UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE INFORMÁTICA

Physics-Guided Deep Learning for Sparse Data-Driven Brain Shift Registration

Tiago Miguel da Silva Assis

Mestrado em Ciência de Dados

Dissertação orientada por:
Prof. Dr. Nuno Cruz Garcia
Prof.ª Dr.ª Inês Prata Machado

2025

Dedicated to my grandparents, Bia and Vitor, and to my girlfriend, Ana.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisors, Prof. Dr. Nuno Garcia and Prof. Dr. Inês Machado, for their constant weekly support throughout my dissertation work, providing invaluable guidance, constructive feedback, and encouragement that pushed me to grow both technically and personally. For the same reasons, I am equally grateful to Dr. Reuben Dorent, whose great ideas and insight shaped this work from its inception to completion, and who kindly provided me with the opportunity to attend MICCAI 2025 in South Korea, where our work was shared. Without their expertise, availability, and genuine engagement, this thesis would not have reached its current level of maturity, nor would I have learned so much about the field of medical image registration. They created a supportive and laid-back atmosphere, which made this journey far more inspiring and enjoyable.

I would also like to thank my friends who made these years a little easier. To my colleagues and friends at FCUL: Rute, Beatriz, and Ana; thank you for making classes and studying more enjoyable, and for the long after-hours talks that I greatly miss. To my lifelong friends, with whom I have shared countless moments and who continue to be a constant presence despite time and life changes, thank you for the memories and the support. A special mention goes to Mateus and Ema, whose friendship has been particularly meaningful.

A heartfelt thank you to my girlfriend Ana, the person whom I love and share my life with daily. She has stood by me through uncertainty, doubt, and achievement, lifting me when I wanted to quit and celebrating every success alongside me. She shares my burdens as I share hers, and together we carry them forward. We will always be ”Sifrang”. I am also grateful and in debt to her parents, whose generosity and support made it possible for me to pursue these studies by providing me with a home.

Os meus mais profundos agradecimentos finais vão para os meus avós, Bia e Vitor, que me tornaram quem eu sou e que sempre me apoiaram de forma incondicional, mesmo nos momentos mais difíceis quando tudo parecia impossível. Que nunca deixaram de me amar, apesar da minha distância ao longo destes anos, tanto física como emocional; obrigado por estarem sempre lá quando eu regresso. Merecem muito mais do que alguma vez poderei retribuir.

This work was partially supported by Fundação para a Ciência e a Tecnologia (FCT) through the LASIGE Research Unit, ref. UID/408/2025.

Resumo

A neurocirurgia de tumores cerebrais depende fortemente de sistemas de neuronavegação para guiar o neurocirurgião na localização precisa de tumores e estruturas anatómicas adjacentes, segundo o seu pré-planeamento cirúrgico. Estes sistemas localizam os instrumentos cirúrgicos em relação à posição real da anatomia do paciente no bloco operatório e às imagens pré-operatórias. No entanto, este alinhamento torna-se progressivamente mais difícil devido ao fenómeno conhecido como *brain shift*, ou desvio cerebral, que corresponde a deformações não lineares do tecido cerebral. Estas deformações são induzidas por fatores físicos, cirúrgicos e biológicos, como a ação da gravidade, a perda de líquido cefalorraquidiano, a remoção de tecido tumoral e alterações na pressão intracraniana. Mesmo desvios da ordem de poucos milímetros podem comprometer significativamente a precisão da neuronavegação, aumentando o risco de remoções incompletas do tumor ou de danos a tecidos saudáveis. Assim, a compensação destes desvios cerebrais constitui um dos principais desafios na neurocirurgia guiada por imagem.

A compensação dos efeitos do desvio cerebral nas imagens utilizadas na neuronavegação é geralmente abordada através de algoritmos de registo (alinhamento) de imagens. Para aumentar a precisão e robustez da neuronavegação e do registo de imagens, modalidades intraoperatórias são utilizadas em alguns centros hospitalares, recorrendo ao uso de ultrassons intraoperatórios ou de ressonância magnética intraoperatória. Apesar da sua eficácia, estas abordagens apresentam limitações práticas importantes, incluindo, no caso do ultrassom, dificuldades associadas à qualidade da imagem e à sua interpretação e, no caso da ressonância magnética, custos elevados, necessidade de infraestruturas especializadas e interrupções ao fluxo cirúrgico.

Métodos de registo baseados em aprendizagem profunda têm demonstrado grandes avanços, permitindo inferências rápidas e eliminando a necessidade de otimizações iterativas para cada paciente, como é comum em métodos tradicionais. Contudo, muitas destas abordagens dependem de dados intraoperatórios de alta qualidade, apresentam dificuldades em cenários de alinhamento entre imagens de modalidades diferentes e podem estimar deformações fisicamente implausíveis quando não são devidamente regularizadas. Uma alternativa consiste em utilizar métodos de registo baseados na obtenção de correspondências esparsas, nos quais um conjunto limitado de pontos anatómicos, ou de interesse, é identificado e emparelhado entre imagens pré- e intraoperatórias. Os desvios cerebrais entre esses pontos correspondentes nas imagens permitem efetuar a interpolação de um volume tridimensional denso com desvios que são a base para efetuar o registo. Estes métodos são particularmente robustos a grandes deformações, alterações topológicas

induzidas pela remoção de tecido tumoral e aos campos de visão limitados proporcionados pelos ultrassons, além de serem mais compatíveis com aplicações de registo em tempo real. No entanto, a interpolação destes desvios esparsos depende, geralmente, de interpoladores geométricos simples, como interpolações lineares ou *thin-plate splines*, que ignoram as propriedades biomecânicas do tecido cerebral. Consequentemente, as deformações resultantes não seguem princípios físicos que levem em consideração a força da gravidade, a flexibilidade e o possível colapso de tecidos e forças de compressão ou extensão.

Sendo assim, esta dissertação propõe uma nova abordagem baseada em aprendizagem profunda guiada pela física com vista a melhorar a precisão e robustez de métodos de registo de imagens médicas quando são usados dados intraoperatórios esparsos. A hipótese central é que um modelo de aprendizagem profunda, treinado com supervisão proveniente de simulações biomecânicas realistas, pode atuar como um "interpolador" avançado capaz de refinar interpolações iniciais de desvios obtidos a partir de correspondências esparsas e interpoladores geométricos em volumes densos de desvios biomecanicamente plausíveis. O objetivo é alcançar um compromisso eficaz entre precisão, realismo biomecânico e eficiência computacional, tornando o método adequado para a integração em sistemas de neuronavegação intraoperatórios.

Para concretizar esta abordagem, foi desenvolvido um fluxo de trabalho completo que combina modelação biomecânica do cérebro, geração de dados sintéticos, deteção de pontos de interesse esparsos e dos seus respetivos desvios e aprendizagem profunda. Primeiramente, foi construída uma base de dados de deformações cirúrgicas sintéticas a partir de imagens de ressonância magnética pré-operatórias reais. Para cada caso, foi gerada uma geometria específica do paciente e atribuídas propriedades mecânicas aos diferentes tecidos e componentes cerebrais (parênquima, ventrículos e tumor), recorrendo a um modelo biomecânico hiperelástico e a um método numérico sem geração de malhas estruturadas (*meshless*). Estas deformações foram simuladas de forma a produzir desvios realistas, induzidas por variações na direção do vetor da gravidade ou na localização do ponto de entrada cirúrgico, servindo como *ground truth*, ou verdades, para a aprendizagem do modelo.

Em segundo lugar, foram automaticamente detetados pontos (voxels, em 3D) de interesse nas imagens pré-operatórias utilizando o algoritmo 3D *scale-invariant feature transform (SIFT)*, um algoritmo de deteção de pontos de interesse invariantes tanto a rotações como a variações de escala da imagem, e associados a estes os seus respetivos desvios obtidos nas simulações biomecânicas. Este procedimento permitiu gerar pares de dados de compostos por um conjunto de desvios conhecidos em locais esparsos da imagem e o correspondente tensor denso de desvios simulados, constituindo assim um conjunto de treino supervisionado adequado para aprendizagem guiada pela física.

O cerne da abordagem proposta é o desenvolvimento de um interpolador profundo guiado biomecanicamente baseado numa arquitetura residual do tipo 3D U-Net. O modelo recebe como sinal de entrada a imagem pré-operatória e um tensor de desvios inicial obtido por interpolação geométrica simples da informação esparsa associada aos pontos de interesse. A rede neuronal é treinada para prever uma correção residual que refina a interpolação inicial, aproximando-a de um

tensor de desvios biomecanicamente realistas. Esta formulação residual simplifica o processo de aprendizagem ao restringir a rede à modelação de correções locais sobre uma estimativa inicial plausível, promovendo uma convergência mais estável e eficiente. Ao explorar explicitamente o conhecimento prévio dos interpoladores geométricos, o modelo concentra-se na incorporação de efeitos biomecânicos não lineares, aumentando a robustez e a plausibilidade física dos desvios estimados.

A avaliação experimental do método foi realizada por comparação quantitativa entre as previsões da rede neuronal e as verdades sintéticas obtidas pelas simulações biomecânicas de um conjunto de dados de validação, enquanto que a avaliação qualitativa foi efetuada entre as simulações biomecânicas e imagens intraoperatórias de ressonância magnética de casos cirúrgicos reais. Além disso, o desempenho do método foi diretamente comparado com interpoladores geométricos simples para avaliar o efeito do processo de refinamento proposto. Os resultados demonstram que o interpolador proposto supera consistentemente a utilização exclusiva de interpoladores simples, tanto em termos de erro médio de desvio como na preservação de propriedades físicas desejáveis. Estudos de ablação evidenciam a importância das escolhas da arquitetura da rede, da formulação residual e dos mecanismos de regularização na obtenção de resultados robustos. Estas melhorias são alcançadas com um custo computacional marginal, mantendo tempos de inferência compatíveis com aplicações intraoperatórias em tempo real.

Do ponto de vista clínico, a abordagem apresentada oferece várias vantagens. Ao depender apenas de informação esparsa, reduz a necessidade de aquisição de imagens intraoperatórias de alta qualidade, tornando-se potencialmente aplicável em ambientes com recursos limitados. A integração de conhecimento biomecânico no processo de aprendizagem da rede neuronal aumenta a plausibilidade física das deformações estimadas, contribuindo para maior confiança na atualização da neuronavegação durante o decorrer da cirurgia quando métodos de registo baseados em pontos de interesse esparsos são empregados para a compensação de desvios cerebrais. E, para além disso, o carácter modular do interpolador permite a sua integração no fluxo de métodos de registo já existentes durante a etapa de interpolação sem necessidade de reformular completamente o processo.

Em suma, esta dissertação introduz um novo paradigma para a interpolação de dados intraoperatórios esparsos em mecanismos de registo e compensação de desvios cerebrais. A principal contribuição reside na demonstração de que é possível para uma rede neuronal aprender, a partir de simulações físicas, um modelo eficiente e preciso que atua como um elo entre interpoladores geométricos simples e abordagens biomecânicas complexas. Este trabalho avança o estado da arte no registo de imagens neurológicas intraoperatórias e estabelece bases sólidas para investigações futuras, incluindo a validação em uma maior quantidade de dados clínicos, a sua integração em métodos de registo já existentes e a sua extensão para cenários multimodais complexos.

Palavras-chave: aprendizagem profunda informada por física, desvio cerebral, interpolação de pontos de interesse esparsos, modelação biomecânica, registo de imagens médicas

Abstract

Accurate neuronavigation in brain tumor surgery relies on the alignment between pre-operative imaging and the patient's intra-operative anatomy. This alignment is severely compromised by brain shift, a complex and non-linear deformation caused by several factors, including gravity, cerebrospinal fluid loss, and tissue resection. Existing image registration approaches often require access to dense and high-quality intra-operative data or computationally expensive biomechanical simulations, limiting their practicality in real-time surgical workflows. Registration methods relying on sparse keypoints offer a promising alternative but typically rely on simple geometric interpolators that ignore the biomechanical behavior of brain tissue to estimate dense displacement fields, resulting in physically implausible deformations.

This thesis proposes a physics-guided deep learning framework that acts as an advanced interpolator of sparse displacement information for brain shift correction. The method integrates data-driven learning with biomechanical modeling by training a residual 3D U-Net-based network to refine initial displacement fields derived from sparse matched keypoints. Supervision is provided by a large-scale dataset of deformations generated through patient-specific biomechanical simulations, ensuring that the learned deformations are biomechanically consistent. The proposed network is designed to be modular, efficient, and compatible with existing keypoint-based registration pipelines that require minimal intra-operative data.

Experimental results evaluated the simulated intra-operative brain deformations and the refinements estimated by the proposed network, demonstrating that the proposed approach significantly outperforms standard interpolation techniques, such as linear and thin-plate spline methods, in terms of displacement accuracy and physical plausibility. The network refines these initial interpolations by incorporating learned physical priors, enabling the generation of biomechanically plausible deformations within an image registration pipeline. These improvements are achieved with negligible computational overhead, supporting real-time applicability. Overall, the work detailed in this thesis advances the state of the art in keypoint-based data-driven brain registration by bridging the gap between biomechanical realism and deep learning efficiency.

Keywords: biomechanical modeling, brain shift, medical image registration, physics-informed deep learning, sparse keypoint interpolation

Contents

List of Figures

List of Tables

List of Algorithms

List of Acronyms

ASD average surface distance.

BC boundary condition.

BITE Brain Images of Tumors for Evaluation.

BraTS-Reg Brain Tumor Sequence Registration.

BRISK Binary Robust Invariant Scalable Keypoints.

ceT1 contrast-enhanced T1-weighted.

CNN convolutional neural network.

CNS central nervous system.

CSF cerebrospinal fluid.

CT computed tomography.

DL deep learning.

DNN deep neural network.

FAST Features from Accelerated Segment Test.

FC fully-connected.

FE finite element.

FFD free-form deformation.

FLAIR fluid-attenuated inversion recovery.

GAN generative adversarial network.

GNN graph neural network.

GP Gaussian process.

GPU graphics processing unit.

HD Hausdorff distance.

IC inverse consistency.

iMRI intra-operative magnetic resonance imaging.

IO instance optimization.

IQR interquartile range.

ISBI International Symposium on Biomedical Imaging.

iUS intra-operative ultrasound.

LeakyReLU leaky rectified linear unit.

LIFT Learned Invariant Feature Transform.

LNCC local normalized cross-correlation.

Max Error maximum Euclidean error.

MHVAE multimodal hierarchical variational autoencoder.

MI mutual information.

MICCAI Medical Image Computing and Computer Assisted Interventions.

MIND Modality Independent Neighbourhood Descriptor.

MRI magnetic resonance imaging.

MSE mean squared error.

MTLED meshless total Lagrangian explicit dynamics.

NCC normalized cross-correlation.

OR operating room.

PDE partial differential equation.

PET positron emission tomography.

PINN physics-informed neural network.

preMRI pre-operative magnetic resonance imaging.

RBF radial basis function.

ReLU rectified linear unit.

ReMIND Brain Resection Multimodal Imaging Database.

RESECT REtroSpective Evaluation of Cerebral Tumors.

SE squeeze-and-excitation.

SIFT Scale-Invariant Feature Transform.

SSD sum of squared distances.

STN Spatial Transformer network.

SURF Speeded-Up Robust Features.

T1 T1-weighted.

T2 T2-weighted.

TCIA The Cancer Imaging Archive.

TLED total Lagrangian explicit dynamics.

TPS thin-plate spline.

TRE target registration error.

UPenn-GBM University of Pennsylvania Glioblastoma.

US ultrasound.

Chapter 1

Introduction

1.1 Neuro-Oncology and Image-Guided Neurosurgery

Brain cancer is a complex and life-threatening condition characterized by the abnormal growth of cells within the brain or surrounding tissues. These tumors can disrupt critical brain functions, making early detection and effective treatment essential. According to recent statistics in the United States [1, 2], 73% of all primary brain and central nervous system (CNS) tumors are non-malignant, while 27% are malignant. The most common form of non-malignant primary brain tumor is meningioma, accounting for 41% of all primary brain tumors and 57% of all non-malignant tumors. In contrast, glioblastoma is the most common form of malignant brain tumor, representing 14% of all primary brain tumors and 52% of all malignant tumors. Although the overall chance that a person will develop a malignant brain tumor in their lifetime is less than 1%, the average five-year relative survival for malignant brain and CNS tumors is 36%. This estimate worsens for glioblastoma, where the five-year relative survival rate is 5%, with a median survival of 12-15 months, classifying it as the most lethal form of brain tumor in adults.

Standard treatment approaches, including surgery in conjunction with chemotherapy and radiation therapy, have significantly improved the survival of these patients. In particular, patients who undergo complete or maximal safe tumor resection significantly increase their survival rate compared to those who do not or those who undergo partial resection, making the extent of tumor resection the main prognostic factor for survival [3–5]. Therefore, discriminating between normal and cancerous tissue and identifying critical regions of interest is an essential step in achieving the finest surgical resection and improved survival rates [6, 7]. This need drives the development and adoption of advanced imaging and surgical technologies, such as neuronavigation.

1.1.1 Neuronavigation of Brain Tumors

Neuronavigation systems were introduced as an evolution of stereotactic surgery [8], which began in the early 20th century as a method to target specific brain regions using rigid frames and external coordinates. With advances in imaging technologies in the late 1980s [9] and 1990s [10], such as computed tomography (CT) and magnetic resonance imaging (MRI), frameless neuronavigation emerged, allowing for more flexible, accurate, and real-time guidance during surgery.

The brain can be conceptualized as a geometric volume divided into three orthogonal spatial planes (coronal, axial, and sagittal), where any point can be defined by its spatial position relative to these planes. Neuronavigation systems exploit this geometric representation by registering (i.e., aligning) the patient's physical anatomy to the corresponding coordinate system derived from three-dimensional medical imaging, which is typically visualized on a computer workstation [11]. Surgeons plan their surgical approach and define resection boundaries pre-operatively, aiming to reduce the risk of damage to surrounding healthy brain tissue while achieving maximal tumor resection [12, 13]. During surgery, these systems serve as point-to-point maps that correlate digital coordinates with physical locations within the brain, providing real-time guidance through continuous tracking of surgical instruments relative to the mapped brain coordinates to improve the precision of tumor localization and resection. The reliability of intra-operative guidance depends on the precise alignment of the pre-operative imaging, where planning was made, and the patient's current anatomical state. However, this alignment is frequently hindered by various dynamic intra-operative changes in brain anatomy, which introduce spatial discrepancies that compromise neuronavigation [14].

1.1.2 Brain Shift

As surgery progresses, the effectiveness of neuronavigation diminishes due to non-linear deformations of the brain tissues, referred to as *brain shift*, which introduce discrepancies between the physical anatomy in the operating room (OR) and the pre-operative planning [14, 15] (Figure 1.1). A typical shift of approximately 1–2 mm at the start of a case can drift to 4–6 mm, and cortical surface displacements exceeding 20 mm have been reported in long resections [16, 17]. This highlights the need for the development of advanced techniques for intra-operative image updating and brain shift compensation.

The phenomenon of brain shift is typically induced by three major factors: (1) physical, (2) surgical, and (3) biological. (1) Physical factors mainly involve patient positioning and the influence of gravity. Although patients are typically supine during surgery, the head is often rotated to provide better access to the surgical site, which may differ from the position adopted during pre-operative imaging acquisition. This difference in positioning can lead to misalignments between pre-operative images and the intra-operative anatomy, as gravity exerts a downward force on brain tissue, contributing to tissue sagging, especially in unsupported areas near resection sites [18, 19]. (2) Surgical factors include both mechanical and procedural influences. Changes in the hardware of the neuronavigation system, such as adjustments to clamps or external markers, can cause displacement. Likewise, the use of surgical tools, such as skin retractors and drills, applies significant force to the head, potentially shifting it relative to the reference imaging [16]. Directly related to the effects of gravity, the extent of tumor resection amplifies brain shift, as tissue removal creates unsupported spaces that can cause the surrounding tissue to sag or expand due to elastic recovery [19]. Fluid loss, particularly cerebrospinal fluid (CSF), also plays a key role. Cerebrospinal fluid evacuation, either intentionally induced by the surgeon or due to natural leakage

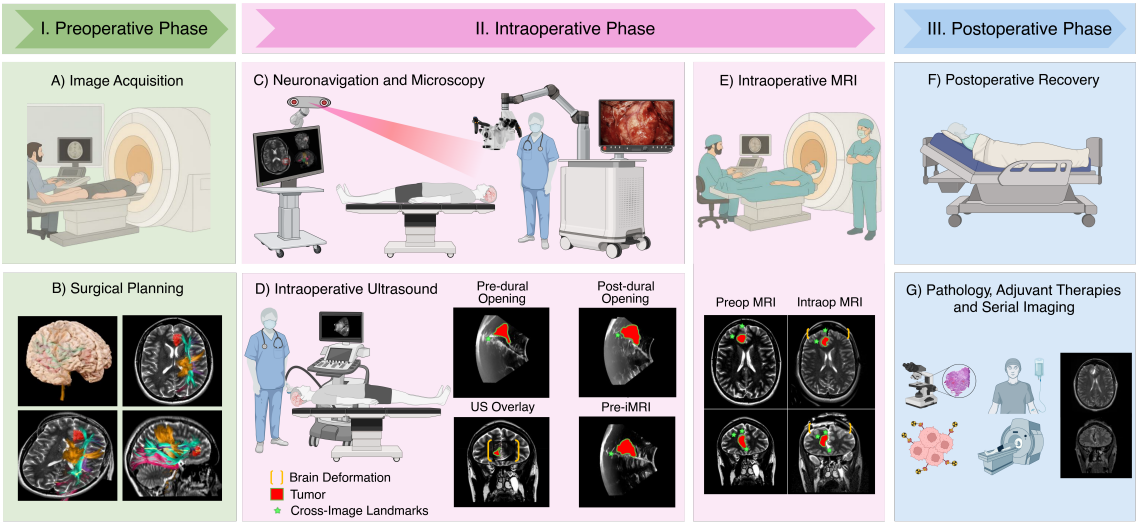


Figure 1.1: Neurosurgical workflow from preoperative assessment through postoperative care, including intraoperative phases. (A) *Image Acquisition*: Following clinical evaluation and discussion of treatment options, preoperative imaging is acquired. (B) *Surgical Planning*: Multisequence magnetic resonance imaging (MRI) is the primary imaging modality used to delineate tumor boundaries and assess the proximity of eloquent and critical brain areas. Additional modalities, such as computed tomography, functional MRI, or diffusion MRI, may be incorporated when needed to guide individualized surgical plans and resection goals. (C) *Neuronavigation and Microscopy*: In the operating room, patient-to-image registration aligns the patient's anatomy with the preoperative imaging, establishing a common coordinate system for neuronavigation. The incision and craniotomy are planned and executed to provide surgical access to the lesion, while neuronavigation and operative microscopy provide continuous image-guided visualization throughout the resection. (D) *Intraoperative Ultrasound*: Ultrasound is employed at critical points: pre-dural to confirm tumor access, post-dural to delineate margins and assess brain shift, and at intervals during resection to monitor for residual tumor. Green stars indicate the same location across imaging modalities and time points. (E) *Intraoperative MRI*: Intraoperative MRI assesses the extent of resection and brain shift. If there is residual tumor present, additional targeted resection is guided by updated MRI data alongside ultrasound and microscopy, before closure. (F) *Postoperative Recovery*: Patients undergo immediate postoperative clinical care and imaging to evaluate the final extent of resection and screen for postoperative complications. (G) *Pathology, Adjuvant Therapies and Serial Imaging*: Pathology informs adjuvant therapy decisions, and serial MRI with ongoing clinical follow-up monitors for recurrence and optimizes long-term management.

after craniotomy, reduces intracranial pressure, allowing the brain to relax and deform under its own weight [20]. (3) Biological factors influencing brain shift include the type of tumor and the use of drugs to manage intracranial pressure during surgery. Different pathologies (e.g., meningiomas vs. glioblastomas) are associated with varying degrees of brain shift due to differences in size, location, and biomechanical impact [21]. The presence or absence of edema also significantly affects the extent and direction of brain shift, as it alters tissue dynamics. In addition, the mannitol drug is commonly administered during neurosurgical procedures to manage intracranial pressure caused by tumor masses or edema [22, 23]. Mannitol effectively reduces pressure by removing excess fluid from the brain, but it can also contribute to brain shift by inducing changes in tissue

properties and fluid balance.

These combined effects make maintaining accurate neuronavigation during surgery increasingly challenging. To alleviate this problem, intra-operative imaging such as ultrasound (US) or MRI is acquired. Intra-operative magnetic resonance imaging (iMRI) provides high-resolution imaging that allows the detection and quantification of brain deformations, allowing easier registration of neuronavigation plans based on pre-operative magnetic resonance imaging (preMRI) data [12, 24, 25]. Despite this, its widespread use is often limited by the high costs, the need for specialized infrastructure, and the disruption of surgical workflows affecting patient safety. Intra-operative ultrasound (iUS) presents a viable alternative to intra-operative magnetic resonance imaging (iMRI), as it is more accessible, cost-effective, and easier to integrate into surgical procedures [26–28]. However, intra-operative ultrasound (iUS) has limitations in image quality, such as noise, artifacts, lower contrast and resolution, and a different image profile that makes it harder to interpret and register to preMRI data. Careful image processing and analysis, and highly trained professionals are required to identify anatomical features that can be matched between these two differing modalities to allow robust registration.

1.1.3 Image Registration for Brain Shift Correction

Image registration is a fundamental task in image-guided surgery, enabling the spatial alignment of pre-operative data with the intra-operative anatomy. To compensate for brain shift, a wide range of registration methods leveraging intra-operative imaging modalities such as iMRI and iUS have been proposed, including learning-based [29–32] and non-learning-based [33–35] methods. These techniques typically align pre- and intra-operative images by optimizing intensity-based similarity metrics (see Section 2.1 for more information on image registration). However, they often struggle in challenging registration scenarios involving (1) large intensity distribution gaps between pre- and intra-operative modalities (e.g., preMRI to iUS), (2) large deformations, and (3) topological changes due to tissue resection. Keypoint-based registration methods are a competitive alternative that try to address these limitations [36–39]. By relying on sparse correspondences rather than voxel-wise similarity, these methods are robust to large deformations, partial fields of view, and topological changes. They also offer interpretable outputs, as matched keypoints can be directly visualized and assessed. However, keypoint-based methods typically rely on simple geometric interpolators, such as splines or linear models, to propagate sparse displacements into dense displacement fields. These interpolators ignore the biomechanical behavior of brain tissues, which can lead to physically unrealistic deformations.

1.2 Motivation and Objectives

The following hypothesis arises: a physics-guided deep learning (DL) framework can act as an advanced interpolator of sparse intra-operative data, refining preliminary displacement estimates into biomechanically consistent deformations for realistic brain shift compensation. This thesis aims to advance image registration for brain shift correction by introducing an efficient method

that can be integrated into keypoint-based registration frameworks as a module capable of denoising an initial interpolated displacement field into a dense displacement field that leads to physically plausible deformations with minimal intra-operative data requirements. Sparse keypoints are particularly advantageous for real-time surgical workflows since they mitigate the dependency on resource-intensive intra-operative data while effectively capturing critical anatomical variations. In contrast to numerous existing methodologies that either require dense and high-quality intra-operative data, depend on computationally expensive iterative optimizations, or rely on biomechanical simulations, this approach utilizes synthetic biomechanically plausible displacement fields as supervision during training. By integrating data-driven learning with physics-based modeling, the proposed method closely aligns with actual tissue behavior, ensuring realistic deformations without compromising computational efficiency. This renders it especially appropriate for incorporation into existing intra-operative neuronavigation systems, where it is crucial to maintain a balance between accuracy, speed, and usability. Additionally, the minimal reliance on extensive intra-operative imaging renders this method adaptable to surgical settings with limited resources.

Thus, in this work, a novel DL framework for estimating dense displacements that produce physically plausible brain deformations from sparse matched keypoints between pre- and intra-operative images is proposed. First, a large-scale dataset of synthetic brain deformations produced by biomechanical simulations was constructed. Second, matched keypoints were simulated by extracting keypoints from the pre-operative imaging using a scale-invariant feature detector and descriptor and pairing them with the corresponding ground-truth displacements. Third, a deep biomechanically-guided interpolator based on a residual 3D U-Net architecture that refines standard interpolation estimates using the pre-operative data was developed. Finally, extensive experiments were conducted on the simulated brain deformations, significantly outperforming standard geometric interpolation methods in terms of displacement error and physical plausibility with negligible computational overhead.

1.3 Contributions of this Thesis

This thesis makes both methodological and scientific contributions to the problem of brain shift correction in image-guided neurosurgery. The central contribution is the formulation and validation of a physics-guided deep learning framework that refines sparse keypoint-derived displacement fields into dense, biomechanically plausible deformations. Unlike existing approaches that either rely on purely geometric interpolation or on computationally expensive biomechanical simulations at inference time, the proposed method introduces a lightweight deep interpolator that integrates learned physical priors while remaining compatible with real-time intra-operative constraints.

