc.2025.107861>.
- [2] Sudhamsh G.V.S., Girisha S. & Rashmi R. (2025) Semi-supervised tissue segmentation from histopathological images with consistency regularization and uncertainty estimation. *Scientific Reports* 15, p. 6506. DOI: <https://doi.org/10.1038/s41598-025-90221-x>.
- [3] Jeong J., Kim K.D., Nam Y., Cho C.E., Go H. & Kim N. (2023) Stain normalization using score-based diffusion model through stain separation and overlapped moving window patch strategies. *Computers in Biology and Medicine* 152, p. 106335. DOI: <https://doi.org/10.1016/j.compbiomed.2022.106335>.
- [4] Feng M. et al. (2020) Automated quantitative analysis of Ki-67 staining and HE images recognition and registration based on whole tissue sections in breast carcinoma. *Diagnostic Pathology* 15, p. 65. DOI: <https://doi.org/10.1186/s13000-020-00957-5>.
- [5] Borovec J. et al. (2020) ANHIR: Automatic Non-Rigid Histological Image Registration Challenge. *IEEE Transactions on Medical Imaging* 39, pp. 3042–3052. DOI: <https://doi.org/10.1109/TMI.2020.2986331>.
- [6] Xiong Z. & Zhang Y. (2010) A critical review of image registration methods. *International Journal of Image and Data Fusion* 1, pp. 137–158. DOI: <https://doi.org/10.1080/19479831003802790>.
- [7] Hoque M.Z., Keskinarkaus A., Nyberg P., Mattila T. & Seppänen T. (2022) Whole slide image registration via multi-stained feature matching. *Computers in Biology and Medicine* 144, p. 105301. DOI: <https://doi.org/10.1016/j.compbiomed.2022.105301>.
- [8] Paknezhad M. et al. (2020) Regional registration of whole slide image stacks containing major histological artifacts. *BMC Bioinformatics* 21, p. 558. DOI: <https://doi.org/10.1186/s12859-020-03907-6>.
- [9] Alotaibi S.R., Alohali M.A., Maashi M., Alqahtani H., Alotaibi M. & Mahmud A. (2025) Advances in colorectal cancer diagnosis using optimal deep feature fusion approach on biomedical images. *Scientific Reports* 15, p. 4200. DOI: <https://doi.org/10.1038/s41598-024-83466-5>.
- [10] Kiran N. et al. (2023) Digital Pathology: Transforming Diagnosis in the Digital Age. *Cureus* DOI: <https://doi.org/10.7759/cureus.44620>.
- [11] Du Z. et al. (2025) Deeply supervised two stage generative adversarial network for stain normalization. *Scientific Reports* 15, p. 7068. DOI: <https://doi.org/10.1038/s41598-025-91587-8>.

- [12] Roy M. et al. (2023) Deep learning based registration of serial whole-slide histopathology images in different stains. *Journal of Pathology Informatics* 14. DOI: <https://doi.org/10.1016/j.jpri.2023.100311>.
- [13] Jiang J., Larson N.B., Prodduturi N., Flotte T.J. & Hart S.N. (2019) Robust hierarchical density estimation and regression for re-stained histological whole slide image co-registration. *PLoS One* 14. DOI: <https://doi.org/10.1371/journal.pone.0220074>.
- [14] Wang C.W., Ka S.M. & Chen A. (2014) Robust image registration of biological microscopic images. *Scientific Reports* 4, p. 6050. DOI: <https://doi.org/10.1038/srep06050>.
- [15] Brown M.V. et al. (2012) Cancer detection and biopsy classification using concurrent histopathological and metabolomic analysis of core biopsies. *Genome Medicine* 4, p. 33. DOI: <https://doi.org/10.1186/gm332>.
- [16] Rivenson Y. et al. (2019) Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nature Biomedical Engineering* 3, pp. 466–477. DOI: <https://doi.org/10.1038/s41551-019-0362-y>.
- [17] Chan J.K.C. (2014) The Wonderful Colors of the Hematoxylin–Eosin Stain in Diagnostic Surgical Pathology. *International Journal of Surgical Pathology* 22, pp. 12–32. DOI: <https://doi.org/10.1177/1066896913517939>.
- [18] Martell M.T. et al. (2023) Deep learning-enabled realistic virtual histology with ultraviolet photoacoustic remote sensing microscopy. *Nature Communications* 14, p. 5967. DOI: <https://doi.org/10.1038/s41467-023-41574-2>.
- [19] ANHIR Challenge Organizers (2020), ANHIR: Automatic Non-rigid Histological Image Registration Challenge. <https://anhir.grand-challenge.org/Data/>. Accessed: 2025-06-09.
- [20] Kumar N., Gupta R. & Gupta S. (2020) Whole Slide Imaging (WSI) in Pathology: Current Perspectives and Future Directions. *Journal of Digital Imaging* 33, pp. 1034–1040. DOI: <https://doi.org/10.1007/s10278-020-00351-z>.
- [21] Ma Y., Jamdade S., Konduri L. & Sailem H. (2025) AI in Histopathology Explorer for comprehensive analysis of the evolving AI landscape in histopathology. *NPJ Digital Medicine* 8, p. 156. DOI: <https://doi.org/10.1038/s41746-025-01524-2>.
- [22] Holub P. et al. (2023) Privacy risks of whole-slide image sharing in digital pathology. *Nature Communications* 14, p. 2577. DOI: <https://doi.org/10.1038/s41467-023-37991-y>.
- [23] Khened M., Kori A., Rajkumar H., Krishnamurthi G. & Srinivasan B. (2021) A generalized deep learning framework for whole-slide image segmentation

- and analysis. *Scientific Reports* 11, p. 11579. DOI: <https://doi.org/10.1038/s41598-021-90444-8>.
- [24] Herrmann M.D. et al. (2018) Implementing the DICOM Standard for Digital Pathology. *Journal of Pathology Informatics* 9, p. 37. DOI: https://doi.org/10.4103/jpi.jpi_42_18.
- [25] National Electrical Manufacturers Association (n.d.), Dicom whole slide imaging. <https://dicom.nema.org/dicom/dicomwsi/>. Accessed: 2025-06-08.
- [26] Pyatov V.A. & Sorokin D.V. (2025) Unsupervised Feature Matching for Affine Histological Image Registration. In: *Lecture Notes in Computer Science*, pp. 34–48. DOI: https://doi.org/10.1007/978-3-031-78201-5_3.
- [27] Gatenbee C.D. et al. (2023) Virtual alignment of pathology image series for multi-gigapixel whole slide images. *Nature Communications* 14, p. 4502. DOI: <https://doi.org/10.1038/s41467-023-40218-9>.
- [28] Trahearn N., Epstein D., Cree I., Snead D. & Rajpoot N. (2017) Hyper-Stain Inspector: A Framework for Robust Registration and Localised Co-Expression Analysis of Multiple Whole-Slide Images of Serial Histology Sections. *Scientific Reports* 7, p. 5641. DOI: <https://doi.org/10.1038/s41598-017-05511-w>.
- [29] Wodzinski M., Marini N., Atzori M. & Müller H. (2024) RegWSI: Whole slide image registration using combined deep feature- and intensity-based methods: Winner of the ACROBAT 2023 challenge. *Computer Methods and Programs in Biomedicine* 250, p. 108187. DOI: <https://doi.org/10.1016/j.cmpb.2024.108187>.
- [30] Elhaminia B. et al. (2025), From Traditional to Deep Learning Approaches in Whole Slide Image Registration: A Methodological Review. Preprint or early access; publication details pending.
- [31] Zou J., Gao B., Song Y. & Qin J. (2022) A review of deep learning-based deformable medical image registration. *Frontiers in Oncology* 12. DOI: <https://doi.org/10.3389/fonc.2022.1047215>.
- [32] Zhang W. (2022) Robust registration of SAR and optical images based on deep learning and improved Harris algorithm. *Scientific Reports* 12, p. 5901. DOI: <https://doi.org/10.1038/s41598-022-09952-w>.
- [33] Strittmatter A., Caroli A. & Zöllner F.G. (2023) A Multistage Rigid-Affine-Deformable Network for Three-Dimensional Multimodal Medical Image Registration. *Applied Sciences* 13, p. 13298. DOI: <https://doi.org/10.3390/app132413298>.