From a methodological standpoint, this work introduces a modular refinement block that can be inserted downstream of any keypoint-based registration pipeline, independently of the keypoint detector or matching strategy. The framework leverages patient-specific biomechanical simulations to provide physically grounded supervision during training, enabling the network to learn

deformation patterns that are difficult to observe directly in clinical settings. Extensive architectural exploration and ablation studies demonstrate how residual learning, squeeze-and-excitation mechanisms, and Jacobian-based regularization jointly contribute to improved anatomical fidelity, stability, and physical plausibility of the resulting displacement fields.

From a data and experimental perspective, this thesis presents the construction of a large-scale dataset of synthetic brain deformations generated using the meshless total Lagrangian explicit dynamics (MTLED) framework, modeling gravity-induced brain shift and tumor resection effects. This dataset enables systematic evaluation of sparse interpolation refinement under controlled conditions and supports quantitative and qualitative comparisons with classical interpolation methods.

Part of this work have been disseminated through peer-reviewed venues. The core methodology and experimental results were published in a Medical Image Computing and Computer Assisted Interventions (MICCAI)-affiliated workshop paper [40], providing initial validation within the medical image computing community. In addition, the literature review done in this work served as a basis for a systematic review paper on novel data-driven brain deformation modeling approaches for image-guided neurosurgery, which is currently under review to be submitted to a peer-reviewed journal. Both these outputs position the contributions of this thesis and any future research stemming from it within the broader research landscape of physics-informed learning for medical image registration and image-guided neurosurgery.

1.4 Thesis Outline

This thesis is organized into seven chapters. The current chapter introduced the clinical context of neurosurgical oncology, outlined the problem of intra-operative brain shift, and motivated the need for physics-informed registration methods. Chapter 2 reviews the theoretical foundations of medical image registration, including classical and learning-based approaches, similarity measures, regularization strategies, and biomechanical modeling of the brain. Chapter 3 surveys prior work on intra-operative, multimodal, and keypoint-based registration, also covering how biomechanical modeling is positioned within the medical image registration literature and establishing the basis for the synthetic simulations used in this thesis. Chapter 4 presents the proposed framework, detailing the clinical datasets, biomechanical simulation pipeline, keypoint extraction strategy, and the design and training of the deep biomechanically guided interpolator. Chapter 5 reports quantitative and qualitative evaluations of both the simulations and the learning-based model, including extensive ablation studies and comparisons with standard interpolators. Chapter 6 analyzes the experimental results in terms of clinical relevance, methodological trade-offs, and practical limitations. Finally, Chapter 7 summarizes the main contributions, discusses limitations, and outlines directions for future research and clinical integration.

Chapter 2

Background

2.1 Medical Image Registration

Addressing brain shift remains one of the most demanding challenges in neurosurgery, as the dynamic and non-linear deformations that occur during surgery can compromise the accuracy of neuronavigation systems. This highlights the need for advanced image registration methods specifically tailored to intra-operative conditions. Image registration constitutes a foundational procedure in medical imaging, enabling the spatial alignment of corresponding anatomical or functional features across two or more images. By aligning images acquired at different time points, from various modalities, viewpoints, or subjects, registration allows the integration of complementary information, thereby enhancing diagnostic precision and therapeutic decision-making. Its importance spans numerous clinical applications, including intervention and treatment planning [41–44], image-guided surgery [45–48], and disease monitoring [49–52].

Image-based registration often relies on a combination of image-to-patient and image-to-image registration strategies. Image-to-patient registration aligns medical images with the patient's physical space. This is typically achieved through optical or electromagnetic tracking systems that identify corresponding anatomical landmarks on the patient and in the image (recall Figure 1.1.II). The number of corresponding points required depends on whether the registration is assumed to be rigid or deformable and the amount of noise in the data collection. Image-to-image registration aligns two volumetric scans (e.g., pre-operative magnetic resonance imaging (preMRI) to intra-operative ultrasound (iUS)) and has been performed using a variety of methods: driven by manually or automatically selected corresponding points [53], image patches [54], and deep learning (DL)-based methods [30, 55–57].

The transformations applied in registration can be rigid, affine, or deformable [58]. Rigid transformations involve only translations and rotations, suitable for aligning structures such as bone or tissues that suffer with minimal shape change. Affine transformations extend this by including scaling and shearing, altering lengths and angles while preserving parallelism. Deformable (also called elastic or non-rigid) transformations capture non-linear anatomical changes and are essential for accurate alignment of soft tissues such as the brain, lungs, or liver, by mapping every voxel from a moving (source) image to a fixed (target) image using dense displacement

vector fields.

Conventional image registration methods estimate spatial transformations by iteratively optimizing a predefined objective function. They rely on hand-crafted similarity metrics, such as mutual information (MI) [59, 60] and normalized cross-correlation (NCC) [61, 62], together with regularization terms designed to enforce smooth and physically plausible deformations [63, 64] (see Section 2.1.5 and Section 2.1.6). Parameter tuning is typically manual and requires substantial domain expertise to balance accuracy and robustness. Because optimization is carried out independently for each image pair, these methods can be time-consuming, computationally expensive, and sensitive to initialization.

Deep learning (DL)-based registration methods fundamentally change this paradigm by learning mappings from image pairs to displacement fields across large datasets. Rather than configuring parameters for every case, these models leverage training data to guide parameter optimization. Supervised approaches depend on ground-truth displacement fields derived from simulations or classic algorithms, which can introduce bias or propagate existing errors, whereas unsupervised and self-supervised frameworks optimize similarity metrics directly against the input image data without the need for explicit ground truths [65]. Once trained, DL models can register unseen image pairs in real time, eliminating the need for instance optimization (IO) during inference. However, they still require careful design of loss functions and architectures to ensure accurate and physically meaningful deformations.

2.1.1 Registration Basis

Traditional image registration methods can be extrinsic or intrinsic, depending on the basis on which the method relies. Extrinsic methods rely on artificial objects directly attached to the patient's anatomy (referred to as fiducials), such as stereotactic frames [66–68] or markers glued or screwed to the skin [69, 70], which are easily detected in any modality and do not include patient-related image information. As such, the registration process is comparatively easier, faster, and usually automated, since no complex optimization algorithms are required [71]. However, the intense pre-operative planning and the invasive nature of the method are major drawbacks to its application in the surgical space. In contrast, intrinsic methods only rely on the patient's image content by directly using and matching voxel properties, limited sets of identified landmarks/keypoints, and segmented structures [72].

Landmarks can be manually identified when they constitute easily identifiable prominent features of the morphology, or automatically when they have prominent geometrical properties, such as corners or local curvature gradient extremes [39, 73–76]. These points are identified in the moving image and matched with their counterparts in the fixed image. Since this set of matched points is sparsely distributed compared to the original image content, it effectively reduces the dimensionality of the deformation parameterization, reducing the computational cost of registration while requiring interpolation methods to propagate a dense displacement field [77, 78].

Voxel property approaches match the intensity patterns in each image, operating directly on

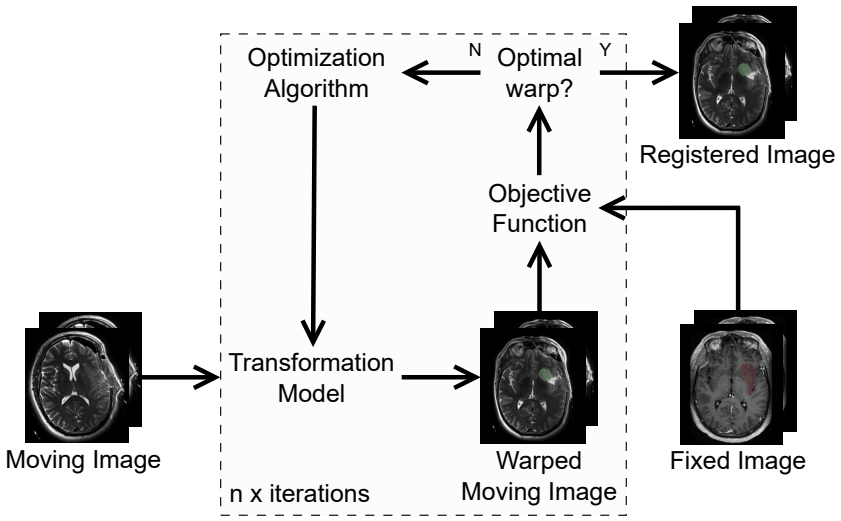


Figure 2.1: Overview of a classic iterative image registration pipeline. A transformation model parametrizes the deformation applied to the moving image (Section 2.1.3 and Section 2.1.4), producing a warped image that is compared to the fixed image through an objective function combining similarity and regularization terms (Section 2.1.5 and Section 2.1.6). An optimization algorithm iteratively updates the transformation parameters to minimize the objective until convergence, yielding the final registered image (Section 2.1.7).

the gray values of the image using mathematical or statistical criteria [61, 79–81]. They measure the intensity similarity between the moving and fixed images and adjust the transformation until the similarity measure is maximized. This method assumes that the image intensities will be the most similar at the optimal registration.

2.1.2 Canonical Formulation

Let $\boldsymbol{I} : \Omega \to \mathbb{R}$ be a medical image, where Ω is the spatial domain in which the image is defined. The primary objective of an image registration procedure between two images in the 3D space $\Omega \in \mathbb{R}^3$ is to estimate a deformation $\phi : \Omega \to \Omega$ so that the moving image $\boldsymbol{I_M}$ warped by ϕ is anatomically aligned with the fixed image $\boldsymbol{I_F}$. The warped moving image is obtained by the composition $\boldsymbol{I_M} \circ \phi$. Figure 2.1 illustrates the traditional image registration pipeline, highlighting all the key steps described later in this chapter.

Registration methods search for an optimal transformation ϕ by optimizing an objective function $\mathcal{L}$ that includes a similarity term $\mathcal{S}$, which quantifies the level of alignment between the warped moving image and the fixed image (see Section 2.1.5), and a regularization term $\mathcal{R}$, which constrains solutions and guides registration towards desirable and plausible outcomes (see Section 2.1.6) [64, 65]. Mathematically, this can be expressed as:

$$\begin{aligned} &\mathcal{L}(\boldsymbol{I_F}, \boldsymbol{I_M}, \phi) = \mathcal{S}(\boldsymbol{I_F}, \boldsymbol{I_M} \circ \phi) - \lambda \mathcal{R}(\phi) \\ &\hat{\phi} = \arg\max_{\phi} \ \mathcal{L}(\boldsymbol{I_F}, \boldsymbol{I_M}, \phi) \end{aligned} \quad , \tag{2.1}$$

where $\lambda \in \mathbb{R}$ weights the regularization term.

2.1.3 Affine Registration Models

For the affine case, the transformation that aligns two voxels $\boldsymbol{p_m} = [x_m, y_m, z_m, 1]^T \in \boldsymbol{I_M}$ and $\boldsymbol{p_f} = [x_f, y_f, z_f, 1]^T \in \boldsymbol{I_F}$ allows rotation, translation, scaling, and shearing, thus any combination of these operations can be captured in a single transformation matrix. This can be expressed as:

$$\begin{bmatrix} x_f \\ y_f \\ z_f \\ 1 \end{bmatrix} = \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & t_x \\ a_{21} & a_{22} & a_{23} & t_y \\ a_{31} & a_{32} & a_{33} & t_z \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} x_m \\ y_m \\ z_m \\ 1 \end{bmatrix}, \tag{2.2}$$

where the transformation matrix is composed by a 3×3 submatrix that contains the parameters a_{ij} representing any combination of rotation, scaling, and shearing, and a 3×1 subvector that contains the translation components [58]. This representation uses homogeneous coordinates [82] to ensure that the transformations can be applied seamlessly with matrix-vector multiplications. The problem is handled by formulating it as a linear system of equations, through which the affine parameters of the transformation matrix are iteratively optimized (recall Section 2.1.7).

2.1.4 Deformable Registration Models

Deformable registration is necessary when rigid or affine models are insufficient to capture anatomical variability, particularly in scenarios involving large or nonuniform deformations, such as in soft tissues like the brain. These transformations can be modeled using parametric or non-parametric models [64, 83]. In parametric transformation models, a select number of parameters that govern the registration process are used and optimized. In contrast, non-parametric transformation models compute a dense displacement field by directly calculating the displacement at each voxel. The displacement field is thus treated as a continuous function of space and is optimized over all possible deformations within the set of constraints.

Parametric Models

Parametric models are usually based on interpolation methods and represent non-rigid deformations with landmarks/control points, basis functions, and splines, reducing the complexity of the transformation by parameterizing it. This category of registration algorithms assumes that a set of matching control points can be identified in both images, and an interpolating function is used to establish correspondences between distant points to yield a dense displacement field. A model relying on control points is parameterized by three components: the set of control points, their associated displacements, and an interpolation strategy.

Radial basis functions. A common approach is to model the transformation as a linear combination of radial basis functions (RBFs) [84], where the value of the displacement at an interpolation point $\boldsymbol{p} \in \mathbb{R}^3$ is given in function of its distance to the set of n known control points $\{\boldsymbol{c_i} \in \mathbb{R}^3\}_{i=1}^n$:

$$\phi(\boldsymbol{p}) = \sum_{i=1}^{n} w_i \theta_i(\|\boldsymbol{p} - \boldsymbol{c_i}\|_2), \tag{2.3}$$

where θ_i is a basis function, and w_i is the corresponding weight coefficient.

Knowing the displacement at one point influences the displacements of points interpolated in the whole image domain. Consequently, the interpolation of sparse areas is feasible but limits the ability to model complex and localized displacements.

Thin-plate splines. Being part of a family of splines that use RBFs, thin-plate spline (TPS) are effective when a set of corresponding control points is available for both images [77, 85]. A spline is a piecewise polynomial function used to approximate a curve or surface by fitting each polynomial to a set of control points, resulting in a smooth and continuous surface. The TPS implementation is derived by minimizing a bending energy function, and it incorporates a global affine transformation in addition to the RBF formulation. This can be expressed as:

$$\phi(\boldsymbol{p}) = \underbrace{\boldsymbol{A} \cdot \boldsymbol{p}}_{affine} + \underbrace{\sum_{i=1}^{n} w_i \theta(\|\boldsymbol{p} - \boldsymbol{c_i}\|)}_{non-linear}, \tag{2.4}$$

where $\boldsymbol{A} \in \mathbb{R}^{4\times4}$ is the affine transformation matrix, and θ is usually defined as $\theta(r) = r^2 \log r$. Applying this equation to the whole set of n control points in the source image and expressing the problem in matrix form with homogeneous coordinates allows us to solve a linear system of equations:

$$\begin{bmatrix} \boldsymbol{D} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Theta} & \boldsymbol{C} \\ \boldsymbol{C^T} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{W} \\ \boldsymbol{A} \end{bmatrix}, \tag{2.5}$$

where $\boldsymbol{D} \in \mathbb{R}^{n\times4}$ is a submatrix of control point displacements, $\boldsymbol{C} \in \mathbb{R}^{n\times4}$ is a submatrix of control points, and $\boldsymbol{W} \in \mathbb{R}^{n\times4}$ is a submatrix of non-linear weight coefficients. The kernel matrix $\boldsymbol{\Theta} \in \mathbb{R}^{n\times n}$ is a kernel submatrix with entries:

$$\Theta_{ij} = \theta(\|\boldsymbol{c_i} - \boldsymbol{c_j}\|) + \lambda_{\text{tps}} \boldsymbol{I}_{ij}, \tag{2.6}$$

where λ_{tps} is a regularization parameter controlling the trade-off between interpolation accuracy and smoothness, and $\boldsymbol{I}_{ij}$ denotes a location in the identity matrix $\boldsymbol{I}$. Solving for $\boldsymbol{A}$ and $\boldsymbol{W}$ using standard algebra yields the parameters that can interpolate any point conditioned on the known displacements of the control points.

B-splines. Under the free-form deformation (FFD) framework [86], B-splines are a powerful modeling tool for 3D deformable objects. Free-form deformations deform an object by manipulating an underlying low resolution $n_x \times n_y \times n_z$ regular mesh with corresponding control point displacement vectors $\boldsymbol{D_{i,j,k}}$ and uniform spacing superimposed on the image. Specifically using cubic B-splines, which are locally polynomial interpolation functions that can perform interpolation on regular grids, the dense displacement for any voxel is given according to a summation of 3D tensor products of piecewise cubic B-splines:

$$\phi = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u) B_m(v) B_n(w) \boldsymbol{D_{i+l,j+m,k+n}}, \tag{2.7}$$

where $i = \lfloor \frac{x}{n_x} \rfloor - 1$, $j = \lfloor \frac{y}{n_y} \rfloor - 1$, $k = \lfloor \frac{z}{n_z} \rfloor - 1$, $u = \frac{x}{n_x} - \lfloor \frac{x}{n_x} \rfloor$, $v = \frac{y}{n_y} - \lfloor \frac{y}{n_y} \rfloor$, $w = \frac{z}{n_z} - \lfloor \frac{z}{n_z} \rfloor$, and B_l is the l-th basis function of the B-spline:

$$\begin{cases} B_0(u) = (1-u)^3/6 \\ B_1(u) = (3u^3 - 6u^2 + 4)/6 \\ B_2(u) = (-3u^3 + 3u^2 + 3u + 1)/6 \\ B_3(u) = u^3/6 \end{cases}, \tag{2.8}$$

where each (l,m,n) corresponds to a control point in the local neighborhood of (i,j,k), and the cubic B-spline basis functions $B_l(u)$, $B_m(v)$, and $B_n(w)$ weight their influence based on their relative position within the grid.

For each cubic B-spline in this formulation, only the four neighboring control points contribute to the displacement of a given voxel. Since B-splines are only defined in the vicinity of each control point, the displacement of a point only affects the transformation in the local neighborhood of that point, allowing for finer displacements and more efficient computations even for a large number of control points. Because of this, B-splines are often referred to as having local support, as opposed to RBFs and TPS, which have global support [86].

Gaussian Process regression. Another approach is Gaussian process (GP) regression, which performs the simultaneous interpolation and uncertainty estimation of the interpolation process [87]. The GP regression model makes the assumption that a Gaussian process with a mean function μ and covariance function κ provides the function ϕ that connects the inputs to the outputs. A starting point for a prior belief could be a Gaussian distribution with zero mean, and an RBF covariance kernel. Observing a set of n control point pairs and their displacements allows the update of a prior belief. These observed displacements of ϕ^* are assumed to be noisy realizations of the true function ϕ, such that $\phi^*(\boldsymbol{p}) = \phi(\boldsymbol{p}) + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma^2)$. To obtain a posterior distribution, conditioned on the observed data, the negative log-likelihood is minimized and any existing GP hyperparameters are learnt. Subsequently, the joint distribution of n observed values and any other point m is Gaussian, and, thus, the displacement at an arbitrary voxel can be directly obtained from the posterior distribution:

$$\begin{cases} p(\phi^* \mid \boldsymbol{P}, \boldsymbol{C}, \boldsymbol{D}) \sim \mathcal{N}(\boldsymbol{M}, \boldsymbol{\Sigma}^2) \\ \boldsymbol{M} = \boldsymbol{K}^{*T}(\boldsymbol{K} + \sigma^2 \boldsymbol{I})\boldsymbol{D} \\ \boldsymbol{\Sigma}^2 = \boldsymbol{K}^{**} - \boldsymbol{K}^{*T}(\boldsymbol{K} + \sigma^2 \boldsymbol{I})^{-1}\boldsymbol{K}^* \end{cases}, \tag{2.9}$$

where $\boldsymbol{P} \in \mathbb{R}^{m \times 3}$ is the matrix of all points in the image; $\boldsymbol{C} \in \mathbb{R}^{n \times 3}$ is the matrix of all control points; $\boldsymbol{D} \in \mathbb{R}^{n \times 3}$ is the matrix of observed displacements; $\boldsymbol{M} \in \mathbb{R}^{m \times 3}$ is a matrix of estimated mean displacements $\mu(\boldsymbol{p_i})$; $\boldsymbol{\Sigma}^2 \in \mathbb{R}^{m \times m}$ is a matrix of covariances between each estimated value, providing the uncertainties associated with each estimation; $\boldsymbol{K} \in \mathbb{R}^{n \times n}$ is the covariance matrix between the values at each control point $\kappa(\boldsymbol{c_i}, \boldsymbol{c_j})$; $\mathbf{K}^* \in \mathbb{R}^{n \times m}$ is the covariance matrix between the values at interpolated points and control points $\kappa(\boldsymbol{p_i}, \boldsymbol{c_j})$; $\boldsymbol{K}^{**} \in \mathbb{R}^{m \times m}$ is the variance matrix between the values at each interpolated point $\kappa(\boldsymbol{p_i}, \boldsymbol{p_i})$; and $\boldsymbol{I} \in \mathbb{R}^{n \times n}$ is the identity matrix that is multiplied by σ^2 to account for the noise assumption.

Non-parametric Models

Non-parametric transformation models assign a displacement vector to every voxel in the image domain, without relying on a compact parameterization or a predefined set of control points [64, 86]. These models offer maximal flexibility but also require strong regularization to ensure anatomically plausible solutions. Many classic non-parametric models are grounded in continuum mechanics and formulated through partial differential equations (PDEs), where the displacements arises from balancing image similarity-derived forces with physically-derived internal forces. This variational formulation inherently embeds smoothness, elasticity, or fluid-like behavior into the transformation, enabling accurate and realistic mappings between images.

Small Deformation Models. A classic example of a small deformation model is linear elastic registration, which tackles the registration problem by modeling the displacements as a physical process that resembles the stretching of an elastic material. In this formulation, the image grid is deformed under the influence of two forces that compete until equilibrium is reached. An external force drives the alignment based on image similarity, while an internal force imposes the elastic properties of the material [88]. The displacements are described by the Navier-Cauchy PDE as:

$$\mu\nabla^2\phi + (\mu + \lambda)\nabla(\nabla \cdot \phi) + \mathbf{F} = 0, \tag{2.10}$$

where ∇^2, ∇, and $\nabla\cdot$ are the Laplace, gradient, and divergence operators, respectively; $\mathbf{F}$ is the force field that drives the registration based on an image similarity criterion; and μ and λ are Lamé's parameters describing the material's rigidity and compressibility, respectively.

Solving this PDE can be achieved using finite differences and successive over-relaxation methods [89], which produce a discrete displacement vector at each voxel location. Linear elastic models perform well for small, smooth deformations, but because the energy acts directly on the displacement magnitude, deformations become penalized proportionally to their size [64]. As a consequence, these models struggle with large or highly localized motions.

Large Deformation Models. To handle larger, non-linear motions, deformations can instead be described through a time-dependent $t \in [0, 1]$ velocity field ν. Time-varying models for image registration can capture complex and large deformations while preserving diffeomorphic mappings that ensure biologically realistic transformations [65, 86]. The deformation $\phi^{(t)}$ evolves according to:

$$\nu^{(t)}(\phi^{(t)}) = \frac{d\phi^{(t)}}{dt}, \qquad \phi^{(0)} = \boldsymbol{I}, \tag{2.11}$$

where $\boldsymbol{I}$ is the identity transform, and integrating this velocity field over time yields the final mapping:

$$\phi^{(1)} = \phi^{(0)} + \int_0^1 \nu^{(t)}(\phi^{(t)})\, dt. \tag{2.12}$$

Because integrating a smooth velocity field produces a diffeomorphism, large-deformation models naturally preserve topology and invertibility.

A classic example of this paradigm is viscous-fluid registration [90, 91], where the model still satisfies Equation 2.10 but instead of solving for the deformation field ϕ, it solves for the velocity field ν.

To simplify its implementation, one can instead assume a time-stationary velocity field formulation, i.e., $\nu^{(t)}(\phi^{(t)})$ becomes $\nu(\phi^{(t)})$, enabling the application of the scaling-and-squaring technique [92, 93] that relates the displacement field to the velocity field as an exponential map:

$$\phi = \exp(\nu) = \exp(2^{-N}\nu)^{2^N}, \tag{2.13}$$

which is equivalent to integrating along the velocity field over the unit time period, in this case, along N time steps of recursive composition of the spatial transformations (N is such that $2^{-N}\nu$ is close to zero).

Demons algorithm. The Demons algorithm [94] represents an example of a different class of non-parametric approaches that arise not from mechanical PDEs but from optical-flow principles combined with diffusion-based regularization. The algorithm iteratively updates the displacement field using an update force f derived from optical flow principles. It uses the gradient of image intensities $\nabla \boldsymbol{I_F}$ and the image intensity differences $\boldsymbol{I_M} \circ \phi - \boldsymbol{I_F}$ to solve the equations:

$$\begin{cases} \mathbf{F} = \frac{(\boldsymbol{I_M}\circ\phi - \boldsymbol{I_F})\nabla \boldsymbol{I_F}}{\|\nabla \boldsymbol{I_F}\|_2^2 + (\boldsymbol{I_M}\circ\phi - \boldsymbol{I_F})^2} \\ \phi_0 = 0 \\ \phi_n^{'} = \phi_{n-1} + K_{\sigma_1} * \mathbf{F} \\ \phi_n = K_{\sigma_2} * \phi_n^{'} \end{cases} \tag{2.14}$$

where ϕ_0 is the initial displacement, which is assumed to be a zero displacement field, n is the number of iterations, K_{σ_i} are Gaussian smoothing filters, and $*$ represents a convolution operation.

During each iteration, a preliminary estimate of displacement is obtained for each voxel. However, if the denominator of the force equation is too small, the equation becomes unstable [95]. Thus, the displacement is regularized using a Gaussian smoothing filter that ensures diffuse estimates across regions with either strong or weak gradients. The final displacement field is obtained when the iterative process converges after n iterations of refinement using the update forces.

Because the original Demons formulation updates the displacement field directly, it does not guarantee diffeomorphic transformations. Nevertheless, Demons-based methods remain widely used due to their speed, simplicity, and robustness, and have been successfully adapted to optimize a variety of similarity metrics and diffeomorphic formulations [96, 97].

2.1.5 Similarity Measures

Similarity measures are integral to image registration, as they quantify the alignment between a fixed image $\boldsymbol{I_F}$ and a warped moving image $\boldsymbol{I_M}\circ\phi$. By evaluating the degree of similarity between images, these measures drive the optimization process to align anatomical structures accurately. The selection of an appropriate similarity measure depends on the image modalities involved, the

nature of the registration, and the application requirements. Similarity measures are commonly divided into intensity-based, landmark-based, and segmentation-based approaches, each relying on different assumptions, strengths, and limitations [64,98]. Table 2.1 summarizes these categories and their main characteristics.

Intensity-based

Intensity-based measures operate directly on the voxel intensities of the images and are particularly effective when there is a consistent relationship between the intensities of corresponding voxels. These measures are widely used and are based on mathematical formulations that capture intensity differences, correlations, or probabilistic dependencies.

Sum of Squared Differences. The sum of squared distances (SSD) is the simplest intensity-based similarity measure, defined as:

$$\mathcal{S}_{SSD} = \sum_{\boldsymbol{p}} \big((\boldsymbol{I_M} \circ \phi)(\boldsymbol{p}) - \boldsymbol{I_F}(\boldsymbol{p})\big)^2, \tag{2.15}$$

where $\boldsymbol{p} \in \mathbb{R}^3$ is a three-dimensional voxel (x, y, z). Sum of squared distances assumes that corresponding structures in the images have identical intensities and relies on the assumption that the noise is Gaussian. However, image similarity does not necessarily imply good registration of the underlying anatomy. Moreover, SSD is sensitive to intensity variations caused by imaging artifacts or modality differences, which limits its applicability in multimodal scenarios. The mean squared error (MSE) is a normalized version of the SSD, differing only by a constant scaling factor corresponding to the number of voxels.

Normalized Cross-correlation. Normalized cross-correlation (NCC) differs from SSD by assuming a linear relationship between image intensities while eliminating the effects of variations in brightness and contrast (i.e., the amplitude of the signal). It is defined as:

$$\mathcal{S}_{NCC} = \frac{\sum_{\boldsymbol{p}} \big((\boldsymbol{I_M} \circ \phi)(\boldsymbol{p}) - \mu_M\big)\big(\boldsymbol{I_F}(\boldsymbol{p}) - \mu_F\big)}{\sqrt{\sum_{\boldsymbol{p}} \big((\boldsymbol{I_M} \circ \phi)(\boldsymbol{p}) - \mu_M\big)^2}\sqrt{\sum_{\boldsymbol{p}} \big(\boldsymbol{I_F}(\boldsymbol{p}) - \mu_F\big)^2}}, \tag{2.16}$$

where μ_M and μ_F are the mean intensities of $\boldsymbol{I_M} \circ \phi$ and $\boldsymbol{I_F}$, respectively. The NCC normalizes intensity differences, making it robust to linear intensity scaling. Similar to SSD, it is primarily used for unimodal registration, as it assumes a direct correlation between the voxel intensities of the two images.

Local normalized cross-correlation (LNCC) is an extension of the standard NCC measure, designed to incorporate spatially localized intensity relationships between images [109]. While NCC computes a global correlation between the voxel intensities of two images, local normalized cross-correlation (LNCC) evaluates similarity within localized regions, making it more robust to intensity variations and non-linear intensity mappings across the entire image domain. The LNCC for a voxel $\boldsymbol{p}$ is computed over a local neighborhood of size w^3 centered on that voxel, which reduces the influence of the overall image domain on these local calculations. This localized approach is particularly useful in deformable image registration, where spatial variations in intensity relationships are common due to non-rigid anatomical deformations or modality differences.

Table 2.1: Comparison between similarity measure categories used in medical image registration.

	Intensity-based	**Landmark-based**	**Segmentation-based**
Principles	Operates directly on voxel intensities and assumes a relationship between intensity patterns of corresponding structures	Assumes reliable landmark or feature descriptor correspondences can be detected and matched across images	Relies on segmented anatomical structures and assumes accurate and consistent labels across images
Strengths and Limitations	+ Simple and general purpose + Effective for unimodal registration − Sensitive to noise, artifacts, and multimodal images	+ Anatomically interpretable + Robust to intensity inconsistencies and large deformations − Performance depends on feature quality and correspondence accuracy	+ Directly enforces anatomical alignment + Interpretable registration evaluation − Does not capture internal deformation errors − Requires prior high-quality annotations
Clinical Relevance	Commonly used when intensity relationships are preserved	Effective and interpretable in multimodal and sparse intra-operative data scenarios	Provides a structure-aware evaluation and guidance
Implementation Examples	MSE [30, 99–101] NCC [31, 102, 103] MI [104, 105]	TRE [100] MIND [37, 38, 57, 106] CNN-based [57, 107]	Dice [30, 56, 108] HD [106]

MSE - Mean squared error; NCC - Normalized cross-correlation; MI - Mutual information;
TRE - Target registration error; MIND - Modality Independent Neighborhood Descriptor;
CNN - Convolutional neural network; HD - Hausdorff distance

Mutual Information. Measures such as MI are particularly effective for multimodal registration, as they do not assume a specific functional relationship between intensities [110, 111]. These measures are based on the joint intensity distribution of the two images, which captures the statistical relationship between their voxel intensities. Derived from Shannon entropy [112], the cross-entropy of two images $\boldsymbol{I_F}$ and $\boldsymbol{I_M} \circ \phi$ can be given by:

$$H = -\sum_{a}\sum_{b} p(a,b)\log p(a,b), \tag{2.17}$$

where $p(a,b)$ is the joint probability of intensities a in $\boldsymbol{I_F}$ and b in $\boldsymbol{I_M} \circ \phi$, which can be estimated [113, 114] using density-estimation methods, such as histograms, Parzen windows, or nearest neighbors. The MI measures the reduction in joint entropy due to alignment and is maximized when the images are aligned, as the joint distribution becomes concentrated and entropy decreases. The MI can be defined as:

$$\mathcal{S}_{MI} = H(\boldsymbol{I_F}) + H(\boldsymbol{I_M} \circ \phi) - H(\boldsymbol{I_F}, \boldsymbol{I_M} \circ \phi). \tag{2.18}$$

To avoid dependence on the amount of image overlap, a normalized MI measure [115] can be used. Similarly to LNCC, measures based on information theory can also be locally computed to improve robustness [116]. Overall, MI tends to underperform in the presence of large modality gaps, such as between ultrasound (US) and magnetic resonance imaging (MRI), because various intensity configurations can produce similarly low joint entropy.

Feature- and Landmark-based

Intensity-based similarity measures can struggle when intensity relationships are non-linear or inconsistent, such as in multimodal contexts. Multimodal registration aligns images acquired from different imaging techniques that follow distinct physical principles, such as computed tomography (CT), MRI, or positron emission tomography (PET). Although multimodal data offer important complementary information, the intensity distributions of different modalities usually vary, leading to complex interrelations [117]. Certain anatomical structures may even appear distorted or missing in one image relative to another. These challenges motivate the use of feature-based measures, which replace raw intensities with more stable geometric or anatomical cues.

To address this, image descriptors have been developed to provide rich structural and contextual information on anatomical regions to improve registration accuracy. These descriptors substitute raw intensity values in similarity measures with feature vectors that encapsulate local or global image properties, offering a more robust representation for aligning complex anatomical structures [117]. Descriptors can be either dense (computed at every voxel) or sparse (computed only at selected keypoints/landmarks), and they leverage diverse techniques from handcrafted to DL-derived features [64, 98]. Automated landmark detection and description are therefore essential, reducing the need for manual annotation and minimizing observer variability. Classic detectors, such as Scale-Invariant Feature Transform (SIFT) [36, 118, 119], Speeded-Up Robust Features (SURF) [120], Features from Accelerated Segment Test (FAST) [121, 122], and Binary Robust Invariant Scalable Keypoints (BRISK) [123], have been adapted to 2D and 3D medical images, extending concepts such as scale-invariant detection and gradient-based description to volumetric data. More tailored advances include Modality Independent Neighbourhood Descriptor (MIND) [124], which was specifically designed for medical image registration and identifies regions based on self-similarity within a local neighborhood around the keypoint. Novel DL approaches have also shown strong performance in detecting keypoints and computing descriptors tailored to specific anatomical contexts, commonly leveraging convolutional neural networks (CNNs). For example, SuperPoint [125] presents a self-supervised framework to train interest point detectors and descriptors by predicting pixel-level interest point locations. Learned Invariant Feature Transform (LIFT) [126] uses a Siamese CNN architecture to detect keypoints, estimate orientation, and compute descriptors in a supervised end-to-end manner.

When a set of corresponding landmarks is available, registration can be explicitly guided by enforcing alignment between paired points [37, 38, 127] (recall Section 2.1.1). Given N corresponding landmark pairs $\{(\boldsymbol{l}^{\boldsymbol{F}}, \boldsymbol{l}^{\boldsymbol{M}})\}_{i=1}^{N}$, where $\boldsymbol{l}^{\boldsymbol{M}}$ denotes a landmark location in the moving image and $\boldsymbol{l}^{\boldsymbol{F}}$ its counterpart in the fixed image, the target registration error (TRE) measures the residual Euclidean distance after warping the moving landmarks. The mean TRE is then expressed as:

$$\mathrm{TRE} = \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{l}^{\boldsymbol{M}} \circ \phi - \boldsymbol{l}^{\boldsymbol{F}} \right\|_2 . \tag{2.19}$$

Segmentation-based

Other anatomical priors, such as segmentation masks, can provide complementary information that offers valuable guidance in aligning complex anatomical structures during the registration process. These segmentations can be manually defined for each image pair or generated from training data in learning-based frameworks.

Segmentation masks $\mathcal{M} : \Omega \rightarrow [0, 1]$ typically provide binary anatomical labels that classify each voxel as belonging to a specific structure. When they are provided, the most common approach is to measure the extent of overlap between warped moving image and fixed image segmentations with specific measures, such as the Dice score [128]. Various structures can be identified in both images, resulting in a set of masks $\{\mathcal{M}^l\}_{l\in\mathcal{L}}$ that localize distinct anatomical regions associated with each label in a set of labels $\mathcal{L}$. The mean Dice score for a pair of images considering all labels can then be defined as:

$$\text{Dice} = \frac{1}{|\mathcal{L}|}\sum_{l\in\mathcal{L}}\frac{2\sum_{\boldsymbol{p}}(\mathcal{M}^l_M\circ\phi)(\boldsymbol{p})\mathcal{M}^l_F(\boldsymbol{p})}{\sum_{\boldsymbol{p}}(\mathcal{M}^l_M\circ\phi)(\boldsymbol{p})+\sum_{\boldsymbol{p}}\mathcal{M}^l_F(\boldsymbol{p})}. \tag{2.20}$$

In addition to overlap-based metrics, distance-based measures are commonly employed to evaluate boundary alignment between corresponding anatomical structures. Among these, the Hausdorff distance (HD) is widely used in medical image registration to quantify the maximum discrepancy between the surfaces of segmented structures after registration. Given two sets of surface points, A and B, extracted from the warped moving and fixed segmentations, respectively, the (directed) HD is defined as:

$$d_H(A, B) = \max_{a\in A}\min_{b\in B}\|a - b\|_2, \tag{2.21}$$

and the symmetric HD is given by:

$$\text{HD}(A, B) = \max\{d_H(A, B), d_H(B, A)\}. \tag{2.22}$$

While the HD captures the worst-case boundary error and is therefore sensitive to outliers, it provides valuable insight into the maximum misalignment between anatomical structures, which is clinically relevant for surgical decision-making. To reduce sensitivity to isolated outliers caused by segmentation noise or minor surface artifacts, the 95th-percentile Hausdorff distance (HD95) is often reported instead. This variant measures the distance below which 95% of the surface points fall, offering a more robust estimate of boundary alignment [65, 129].

Complementarily, the average surface distance (average surface distance (ASD)) is frequently reported in conjunction with HD to assess boundary alignment. Unlike HD, which focuses on extreme discrepancies, the ASD measures the mean distance between corresponding surfaces, providing a global assessment of registration accuracy. The ASD is less sensitive to outliers and reflects the typical boundary error between anatomical structures. Formally, the directed ASD from surface A to surface B is defined as:

$$d_{ASD}(A, B) = \frac{1}{|A|}\sum_{a\in A}\min_{b\in B}\|a - b\|_2, \tag{2.23}$$

and the symmetric ASD is obtained by averaging the two directed distances:

$$\mathrm{ASD}(A, B) = \frac{1}{2}\big(d_{ASD}(A, B) + d_{ASD}(B, A)\big). \tag{2.24}$$

Distance-based metrics complement overlap measures like the Dice, as a high overlap can still coincide with localized boundary errors, particularly for small or thin structures such as ventricles or tumor margins. Consequently, reporting both overlap-based and distance-based metrics is considered best practice for a comprehensive evaluation of registration accuracy when segmentation masks are available.

2.1.6 Regularization Terms

Because registration is an ill-posed problem, the maximum of Equation 2.1 may correspond to anatomically or physically implausible deformations. To address this, both the transformation model and the optimization framework must incorporate mechanisms through regularization that ensure solutions remain realistic, reliable, and robust. Regularization terms introduce prior knowledge about desirable deformation properties (e.g., smoothness, diffeomorphism, or inverse consistency (IC)) and may stem from analytical formulations (gradient-based penalties, Jacobian constraints) [30–32, 55, 101, 130], from biomechanical or organ-specific models (e.g., hyperelastic or viscous-fluid PDEs) [131–133], or from learned priors (e.g., adversarial losses and latent-space constraints) [134–136]. To preserve the continuity and differentiability of anatomical structures, transformations should ideally be diffeomorphic, meaning they are smooth and invertible. Diffeomorphic mappings ensure smooth deformation of tissues without introducing folds and discontinuities across the registering images [137]. Table 2.2 summarizes the main categories of regularization.

Additionally, the problem may be regularized through implicit and explicit regularization [63, 64]. Implicit regularization can be accomplished by parameterizing the deformations using smooth functions, whereas explicit regularization can be realized by implementing hard and soft constraints. Hard constraints force the solution to satisfy certain criteria for the registration to be successful. In contrast, soft constraints encode preferences regarding specific configurations, guiding the model to desired results. In the latter, deviations from these preferences are allowed to a certain degree. Thus, the regularization term $\mathcal{R}$ in Equation 2.1 penalizes implausible deformations by using explicit soft constraints to guide the model.

Diffusion Regularization. Diffusion penalizes the first-order spatial derivatives of the deformation field $\nabla\phi$, encouraging smooth and continuous deformations throughout the domain. It computes the squared L2-norm of the gradients, effectively penalizing disparities between adjacent deformations, and can be formulated as:

$$R_{diffusion} = \sum_{\boldsymbol{p}} \|\nabla\phi(\boldsymbol{p})\|_2^2. \tag{2.25}$$

Bending Energy. Bending energy extends diffusion regularization by instead penalizing the second-order derivatives $\nabla^2\phi$, promoting curvature smoothness [145].

Table 2.2: Comparison between regularization paradigms in medical image registration.

	Analytical Regularization	**Physics-informed Regularization**	**Learned Regularization**
Principles	Explicit mathematical penalties derived from differential formulations	Biomechanical and organ-specific constraints to tissue motion	Implicit deformation priors learned from data distributions
Strengths and Limitations	+ Simple and computationally efficient + Promotes diffeomorphism − Limited expressiveness for complex tissue behavior − May oversmooth deformations	+ Produces physically realistic deformations + Models resection, gravity, tissue relaxation, fluid loss − Sensitive to parameterization of tissue properties − Higher computational cost	+ Captures complex nonlinear deformation patterns + Adapts to data variability − Dependent on training data quality and diversity − Limited interpretability
Clinical Relevance	Prevents folding or unrealistic distortions	Ensures realistic deformation behavior of tissues in surgical settings	Improves robustness and generalization in heterogeneous clinical data
Examples	Diffusion [30, 55, 138] Jacobian determinant [130, 139] Inverse-consistency [32, 101, 140] Diffeomorphic integration [31, 141]	Hyperelastic models [131] Viscous-liquid models [132] Divergence-free flows [133]	Adversarial losses [134, 142] Latent-space priors [143, 144] Uncertainty maps [135, 136]

Invertibility and Topology Preservation. Invertibility guarantees that the transformation relating two images can be reversed [146]. Mathematically, a transformation ϕ is invertible if it is injective, ensuring every point in the fixed image corresponds uniquely to a point in the moving image. Noninvertible deformations imply multiple points in the fixed image map to the same location in the moving image, causing tissue folding, loss of anatomical integrity, or ambiguities in downstream clinical analyses. This condition is commonly evaluated by analyzing the Jacobian determinant of the deformation field, $|\boldsymbol{J}_\phi|$, at each voxel. Values of $|\boldsymbol{J}_\phi| > 0$ indicate local invertibility (i.e., no folding or tearing) [146], while $|\boldsymbol{J}_\phi| = 1$ implies local volume preservation [147]. To enforce these properties, the regularization terms that are often added to the loss function penalize negative Jacobian values or deviations from unit volume:

$$\mathcal{R}_{Jdet} = \frac{1}{|\Omega|} \sum_{\mathbf{p} \in \boldsymbol{\Omega}} \sigma\big(|\boldsymbol{J}_\phi|(\boldsymbol{p})\big), \tag{2.26}$$

where σ is a function that penalizes the objective function for certain values of $|\boldsymbol{J}_\phi|$. These techniques help prevent artifacts such as overlapping regions or implausible tissue compression or expansion.

Inverse Consistency. Inverse consistency addresses the symmetry of the registration process, ensuring that the order of registration (i.e., whether $\boldsymbol{I_M}$ is registered to $\boldsymbol{I_F}$ or vice versa) does not introduce directional biases [148]. This is enforced by penalizing deviations from the identity $\boldsymbol{I}$

when composing forward and backward transformations:

$$\phi_{M,F} \circ \phi_{F,M} = \boldsymbol{I}. \tag{2.27}$$

Such constraints preserve symmetry and encourage physical plausibility, ensuring $\phi_{M,F}$ and $\phi_{F,M}$ behave as true inverses [101, 149]. An IC regularizer term may then be written as:

$$R_{IC} = \|\phi_{M,F} \circ \phi_{F,M} - \boldsymbol{I}\|_2^2 + \|\phi_{F,M} \circ \phi_{M,F} - \boldsymbol{I}\|_2^2, \tag{2.28}$$

where $\phi_{M,F}$ is the forward displacement field to register $\boldsymbol{I_M}$ to $\boldsymbol{I_F}$, and $\phi_{F,M}$ is the backwards displacement field with reverse inputs.

2.1.7 Optimization

Image registration is an inherently ill-posed problem (as defined by Hadamard [150] for mathematical problems), since multiple transformations may map one image to another with comparable accuracy. In practice, no unique or perfect transformation exists that fully aligns the fixed and moving images, and the task typically requires optimization over a high-dimensional and non-convex parameter space. This ambiguity arises due to factors like missing correspondences, noise, intensity variations, and the under-constrained nature of the problem, which is often addressed by guiding the solution toward a meaningful deformation space through the use of regularization terms or prior knowledge. Therefore, the transformation parameters can be directly computed from the available data or searched for by optimizing an objective function dependent on these parameters [64, 151]. In the latter case, the maximization of Equation 2.1 generally does not have a closed-form solution; consequently, the displacements are estimated by resorting to numerical optimization methods [152].

In deformable registration applications, choosing an optimizer can be difficult because the more flexible the transformation model, the more parameters are generally required to describe it [64]. Common approaches employ continuous optimization methods [153], such as gradient descent [154], conjugate gradient [155], Powell's conjugate directions [156], quasi-Newton [155], and Levenberg-Marquardt [92], that estimate the optimal parameters following an update rule of the form:

$$\theta_{t+1} = \theta_t + \alpha_t g_t(\mathcal{S}, \mathcal{R}, \theta_t), \tag{2.29}$$

where θ denotes the vector of parameters of the transformation model, t is the index of the current optimization iteration, α is the step size, and g computes the search direction in the parameter space by taking into account both the similarity and regularization terms.

The specific continuous optimization algorithm defines the way α and g are defined. Alternatively, the problem can also be approached from a discrete optimization point of view after quantization of the displacement space into discrete intervals [157–159]. More recent approaches that employ deep neural networks (DNNs) to parameterize the displacements are typically optimized with tailored optimization paradigms, such as stochastic gradient descent with momentum [160, 161], or adaptive optimizers such as Adam and Adamax [162] which leverage adaptive step sizes to accelerate convergence.

2.2 Deep Learning in Medical Image Registration

Traditional image registration methods, which rely on iteratively solving an optimization problem, are widely recognized but have significant drawbacks. Primarily, these methods are computationally demanding and slow, as the optimization process must be repeated for each distinct moving and fixed image pair, resulting in redundancy. Although many improvements to traditional methods have been made [64, 83, 98], DL-based methods have demonstrated remarkable performance in image registration [65, 163] and various other medical image analysis tasks [164–166].

While classic iterative methods have established a strong foundation for deformation modeling, their need to solve an optimization problem for each new image pair and their reliance on handcrafted similarity metrics have motivated a paradigm shift toward DL. Unlike iterative approaches, DL-based models learn an implicit function that can directly infer a displacement field in a single forward pass, eliminating the need for case-specific optimization. This enables near-instantaneous registration, a critical advantage in time-sensitive surgical workflows. Through large-scale training, such models learn data-driven representations that replace handcrafted metrics and improve robustness and generalization across heterogeneous anatomies and pathologies.

Deep learning methods are trained by optimizing a global objective function (similar to Equation 2.1) over a dataset, additionally benefiting from implicit regularization that arises from the diversity of training samples, which smooths the loss landscape and reduces susceptibility to suboptimal minima [65]. Once trained, these networks can process unseen image pairs in one pass without additional optimization steps, which facilitates clinical deployment. However, they also present challenges: they may overfit to the training distribution, leading to reduced generalization at inference time, and their performance is fundamentally constrained by model capacity as well as the quantity and quality of the available training data [65].

Early image registration DNNs mainly leveraged encoder architectures, which mainly served as upstream feature extraction for conventional techniques [167] or as regressors for estimating affine transformation parameters [168, 169]. The success of U-net-like architectures [170–172] in the field of medical imaging prompted the application of encoder-decoder frameworks to learning-based deformable registration, implemented within supervised learning frameworks. In medical image registration, supervised learning-based methods use ground-truth displacements, landmark correspondences, or anatomical labels as target output during the training process, comparing the network predictions with the ground truths. In contrast, unsupervised methods refer to those that do not require such ground truths (Figure 2.2). The ground-truth data for supervised methods often comes from traditional registration approaches [33, 35, 173] or is manually annotated by experts, a process that is time-consuming and can restrict registration performance as it relies on the quality of the generated labels. The advent of Spatial Transformer networks (STNs) [174] (not to be confused with Transformers [175]) marked a shift towards unsupervised methods, which eliminate the dependency on ground truths. The Spatial Transformer applies the predicted deformation to warp the moving image, subsequently evaluating it against the fixed image itself. When anatomical label maps are accessible for both images, the Spatial Transformer can also generate the warped

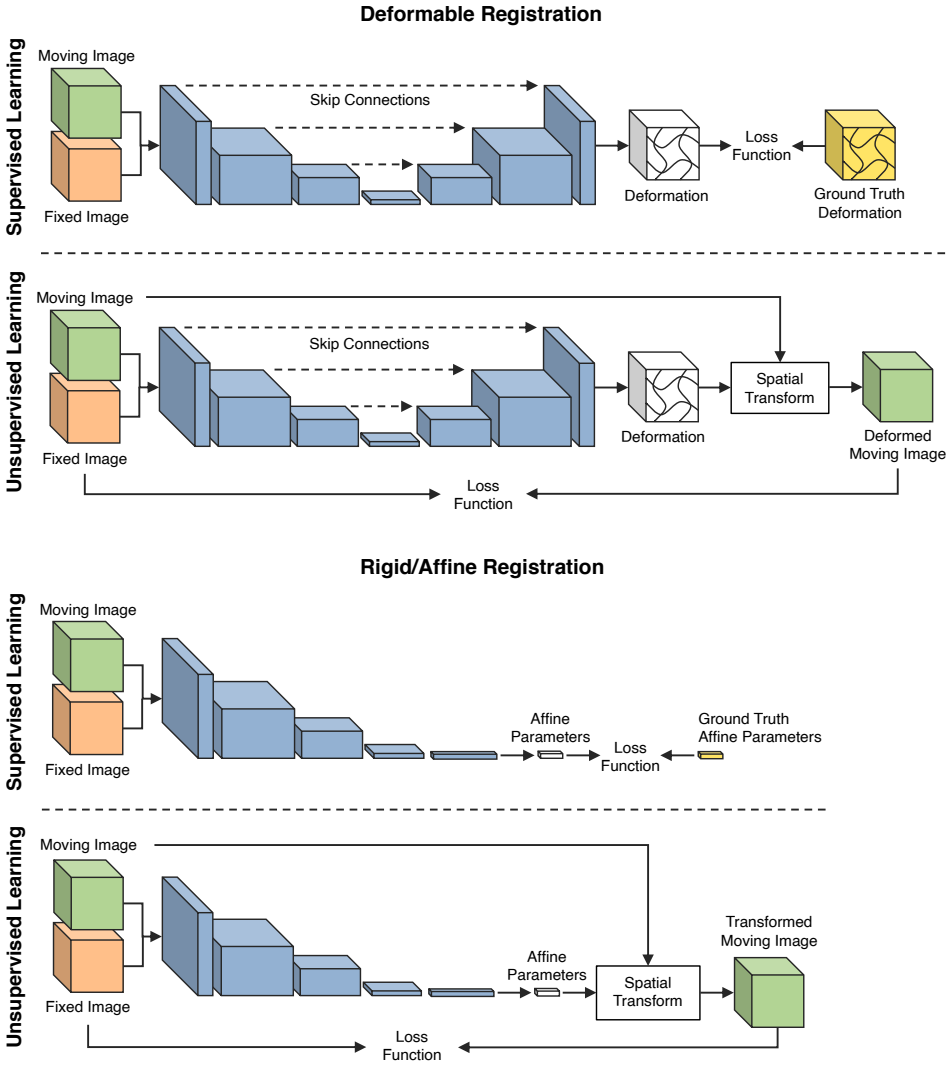


Figure 2.2: Overview of common learning-based frameworks for medical image registration. The top panels showcase the mechanism of deformable image registration, where a dense displacement field is predicted, commonly utilizing U-Net architectures. The bottom panels demonstrate rigid/affine registration, which generally relies on an encoder and fully connected layers to predict a set of transformation parameters. Adapted from [65].

moving label map using the predicted displacement. Anatomy-based guidance can be employed during network training by using these warped label maps [105, 176–178]. In practice, any of the parameterizations discussed in Section 2.1.3 (affine models) and Section 2.1.4 (deformable models) can be adapted into a DNN framework. The networks are trained through backpropagation [179] by globally optimizing a loss function similar to Equation 2.1 using the training data. Unseen testing images are then fed into the trained networks for inference.

2.2.1 Network Architectures

Generative Adversarial Networks. Generative adversarial networks (GANs) [180] applied to image registration usually employ a generator, producing a deformation field to transform the moving

image, and a discriminator, that subsequently assesses the similarity between the deformed moving image and the fixed image, functioning as a learned similarity mechanism. This configuration alleviates the need for a handcrafted similarity metric, making the approach adaptable to complex multimodal registration scenarios [104, 106, 134].

Contrastive Learning. Contrastive learning [181] focuses on learning representations by distinguishing between positive pairs (similar images) and negative pairs (differing images). The objective is to bring representations of positive pairs closer in the feature space while pushing negative pairs apart. Using contrastive learning, the network can learn to determine whether two images are properly aligned without relying on a handcrafted similarity measure, which is often less robust [182, 183]. Siamese networks [184] are particularly well-suited for this task, as they consist of twin networks that share weights and learn to compare image pairs. When combined with contrastive learning, Siamese networks can effectively perform image registration by learning a robust similarity metric, making them effective in multimodal registration scenarios [185, 186].

U-Net. The U-Net architecture [170–172] has shown impressive performance in various dense prediction tasks, including segmentation and registration. Its design, featuring symmetric encoder and decoder branches connected by skip connections, effectively fuses coarse and fine-grained information, making it particularly suitable for tasks requiring spatial accuracy. In image registration, U-Net serves as the backbone of many models [30, 31, 187, 188], as its modular nature allows researchers to adapt its architecture for specific applications with domain-specific enhancements. Similarly, encoder-decoder architectures without skip connections can also be employed to learn low-dimensional representations by forcing the network to condense information into a compressed latent space. Such designs inherently regularize the learning process by restricting the model to focus on global rather than local patterns.

Attention Blocks. Transformers [175], originally introduced for natural language processing, have recently been adopted in medical image registration to address limitations of convolutional architectures in modeling long-range spatial relationships. By leveraging self-attention, these methods explicitly model long-range dependencies between fixed and moving images, which are difficult to represent using standard convolutional architectures. Attention mechanisms have been incorporated either as modules within convolutional encoders [56,103] or through fully Transformer-based designs [189]. While these approaches improve sensitivity to global context and facilitate the modeling of large deformations, they typically incur increased computational and memory costs, which may limit their practicality in time-constrained or intra-operative settings.

Dual Encoders. Dual-encoder architectures can independently process fixed and moving images before merging their representations in a deeper latent stage of the network. For example, the encoder can be used to generate descriptors for both images, which are then fed to a decoder to predict a relevant output, such as displacement fields [190] or feature extraction [57, 107]. This formulation allows for more tailored feature extraction from each image modality, enhancing the registration's flexibility in multimodal cases.

Multiresolution. Multiresolution architectures address the challenges of large deformations by

decomposing the predicted displacement into hierarchical levels. Each level refines an upscaled displacement field predicted at the previous resolution level while producing feature maps for subsequent refinement [31, 186, 191, 192]. This coarse-to-fine strategy facilitates accurate alignment by first resolving large-scale displacements and subsequently refining smaller details.

Instance Optimization. To overcome some intrinsic limitations of DL-based methods, especially the lack of large datasets for learning, many DL methods rely on IO. As the problem is highly non-convex, this refers to employing a two-step approach where the deep registration model predictions are utilized as initializations for traditional IO techniques to refine and adapt the predicted displacement field to the individual geometry and pathology of each case [102, 193–195].

Dynamic Regularization. Dynamic regularization addresses the limitation of fixed regularization weights in DNNs by allowing the strength of deformation constraints to adapt at inference time. This is commonly implemented using a lightweight auxiliary network, a hypernetwork, that modulates parameters or feature statistics of the main registration model based on a conditioning input. In this setting, the model learns how changes in regularization affect the deformation field and adapts its internal feature representations accordingly. This enables continuous control over deformation smoothness without retraining, supporting efficient patient-specific tuning, similar to IO, while preserving the runtime advantages of learning-based registration.

2.3 Biomechanical Modeling of the Brain

Biomechanical modeling of the brain provides a physics-based framework for understanding and predicting brain deformation during neurosurgical procedures [196, 197]. This modeling approach is grounded in continuum mechanics, allowing the brain to be represented as a deformable body governed by mathematical models that describe deformation, stress, and strain to study phenomena like elasticity, plasticity, and fluid flow.

Brain biomechanics in the scope of deformation modeling requires the construction of a physical model and a mathematical model to accurately simulate the appropriate behaviors of interest. The physical model of brain deformation is defined by four key components: geometry, material properties, boundary conditions (BCs), and loading [196]. Once the physical model is defined, the governing equations that describe its behavior need to be considered. These equations of continuum mechanics (that include mass, momentum, and energy) follow a Lagrangian framework [198–201], where the modeling tracks the displacement of objects over time instead of at fixed points in space. These PDEs, together with BCs, define the mathematical model associated with the biomechanical model.

2.3.1 Brain Geometry

A biomechanical model of the brain requires detailed geometric information of the surfaces representing its boundaries, the ventricles, and other relevant anatomical structures. This geometry is typically reconstructed from segmentations of preMRI, including the parenchyma, ventricles, and any lesions or tumors. Although this data can be acquired from atlas-based geometries [202–204],

it does not suffice for the tasks of surgery planning and image registration for intra-operative use. These tasks require patient-specific data acquired from radiological images, as anatomy and pathology varies between individuals. For instance, although the external brain surface can be easily extracted via skull-stripping algorithms [205, 206], the internal structures are harder to segment automatically, especially in the presence of pathologies such as tumors [207–209]. For these cases, alternatives to segmentations have been studied, such as meshless discretization and fuzzy tissue classification [210, 211].

In modern neurosurgical frameworks, this geometric information not only defines the simulation domain but also serves as the reference configuration for intra-operative registration. In most cases, the pre-operative anatomical model reconstructed from MRI is used to provide the baseline geometry against which intra-operative deformations caused by gravity, cerebrospinal fluid (CSF) drainage, or resection are measured and corrected.