- [34] Auer M., Regitnig P. & Holzapfel G.A. (2005) An automatic nonrigid registration for stained histological sections. *IEEE Transactions on Image Processing* 14, pp. 475–486. DOI: <https://doi.org/10.1109/TIP.2005.843756>.

- [35] Venet L., Pati S., Feldman M.D., Nasrallah M.P., Yushkevich P. & Bakas S. (2021) Accurate and robust alignment of differently stained histologic images based on greedy diffeomorphic registration. *Applied Sciences (Switzerland)* 11, pp. 1–18. DOI: <https://doi.org/10.3390/app11041892>.
- [36] Abdel-Basset M., Fakhry A.E., El-henawy I., Qiu T. & Sangaiah A.K. (2017) Feature and Intensity Based Medical Image Registration Using Particle Swarm Optimization. *Journal of Medical Systems* 41, p. 197. DOI: <https://doi.org/10.1007/s10916-017-0846-9>.
- [37] Hisham M.B., Yaakob S.N., Raof R.A.A., Nazren A.B.A. & Wafi N.M. (2015) Template Matching using Sum of Squared Difference and Normalized Cross Correlation. In: 2015 IEEE Student Conference on Research and Development (SCoReD), IEEE, pp. 100–104. DOI: <https://doi.org/10.1109/SCORED.2015.7449303>.
- [38] Ching W.S. (1995) Normalized cross-correlation: the contrast-dependent problem and its solution. *Journal of Electronic Imaging* 4, p. 278. DOI: <https://doi.org/10.1117/12.210718>.
- [39] Kosiński W., Michalak P. & Gut P. (2012) Robust Image Registration Based on Mutual Information Measure. *Journal of Signal and Information Processing* 3, pp. 175–178. DOI: <https://doi.org/10.4236/jsip.2012.32023>.
- [40] Doyle J. et al. (2023) Whole-Slide Imaging, Mutual Information Registration for Multiplex Immunohistochemistry and Immunofluorescence. *Laboratory Investigation* 103, p. 100175. DOI: <https://doi.org/10.1016/j.labinv.2023.100175>.
- [41] Karami E., Prasad S. & Shehata M. (2017), Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images. Publication details unavailable; possibly a preprint or workshop paper.
- [42] Bay H., Ess A., Tuytelaars T. & Van Gool L. (2008) Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding* 110, pp. 346–359. DOI: <https://doi.org/10.1016/j.cviu.2007.09.014>.
- [43] Zitová B. & Flusser J. (2003) Image registration methods: a survey. *Image and Vision Computing* 21, pp. 977–1000. DOI: [https://doi.org/10.1016/S0262-8856\(03\)00137-9](https://doi.org/10.1016/S0262-8856(03)00137-9).
- [44] Zhang C. et al. (2022), A Hybrid Deep Feature-Based Deformable Image Registration Method for Pathology Images. Publication details unavailable; possibly a preprint or workshop paper.
- [45] DeTone D., Malisiewicz T. & Rabinovich A. (2017), SuperPoint: Self-Supervised Interest Point Detection and Description. Available as preprint.
- [46] Song G., Han J., Zhao Y., Wang Z. & Du H. (2017) A Review on Medical Image Registration as an Optimization Problem. *Current Medical Imaging Reviews* 13. DOI: <https://doi.org/10.2174/1573405612666160920123955>.

- [47] Dong J., Lu K., Xue J., Dai S., Zhai R. & Pan W. (2018) Accelerated nonrigid image registration using improved Levenberg–Marquardt method. *Information Sciences* 423, pp. 66–79. DOI: <https://doi.org/10.1016/j.ins.2017.09.059>.
- [48] LeCun Y., Bengio Y. & Hinton G. (2015) Deep learning. *Nature* 521, pp. 436–444. DOI: <https://doi.org/10.1038/nature14539>.
- [49] Boor P. (2024) Deep learning applications in digital pathology. *Nature Reviews Nephrology* 20, pp. 702–703. DOI: <https://doi.org/10.1038/s41581-024-00870-w>.