2.3.2 Boundary Conditions and Loading

Boundary conditions define the interaction between the brain and its surrounding environment. One of the most critical and challenging aspects of biomechanical modeling is how to model the brain-skull interface [212, 213]. Early models assumed a fixed boundary, constraining all motion at the inner skull surface [214, 215]. However, experimental and computational analyses demonstrated that the brain can slide relatively to the skull, particularly after craniotomy and dural opening, allowing it to be modeled as a frictionless sliding contact [216–218]. The skull itself is much stiffer than the parenchyma, so it is assumed to be rigid and is modeled as such. Constraints of the spinal cord on brain motion can also be simulated by constraining the spinal end of the model in the same way.

Loading forces represent the physiological and surgical factors that induce brain deformation during neurosurgery. These include gravitational forces associated with patient positioning, changes in intracranial pressure following CSF drainage and tumor resection, or external constraints imposed by surgical tools [14]. Intra-operative simulations often take boundary motions derived from measurable displacements as the primary loading mechanism [219, 220]. For instance, cortical surface motion captured by intra-operative imaging modalities (e.g., stereo vision, US, or MRI) can be directly applied as BCs [221]. On the other hand, physiological effects like gravity, head orientation, or CSF loss represent distributed forces and pressures [222–225]. The integration of such physics-informed data allows the model to predict realistic brain deformations from limited intra-operative cues.

2.3.3 Material Properties

The brain's material properties are required quantitative physical parameters for the biomechanical model, as they describe how the brain tissue responds under certain mechanical loads. They include measures of tissue stiffness (e.g., elastic modulus), compressibility (e.g., Poisson's ratio), density, viscosity, and other non-linear constitutive parameters specific to the chosen material

model [226–229]. The brain exhibits a very complex mechanical response, as its stiffness and stress-strain relationship vary significantly under tension and compression [230, 231]. However, modeling results are weakly dependent on the material properties when the loading is prescribed as imposed motion on the boundaries of the geometry [232, 233]. Due to this, the brain is generally modeled as incompressible and isotropic at the macroscopic levels relevant to neurosurgical simulation, using a simplified hyperelastic material model, such as the Neo-Hookean formulation [220, 234, 235]. For scenarios involving large, non-linear deformations and heterogeneous strain fields, a more advanced constitutive model, such as the Ogden formulation [225, 236], is recommended. This is particularly relevant in tumor resection simulations, where accurately capturing the mechanical response near resection margins requires modeling non-linear stress distributions and the asymmetric behavior of brain tissue under tension and compression [224, 236].

2.3.4 Mathematical Model

The computational efficiency and accuracy of biomechanical brain modeling depend on the numerical formulation adopted for solving the governing continuum mechanics equations. The continuous differential equations of motion that govern tissue deformation are solved step by step by time integration [237, 238], and, under the applied loading conditions, the model evolves from an initial configuration towards a steady state [239, 240].

Most commercial algorithms use the updated Lagrangian formulation [241], where all model variables refer to the system's configuration at the previous iteration step. This simplifies the implementation of some physical behaviors but is computationally expensive as it requires all derivatives with respect to the spatial coordinates to be recomputed at each time step. Conversely, in a total Lagrangian formulation [201, 238, 242], all spatial derivatives are expressed with respect to the initial undeformed configuration of the system and can therefore be precomputed. This formulation drastically reduces the computational burden per step, making it particularly important for time-critical applications such as surgical simulation and intra-operative image registration [196].

Time integration can be performed using either implicit or explicit methods [239, 240]. While implicit methods [243] are unconditionally stable, they require the repeated solution of large nonlinear systems of equations at each time step. This makes implicit methods computationally expensive (up to three orders of magnitude larger than explicit methods [237]) and unsuitable for intra-operative use [196]. Conversely, explicit integration methods [242, 244] are only conditionally stable but highly efficient since they do not require iterative solvers or full matrix computations at each step.

To obtain a numerical solution, the brain volume is represented by a computational grid created through spatial discretization [245]. This process divides the continuous domain into discrete nodes that approximate the continuum equations numerically. Each node corresponds to a point in the brain's geometry where quantities such as displacement, velocity, and stress are computed. The finite element (FE) method [246, 247] remains the most widely used numerical approach for modeling brain biomechanics. In neurosurgical simulation and intra-operative registration, the

computational grid usually takes the form of a FE mesh with low-order elements (e.g., 8-node hexahedra or 4-node tetrahedra) to minimize computational cost. These elements define a fixed mesh topology connecting nodes, and each element has shape functions that interpolate displacement within its volume. Consequently, hybrid meshes combining both hexahedral and tetrahedral elements are often employed, offering a practical balance between patient-specific accuracy and computational speed [248–250].

An alternative to the FE method is the use of meshless methods [225, 235, 242, 244], which eliminate the need for explicit mesh generation. Instead of constructing a structured element-based grid, the computational domain is represented by a point cloud (i.e., a set of scattered nodes) distributed within the brain volume. This approach removes the need for strict mesh creation, allowing straightforward automatic generation of patient-specific models directly from medical images. In meshless formulations, field variables are approximated using shape functions defined solely by nodal positions, without any strict element connectivity [210, 251]. Because node placement is almost arbitrary, meshless methods are particularly robust under large deformations and topological changes, such as those occurring during tumor resection or cortical collapse.

2.3.5 Biomechanics-constrained Intra-operative Registration

Within the context of intra-operative image registration, a biomechanical model can act as a regularizer for the registration to enforce biomechanical consistency [249, 252] when intra-operative data is available (typically MRI, US, or stereovision). In the absence of complete intra-operative imaging, biomechanical models may instead incorporate sparse intra-operative cues such as head orientation, brain shift due to surgical retraction, or amount of CSF loss to update the pre-operative anatomy [220, 253–256]. In these cases, their role shifts from registration to prediction, requiring more advanced modeling and tissue characterization. Hybrid methods have also been proposed that leverage partial volumetric data or surface information, using manually delineated cortical surfaces [257], laser range scanning [258, 259], or 2D microscope images [260, 261]. In such scenarios, biomechanical models serve as a regularization prior where data is observed, and play a predictive role to estimate the unobserved deformation field.

In the frameworks mentioned above, these pipelines reconstruct a patient-specific biomechanical model from preMRI, which defines the undeformed reference configuration. Intra-operative measurements are then incorporated as BCs or external loadings. Solving the resulting biomechanical system yields nodal displacement vectors that must be interpolated onto the image voxel grid using robust schemes, such as multilevel B-splines [262, 263]. This interpolation step enables real-time visualization of the warped pre-operative volume into the current intra-operative state [220, 263] (Figure 2.3).

Validation studies [220, 264, 265] have shown that biomechanics-based registration outperforms rigid and spline-based approaches, and achieves comparable results to intensity-based deformable registration. Due to these strengths, recent advances have extended biomechanical modeling of the brain toward physics-guided DL models, including physics-informed neural networks

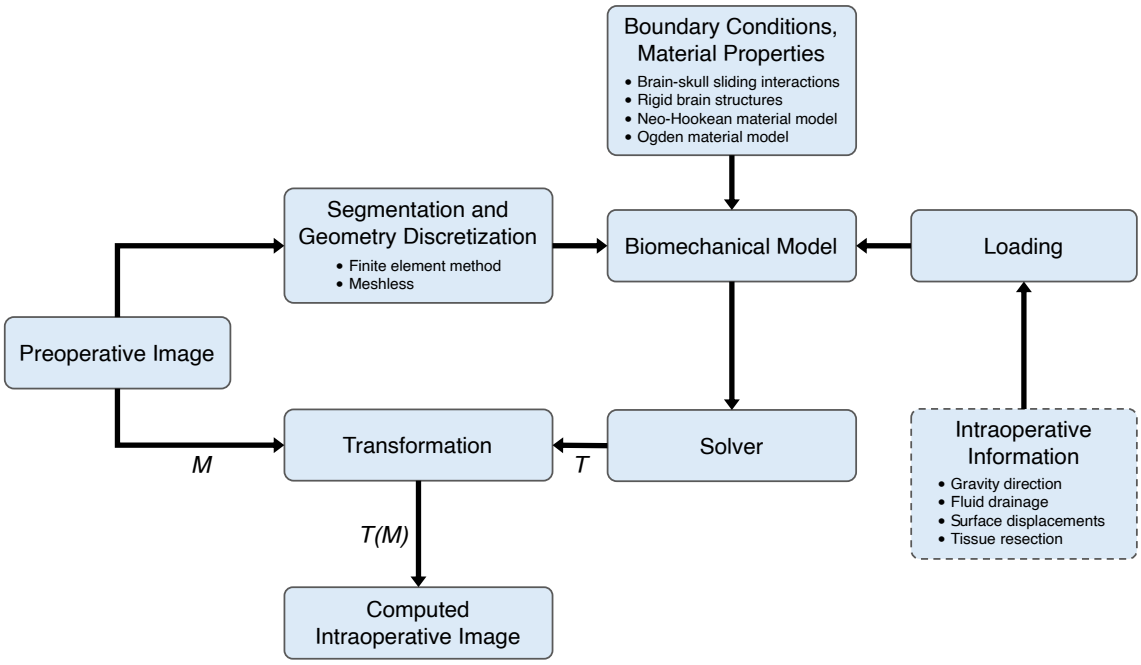


Figure 2.3: Registration based on biomechanical modeling of the brain. Pre-operative image segmentation and geometry discretization are used to construct a patient-specific biomechanical model, which is updated intra-operatively using boundary conditions, material properties, and surgical cues (e.g., head orientation, tissue resection, cortical surface displacements, cerebrospinal fluid amount) to compute a transformation T that maps the pre-operative anatomy M to the intra-operative state $T(M)$. Adapted from [220].

(PINNs) [131, 266, 267] and graph neural networks (GNNs) [268]. These approaches embed biomechanical PDEs directly into neural architectures, allowing deformation prediction from limited intra-operative data with physical consistency and near real-time performance. Common strategies also include using biomechanical models to generate synthetic ground-truth displacements for supervised DL-based registration approaches [40] (such as this thesis), incorporating PDE-based loss terms [131], using FE method outputs as priors or regularizers [142], or, alternatively, directly driving FE-based models with sparse intra-operative information such as matched keypoints [265].

Chapter 3

Related Work

This thesis focuses on the task of aligning pre-operative and intra-operative magnetic resonance imaging (iMRI) following tumor resection by predicting biomechanically plausible deformations using sparse keypoints within a supervised learning framework. To achieve this, biomechanical brain simulations were employed as ground truth to provide implicit, physically meaningful constraints to the learning process. Consequently, it is essential to review the state of the art in intra-operative and longitudinal image registration, examine how sparse keypoint interpolation is currently applied in this context, and explore how biomechanical modeling of brain deformation under tissue resection can serve as a source of physically-informed guidance for more realistic and robust registration methods.

3.1 Intra-operative and Longitudinal Image Registration

Historically, image registration has been dominated by classical iterative methods that perform case-specific optimization to align each individual image pair [269–272]. More recently, the field has shifted toward deep learning (DL) frameworks that learn complex deformation mappings directly from data [30, 31, 56, 188, 273]. The following subsections review representative learning-based approaches published in recent years (from 2020 to mid-2025), grouped according to how they frame registration and handle clinical challenges such as multimodality and missing correspondences. An overview of the main DL-based methods and datasets discussed here is provided in Table 3.1 and Table 4.2.

3.1.1 Direct Displacement Field Regression

The most common DL formulation for registration directly regresses dense displacement fields from image pairs using U-Net-like architectures [29, 30, 187, 273] (Figure 3.1-A). These models take image pairs as input and directly predict a displacement vector for each voxel in the image domain.

An example is iRegNet by Zeineldin *et al.* [274], originally proposed as a supervised U-Net for pre-operative magnetic resonance imaging (preMRI) to intra-operative ultrasound (iUS) registration with real-time inference on the Brain Images of Tumors for Evaluation (BITE) [282] and

Table 3.1: Summary of similarity losses, regularization strategies, and evaluation metrics used in novel deep learning-based neurosurgical registration methods. The datasets used during model training are also shown (= BITE, = RESECT, = ReMIND, = BraTS-Reg, = Private).

Reference	Dataset	Similarity Term					Regularization Term				Evaluation Metric						
		MSE	NCC	MI	MIND	Adversarial	Diffusion	Bending	Jacobian	IC	TRE	Robustness	MAE	Dice	HD	Jacobian	RMSE
MRI-CT																	
[106]			•		•	•	•				•			•	•	•	
US-US																	
[134]						•					•						
MRI-US																	
[103]			•					•			•					•	
[104]	/			•		•					•						
[274]	/	•	•				•				•						
MRI-MRI																	
[32]			•				•			•	•	•				•	
[275]			•				•			•	•						
[276]			•				•			•	•	•	•			•	
[277]			•				•				•					•	•
[195]	/	•	•				•				•						
[135]			•				•			•	•	•				•	
[136]			•				•			•	•	•				•	
[278]			•				•		•							•	
[279]		•					•					•	•				
[280]			•				•				•	•	•				
[105]			•	•			•					•	•				
[102]			•				•					•	•				
[100]		•	•				•		•			•	•			•	
[281]			•				•				•	•		•	•	•	

MSE - Mean squared error; NCC - Normalized cross-correlation; MI - Mutual information; MIND - Modality Independent Neighborhood Descriptor; IC - Inverse consistency; TRE - Target registration error; MAE - Maximum absolute error; HD - Hausdorff distance; RMSE - Root mean squared error.

REtroSpective Evaluation of Cerebral Tumors (RESECT) [283] datasets. The same backbone was later adapted [280] in a self-supervised setting for pre-operative to follow-up magnetic resonance imaging (MRI) alignment in the Brain Tumor Sequence Registration (BraTS-Reg) (2022) challenge, showing how simple convolutional neural network (CNN) architectures can be repurposed across modalities and tasks with appropriate loss function design. Supervised strategies have also been used to learn deformation patterns directly from outputs of classic registration algorithms. For instance, Shimamoto *et al.* [284] trained a dual U-Net (W-Net) to predict brain shift after dural opening using displacement fields generated by the Demons [94] algorithm as supervision. Although it achieved good mean target registration error (TRE) results when compared to standard affine approaches, the upper bound of supervised models is constrained by the quality and biases of the ground truth used during training, a general limitation of supervised methods.

Table 3.2: Comprehensive comparison of key publicly available datasets for benchmarking neurosurgical image registration.

Dataset	**RESECT** [283]	**BITE** [282]	**ReMIND** [285]	**BraTS-Reg** [286]
Clinical Data				
# Patients (M:F)	23[a]	14 (9:5)	114 (61:53)	259[a]
Age (years)	> 18[a]	52.00 ± 17.05 (23–76)	46.83 ± 14.76 (20–76)	>18[a]
Tumor Volume[b] (cm^3)	40.0 ± 44.09 (1.4–165.9)	34.65 ± 23.62 (0.2–79.2)	24.08 ± 23.31 (0.1–138.5)	–
Glioma Type	LGG (23)	LGG (4), HGG (10)	LGG (34), HGG (60)	HGG (259)
Magnetic Resonance				
Scanner	3 T Magnetom Skyra 1.5 T Magnetom Avanto (Siemens)	1.5 T Signa EXCITE (GE Healthcare) 3 T Magnetom Trio (Siemens)	3 T Magnetom Verio (Siemens)	–
Sequences	ceT1, T2-FLAIR	ceT1, T2	[c] T1, ceT1, T2, T2-FLAIR	T1, ceT1, T2, T2-FLAIR
Voxel Size	1 mm^3	1 mm^3	Variable	1 mm^3
Navigation				
Unit	Sonowand Invite (Sonowand AS)	IBIS NeuroNav [287]	Cranial Navigation (Brainlab)	–
Tracking	Polaris Spectra (Northern Digital Inc.)	Polaris (Northern Digital Inc.)	Polaris Spectra (Northern Digital Inc.)	–
Ultrasound				
Processing	Built-in	HDI 5000 (Philips ATL)	bk5000 (BK Medical)	–
Probe	12FLA-L (linear) (Sonowand AS)	P7-4 (phased) (Philips ATL)	N13C5 (curved) (BK Medical)	–
Frequency	12–6 MHz	7–4 MHz	13–5 MHz	–
Resolution	0.14 mm^3 to 0.24 mm^3	0.3 mm^3	$0.1 \times 0.1 \times 0.5$ mm	–
Annotations				
Landmarks (per patient)	16–34 (US-US) 9–16 (MRI-US)	19–40 (MRI-US)	Not yet available	6–50 (MRI-MRI)
Segmentations	Tumor, parenchyma, resection cavity [288] (manual)	Tumor (manual)	[c] Tumor, resection cavity (manual) [c] Parenchyma, ventricles (automatic)	–
SotA Accuracy (TRE)	$<$ 1–2 mm [104, 274, 289]	1–2 mm [104, 274]	3–5 mm [290]	1–3 mm [286]

TCIA - The Cancer Imaging Archive; LGG - Low-grade glioma; HGG - High-grade glioma; T1 - T1-weighted; ceT1 - Contrast-enhanced T1-weighted; T2 - T2-weighted; FLAIR - Fluid-attenuated inversion recovery; MRI - Magnetic resonance imaging; iUS - intra-operative ultrasound.

[a] This dataset does not provide data regarding patient age or gender.

[b] Tumor volume is quantified based on manually drawn boundaries in the pre-operative MRI.

[c] Data not available for every patient.

3.1.2 Landmark- and Feature-based Registration

Landmark- and feature-based methods adopt a different perspective, focusing on sparse salient structures or anatomical label maps to drive the registration rather than relying on dense voxel-wise alignment [37–39, 291] (Figure 3.1-B). By detecting and matching keypoints or features across images, these approaches tend to be more robust to large deformations, partial fields of view, and topological changes. They also offer increased interpretability, as correspondences can be visualized and quantitatively inspected. Such properties make these methods particularly well-suited for clinical applications involving complex anatomical changes (e.g., tumor resection), sparse annotations, or multimodal data.

Pirhadi *et al.* [107] illustrates this paradigm with a two-stage Siamese CNN for iUS-to-iUS registration inspired by object-tracking tasks. Their pipeline uses 2.5D patches across orthogonal planes to learn cross-modal correspondences, followed by affine transformation estimation via iterative reweighted least squares [292]. The model was pretrained on natural images and fine-tuned on medical data, showing strong performance with lower complexity compared to fully convolutional networks. However, the approach remains semi-automatic as it relies on manual annotation of reference landmarks. Moving beyond purely geometric keypoints, Nasser *et al.* [105] proposed WSSAMNet, which first segments regions of interest around manually annotated landmarks and then uses these segmentations as attention maps to guide a subsequent registration step. These maps guide the registration network toward relevant anatomical structures and help manage scenarios with missing correspondences, outperforming purely intensity-based baselines like VoxelMorph [30] and SyN [33] on the BraTS-Reg (2022) challenge.

3.1.3 Transformer-based Registration

Motivated by the success of Transformers in both natural language processing [175] and computer vision [293], several registration methods have explored attention mechanisms to capture long-range spatial dependencies [56, 189] (Figure 3.1-C). However, standard 3D Transformer models are computationally demanding, which limits their practicality for high-resolution neuroimaging.

Aziz *et al.* [103] addressed this limitation with EfficientMorph, a parameter-efficient Transformer based model for both uni- and multimodal registration. Their innovations include plane-wise attention mechanisms, where the model decomposes the problem into a sequence of 2D attentions across orthogonal planes, and efficient high-resolution tokenization by grouping and concatenating the features of adjacent non-overlapping voxel token blocks. This design significantly reduces memory and parameter count while retaining competitive performance, yielding state-of-the-art results on the ReMIND2Reg[1] dataset [129, 285, 294] with lower TRE, $16\times$ fewer parameters, and $5\times$ faster convergence than TransMorph [56]. This study shows that Transformer performance and clinical feasibility are not mutually exclusive, though plane-wise attention may be slightly less effective at capturing complex 3D interactions compared to true 3D attention.

[1] https://doi.org/10.5281/zenodo.12700312

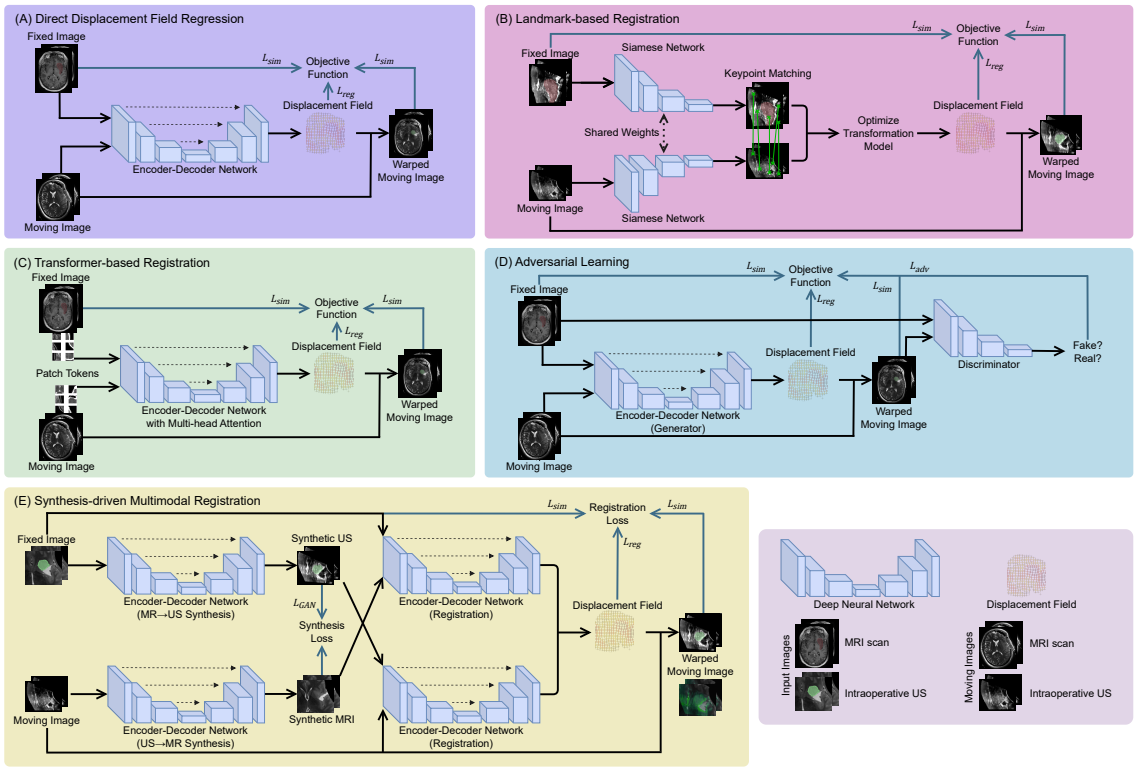


Figure 3.1: Representative deep learning frameworks for medical image registration. Blue arrows show the inputs to the objective function that guides backpropagation. (A) *Direct Displacement Field Regression* (Section 3.1.1): convolutional encoder-decoder networks directly predict dense voxel-wise displacement vectors between moving and fixed images by minimizing a composite loss comprising image similarity (L_{sim}) and regularization (L_{reg}) terms. (B) *Landmark-based Registration* (Section 3.1.2): these networks detect, match, and/or take sparse anatomical landmarks as input, using these correspondences to derive the displacement field. Green lines and stars represent matched keypoints between images. (C) *Transformer-based Registration* (Section 3.1.3): networks employ self-attention mechanisms to capture long-range spatial dependencies and global context, improving structural coherence and alignment across distant anatomical regions. (D) *Adversarial Learning* (Section 3.1.4): registration is guided by a discriminator network that distinguishes between realistic and unrealistic deformations, enforcing anatomical plausibility via a learned similarity measure, or adversarial loss (L_{adv}). (E) *Synthesis-driven Multimodal Registration* (Section 3.1.4): modality translation networks (e.g., MRI↔US synthesis) generate synthetic counterparts to facilitate cross-modality alignment. Registration and synthesis are jointly optimized with an added synthesis or adversarial losses (L_{GAN}).

3.1.4 Generative Adversarial Networks-based and Synthesis-driven Multimodal Registration

Multimodal registration, such as MRI-to-ultrasound (US) or MRI-to-computed tomography (CT), remains one of the most challenging problems in neurosurgical image analysis due to the large differences in image characteristics across modalities. Traditional similarity metrics employed in multimodal scenarios, such as normalized cross-correlation (NCC) or mutual information (MI), may still fail due to these fundamental differences in intensities, gradients, and noise. To address this, recent research explored the use of generative adversarial networks (GANs) [180] and

multimodal hierarchical variational autoencoders (MHVAEs) [295], which either implicitly learn modality-invariant representations through adversarial training (Figure 3.1-D) or synthesize cross-domain images into a common appearance space to facilitate alignment (Figure 3.1-E).

Han *et al.* [106] exemplify a synthesis-driven design by combining a CycleGAN for bidirectional MRI-CT translation with a dual-channel registration network that operates on real pre-operative and synthetically deformed intra-operative images. Uncertainty maps learned during synthesis are used to weight the contributions of each modality, ensuring that more reliable registration paths dominate where confidence is higher. Training incorporated standard cycle consistency and adversarial losses, along with a unique structure-consistency loss that measures the L1-norm between Modality Independent Neighbourhood Descriptor (MIND) features of the synthesized and input images. Similarly, Rahmani *et al.* [104] introduced D2BGAN, a Dual Discriminator Bayesian GAN for preMRI-to-iUS registration incorporating probabilistic priors in the form of a Bayesian term in the cycle-consistency loss and a standard MI-based loss to enforce cross-modal structural consistency. This probabilistic formulation does not rely on voxel intensities, making it efficient and informative for multimodal cases.

Rather than relying solely on explicit similarity losses, some adversarial frameworks use the discriminator itself as a learned surrogate for alignment quality. Wodzinski and Skalski [134] proposed a real-time GAN-based affine registration network for iUS-to-iUS alignment, where the generator directly predicts affine transformation parameters and the discriminator learns to distinguish aligned from misaligned volume pairs. The model achieved competitive TRE on the RESECT dataset with submillisecond inference times. This setup reframes registration as a learned assessment of alignment quality, allowing the discriminator to encode complex notions of correspondence that are difficult to express with traditional similarity metrics. Other work, such as that by Dorent *et al.* [295], uses MHVAEs to synthesize MRI from US and demonstrate that synthesis substantially enhances registration accuracy and feature matching, as the registration then occurs between the same modality. This framework was later leveraged to train a cross-modal feature matcher between MRI and US [39, 296], further underscoring the value of generative synthesis in reducing cross-modality complexity.

3.1.5 Multimodality Fusion

In clinical practice, medical imaging often involves multiple modalities or sequences, each capturing different aspects of brain anatomy or pathology. However, the heterogeneity and anisotropy of these inputs, such as varying spatial resolutions or contrasts, pose significant challenges for learning-based models. Instead of relying on a single input, some novel approaches employ modality fusion strategies that aim to combine data from multiple sources during the registration process.

One example of view-wise fusion is the MSF-AR Net proposed by Liu *et al.* [278], which addresses the resolution anisotropy commonly found in MRI by processing axial, sagittal, and coronal views in parallel streams. Its cross-attention-guided fusion module dynamically fuses the outputs of each stream by learning spatial attention weights that select the most trustworthy

deformation information from each plane. Abderezaei *et al.* [279] adopted a different strategy, employing an Inception-style [297] network to fuse multiparametric MRI into one channel, learning and encoding modality-specific features. This fused input is processed by a TransMorph-based registration network with a Swin Transformer [298] encoder for long-range context modeling. On the BraTS-Reg dataset, this approach notably outperformed the baseline TransMorph in cases with complex tumor changes.

3.1.6 Handling Absent Correspondences in Longitudinal Registration

One of the most difficult aspects of brain tumor registration is the presence of non-corresponding regions, such as resection cavities or newly formed lesions, which are present in one image but not in the other. This violates standard assumptions of smooth and invertible deformations. Failing to account for these changes can lead to implausible deformations, high errors, and distorted anatomy near the lesion. Recent studies have proposed to tackle this issue by learning to detect and exclude non-corresponding regions during registration and by adopting multiresolution pyramidal network architectures [31, 130, 188, 192].

The DIRAC framework by Mok and Chung [32] is a prominent example that focuses on consistency-based masking. Built on a bidirectional multiresolution backbone (LapIRN [31, 55]), DIRAC registers image pairs in both forward and backward directions and computes a voxel-wise consistency error. Regions with high inconsistency are assumed to represent missing or altered anatomy and are excluded from the similarity loss through a learned mask. This prevents the network from forcing unrealistic correspondences near tumor cavities or resected regions. DIRAC achieved the top rank in the BraTS-Reg (2022) challenge (MICCAI track) and strongly influenced subsequent models. Zhang *et al.* [276] extended this idea to allow multiparametric MRI integration by incorporating parallel contrast-specific pyramid streams and refining the inconsistency mask using morphological post-processing. This refinement proved effective in improving the sensitivity and specificity of non-correspondence detection, achieving comparable performance to DIRAC while incorporating richer multimodal cues.

Other recent studies enhance correspondence handling with explicit priors or focused attention. Feng *et al.* [275] introduced a stepwise corrected-attention mechanism that modulates network focus at each resolution level. This mechanism uses the displacement field from previous resolution levels and the corresponding warped segmentation of the pre-operative lesion to emphasize stable tissue while down-weighting pathological areas. This approach improves alignment around resection cavities, but is sensitive to the quality of segmentations. A different strategy is taken by Meng *et al.* [100], which presents NICE-Net, a noniterative cascade network, that combines an unsupervised NCC loss with sparse landmark supervision, providing weak geometric guidance in regions where intensity-based matching is unreliable (i.e., the resection cavity). Their model additionally incorporates inter-subject pretraining and case-specific fine-tuning to better adapt to individual patient anatomy.

Beyond geometry-based masking, Tang *et al.* [135] proposed an auxiliary-image-based in-

tensity consistency constraint. Instead of relying on consistency calculations between just two images, the model introduces a third, auxiliary image, and compares the direct registration path with an indirect one via the auxiliary scan. Differences between these two paths form a soft weight map that down-weights unreliable regions. This approach is particularly appealing for longitudinal studies with multiple time points. Similarly, Wu *et al.* [136] proposed a self-explainable approach using internal model behavior to detect inconsistencies without requiring explicit masks. Their method introduces noise-removed inconsistency activation maps, which highlight regions where the model struggles to reduce the similarity loss. The resulting maps serve as soft masks to weigh the loss function, improving robustness to unreliable or missing correspondences.

In contrast to all masking-based approaches, Joshi and Hong [99] introduced MetaRegNet, a metamorphic image registration framework that explicitly models both deformation and appearance changes. Their architecture uses dual branches to learn a smooth spatial deformation and an additive intensity source term representing new or missing structures (e.g., resection cavity or tumor recurrence). Built employing Lipschitz continuous residual blocks [299], the model encourages physically plausible deformations while accounting for lesion-induced mass effects.

3.1.7 Hybrid Learning

Even though DL-based registration models enable fast inference, they may fail to capture patient-specific details or generalize to out-of-distribution data. To mitigate this, hybrid registration frameworks have been proposed that integrate classic instance optimization (IO) with learning-based models at different stages of the pipeline. In these approaches, IO may be used to provide robust initial alignment, to refine deep network predictions at test time, or to be interleaved with learning-based components across multiple stages. In doing so, these frameworks iteratively tailor the deformations to the patient's specific anatomy and pathology.

A key example is the work of Wodzinski *et al.* [102], who proposed a multistage pipeline that first performs affine registration via IO, followed by non-rigid deformation prediction using a modified LapIRN. The predicted displacement field is then refined through a multiresolution non-rigid optimization stage guided by inverse consistency (IC). During this refinement, voxel-wise weights derived from forward-backward consistency modulate each term of the objective function, enabling precise corrections in challenging regions. This hybrid strategy won first place in the BraTS-Reg (2022) challenge (IEEE ISBI track), highlighting the benefits of combining global, data-driven predictions with local, optimization-based refinements. Similarly, Waldmannstetter *et al.* [195] extended VoxelMorph with a test-time IO refinement phase, showing systematic improvements near tumor margins. The model was trained with a composite mean squared error (MSE) and NCC loss to balance global intensity consistency with local structural alignment. Applied to pre-to-iMRI registration, this step consistently improved TRE over baseline VoxelMorph. The study also demonstrated that combining complementary similarity metrics improves convergence and robustness, suggesting hybrid loss designs can outperform single-metric formulations. In multimodal registration, Ha and Heinrich [194] proposed a modality-agnostic, self-supervised

framework that pairs deep feature learning with classic optimization for preMRI-iUS registration. A 3D CNN is trained to predict MIND descriptors, producing dense, modality-invariant features used to estimate probabilistic displacements on a sparse control grid, which are then interpolated and refined in a fast IO step.

As the field matures, these two components are no longer viewed as mutually exclusive but rather as synergistic tools: DL provides powerful priors and initial estimates while classic IO remains valuable for robustness, especially in complex and unique pathological settings.

3.2 Sparse Keypoint Interpolation for Deformable Registration

As discussed in Section 2.1.4, parametric registration methods represent non-rigid deformations using a sparse set of landmarks/keypoints combined with basis functions or splines. In this setting, the core idea is to estimate displacements only at a limited number of salient locations and then interpolate a dense displacement field over the entire volume. Sparse keypoint interpolation typically proceeds in three steps: (i) detect or provide keypoints in both fixed and moving images (often along with local descriptor vectors), (ii) establish correspondences between these points, and (iii) use the resulting sparse displacement vectors as conditioning for an interpolation model, such as splines [77], or Gaussian process (GP) regression [87].

Approaches driven by keypoints typically rely on feature detectors and descriptors (previously reviewed in Section 2.1.5), such as Scale-Invariant Feature Transform (SIFT) [119], Speeded-Up Robust Features (SURF) [120], MIND [124], Learned Invariant Feature Transform (LIFT) [126], or SuperPoint [125], to identify meaningful and salient locations. Once correspondences are established between the moving and fixed images, geometric interpolators propagate the sparse displacements at these keypoints across the image volume. Although effective in many applications, these interpolators impose global smoothness and continuity assumptions that limit their ability to model the complexities of neurosurgery. They often produce globally smooth deformations that do not account for sliding motions, tissue stiffness, or discontinuities caused by resection. Although the interpolated displacement fields are mathematically regular, they can be biomechanically implausible. These limitations motivate hybrid frameworks where sparse keypoints provide robust correspondences, while additional physical or learned priors refine their interpolation.

Several recent works have continued to focus on designing descriptors and networks that are contrast-invariant and anatomically meaningful for multimodal neurosurgical applications. Rasheed *et al.* [39] proposed learning contrast-invariant descriptors for MRI-US registration using a synthesis-based framework, in which synthetic ultrasound is generated from MRI to serve as a training target for contrastive learning. Morozov *et al.* [296] extend this concept into a 3D setting, learning descriptors for MRI-US volumes that enable patch-based rigid registration using a sparse set of keypoint matches. Similarly, Salari *et al.* [300] further demonstrate how contrastive learning can produce MRI-US landmark detectors that align with clinically meaningful structures such as vessels and sulci.

Recent learning-based approaches take a keypoint-driven registration approach by integrat-

ing feature detection, matching, and dense interpolation into end-to-end registration frameworks. Wang *et al.* [38] exemplifies this with KeyMorph, a network that learns corresponding keypoints and estimates an affine transformation directly from them in an end-to-end framework. Rather than relying on fixed hand-crafted descriptors, the produced keypoints adapt to anatomical structures during learning. This combination provides both interpretability and robustness to large misalignments. The same authors extend this principle to a much larger scale by training BrainMorph [291], a foundational keypoint model, trained on over 100K brain MRI scans. It supports multimodal, pairwise, and groupwise registration and allows flexible parameterizations of the deformation model, making it one of the first keypoint-driven foundation models for neurosurgical alignment. Heinrich *et al.* [37] propose a hybrid strategy that integrates sparse keypoint detection within a U-Net registration backbone. The method predicts probabilistic heatmaps for candidate keypoint locations and estimates corresponding matches through local correlation volumes. This yields a hybrid model that benefits from deep features while retaining the interpretability of keypoint registration.

Overall, sparse keypoint interpolation offers a robust and interpretable alternative to dense similarity-based registration, particularly when anatomy undergoes significant deformation or when only limited intra-operative data are available or large topological changes are present [117]. Classic interpolators remain fast and simple but lack biomechanical realism, limiting their ability to capture complex brain motion. Recent learning-based methods mitigate some of these issues by integrating keypoint detection in a learning framework with more flexible deformation models that yield reliable alignment in challenging neurosurgical settings. Yet, ensuring that interpolated fields remain biomechanically plausible remains an open and critical challenge given the reliance on sparse measurements during surgery. Addressing this need is a central motivation of this thesis, which seeks to integrate biomechanical constraints into keypoint-driven deformation estimation.

3.3 Biomechanical Modeling of the Brain in Neurosurgery

Constructing a biomechanical model of the brain serves a wide range of neurosurgical tasks: surgical training, pre-operative planning, and intra-operative registration. Within the context of the last, biomechanical modeling has long been explored as a means to compensate for brain shift brought on by craniotomy, gravity, cerebrospinal fluid (CSF) loss, and tumor resection [253–256].

The most widely adopted approaches rely on patient-specific finite element (FE) simulations to estimate whole-brain deformation under intra-operative loads. A representative example includes the work of Mostayed *et al.* [220], who developed fast explicit solvers to generate clinically usable biomechanical model predictions for craniotomy-induced shift correction. Their workflow constructs individualized brain meshes from preMRI, segmenting parenchyma, tumor, and ventricles, and assigning them specific material properties. Cortical displacements derived by rigidly aligning preMRI and iMRI act as the boundary conditions (BCs). To satisfy the strict intra-operative time constraints, the authors adopt a total Lagrangian explicit dynamics (TLED) formulation implemented for graphics processing units (GPUs) [301], avoiding iterative solvers

and producing deformation updates in seconds. Evaluated on a large retrospective database of glioma patients available at the Children's Hospital in Boston, the method consistently outperforms classic B-spline registration, particularly in regions undergoing large shifts where image registration approaches struggle to maintain physical plausibility.

As intra-operative imaging availability varies, some methods often seek to incorporate only sparse but reliable structural cues. Morin *et al.* [249] proposed a biomechanical model driven by vascular information extracted from preMRI and iUS. By matching vessel trees using an iterative closest point algorithm [302] and enforcing sliding or directional constraints, their method captures both craniotomy- and resection-related deformation without explicitly modeling the resection cavity. Validated on retrospective studies with near-intra-operative runtimes, this work offers a computationally efficient solution when only limited data are available.

Other frameworks integrate richer but still incomplete intra-operative measurements into their models, such as stereo camera reconstructions. Fan *et al.* [303] combine intra-operative stereovision with a patient-specific FE model guided by sparse displacement constraints. These displacements are measured before resection using textured cortical surfaces and vascular features, and during resection using optical-flow tracking. This enables the generation of updated MRI volumes that qualitatively match intra-operative imaging in proof-of-concept demonstrations.

Tumor resection introduces topological changes that challenge static mesh representations, motivating the development of adaptive biomechanical frameworks. Drakopoulos *et al.* [224] address this by dynamically removing mesh elements corresponding to resected tissue and incorporating sparse displacement vectors extracted from iMRI. This adaptive FE domain evolves with the surgical cavity and improves deformation accuracy relative to rigid and B-spline baselines while maintaining runtimes compatible with surgical workflows.

Atlas-based methods provide yet another alternative by avoiding full FE re-solving altogether. Sun *et al.* [259] construct a large deformation atlas that captures expected variations in patient positioning, CSF drainage, and edema growth. Sparse intra-operative measurements are then used to search the deformation atlas for the best-matching global displacement field. While less flexible than fully dynamic FE approaches, an atlas-based solution achieves faster updates and is useful when intra-operative data are limited.

Finally, to reduce the manual burden on surgeons associated with generating patient-specific models, automated and meshless methods have emerged. Yu *et al.* [225, 236] employ a meshless TLED [235, 242] framework with automated preprocessing and fuzzy tissue classification to construct patient-specific biomechanical models from preMRI, eliminating the need for strict FE meshing or hard segmentations. The loading relies on gravity forces guided by the incision vector direction, along with resection-induced reaction forces estimated at the tumor boundaries. This fully automated pipeline significantly decreases preprocessing effort while producing accuracy on par with established B-spline registration.

Chapter 4

Methods

4.1 Overview and Problem Setting

The proposed framework assumes access to a pre-operative magnetic resonance imaging (preMRI), $\boldsymbol{I_{pre}} \in \mathbb{R}^{D \times W \times H}$, where D denotes the depth, W the width, H the height, and a sparse set of M matched keypoints $\{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i=1}^{M}$, where $(\boldsymbol{x}_i, \boldsymbol{y}_i) \in \mathbb{R}^3 \times \mathbb{R}^3$ represent corresponding 3D anatomical locations in the pre-operative and intra-operative spaces, respectively. The matched keypoints may be obtained either manually or automatically from any intra-operative imaging modality, such as intra-operative ultrasound (iUS). At each keypoint $(\boldsymbol{x}_i, \boldsymbol{y}_i)$, the displacement vector $\boldsymbol{d}_i = \boldsymbol{y}_i - \boldsymbol{x}_i$ captures the local brain displacement occurring during surgery. The objective is to estimate a *dense* and *physically plausible* displacement field $\phi \in \mathbb{R}^{3 \times D \times W \times H}$ that estimates the surgical brain deformations. To this end, a deep interpolator f_θ is trained, parameterized by learnable parameters θ, using biomechanical simulations (Figure 4.1).

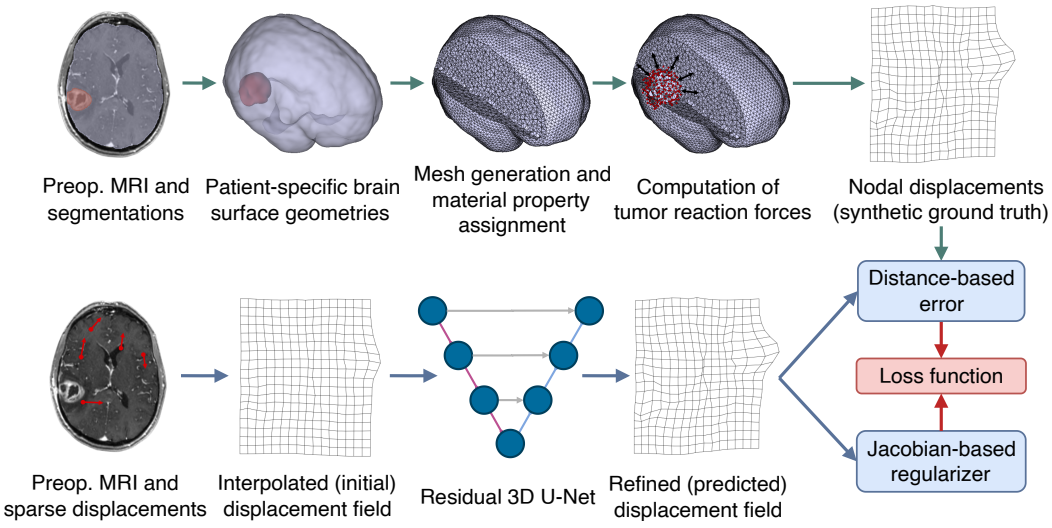


Figure 4.1: Overview of the proposed framework. A biomechanical simulation generates synthetic ground-truth displacement fields using pre-operative MRI and joint segmentation of tumor and surrounding structures. Sparse intra-operative keypoint displacements are interpolated to form an initial estimate, which is refined by a residual 3D U-Net. The final displacement field is supervised using voxel-wise error and a Jacobian-based regularization loss. Adapted from [40].

4.2 Clinical Data

This work leverages large-scale, publicly available collections of multimodal imaging data from brain tumor surgeries. This data has been rigorously curated, fully de-identified, and distributed via The Cancer Imaging Archive (TCIA) in standardized formats, making it a valuable benchmark to support research in brain shift compensation, image-guided neurosurgery, and computational method development. A description of the final processed data actually used for training and evaluation can be found in Section 4.5.

4.2.1 The Brain Resection Multimodal Imaging Database (ReMIND)

The Brain Resection Multimodal Imaging Database (ReMIND) [285] database is the largest publicly available collection of multimodal imaging data from brain tumor surgeries. It includes 114 patients treated at the Advanced Multimodality Image Guided Operating suite at Brigham and Women's Hospital between 2018 and 2022, covering gliomas ($n = 92$), metastases ($n = 11$), and other non-glioma tumors ($n = 11$). The dataset contains 369 multiparametric preMRI series (T1-weighted (T1), contrast-enhanced T1-weighted (ceT1), T2-weighted (T2), and T2-fluid-attenuated inversion recovery (FLAIR)), 320 three-dimensional iUS volumes, 301 intra-operative magnetic resonance imaging (iMRI) series, and 356 expert-validated segmentations (Figure 4.2). Ultrasound data were reconstructed from tracked sweeps at up to three surgical stages (before dural opening, after dural opening, and after tumor resection), while iMRI was performed to assess residual tumor. Segmentations include pre-operative and residual tumor, cerebrum, ventricles, and prior resection cavities. Detailed demographic, histopathological, and surgical metadata are also provided.

For this work, all ReMIND images were resampled to 1 mm^3 isotropic resolution, and intra-patient rigid coregistration was performed using the "General Registration (BRAINS)" module in 3D Slicer [304], a free open source software for medical image analysis and visualization.

4.2.2 University of Pennsylvania Glioblastoma (UPenn-GBM) Dataset

The University of Pennsylvania Glioblastoma (UPenn-GBM) dataset [305] is a multiparametric magnetic resonance imaging (MRI) collection of 630 glioblastoma patients imaged at the Hospital of the University of Pennsylvania between 2006 and 2018. It includes T1, ceT1, T2, and T2-FLAIR scans acquired at pre-operative ($n = 611$), follow-up ($n = 19$), or both ($n = 41$) time points. Imaging was acquired on scanners from multiple manufacturers, co-registered to the SRI24 template [306], skull-stripped, and resampled to 1 mm^3 isotropic resolution. Both manual and automatic tumor masks delineating the enhancing tumor, necrosis, and peritumoral edema at each time point are also provided and curated by experts. Additional metadata included patient demographics, molecular markers, and treatment timelines, providing clinical context for imaging analyses in machine learning.

Both of these datasets provide complementary strengths for this work: ReMIND offers high-

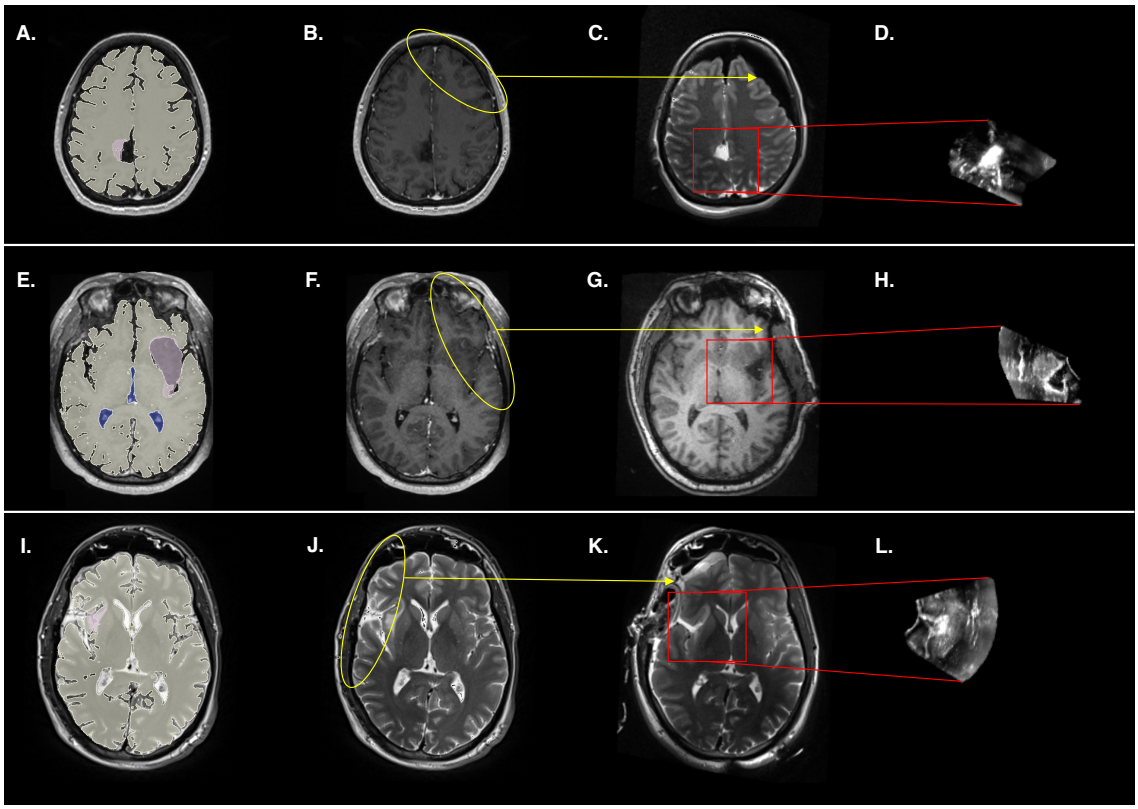


Figure 4.2: Illustrative examples of imaging data drawn from ReMIND are presented for three distinct patients: case #51 (A-D), case #6 (E-H), and case #81 (I-L). (A, E, I) Segmentations are superimposed on the pre-operative MRI, delineating the cerebrum in white, the ventricles in blue, and the whole tumor in pink. (B, F, J) Pre-operative MRI is shown with yellow circles and arrows highlighting regions that suffered notable brain shift during surgery. (C, G, K) Intra-operative MRI is annotated with red boxes specifying the field of view of the intra-operative US. (D, H, L) Intra-operative US shows a focused view of the resection cavity during resection.

quality intra-operative multimodal imaging for studying brain shift and evaluating the simulated surgical deformations against real cases, while UPenn-GBM contributes a large longitudinal cohort for training and evaluating the proposed model across a large and wide range of glioblastoma cases. Their combination enables realistic modeling of intra-operative changes while supporting generalization across institutions and pathologies.

4.3 Synthetic Ground Truth Generation

The goal of the deep interpolator f_θ is to estimate a dense displacement field from a sparse set of M displacement vectors $\{\boldsymbol{d}_i\}_{i=1}^{M}$. In the absence of ground-truth dense displacements in clinical data, the interpolator is trained using synthetic displacements generated through biomechanical simulations. Specifically, the biomechanical framework introduced in [225, 236] was adopted, which simulates brain deformations induced by tumor resection using the meshless total Lagrangian explicit dynamics (MTLED) algorithm [235, 242]. This approach employs a total Lagrangian formulation with explicit time integration to realistically model intra-operative tissue deformation.

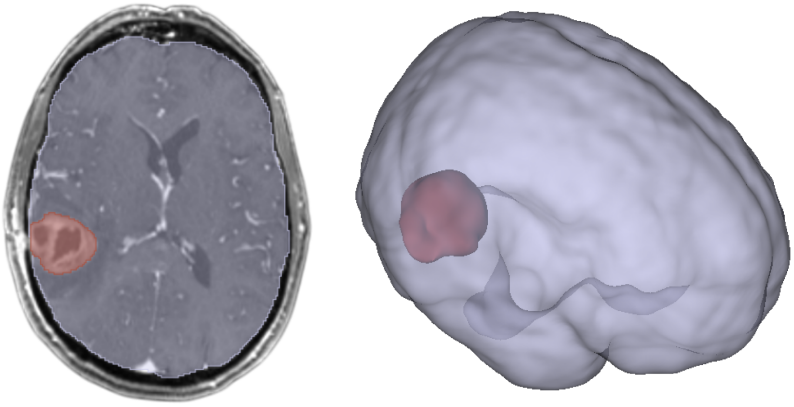


Figure 4.3: Pre-operative brain magnetic resonance imaging (MRI) with tumor segmentation and the corresponding volumetric surface model. (Left) Axial MRI slice showing the segmented brain (gray) and tumor (red). (Right) Volumetric surface model reconstructed from segmentations, used for geometry discretization in the biomechanical modeling pipeline.

4.3.1 Patient-Specific Geometry

The biomechanical framework requires patient-specific brain geometry, including the surfaces of the tumor core and surrounding structures: brain parenchyma, cerebrospinal fluid (CSF), and skull, which must be extracted from the preMRI data. While the segmentation of brain regions in the presence of tumors can be obtained using dedicated frameworks [307, 308], these methods do not delineate the CSF and skull. Instead, SynthSeg [309] was employed on ceT1 images to obtain the required segmentations. Although not specifically designed for pathological data, its outputs include the segmentation of the parenchyma with the surrounding CSF, which allows us to delineate these spaces with sufficient accuracy to approximate both the brain-CSF interface and the skull boundary. These approximations provide the necessary data to define the boundary conditions (BCs) for the biomechanical model. The SynthSeg segmentations were then merged with the tumor core segmentation into a binary labelmap, thus ensuring that complete tumor representations are present even in cases where SynthSeg outputs incomplete or partially missing tumor structures due to the previously mentioned limitations. This labelmap is subsequently converted into volumetric surface (hollow) models using the "Model Maker" module in 3D Slicer, generating the primary geometric input for the biomechanical simulation pipeline (Figure 4.3).

4.3.2 Biomechanical Simulations

Surgical brain deformations in the frame of this thesis are assumed to be primarily driven by gravity and tissue resection [225, 236]. Before craniotomy, the brain is in an unloaded state in which gravitational forces are balanced by the buoyancy of intracranial fluids. Craniotomy disrupts this equilibrium due to pressure release, resulting in gravity-induced brain deformation. Subsequently, tissue resection further amplifies this process by removing structural support within the brain parenchyma, creating a resection cavity that allows surrounding tissue to relax, collapse, or shift toward the cavity under gravity and residual stresses.

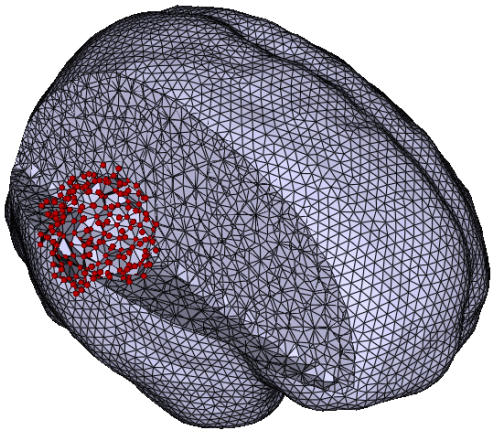


Figure 4.4: Discretized brain geometry with a tetrahedral integration grid. The brain domain is shown in a sliced view to expose the interior, with points associated with the tumor highlighted in red.

The meshless approach uses a cloud of points for spatial discretization and tetrahedral background cells with one integration point per cell, simplifying grid construction compared to traditional finite element (FE) methods. As stated in the previously cited works of Yu *et al.*, these "tetrahedral integration cells are not finite elements and they do not need to meet the stringent quality requirements demanded by a finite element biomechanical brain model". Each brain geometry was composed of 20K–25K nodes. Two biomechanical brain models are constructed: a preresection model that includes the tumor, and a post-resection model in which tumor nodes and their connectivity are removed to simulate the resection cavity (Figure 4.4).

The assumed parameters for the Ogden hyperelastic material model that describe the nonlinear stress-strain behaviour of the brain were used in [236] and are reported in Table 4.1. These parameters, including shear modulus μ, Poisson's ratio ν, material constant α, and material constant $D = \frac{3(1-2\nu)}{\mu(1+\nu)}$ describe physical properties of the brain, and have been previously calculated and applied in the literature [210, 229, 232]. The parenchyma is modeled as nearly incompressible ($\nu = 0.49$), whereas CSF is modeled as highly compressible ($\nu = 0.10$) to reflect fluid drainage dynamics. The tumor is modeled as being stiffer than healthy tissue ($\mu_0 = 842$ Pa), with a shear modulus three times higher ($\mu_0 = 2526$ Pa). A density value of 1000 kg/m^3 was additionally used in all situations.

Table 4.1: Ogden hyperelastic material parameters assigned to brain parenchyma, tumor tissue, and cerebrospinal fluid (CSF) for the biomechanical brain model.

Tissue Type	**Shear Modulus** μ (Pa)	**Poisson's Ratio** ν	**Material Constant** α	**Material Constant** D (Pa^{-1})
Parenchyma	842	0.49	-4.7	4.78×10^{-5}
Tumor	2526	0.49	-4.7	1.59×10^{-5}
CSF	4.54	0.10	2.0	4.81×10^{-1}

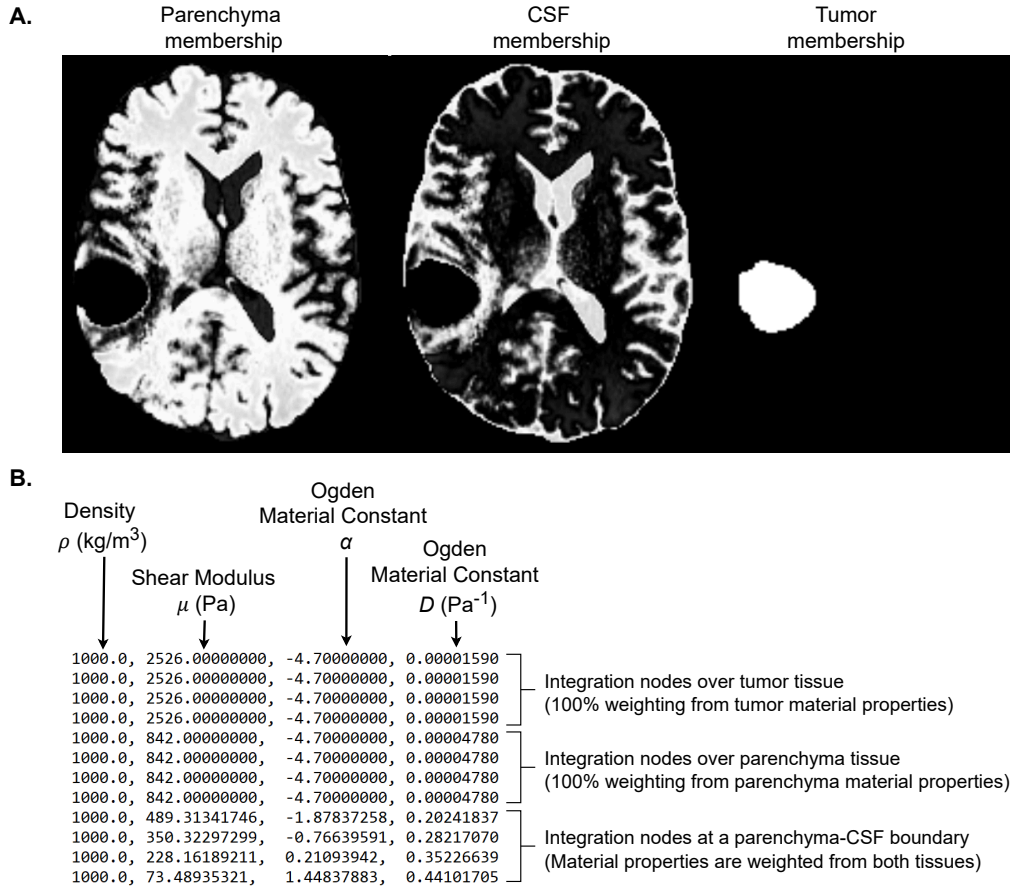


Figure 4.5: Fuzzy tissue classification and material property assignment used in the biomechanical model. (A) Example of fuzzy clustering memberships derived from pre-operative MRI, showing voxel-wise memberships for brain parenchyma, ventricles/CSF, and tumor. (B) Excerpt from the material properties file generated from the clustering results, listing the material properties assigned at integration points. Integration points located within homogeneous tissue regions receive 100% weighting from the corresponding tissue properties, while points near tissue interfaces are assigned weighted properties based on the membership functions.