- [50] Li Y. et al. (2024) Virtual histological staining of unlabeled autopsy tissue. *Nature Communications* 15, p. 1684. DOI: <https://doi.org/10.1038/s41467-024-46077-2>.
- [51] Jewsbury R., Wang R., Bhalerao A., Rajpoot N. & Vu Q.D. (2024), StainFuser: Controlling Diffusion for Faster Neural Style Transfer in Multi-Gigapixel Histology Images. Publication venue not specified; possibly a preprint or under review.
- [52] Wei X., Ge L., Huang L., Luo J. & Xu Y. (2025) Unsupervised Non-Rigid Histological Image Registration Guided by Keypoint Correspondences Based on Learnable Deep Features With Iterative Training. *IEEE Transactions on Medical Imaging* 44, pp. 447–461. DOI: <https://doi.org/10.1109/TMI.2024.3447214>.
- [53] Wodzinski M. & Müller H. (2021) DeepHistReg: Unsupervised Deep Learning Registration Framework for Differently Stained Histology Samples. *Computer Methods and Programs in Biomedicine* 198, p. 105799. DOI: <https://doi.org/10.1016/j.cmpb.2020.105799>.
- [54] Wei X., Ge L., Huang L., Luo J. & Xu Y. (2025) Unsupervised Non-Rigid Histological Image Registration Guided by Keypoint Correspondences Based on Learnable Deep Features With Iterative Training. *IEEE Transactions on Medical Imaging* 44, pp. 447–461. DOI: <https://doi.org/10.1109/TMI.2024.3447214>.
- [55] Faryna K., Van Der Laak J. & Litjens G. (2021), Tailoring automated data augmentation to H&E-stained histopathology. Available online: <https://github.com/DIAGNijmegen/pathology-he-auto-augment>.
- [56] Saha A., Prasad P. & Thabit A. (2020) Leveraging Adaptive Color Augmentation in Convolutional Neural Networks for Deep Skin Lesion Segmentation. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE, pp. 2014–2017. DOI: <https://doi.org/10.1109/ISBI45749.2020.9098344>.
- [57] Shorten C. & Khoshgoftaar T.M. (2019) A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data* 6, p. 60. DOI: <https://doi.org/10.1186/s40537-019-0197-0>.

- [58] Sebai M., Wang X. & Wang T. (2020) MaskMitosis: a deep learning framework for fully supervised, weakly supervised, and unsupervised mitosis detection in histopathology images. *Medical & Biological Engineering & Computing* 58, pp. 1603–1623. DOI: <https://doi.org/10.1007/s11517-020-02175-z>.
- [59] Golestani N., Wang A., Moallem G., Bean G.R. & Rusu M. (2025) PViT-AIR: Puzzling vision transformer-based affine image registration for multi histopathology and faxitron images of breast tissue. *Medical Image Analysis* 99, p. 103356. DOI: <https://doi.org/10.1016/j.media.2024.103356>.
- [60] Khan A.M., Rajpoot N., Treanor D. & Magee D. (2014) A Nonlinear Mapping Approach to Stain Normalization in Digital Histopathology Images Using Image-Specific Color Deconvolution. *IEEE Transactions on Biomedical Engineering* 61, pp. 1729–1738. DOI: <https://doi.org/10.1109/TBME.2014.2303294>.
- [61] Salehi P. & Chalechale A. (2020) Pix2Pix-based Stain-to-Stain Translation: A Solution for Robust Stain Normalization in Histopathology Images Analysis. In: 2020 International Conference on Machine Vision and Image Processing (MVIP), IEEE, pp. 1–7. DOI: <https://doi.org/10.1109/MVIP49855.2020.9116895>.
- [62] Hoque M.Z., Keskinarkaus A., Nyberg P. & Seppänen T. (2024) Stain normalization methods for histopathology image analysis: A comprehensive review and experimental comparison. *Information Fusion* 102, p. 101997. DOI: <https://doi.org/10.1016/j.inffus.2023.101997>.
- [63] Macenko M. et al. (2009) A method for normalizing histology slides for quantitative analysis. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE, pp. 1107–1110. DOI: <https://doi.org/10.1109/ISBI.2009.5193250>.