The material properties are assigned at each integration cell of the brain model using a fuzzy C-means clustering algorithm [310]. This probabilistic labeling avoids the need for explicit segmentations and provides greater flexibility by assigning each integration cell a probability of belonging to each tissue class rather than a fixed label. Three clusters (i.e., tissue classes; $C = 3$) were used for the clustering algorithm, namely the parenchyma (i.e., healthy tissue), ventricles/CSF, and tumor. The material properties MP at each location $j \in \Omega$ of the image domain are then interpolated by computing a weighted sum of the properties of each tissue class C, using the membership functions $\{u_{jk}\}_{k=1}^{C}$ derived from clustering voxel intensity data:

$$MP_j = \sum_{k=1}^{C} u_{jk} \times MP_k, \tag{4.1}$$

i.e., the material property values at each integration node are weighted by the probability of that node being assigned to one of the three tissue types. An example of clustering results, along with

the material properties computed for a sample of integration nodes, is illustrated in Figure 4.5-A and Figure 4.5-B, respectively.

Gravity-induced deformations are then simulated using the MTLED solver by first computing the internal forces at the tumor-parenchyma interface, defined as the set of nodes shared between the tumor and surrounding healthy tissue in the preresection model. These internal reaction forces are subsequently applied in the opposite direction to the post-resection model (without the tumor). Since the applied loading is determined by gravity forces, the solver requires the intra-operative gravity direction vector as input. In contrast to the original framework, where this vector was manually specified by the surgeons, access to the surgical entry point direction was not available. Thus, an automated procedure to estimate potential surgical entry points and derive the corresponding gravity direction vectors was developed.

4.3.3 Surgical Entry Point Estimation

The gravity vector was estimated from the nearest surface point to the tumor center, under the assumption that surgeons typically select the shortest path to the tumor, with the head positioned such that gravity and the access path are aligned to minimize blood loss. The tumor center coordinates are obtained from its segmentation, and the brain surface is extracted using a marching cubes algorithm [311]. The Euclidean distance between each surface voxel and the tumor center is computed, and the closest voxel is designated as the primary entry point (Figure 4.6). Constraints are applied to exclude voxels located under the brain or at angles not physically allowed during surgery.

Recognizing that an entry point at the closest surface point to the tumor may not coincide with the actual surgical entry point, for instance, when alternative trajectories are required to preserve eloquent brain regions, variability was introduced to model differences in patient positioning and

Algorithm 1: Gravity vectors generation procedure.

Input: Binary brain mask B, binary tumor mask T, angle range Δ, number of vectors K
Output: Gravity vectors $\{g_1, \ldots, g_K\}$

$tumor\_center \leftarrow$ mean(coordinates where $T > 0$) ;
$surface \leftarrow$ marching_cubes(B) ;
$entry \leftarrow p \in surface$ with minimum $\|p - tumor\_center\|$, s.t. spatial constraints ;
$main\_vector \leftarrow$ normalize_vector($entry - tumor\_center$) ;

$vectors \leftarrow \{main\_vector\}$;
$(\theta_0, \phi_0) \leftarrow$ get_spherical_angles($main\_vector$) ;
while $|vectors| < K$ **do**
    $(\delta_\theta, \delta_\phi) \leftarrow$ sample_angle_variation($[-\Delta, +\Delta]$) ;
    $(\theta, \phi) \leftarrow (\theta_0 + \delta_\theta, \phi_0 + \delta_\phi)$;
    $new\_vector \leftarrow$ get_cartesian_coordinates(θ, ϕ) ;
    if new_vector *is valid* **then**
        add new_vector to $vectors$;

return $vectors$

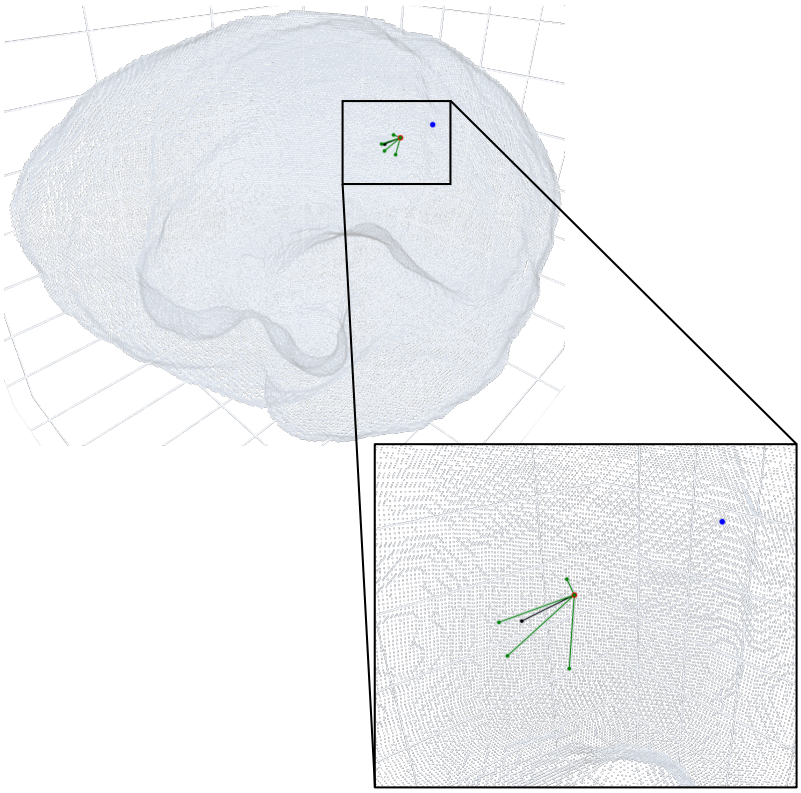


Figure 4.6: Estimation of the gravity vector direction based on tumor location and potential surgical entry points. The tumor center is indicated in red, and the nearest point on the brain surface is shown in blue. The primary gravity vector is computed as the inverted vector connecting the tumor center to the nearest surface point. Random perturbations (green) of this primary vector are generated to account for uncertainty in the actual surgical entry point and head orientation during surgery.

surgical approach. For each case, K plausible gravity vectors are generated by perturbing the base direction by up to $\pm 10^{\circ}$ along each spatial axis, while maintaining compliance with the previously defined constraints and guaranteeing an even spatial distribution of the resulting vectors. See Algorithm 1 for a simplified overview of the algorithm procedure.

As the final step of the biomechanical simulations, the simulated nodal displacements were interpolated into a dense displacement field using a multilevel cubic B-spline via the "Scattered Transform" module [263] in 3D Slicer, producing the synthetic ground truth used for supervised learning . This leads to a dataset $\mathcal{D}_{\text{total}} = \{(\boldsymbol{I}_{\boldsymbol{pre}}^{(n)}, \phi_{\text{gt}}^{(k,n)})_{k=1}^{K}\}_{n=1}^{N}$ containing K distinct displacement fields for each of the N preMRI.

4.4 Synthetic Matched Keypoints Strategy

To simulate the intra-operative acquisition of sparse sets of M displacement vectors $\{\mathbf{d}_i\}_{i=1}^{M}$, which in a real procedure would be obtained by comparing pre-operative and intra-operative images, the following strategy was adopted: (1) keypoints from ceT1 preMRI, $\boldsymbol{I}_{\boldsymbol{pre}}$, were extracted using the widely used 3D SIFT-Rank algorithm [312], and (2) their associated displace-

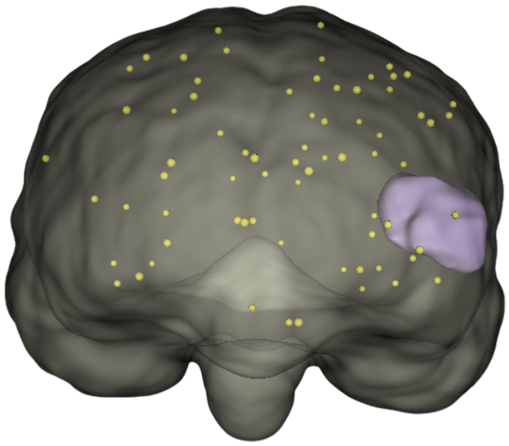


Figure 4.7: Example of the distribution of randomly sampled keypoints within the brain geometry. The segmented tumor is highlighted in purple, and yellow markers indicate the keypoints automatically detected using the 3D SIFT-Rank algorithm.

ment vectors were retrieved using the synthetic ground-truth displacement fields ϕ_{gt}. The 3D Scale-Invariant Feature Transform (SIFT)-Rank algorithm employs a ranking strategy to automatically select the most salient, distinctive, and stable landmarks, which are robust to variations in intensity and structure. Since SIFT typically generates thousands of keypoints, M keypoints from the detected set were randomly sampled (Figure 4.7).

4.5 Neuroimaging Dataset for Physics-Informed Learning

A custom neuroimaging dataset [313] was constructed by combining cases from ReMIND and UPenn-GBM cohorts, augmented with synthetic ground-truth displacement fields generated via patient-specific biomechanical modeling and extracted keypoints via 3D SIFT-Rank (Table 4.2 and Figure 4.8). In total, 394 displacement fields for 207 unique patients were successfully generated: 190 simulations from 45 ReMIND cases ($K = 3$–5 simulations per case) and 204 simulations from 162 UPenn-GBM cases ($K = 1$–3 simulations per case). Each entry in the dataset includes the synthetic displacement fields, unstripped pre-operative ceT1 and T2 images, manually refined anatomical segmentations, and extracted keypoints. This structured resource was used in this work and is tailored for the development and evaluation of learning-based registration methods that aim to integrate biomechanical priors or to generalize under sparse supervision.

The dataset served distinct purposes in the evaluation of the framework. The ReMIND cases were exclusively used to assess the accuracy and plausibility of the biomechanical simulations by comparing pre-operative images warped with the synthetic ground-truth displacement fields with the corresponding intra-operative MRI scans. The ReMIND data was not used for training the deep learning (DL) model because its cases include heterogeneous or missing MRI modalities (i.e., not all subjects have both ceT1 and T2 sequences), which would have required the network to have additional mechanisms to handle this. In contrast, the UPenn-GBM dataset was used for training and evaluating the deep interpolator, as well as for all ablation studies. Although UPenn-GBM

Table 4.2: Descriptive information of the data subsets as they were used in the creation of the custom neuroimaging database employed for evaluation of both the biomechanical simulations and network performance.

	UPenn-GBM [305] (subset)	**ReMIND [285] (subset)**
Original Source	University of Pennsylvania Health System	Brigham and Women's Hospital
No. of Patients (M:F)	162 (74:88)	45 (24:21)
Age (years)	63.1 ± 11.7 (Range: 23–88)	45.9 ± 13.5 (Range: 20–73)
Pathology	HGG (162)	HGG (13), LGG (25), MET (3), NA (4) Previous craniotomy (18)
Tumor volume (cm^3)	18.9 (IQR: 7.8–32.6) (Range: 0.2–129.2)	13.3 (IQR: 6.0–31.5) (Range: 0.1–80.6)
Imaging Modalities	ceT1	ceT1, T2
No. of Biomechanical Simulations	204 (1–3 per patient)	190 (3–5 per patient)
Roles	Primary dataset for supervised learning; Training, validation, and testing of the deep interpolator; Not used for evaluation of biomechanical simulations due to missing iMRI	Evaluation of biomechanical simulations versus real iMRI; Not used for network training due to missing modalities

HGG - High-grade glioma; LGG - Low-grade glioma; MET - Metastasis; NA - Not available; IQR - Interquartile range; ceT1 - Contrast-enhancing T1-weighted; T2 - T2-weighted; iMRI - Intraoperative magnetic resonance imaging

lacks intra-operative scans and therefore cannot be used to validate the biomechanical simulations in the same manner as ReMIND, it provides a sufficiently large and uniform set of pre-operative images and tumor masks suitable for supervised learning. To maintain a consistent and streamlined pipeline, the network was trained solely on ceT1 images, which are routinely available in brain tumor imaging. The full experimental details are presented in Section 5.1 and Section 5.2.

4.6 Deep Biomechanically-guided Interpolator

To design the deep physically-guided interpolator, a denoising approach was taken. Given a sparse set of displacement vectors $\{\boldsymbol{d}_i\}_{i=1}^{M}$, an initial dense displacement field $\phi_{\text{init}} \in \mathbb{R}^{3\times D\times W\times H}$ was computed using a standard interpolation technique such as linear or thin-plate spline (TPS) interpolation. This initial estimate ϕ_{init} is then refined by the model f_θ conditioned on the pre-operative image $\boldsymbol{I_{pre}}$ to approximate the ground-truth displacement field ϕ_{gt}, i.e., $f_\theta(\boldsymbol{I_{pre}}, \phi_{\text{init}}) \approx \phi_{\text{gt}}$, according to the mapping $f_\theta : \left(\mathbb{R}^{D\times W\times H}, \mathbb{R}^{3\times D\times W\times H}\right) \mapsto \mathbb{R}^{3\times D\times W\times H}$.

4.6.1 Residual Network Architecture

Motivated by denoising diffusion models [314], which have shown that learning to predict the noise in a noisy signal leads to better performance than predicting the clean signal directly, the network predicts a residual displacement ϵ_θ, such that $f_\theta(\boldsymbol{I_{pre}}, \phi_{\text{init}}) = \phi_{\text{init}} + \epsilon_\theta(\boldsymbol{I_{pre}}, \phi_{\text{init}})$.

The residual network ϵ_θ is a 3D U-Net architecture variant of [315]. After ablation studies

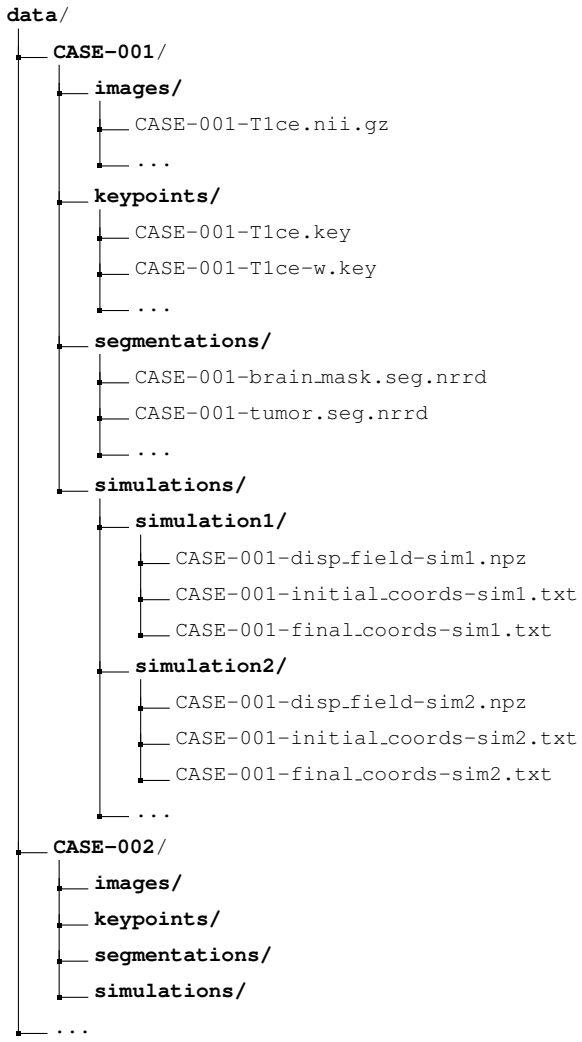


Figure 4.8: Schematic overview of the directory structure of the custom neuroimaging dataset for physics-informed learning. Each patient case contains pre-operative MRI, extracted 3D SIFT keypoints, manually refined anatomical segmentations, and multiple biomechanical simulations providing initial and displaced nodal coordinates and the corresponding displacement fields.

on several network components (see Section 5.2.1), the final architecture is described next. At each resolution level, ResNet blocks [316] are employed with additional spatial- and channel-wise squeeze-and-excitation (SE) [317] modules that encourage feature and spatial recalibration. Spatial SE performs a global pooling step across each channel's spatial dimensions, which results in a vector descriptor with a single scalar per channel. This vector is then passed through fully-connected (FC) layers to generate a set of channel-wise weights used to recalibrate the importance of feature channels. Similarly, a channel-wise SE performs squeezing across each spatial location such that at each location a linear combination of all channels is computed. This projection is then used to recalibrate the importance of each spatial location (Figure 4.9). Downsampling in the encoder path is performed using max-pooling, while upsampling in the decoder path is achieved via transposed convolutions. Same-size feature maps from the encoder are merged with decoder

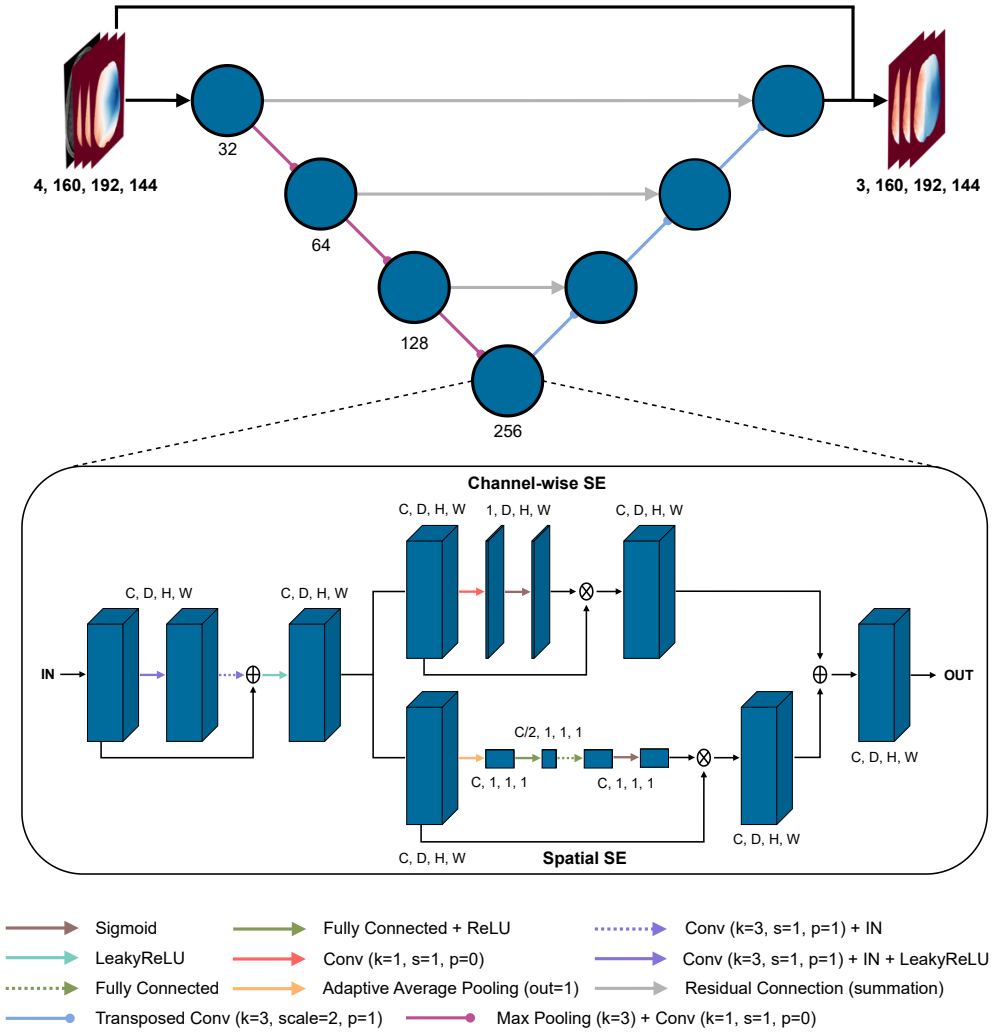


Figure 4.9: Architecture of the proposed deep biomechanically-guided interpolator. The network is a residual 3D U-Net that takes as input the pre-operative MRI and an initial dense displacement field obtained from sparse keypoint interpolation, and outputs a residual displacement field that is added to the input to correct it. The network comprises four resolution levels with 3D convolutional ResNet blocks, instance normalization, and LeakyReLU activation functions, using max-pooling for downsampling and transposed convolutions for upsampling. Skip connections are implemented via element-wise summation, and spatial and channel-wise squeeze-and-excitation modules are incorporated within each residual block. The network predicts a residual displacement field that is added to the initial interpolation, correcting it into biomechanically plausible deformations.

features through element-wise summation rather than concatenation. The network comprises 4 resolution levels, starting with 32 feature channels and doubling at each downsampling stage up to a maximum of 256 channels, while the spatial resolution is halved. Each convolution within the ResNet blocks is followed by instance normalization and a leaky rectified linear unit (LeakyReLU) activation function with a negative slope of 10^{-2}, except for the final convolution in each block, where activation is applied after the residual summation. The SE modules are then applied at the end of each block, resulting in a network with 7.3M parameters.

4.6.2 Training Procedure

The UPenn-GBM portion of the custom dataset was split into 121 training, 16 validation, and 25 test cases, following a 75:10:15 ratio. The deep interpolator f_θ is trained under full supervision using the synthetic ground-truth displacement fields. All inputs were cropped to a fixed size of $160 \times 192 \times 144$, which calls for 2.64 TFLOPs per forward pass at these tensor sizes. Pre-operative ceT1 images were normalized by subtracting the mean and dividing by the standard deviation. Data augmentation was applied only to images during training and included Gaussian noise and blur, intensity adjustments (brightness, contrast, gamma), and simulated low resolution. These augmentations followed the strategies used in nnU-Net [172] to improve generalization capabilities (see Annex A). The network was trained using the Adam optimizer with a fixed learning rate of 1×10^{-4}, a batch size of 1, and for 100 epochs. To ensure anatomical relevance for all methods, interpolated displacements in the background or skull are set to zero.

At each training iteration, a training pre-operative image $\boldsymbol{I_{pre}}$ with its SIFT keypoints and a precomputed ground-truth displacement field ϕ_{gt} from the set of simulations with different gravity-induced brain shifts were sampled. Then, a random set of M sparse displacements $\{\boldsymbol{d}_i\}_{i=1}^{M}$ were sampled and an initial dense displacement field ϕ_{init} was computed on-the-fly using a standard interpolation technique (linear or TPS). The network f_θ is trained to minimize the mean squared error (MSE) between the predicted $\phi_{\text{pred}} = f_\theta(\boldsymbol{I_{pre}}, \phi_{\text{init}})$ and true displacement fields ϕ_{gt} over the image domain. To encourage smooth displacements in non-resected brain regions, an additional Jacobian determinant regularization was introduced to encourage a local orientation consistency constraint on the estimated displacement field. The total loss function $\mathcal{L}$ then becomes:

$$\mathcal{L} \triangleq \frac{1}{|\Omega_{\text{h}}|}\left(\sum_{\boldsymbol{x}\in\Omega_{\text{h}}}\big(\phi_{\text{pred}}(\boldsymbol{x}) - \phi_{\text{gt}}(\boldsymbol{x})\big)^2 + \lambda_{\text{reg}}\sum_{\boldsymbol{x}\in\Omega_{\text{h}}}\text{ReLU}\big(-|\boldsymbol{J}_{I+\phi_{\text{pred}}}|(\boldsymbol{x})\big)\right), \tag{4.2}$$

where λ_{reg} weights the regularization term, Ω_{h} denotes the non-tumorous brain area, and $\boldsymbol{J}_{I+\phi_{\text{pred}}}$ is the Jacobian matrix of the deformation field. The rectified linear unit (ReLU) operation ensures that only negative Jacobian determinants are penalized. Since the biomechanical model computes the displacements accounting for the absence of the tumor (i.e., there are no valid displacements inside the tumor cavity), this region was masked from the loss computation during optimization.

Chapter 5

Experiments and Results

For all experiments, statistically significant (*****) or non-significant (**ns**) results are reported in each graph using either paired t-tests (validation set) or Bonferroni-corrected paired Wilcoxon signed-rank tests (test set) with a significance threshold of $\alpha = 0.01$.

Training and evaluation of the proposed framework were conducted using the University of Pennsylvania Glioblastoma (UPenn-GBM) data, which aggregates standardized imaging data from multiple institutions. Tumor volumes in this subset ranged from 0.2 mL to 129.2 mL, with a median of 18.9 mL (interquartile range (IQR): 7.8–32.6). The Brain Resection Multimodal Imaging Database (ReMIND) data, used exclusively for evaluating the biomechanical simulations, comprises single-institution data with tumor volumes spanning 0.1 mL to 80.6 mL and a median of 13.3 mL (IQR: 6.0–31.5). Overall, ReMIND cases tended to present smaller tumors, whereas UPenn-GBM exhibited a broader distribution with larger volumes (Table 5.1). This is consistent with the metadata, as ReMIND comprises low-grade gliomas, high-grade gliomas, and metastases (which are usually smaller) whereas UPenn-GBM only comprises high-grade gliomas (which are typically larger) [318].

5.1 Intra-operative Brain Shift Simulations

The accuracy of the biomechanical simulations of intra-operative brain deformation was evaluated on ReMIND using both quantitative and qualitative analyses. Quantitative evaluation focused on standard overlap and surface-based metrics that capture the accuracy of boundary alignment, computed for the brain parenchyma and ventricles, including the Dice score, average surface distance (ASD), and the 95th-percentile Hausdorff distance (HD) (Figure 5.1).

Table 5.1: Summary statistics of tumor volumes (in mL) for the datasets used in this study. The mean, standard deviation (Std), minimum (Min), and maximum (Max) values, as well as the first (Q1), second (Median), and third (Q3) quartiles of the tumor volume distributions are reported.

Dataset	**Mean**	**Std**	**Min**	**Q1 (25%)**	**Median (50%)**	**Q3 (75%)**	**Max**
UPenn-GBM	24.1	21.7	0.2	7.8	18.9	32.6	129.2
ReMIND	13.3	20.3	0.1	6.0	13.3	31.5	80.6

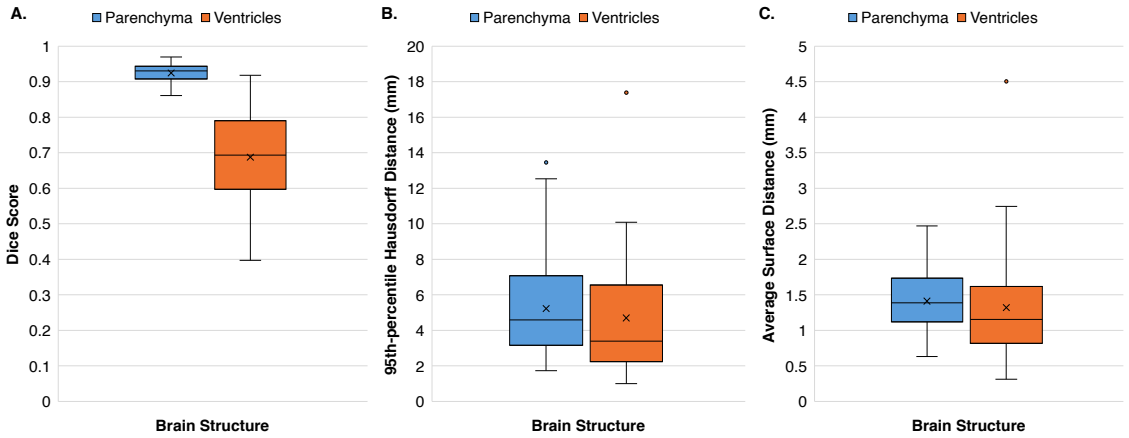


Figure 5.1: Quantitative evaluation of biomechanical brain shift simulations on ReMIND. Dice score (A), average surface distance (B), and 95th-percentile Hausdorff distance (C) are reported for the brain parenchyma (blue) and ventricles (orange). Statistical significance was determined using pair-wise t-tests, with all cases being significant ($p < 0.01$).