- [64] Reinhard E., Adhikhmin M., Gooch B. & Shirley P. (2001) Color transfer between images. *IEEE Computer Graphics and Applications* 21, pp. 34–41. DOI: <https://doi.org/10.1109/38.946629>.
- [65] Kang H. et al. (2021) StainNet: A Fast and Robust Stain Normalization Network. *Frontiers in Medicine (Lausanne)* 8. DOI: <https://doi.org/10.3389/fmed.2021.746307>.
- [66] Ehteshami Bejnordi B. et al. (2016) Stain Specific Standardization of Whole-Slide Histopathological Images. *IEEE Transactions on Medical Imaging* 35, pp. 404–415. DOI: <https://doi.org/10.1109/TMI.2015.2476509>.
- [67] Swiderska-Chadaj Z. et al. (2020) Impact of rescanning and normalization on convolutional neural network performance in multi-center, whole-slide classification of prostate cancer. *Scientific Reports* 10, p. 14398. DOI: <https://doi.org/10.1038/s41598-020-71420-0>.

- [68] Zheng Y., Jiang Z., Zhang H., Xie F., Shi J. & Xue C. (2019) Adaptive color deconvolution for histological WSI normalization. *Computer Methods and Programs in Biomedicine* 170, pp. 107–120. DOI: <https://doi.org/10.1016/j.cmpb.2019.01.008>.
- [69] Ho J., Jain A. & Abbeel P. (2020), Denoising Diffusion Probabilistic Models. Available at: <https://arxiv.org/abs/2006.11239>.
- [70] Wu K. et al. (2021) Improvement in Signal-to-Noise Ratio of Liquid-State NMR Spectroscopy via a Deep Neural Network DN-Unet. *Analytical Chemistry* 93, pp. 1377–1382. DOI: <https://doi.org/10.1021/acs.analchem.0c03087>.
- [71] Rombach R., Blattmann A., Lorenz D., Esser P. & Ommer B. (2021), High-Resolution Image Synthesis with Latent Diffusion Models. Available at: <https://arxiv.org/abs/2112.10752>.
- [72] Podell D. et al. (2023), SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis. Preprint; available at: <https://arxiv.org/abs/2307.01952>.
- [73] Shen Y. & Ke J. (2023) StainDiff: Transfer Stain Styles of Histology Images with Denoising Diffusion Probabilistic Models and Self-ensemble. In: *Lecture Notes in Computer Science*, pp. 549–559. DOI: https://doi.org/10.1007/978-3-031-43987-2_53.
- [74] Borovec J. (2019), BIRL: Benchmark on Image Registration methods with Landmark validation. Preprint or project report; publication details not specified.
- [75] Shaban M.T., Baur C., Navab N. & Albarqouni S. (2019) Staingan: Stain Style Transfer for Digital Histological Images. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE, pp. 953–956. DOI: <https://doi.org/10.1109/ISBI.2019.8759152>.
- [76] Zhu J.Y., Park T., Isola P. & Efros A.A. (2017) Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In: 2017 IEEE International Conference on Computer Vision (ICCV), IEEE, pp. 2242–2251. DOI: <https://doi.org/10.1109/ICCV.2017.244>.
- [77] Sara U., Akter M. & Uddin M.S. (2019) Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. *Journal of Computer and Communications* 7, pp. 8–18. DOI: <https://doi.org/10.4236/jcc.2019.73002>.
- [78] Klein S., Staring M., Murphy K., Viergever M.A. & Pluim J. (2010) elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Transactions on Medical Imaging* 29, pp. 196–205. DOI: <https://doi.org/10.1109/TMI.2009.2035616>.

- [79] Avants B.B., Tustison N.J., Song G., Cook P.A., Klein A. & Gee J.C. (2011) A reproducible evaluation of ants similarity metric performance in brain image registration. *NeuroImage* 54, pp. 2033–2044.
- [80] Ehteshami Bejnordi B. et al. (2016) Stain Specific Standardization of Whole-Slide Histopathological Images. *IEEE Transactions on Medical Imaging* 35, pp. 404–415. DOI: <https://doi.org/10.1109/TMI.2015.2476509>.