To further assess boundary alignment, percentile-based Hausdorff distance curves were computed across all cases. For a given percentile P, the corresponding distance D indicates that $P\%$ of surface voxel pairs have an HD below D [196, 220]. This allows the estimation of the percentage of boundary voxels whose displacements have been computed with clinically acceptable tolerance (Figure 5.2). Given that manual neurosurgery is limited by the image resolution and generally has an accuracy of at best 1.00 mm [319–322], an alignment error smaller than two times the original in-plane resolution of the intra-operative image is difficult to avoid. Thus, for the cases analyzed here, any voxel pairs with an HD value less than 2.00 mm are considered successfully registered.

Across the cases in the evaluation data, simulated deformations demonstrated strong agreement with intra-operative anatomy. For the parenchyma, Dice scores were consistently high (mean 0.92 ± 0.02, IQR: 0.91–0.94), indicating that the simulated post-resection configuration closely matched the observed intra-operative state. In contrast, ventricular alignment was more variable, with a mean Dice of 0.69 ± 0.13 (IQR: 0.60–0.79), consistent with the smaller size and higher sensitivity of ventricular structures to boundary misalignment.

Distance-based metrics revealed similar complementary results. The parenchyma exhibited a mean HD95 of 5.23 ± 2.34 mm, whereas ventricles achieved a slightly lower mean HD95 of 4.70 ± 3.38 mm. This indicates that, despite lower volumetric overlap for the ventricles, the largest surface deviations were not systematically worse than those observed for the parenchyma. Similar patterns were observed for ASD, with mean values of 1.41 ± 0.40 mm for parenchyma and 1.32 ± 0.77 mm for ventricles.

Percentile HD analysis shows a slow increase in the percentage of surface points with an HD up to 2.00 mm, with fewer than 35% having an HD below 1.00 mm, in both parenchyma and ventricles. At the 2.00 mm HD threshold for successful registration, approximately 77% of parenchymal and 73% of ventricular surface points were aligned. Beyond the 90th percentile, a sharp increase in HD shows that points above this area, and especially over the 95th percentile,

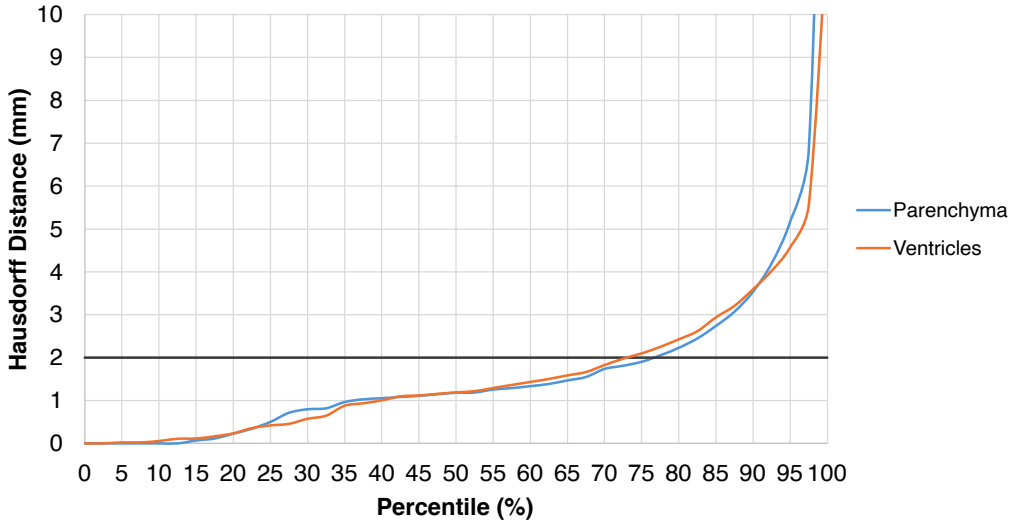
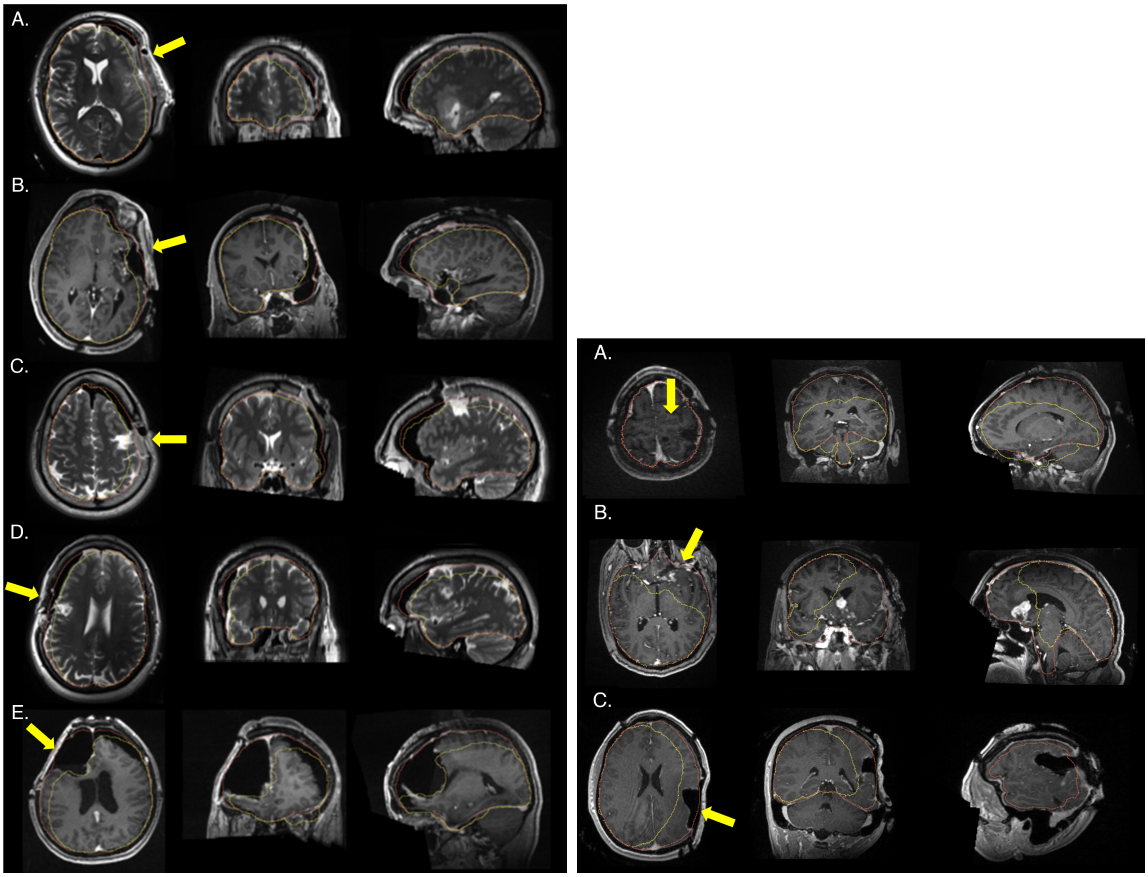


Figure 5.2: Hausdorff distance percentile analysis for biomechanical simulation accuracy. Mean Hausdorff distance is plotted as a function of percentile for parenchyma (blue) and ventricle (orange) surfaces. A threshold line for clinically acceptable alignment is drawn at 2.00 mm.

may correspond to regions with weak or absent anatomical correspondence (e.g., near resection boundaries). Although the parenchyma is more aligned than the ventricles throughout the graph, the most misaligned points (over the 90th percentile) correspond to points in the parenchyma over points in the ventricles, which verifies the results previously found.

Qualitatively, the accuracy of the biomechanical framework was assessed by visualizing the alignment between the simulated and actual intra-operative states (Figure 5.3). The pre-operative brain was warped with the simulated displacement fields (contours in yellow) and overlaid on intra-operative magnetic resonance imaging (iMRI) scans, allowing direct visual comparison between predicted and real intra-operative deformations. For reference, the original pre-operative state (contours in red) was also overlaid to illustrate the extent of brain shift following resection. Visualizations are shown in the axial, coronal, and sagittal planes (in this order) with a focus on the resection cavity, where deformations are most pronounced. Only a small subset of randomly selected cases was evaluated in this manner (Figure 5.3.I), as a detailed qualitative analysis of hundreds of cases would be extremely time-consuming. Additionally, cases with the worst quantitative metrics were specifically selected for analysis (Figure 5.3.II).

Visual comparison of brain contours reveals several characteristic behaviors of the biomechanical simulations. In case I-A, a clear sagging of the brain is observed in the approximate direction of the simulated gravity vector (yellow arrow), as evidenced by the displacement between the pre-operative contours (red) and the intra-operative image. The biomechanical model successfully captures this trend, although a mild overshoot is apparent when comparing the simulated contours (yellow) to the true intra-operative anatomy. Cases I-B and I-D represent the most accurate predictions among the examples, with the simulated post-resection anatomy closely matching the intra-operative state. In these cases, the model effectively reproduces the contraction of tissue



(5.3.I) (5.3.II)

Figure 5.3: Qualitative evaluation of biomechanically simulated intra-operative brain deformations. Pre-operative brain contours (red) and simulated post-resection contours (yellow) are overlaid on intra-operative MRI for axial, coronal, and sagittal views, respectively. Yellow arrows show the gravity direction estimation given to the biomechanical model. (I) Randomly selected cases illustrate typical deformation patterns and gravity-induced sagging near the resection cavity with good predictions. (II) Worst-performing cases highlight limitations related to gravity direction estimation, minimal real brain shift, or recurrent surgeries with pre-existing cavities.

around the resection cavity induced by cranial opening and tumor removal, with only minor overestimation along the superior-inferior axis, as visible in the sagittal views. In contrast, case I-C exhibits a dominant anterior-posterior sagging pattern, visible in the sagittal plane, despite the surgical entry point being located laterally. This suggests that the patient was positioned supine with the head facing upward during surgery, rather than oriented laterally as inferred by the estimated gravity vector. As a result, the gravity direction provided to the biomechanical model does not align with the true physical orientation, leading to the observed discrepancy. Case I-E illustrates a more aggressive resection, with substantial resected tissue near the tumor cavity that is correctly predicted by the model. However, slight errors in the estimated head orientation result in asymmetric sagging, with underestimation on the right side and compensatory overshooting on the left in the axial view, indicating a misaligned gravity component along this axis. Overall, these examples demonstrate that the biomechanical framework captures the dominant deformation patterns

associated with tumor resection, particularly in the vicinity of the cavity, while most residual errors can be attributed to uncertainty in gravity direction estimation due to the lack of intra-operative positional information.

In most of the worst-performing cases, minimal brain shift is observed, as indicated by the close overlap between pre-operative contours (red) and intra-operative anatomy, suggesting that little to no gravity-induced deformation occurred. Two of these cases, II-A and II-B, correspond to recurrent surgeries, in which pre-existing resection cavities are present. These anatomical alterations are not explicitly modeled in the simulation pipeline and likely contribute to the unrealistic and overly aggressive simulated deformations observed in these cases. In case II-C, the simulated deformation is qualitatively more plausible but still substantially overestimates the true displacement. Notably, despite the relatively large resection volume, the intra-operative anatomy exhibits limited brain shift, likely due to patient positioning that restricts gravity-driven sagging. These examples highlight scenarios in which the current biomechanical assumptions break down, particularly when true deformations are minimal or when prior surgical alterations are present.

5.2 Deep Interpolator Ablation Studies

Ablation studies were performed to evaluate the impact of key components of the proposed method: (1) network architecture, (2) regularization term, (3) interpolator for displacement field initialization, and (4) number of matched keypoints sampled for displacement field initialization.

To assess the performance of the deep interpolator, the predicted displacement fields were evaluated using several complementary metrics. The mean squared error (MSE) in mm^2 was computed between the predicted and ground-truth displacement fields, both within the whole brain and specifically within the edematous tumor region, to quantify the overall accuracy and the accuracy near the resected area, respectively. The maximum Euclidean error (Max Error) in mm was additionally reported, capturing the worst-case deviation in the predicted displacements, which is critical for identifying potentially clinically unsafe local misalignments. To assess the geometric alignment of brain structures, the 95th percentile HD between the brain segmentations warped by the predicted and ground-truth displacement fields was computed. To evaluate the anatomical plausibility of the deformations, the percentage of voxels with non-positive Jacobian determinant values ($\%|\boldsymbol{J}_\phi| < 0$) was computed. Finally, the inference time to evaluate the computational efficiency of the proposed method in comparison to the baselines was reported, employing graphics processing unit (GPU) acceleration whenever possible.

All experiments, except when explicitly stated, used a fixed set of 20 keypoints per case for the interpolated displacement field initialization. Each configuration was trained for 100 epochs without regularization, using either thin-plate spline (TPS) or linear interpolation to initialize the displacement field, in order to evaluate performance across the two interpolation strategies most commonly used in sparse keypoint-based registration. Ablation studies for the network design (Section 5.2.1 and Section 5.2.2) were evaluated using the **validation set**, whereas direct comparisons with baseline interpolators (Section 5.2.3) were evaluated using the **test set**.

5.2.1 Network Design

Preliminary Architecture Search

As a starting point in the choice of a network architecture, a U-Net-like architecture with ResNet blocks (Res-UNet), four resolution levels, $\{16, 32, 64, 128\}$ features, and employing residual learning was adopted. Various experiments evaluating the impact of architectural design choices, including the type of squeeze-and-excitation (SE) blocks, network depth, and the number of feature channels per convolutional level were performed. Table 5.2 summarizes the results obtained.

Across both linear and thin-plate spline (TPS) initializations, architectures employing combined spatial and channel-wise SE blocks consistently achieved the best performance overall. Specifically, these models yielded lower whole-brain and edema-region MSE, as well as competitive or lower Max Errors compared to architectures using only spatial SE or only channel-wise SE. With respect to network depth, a four-level architecture outperformed a five-level counterpart across all evaluated metrics in both interpolation strategies. Increasing the depth to five levels resulted in higher errors, indicating an excessive loss of spatial detail due to the additional downsampling steps. Additionally, the number of feature channels used at each convolutional level was revealed to be a balance between global and local accuracy. Increasing the number of features to $\{64, 128, 256, 512\}$ led to statistically significant improvements in whole-brain MSE for both interpolator initializations ($p < 0.01$). However, this configuration did not improve performance within the edematous tumor regions and produced comparable or slightly worse Max Errors

Table 5.2: Architectural search results for various network configurations. Whole-brain and edema-region mean squared error (MSE) and maximum Euclidean error (Max Error) are reported for variations in squeeze-and-excitation (SE) block type, network depth, and number of feature channels per convolutional level. Results are shown for both linear and thin-plate spline (TPS) initializations. The mean and standard deviation are reported. Results in **bold** show the best-performing architecture in each metric per section. Statistical significance was determined using pair-wise t-tests, with * indicating significantly better results ($p < 0.01$) when compared to the chosen architecture (highlighted in green).

Architecture	Linear			TPS		
	MSE (mm^2)↓		Max Error↓ (mm)	MSE (mm^2)↓		Max Error↓ (mm)
	Brain	Edema		Brain	Edema	
SE Blocks						
scSE-4lvl-16	**5.74 (0.14)**	16.77 (0.49)	**58.72 (0.39)**	**10.78 (0.33)**	**17.27 (1.09)**	**67.11 (0.96)**
sSE-4lvl-16	**5.69 (0.11)**	**16.17 (0.65)**	59.67 (0.26)	13.79 (0.39)	**17.27 (1.13)**	72.87 (0.67)
cSE-4lvl-16	6.04 (0.17)	17.54 (0.55)	61.55 (0.34)	11.73 (0.34)	16.06 (0.54)	**65.42 (1.27)**
Depth						
scSE-**4lvl**-16	**5.74 (0.14)**	**16.77 (0.49)**	**58.72 (0.39)**	**10.78 (0.33)**	**17.27 (1.09)**	**67.11 (0.96)**
scSE-**5lvl**-16	6.59 (0.09)	17.60 (0.72)	60.00 (0.30)	14.80 (0.34)	19.02 (0.76)	70.58 (0.73)
# Features						
scSE-4lvl-**16**	5.74 (0.14)	16.77 (0.49)	**58.72 (0.39)**	10.78 (0.33)	17.27 (1.09)	67.11 (0.96)
scSE-4lvl-**32**	5.30 (0.07)	**15.33 (0.94)**	**58.83 (0.73)**	9.18 (0.35)	**15.33 (0.70)**	62.72 (1.37)
scSE-4lvl-**64**	**5.05* (0.10)**	16.86 (0.82)	**59.15 (1.04)**	**8.19* (0.20)**	16.16 (0.69)	**61.01* (0.85)**

SE - squeeze-and-excitation; scSE - spatial and channel-wise SE; sSE - spatial SE; cSE - chanel-wise SE

relative to the $\{32, 64, 128, 256\}$ feature configuration. The latter achieved a much lower edema-region MSE, which are very important localized regions of interest. Finally, the doubling of the number of feature channels also increases the number of parameters and the amount of memory allocated by the model, which turns it very computationally expensive. Thus, for subsequent experiments, a balance between global accuracy, local detail, and model complexity was achieved by selecting a spatial and channel-wise SE model with four levels and a feature progression of $\{32, 64, 128, 256\}$.

Ablation Studies on Main Network Components

To ensure controlled comparisons, all models evaluated in the following ablation studies were configured with the same architectural backbone selected in the last section, comprising four depth levels, feature dimensions of $\{32, 64, 128, 256\}$, instance normalization, and leaky rectified linear unit (LeakyReLU) activation functions. The ablations presented here focus on assessing the impact of three key design choices: (i) replacing standard convolutional blocks with ResNet blocks at each level (UNet vs. Res-UNet); (ii) incorporating SE blocks for adaptive spatial and channel-wise feature weighting (SE vs. non-SE); and (iii) employing residual learning (w/ R vs. w/o R), where the network predicts a correction to the initial interpolated displacement field rather than directly regressing the full ground-truth field. Relevant metric profiles are visualized throughout training using line plots complemented by box-and-whisker plots summarizing mean performance over the validation test in the final 20 epochs, during which the loss often converged (Figures 5.4 and 5.5).

Across all ablation experiments, configurations that did not employ residual learning consistently resulted in inferior initialization, slower optimization, and substantially worse final performance across all evaluated metrics, with differences remaining statistically significant at convergence ($p < 0.01$). Specifically, the standard U-Net architecture struggled to learn meaningful refinements of the displacement field, so it was omitted from the box-and-whisker plots for clarity.

When using TPS for displacement initialization, the Res-UNet-SE architecture (combining ResNet blocks and SE modules) achieved the lowest whole-brain MSE ($p < 0.001$), with a mean value of 10.78 ± 0.33 mm^2 (Figure 5.4-A and Figure 5.4-B). In contrast, performance within the edematous tumor region was comparable across all models employing residual learning, with no statistically significant differences observed ($p > 0.01$) and mean values around 17.15 ± 0.88 mm^2 (Figure 5.4-C). With respect to the Max Error, both Res-UNet variants outperformed other configurations, with Res-UNet achieving just slightly but significantly ($p < 0.01$) better results than Res-UNet-SE, attaining a mean of 59.50 ± 0.32 mm (Figure 5.4-D).

Similar trends were observed for linear interpolation initialization. Res-UNet-SE again yielded the best whole-brain MSE ($p < 0.001$), with a mean of 5.74 ± 0.14 mm^2 (Figures 5.5-A and 5.5-B). As with the TPS initialization, all architectures employing residual learning demonstrated comparable accuracy within the edema region ($p > 0.01$), with mean MSE values of approximately 17.02 ± 0.59 mm^2 (Figure 5.5-C). The Max Error further showed no statistically significant dif-

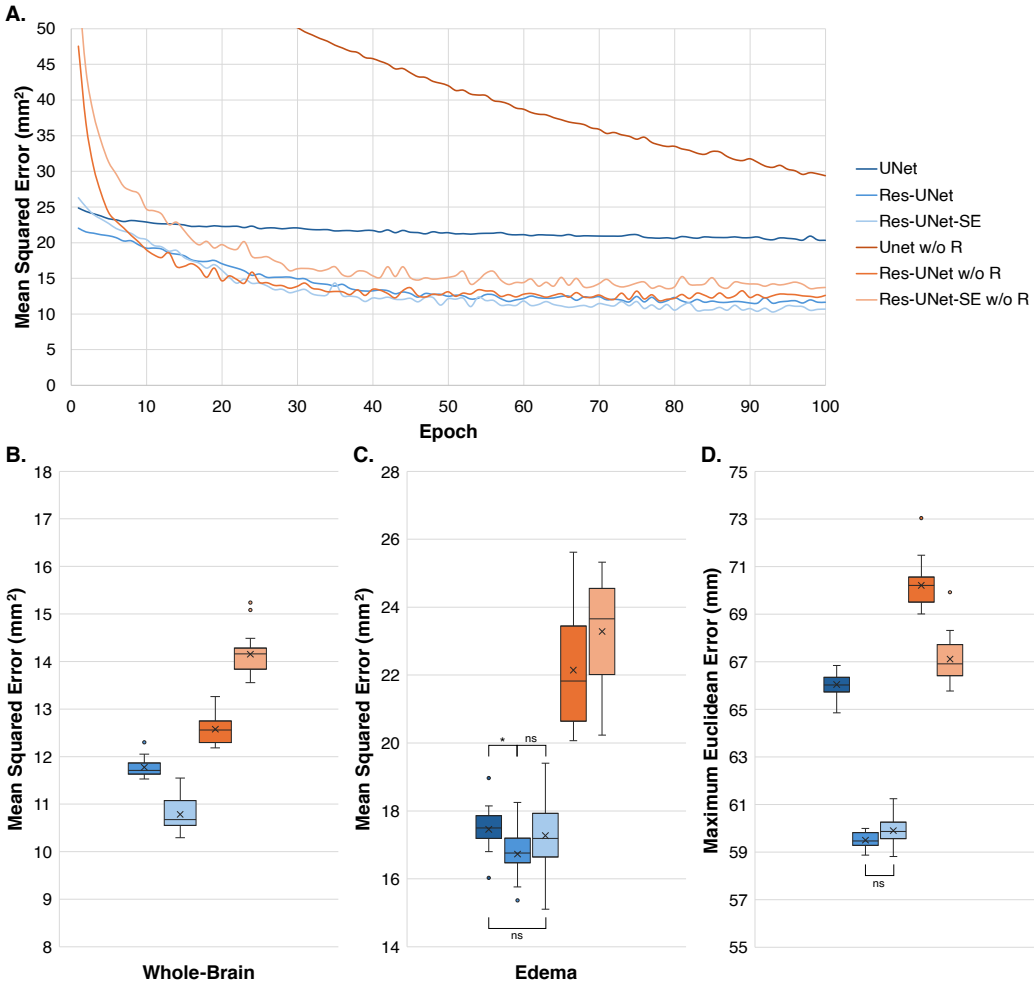


Figure 5.4: Ablation study of network architectures using thin-plate spline initialization. Performance trends are shown for (A,B) whole-brain mean squared error, (C) edema-region mean squared error, and (D) maximum Euclidean error. Statistical significance was determined using pair-wise t-tests, with * indicating significant results ($p < 0.01$) and **ns** indicating non-significant results ($p > 0.01$). Any unmarked comparisons are assumed to be significant or were not considered relevant.

ferences between Res-UNet and Res-UNet-SE ($p > 0.01$), both achieving mean values around 58.77 ± 0.47 mm (Figure 5.5-D).

Overall, Res-UNet architectures incorporating residual learning consistently delivered the strongest performance across all experimental settings. The final architectural choice therefore reduced to Res-UNet variants with or without SE blocks. Considering the better performance in whole-brain MSE achieved by Res-UNet-SE, and the comparable results for other metrics, this architecture was chosen for further studies.

5.2.2 Regularization

The weight coefficient λ_{reg} in Equation 2.1 controls the trade-off between data fidelity and deformation plausibility enforced through the regularization term. To determine an optimal balance, a

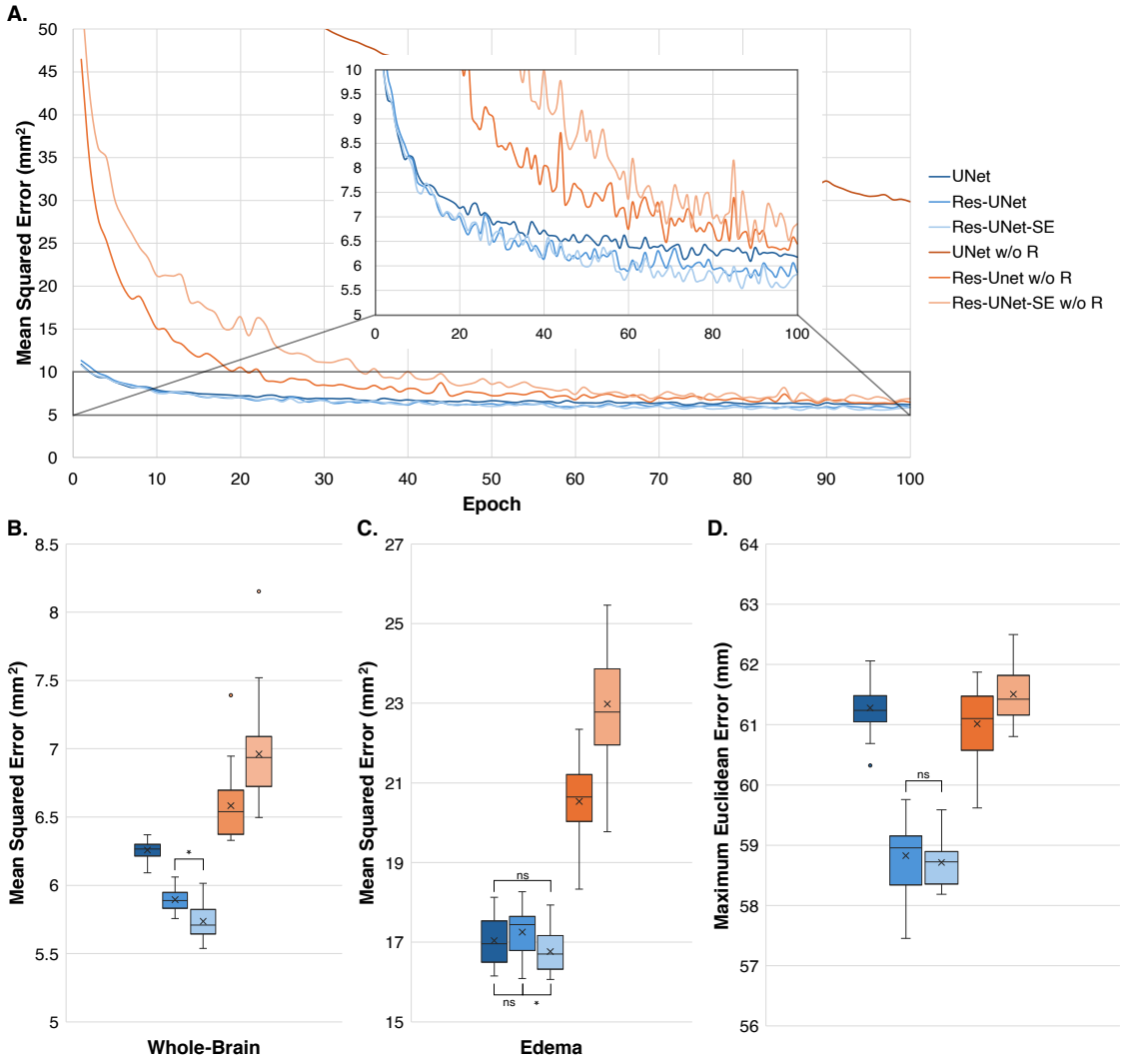


Figure 5.5: Ablation study of the main network architectures using linear interpolation initialization. Performance trends are shown for (A,B) whole-brain mean squared error, (C) edema-region mean squared error, and (D) maximum Euclidean error. Statistical significance was determined using pair-wise t-tests, with * indicating significant results ($p < 0.01$) and **ns** indicating non-significant results ($p > 0.01$). Any unmarked comparisons are assumed to be significant or were not considered relevant.

grid search across $\lambda_{\text{reg}} \in [0, 75]$ was performed.

The results show that $\lambda_{\text{reg}} = 50$ provides the most favorable trade-off between accuracy and regularization. This setting achieved the lowest mean and median MSE among all tested values ($p < 0.01$), with a mean of 5.04 ± 0.12 mm^2 (Figure 5.6-A). At the same time, it maintained a Max Error comparable to other high regularization weights (60.04 ± 0.35; $p > 0.01$), while exhibiting reduced variance, mm(Figure 5.6-B).

Increasing λ_{reg} led to a sharp reduction in the proportion of folded voxels up to $\lambda_{\text{reg}} \approx 30$, beyond which improvements began to become small, indicating diminishing returns for stronger regularization (Figure 5.7). Although the reduction in folding between $\lambda_{\text{reg}} = 50$ and $\lambda_{\text{reg}} =$

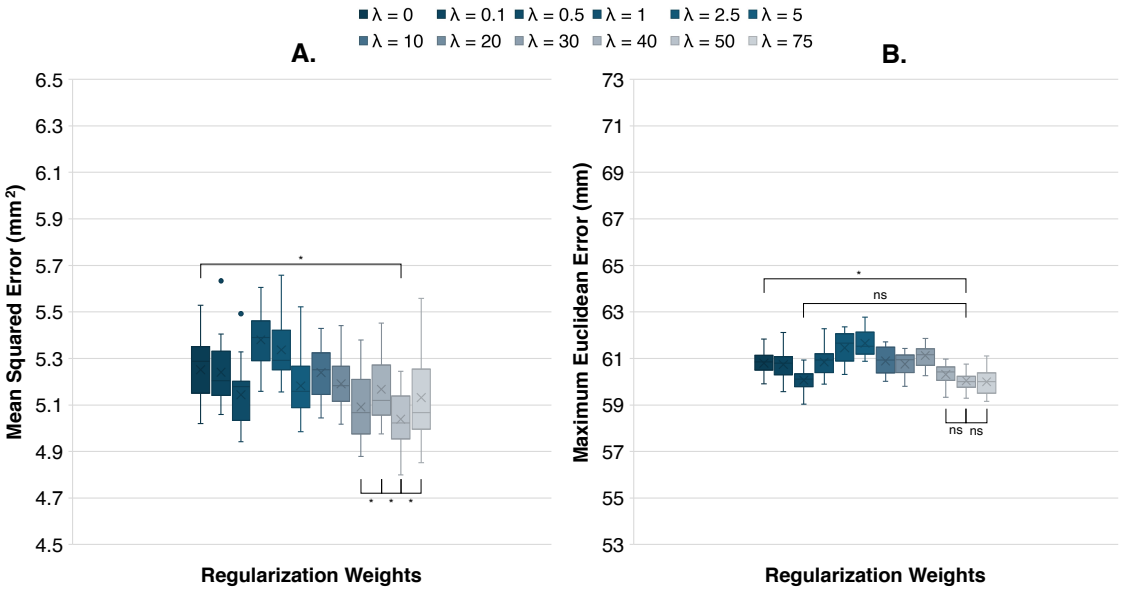


Figure 5.6: Effect of Jacobian regularization strength on interpolation accuracy and stability. (A) Mean squared error and (B) maximum Euclidean error are shown for different values of the regularization weight λ_{reg}. Statistical significance was determined using pair-wise t-tests, with * indicating significant results ($p < 0.01$) and **ns** indicating non-significant results ($p > 0.01$). Any unmarked comparisons are assumed to be significant or were not considered relevant.

75 remained statistically significant ($p < 0.01$), the marginal gains came at the cost of slightly reduced accuracy in terms of whole-brain MSE. Based on this analysis, $\lambda_{\text{reg}} = 50$ was selected for all subsequent experiments, as it achieves a strong compromise between registration accuracy and anatomically plausible, stable deformation fields.

5.2.3 Baseline Comparisons

Two widely-used interpolation methods were used in these experiments for comparison and initialization: (1) linear interpolation via a 3D Delaunay triangulation approach using a publicly available differentiable implementation [76] and (2) a TPS approach using a public implementation [38]. The linear method implementation produces a smooth, piecewise-linear vector field by interpolating values within tetrahedral simplices, while TPS yields a globally smooth field by minimizing the bending energy[1]. Both interpolation methods serve as non-learning-based baselines and rely solely on the spatial distribution and displacement of the sparse keypoints. They do not incorporate anatomical context or physical priors, providing a meaningful comparison to the learning-based framework's refined displacement fields. For clarity during comparisons, the proposed method is referenced as **DBGI** in all subsequent experiments.

The proposed method consistently outperformed baseline interpolators (Table 5.3), reducing whole-brain MSE between the ground truth and the interpolated displacement field significantly by up to 48% (-3.1 mm^2) with TPS and 65% (-7.0 mm^2) with linear interpolation. Improvements in

[1] A regularization weight λ_{tps} of 0.1 was used in the TPS implementation, which achieved the best empirical results on the validation set.

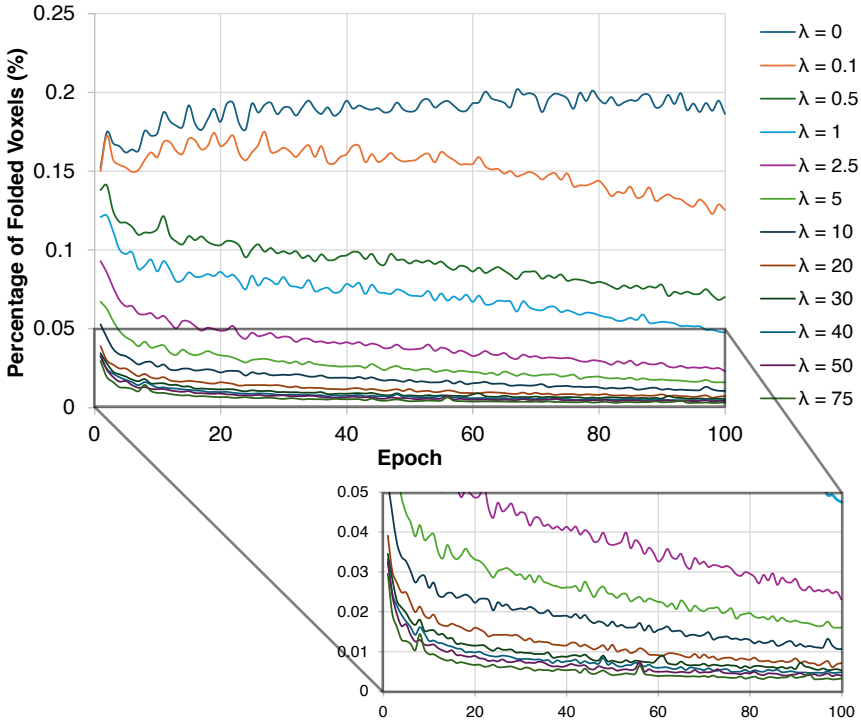


Figure 5.7: Evolution of the percentage of voxels with non-positive Jacobian determinant values across training epochs for different Jacobian regularization weights.

the edema were more modest but with reduced variance, suggesting a more stable approach. The residual architecture enabled the model to learn finer corrections and improved accuracy across the board, especially in the edema regions. Notably, omitting residual learning with TPS initializations led to increased voxel folding (+0.88 pp), which can be attributed to TPS's lack of flexibility to adapt to fine-grained deformations, leading the network to overcompensate. Adding a Jacobian

Table 5.3: Quantitative evaluation of different approaches using linear and thin-plate splines interpolation. R denotes the residual architecture and J the Jacobian regularization term. Mean and standard deviation are reported. Statistical significance was determined using a Bonferroni-corrected paired Wilcoxon signed-rank test, with * indicating statistically significant improvements ($p < 0.01$) for DBGI. Reproduced from [40].

Method	**MSE** (mm^2)↓		**Max Error**↓	**HD95**↓	$\%\lvert J_\phi\rvert < 0$↓	**Time**↓
	Brain	Edema	(mm)	(mm)	(%)	(s)
L (baseline)	10.7*$_{(5.4)}$	7.6 $_{(5.6)}$	28.3 $_{(8.0)}$	3.7*$_{(1.1)}$	0.74*$_{(0.11)}$	1.81 $_{(0.02)}$
DBGI (L)$_{\text{w/o R+J}}$	4.2 $_{(2.2)}$	10.9*$_{(6.6)}$	27.8 $_{(8.5)}$	2.8 $_{(0.7)}$	0.97*$_{(0.20)}$	–
DBGI (L)$_{\text{w/o J}}$	3.7 $_{(1.7)}$	9.4 $_{(7.1)}$	27.3 $_{(8.4)}$	2.8 $_{(0.5)}$	1.16*$_{(0.25)}$	–
DBGI (L)	**3.7** $_{\mathbf{(1.6)}}$	6.4 $_{(3.0)}$	26.2 $_{(8.2)}$	**2.7** $_{\mathbf{(0.6)}}$	**0.64** $_{\mathbf{(0.21)}}$	1.81 $_{(0.02)}$
TPS (baseline)	6.5*$_{(2.6)}$	6.4 $_{(5.5)}$	25.4*$_{(6.9)}$	2.8 $_{(0.7)}$	0.64*$_{(0.21)}$	0.58 $_{(0.01)}$
DBGI (TPS)$_{\text{w/o R+J}}$	4.6*$_{(2.1)}$	10.6*$_{(7.5)}$	26.4 $_{(8.7)}$	2.9 $_{(1.0)}$	1.47*$_{(0.36)}$	–
DBGI (TPS)$_{\text{w/o J}}$	3.5 $_{(1.6)}$	7.2 $_{(5.2)}$	23.4 $_{(5.2)}$	3.3 $_{(0.7)}$	0.99*$_{(0.14)}$	–
DBGI (TPS)	**3.4** $_{\mathbf{(1.6)}}$	5.9 $_{(3.3)}$	**22.7** $_{\mathbf{(4.9)}}$	3.1 $_{(0.5)}$	**0.59** $_{\mathbf{(0.22)}}$	0.59 $_{(0.01)}$

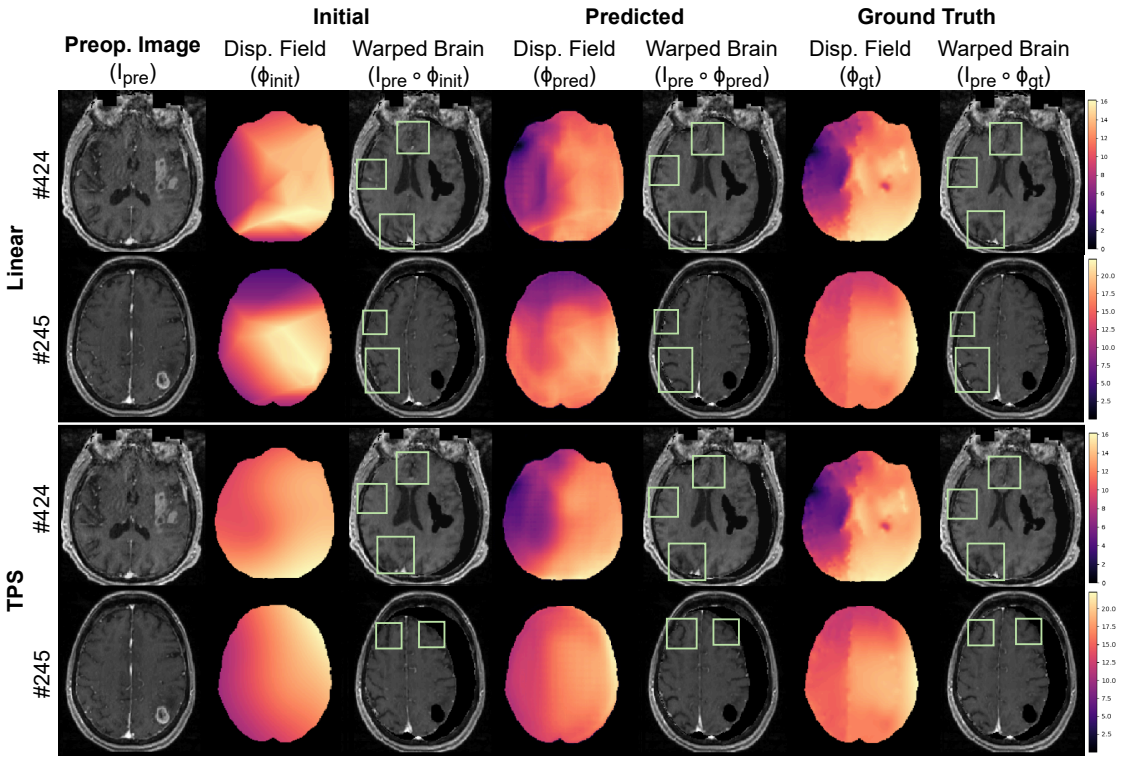


Figure 5.8: Qualitative comparison of displacement field interpolations and corresponding warped brain anatomies. Initial displacement fields obtained via linear and thin-plate spline interpolation are compared with the refined predictions of the proposed deep interpolator and the biomechanical ground truth. Displacement fields are colored by the vector magnitudes in the axial plane, and green squares highlight corresponding regions with the most noticeable improvements over the baselines. Adapted from [40].

regularizer (J) improved deformation smoothness, reducing noninvertible mappings by 40 to 45% (-0.40 pp to -0.52 pp), without compromising accuracy. Inference time increased negligibly compared to baseline methods ($+10$ ms).

Qualitative comparison examples are shown in Figure 5.8 using both linear and TPS initial interpolations. Baseline interpolators produce overly smooth displacements, as seen in the displacement field vector magnitudes: TPS is globally smooth, and linear interpolation uses a piece-wise formulation resulting in piece-wise smooth regions. In contrast, the deep interpolator approaches the biomechanical realism of the biomechanical simulations, predicting, in all cases, finer local deformations that result in better alignment of the brain, notably in the sulci and the longitudinal fissure that separates the brain's hemispheres, and reducing folding artifacts by correcting the initial interpolations.

Impact of Number of Keypoints

Finally, the impact of the number of input keypoints M, varying it from 5 to 50 (Table 5.4), was analyzed. As expected, increasing the number of keypoints led to lower MSE, as the interpolation benefits from more accurate and localized displacement observations. At low keypoint

Table 5.4: Impact of the number of M matched keypoints in terms of mean squared error (mm^2) for linear interpolation (L) and thin-plate splines (TPS). Statistical significance was determined using a Bonferroni-corrected paired Wilcoxon signed-rank test, with all cases being significant ($p < 0.01$) for DBGI. Adapted from [40].

Method	$M = 5$	$M = 10$	$M = 15$	$M = 20$	$M = 50$
L (baseline)	17.2 (9.1)	13.3 (6.5)	11.4 (5.9)	10.7 (5.4)	6.4 (3.1)
DBGI (L)	**7.8 (4.3)**	**4.8 (2.5)**	**4.7 (2.7)**	**3.7 (1.6)**	**2.3 (1.1)**
TPS (baseline)	18.0 (12.4)	10.2 (4.8)	7.2 (2.9)	6.5 (2.6)	3.8 (1.6)
DBGI (TPS)	**11.0 (8.4)**	**5.7 (3.3)**	**4.4 (2.5)**	**3.4 (1.6)**	**2.1 (1.1)**

counts (e.g., 5 keypoints), TPS interpolation performed poorly, likely due to instability with limited control points. In contrast, linear interpolation demonstrated greater robustness in such settings. However, with a higher number of keypoints (e.g., 50), TPS produced smoother and more accurate results, outperforming linear interpolation. This illustrates a trade-off between robustness and smoothness that depends on the spatial density of keypoints. Notably, in all cases, the deep interpolator significantly improved upon the initial interpolation, reducing error regardless of the number of points M and the interpolation method.

Chapter 6

Discussion

The goal of this work was to develop a deep biomechanically-guided interpolator that refines sparse keypoint-derived displacement fields into anatomically and biomechanically plausible deformations for brain shift compensation. The proposed framework does not perform registration directly; rather, it acts as a correction module that can be seamlessly inserted into any keypoint-based pipeline that relies on geometric interpolation. This distinction is critical when situating the contribution relative to related work: while previous methods focus on keypoint detection/matching or on full intensity-based registration, none, to our knowledge, have addressed the weaknesses of the interpolation stage itself, nor have they used biomechanical simulations to supervise a deep refinement model. The results presented in this work directly address this and support the objectives and hypotheses outlined in Section 1.2 of this thesis. (1) The construction of a large-scale dataset of synthetic, biomechanically realistic brain deformations was successfully achieved through patient-specific biomechanical simulation with meshless total Lagrangian explicit dynamics (MTLED), enabling the generation of synthetic ground-truth displacement fields that can rarely be obtained in real surgical scenarios. (2) The strategy for deriving matched keypoints from synthetic deformations proved effective, allowing the model to operate under the sparse-data constraints. (3) The proposed deep biomechanically-guided interpolator fulfilled its intended role as an advanced correction mechanism for classic geometric interpolation. Across all experiments, the deep interpolator reduced whole-brain displacement error by 48%–65% relative to the baseline interpolators across the validation and test sets. The model does not simply memorize simulation-specific patterns but learns how to refine the initial displacements of geometric interpolators in a generalized manner, which is evidenced by the fact that the model improves both thin-plate spline (TPS) and linear initializations even at low keypoint densities (e.g., 5–10 keypoints). The addition of the pre-operative magnetic resonance imaging (preMRI) as an input to the model drives the network to learn a mapping that is conditioned on structural magnetic resonance imaging (MRI) features rather than relying solely on displacement patterns. This real anatomical context is important for the model's ability to correct local errors from geometric interpolators, particularly in regions with complex geometry. It corrects unrealistic local gradients and better aligns finer structures (e.g., near deep sulci or the longitudinal fissure), as illustrated in the qualitative results shown in Figure 5.8. Finally, (4) the model is computationally

efficient at inference, adding negligible overhead to standard interpolators. This meets the real-time requirements of surgical procedures and allows its integration into existing keypoint-based workflows.

The superiority of the method over both TPS and linear interpolation highlights different limitations of these classic approaches: linear interpolation produces piecewise-linear artifacts that cannot capture smooth biomechanical displacement fields, and TPS interpolation produces over-smoothed displacement fields that fail to represent local variations in deformations. Notably, these findings validate the hypothesis that deep learning (DL) can function as a physics-guided interpolator to denoise and biomechanically constrain the field produced from sparse intra-operative information. They also show that the method effectively bridges data-driven learning with biomechanical priors, two components that have long been viewed as somewhat exclusive approaches until the advent of DL and their integration as synergistic tools.

6.1 Biomechanical Simulations

This work relied on synthetic displacement fields generated through a biomechanical simulation pipeline based on the MTLED framework [235, 242]. This simulator models the brain as a non-linear, nearly incompressible soft tissue and accounts for gravity-induced sagging as well as tissue relaxation after tumor resection. The biomechanically simulated deformations provide a strong supervisory signal for the deep interpolator by encoding anatomically plausible deformation patterns governed by continuum mechanics and patient-specific anatomy. This physics-based supervision allows the network to learn refinement behaviors from realistic deformations with known correspondences, implicitly capturing essential biomechanical properties, such as how deformations depend on tumor location and the local topology of the brain, the differing characteristics of healthy untouched tissue versus tumorous tissue, or the collapse patterns that occur around the resection cavity.

Despite this, biomechanical simulations of the brain under tumor resection rely on assumptions about tissue properties, boundary conditions (BCs), and direction and magnitude of gravitational forces that may not be correct, greatly limiting the realism of each simulation. While this problem was tackled by exploring multiple angles of incision via added variability to the angle of each base gravity vector, the simulation may still not fully represent the diversity of patient-specific intra-operative conditions. The head of the patient was assumed to be positioned perpendicular to each entry point/gravity vector; however, variations in craniotomy size, head fixation angle, surgical approach, cerebrospinal fluid (CSF) drainage rate, and other complex resection dynamics limit this approach. Additionally, the biomechanical model assumes idealized material properties derived from the literature [226, 229], which may not reflect pathological or inter-patient variability, as tumor consistency can vary substantially depending on necrosis, edema content, and vascularization [207–209]. As a result, these synthetic deformations, while robust, may not generalize perfectly to all real cases and introduce subtle biases into the learned model. This is the main explanation for most of the deviations of the biomechanical model when compared to the

real intra-operative cases seen in the Brain Resection Multimodal Imaging Database (ReMIND). Furthermore, harsher errors, such as the ones shown in the worst-performing cases in Figure 5.3, can be attributed to any (or all) of these limitations, although several unexplainable convergence errors were also encountered for many cases of both ReMIND and the University of Pennsylvania Glioblastoma (UPenn-GBM) dataset during the generation of the simulated displacement fields with MTLED for the custom dataset (recall Section 4.5). Due to this, only 45 simulated unique cases from ReMIND (out of 114; 39.5%) and 162 from UPenn-GBM (out of 630; 25.7%) were successful. Five displacement fields for each case were expected to be simulated, but only 3–5 simulations per case were successful for ReMIND, and 1–3 simulations per case were successful for UPenn-GBM. Several approaches were tested to find the culprit, such as replacing tumor segmentations from unsuccessful cases with segmentations from successful cases to exclude the possibility of errors coming from bad tumor geometry, greatly smoothing out the initial 3D brain model before discretization to avoid sharp angles, and tinkering with MTLED internal parameters, achieving very inconsistent results with varying success rates.

6.2 Architectural Choices

Detailed architectural analysis revealed that combined spatial and channel-wise squeeze and excitation (SE) blocks consistently outperformed spatial-only and channel-only variants across both initializations. This suggests that deformation refinement benefits from jointly modeling *where* corrections should occur (spatial) and *which* components of the displacement vectors should be emphasized (channel).

Architectural depth was also an important factor. A four-level U-Net outperformed a deeper five-level variant across all evaluated metrics, as additional downsampling degraded performance by limiting the preservation of fine anatomical detail. Additionally, varying the number of feature channels per convolutional level revealed a trade-off between global accuracy and localized performance. Doubling the number of features improved whole-brain mean squared error (MSE) and yielded statistically significant gains in some configurations, indicating increased capacity to model global deformation trends. However, architectures with fewer feature channels achieved better performance within the edematous tumor regions and comparable maximum Euclidean errors (Max Errors), suggesting improved sensitivity to localized, clinically relevant deformations.

The ablation studies on the key network components showed that the residual formulation was crucial, as models trained to directly regress the full displacement field had significantly worse convergence, slower learning, and inferior final performance. This is consistent with findings from denoising diffusion models, where predicting a correction (noise) rather than an absolute target leads to more stable learning [314]. Residual learning aligns well with the intended purpose of the framework, which is refining an initial approximation rather than generating a field from scratch, reducing the burden on the network to learn large global displacements. This is especially visible when TPS is used as initialization because omitting the residual pathway increased voxel folding due to the overly smooth global behavior of TPS interpolations that caused the network to

overcompensate.

Moreover, the inclusion of both ResNet blocks and squeeze-and-excitation (SE) modules resulted in the best whole-brain MSE all around and provided the most stable architecture for the final model. This stems from the residual connections improving gradient flow and enabling deeper non-linear corrections. The SE modules, adaptively weighing channels and spatial locations, are beneficial and an important point when conditioning the model on both the pre-operative image and the interpolated displacement fields. As the model input concatenates both of these, it is important that the model learns the magnitude and behavior of deformations at specific locations in the volume (e.g., the skull does not deform, and regions near the tumor suffer most of the deformations), and what components of each displacement vector require correcting. Interestingly, the SE blocks were strongly dependent on the residual learning formulation. When the network directly predicted the full displacement field, SE blocks degraded performance by reinforcing dominant global patterns already captured by the initial interpolation, while suppressing fine local deformations. When residual learning was used instead, SE blocks improved performance by helping the network focus on complex regions requiring refinement. This suggests that SE mechanisms are better suited for correcting existing deformation estimates, where the feature space already contains useful signals that need to be selectively emphasized, than for estimating full displacement fields, which is a high-variance regression problem.

Finally, the Jacobian-based regularizer had a clear positive effect on anatomical plausibility. Increasing the regularization weight produced a sharp reduction in noninvertible mappings by 40–45%, suggesting that supervision with synthetic biomechanical data alone is insufficient. An explicit regularization term was necessary to constrain gradient behavior during training and prevent deviations to physically implausible deformation patterns.

6.3 Comparison with Standard Interpolators

The final experimental comparisons demonstrate that the proposed deep refinement framework consistently improves upon standard geometric interpolators, namely linear interpolation and thin-plate splines (TPS), across all evaluated metrics. These improvements were observed for both global accuracy, as measured by whole-brain MSE, and localized accuracy in clinically relevant regions such as the edematous tumor area. Importantly, these gains were achieved without increasing inference time, preserving the computational efficiency required for intra-operative use.

Three-dimensional linear interpolation implementations, while robust to sparse and unevenly distributed keypoints, inherently produce piecewise-linear displacement fields that introduce discontinuities in deformation gradients. These artifacts are particularly problematic in regions with complex geometry, where brain tissue suffers heterogeneous deformations. The proposed refinement model effectively reduces these artifacts and enforces anatomically consistent transitions in the displacement field. This is reflected in the substantial reduction in whole-brain displacement error and improved alignment of fine anatomical structures in qualitative evaluations. In contrast, TPS interpolation inherently yields globally smooth deformation fields. While this property is

advantageous for avoiding sharp discontinuities, it also leads to oversmoothing and an inability to represent localized deformation patterns. The proposed method inherently addresses this limitation by predicting refinements that reintroduce local deformation variability supported by the anatomical context of preoperative imaging.

Crucially, these results highlight a previously unaddressed limitation in keypoint-based registration pipelines, where geometric interpolation is typically treated as a final and immutable step. By demonstrating that a learned, simulation-supervised refinement model can bridge the gap between sparse correspondence information and dense, biomechanically plausible deformations, this work establishes that the interpolation stage itself is a critical opportunity for improvement.

6.4 Clinical Implications

Given that the proposed method can rapidly refine a displacement estimate with low inference time and computational overhead, it can be integrated into existing navigation workflows without disruptions. Because these workflows rely on sparse matched keypoints, they avoid the need for computationally expensive intra-operative imaging or high-end neuronavigation systems. This has important implications for global neurosurgery in places that lack resources for advanced imaging, provided that a low-cost modality or manually selected anatomical points can be used to produce corresponding keypoint displacements. However, a major consideration is the reliability of the keypoints used. In a clinical setting, these keypoints can be either automatically detected and matched by upstream automated methods, but their performance and reliability vary substantially. Some degree of surgeon input, either through manual keypoint selection or verification, may still be necessary to ensure that robust correspondences are given as initialization to the model. Additionally, a minimal amount of intra-operative imaging needs to be acquired to calculate the displacements at these keypoints for the initialization of the interpolated displacement field. Real clinical workflows must still accommodate this, and variability in time constraints, tissue exposure, and surgical techniques all may impact the feasibility of obtaining sufficiently informative keypoints. The number of the keypoints and their placement directly influence the stability and accuracy of the refined displacement field. Although the method performed well even under relatively sparse keypoint configurations (see Table 5.4), the variance in the observed results demonstrates that keypoints that are poorly placed, mismatched, or too clustered may propagate misleading trends into the initial field, which the refinement model may only partially correct.

A further practical implication is interpretability. Unlike intensity-based registration, keypoint-based methods allow for direct visualization of what is driving the registration process, and the interpolation refinement is easier to understand as a correction relative to those keypoints. This improves user trust and clinical validation in situations where surgeons rely on these deformation estimates to guide resection margins or determine the location of functional brain structures.

6.5 Contextualization within Related Work

While the proposed method does not perform registration directly, its contribution sits at a critical point in keypoint-based registration pipelines by enhancing the interpolation step with a physics-guided refinement mechanism. This situates it in a unique spot, because current keypoint-based methods, whether they rely on Scale-Invariant Feature Transform (SIFT) keypoints and descriptors or deep feature extractors, share a common dependency on geometric interpolators to convert sparse matches into dense displacement fields. Yet, these interpolators fail to represent complex, heterogeneous brain shift. This may be a reason why current DL-based registration methods rely on instance optimization (IO) post-processing steps (recall Section 3.1.7) to achieve state-of-the-art performance, since commonly employed methods disregard finer biomechanical behaviors of the brain under craniotomy, tissue resection, and multimodal registration. On the other hand, biomechanical modeling of the brain offers physical realism but remains too slow for intra-operative integration in surgical workflows for continuous image updating.

The proposed approach thus establishes a middle ground that is currently missing in the literature: a biomechanically supervised refinement module that can be inserted directly into any sparse keypoint-based method without modifying the upstream pipeline, providing biomechanically informed deformations in real time.

Chapter 7

Conclusion

This thesis presented a physics-guided deep learning (DL) framework for refining sparse keypoint-derived displacement fields into dense, biomechanically plausible deformations for brain shift compensation. Motivated by the limitations of classical geometric interpolators, which fail to capture the true mechanical behavior of the brain during tumor resection, the proposed method addressed a critical yet underexplored component of keypoint-based registration pipelines. Rather than proposing a full registration solution, this work focused on improving the interpolation stage itself, positioning the framework as a modular refinement block that can be seamlessly integrated into existing sparse correspondence workflows. This design choice enables compatibility with a wide range of upstream keypoint extraction and matching strategies while remaining suitable for intra-operative integration.

This work combined patient-specific anatomical information from pre-operative magnetic resonance imaging (preMRI) and synthetic displacement fields generated through a biomechanical simulation pipeline. Biomechanical simulations served as physically grounded supervision, enabling the network to learn deformation patterns that are difficult or impossible to observe directly in real surgical settings. Through extensive experiments, the proposed deep interpolator demonstrated significant improvements over classical interpolators in accuracy, anatomical fidelity, and deformation plausibility. These gains were achieved with negligible computational overhead, maintaining inference times compatible with real-time neuronavigation. Overall, these results validate the central hypothesis of the thesis that a physics-guided DL model can correct and constrain the outputs of geometric interpolators, producing displacement fields that more accurately reflect the biomechanics of the brain.

This thesis provides several novel contributions to the field of neurosurgical image registration and brain shift correction that together establish a new direction for enhancing keypoint-based registration pipelines by integrating biomechanical realism into the interpolation step:

- A deep, biomechanically guided interpolation module that refines classic geometric interpolators using physics-driven supervision. This module is explicitly designed to operate downstream of keypoint matching, making it agnostic to the choice of keypoint detector or descriptor and broadly applicable across pipelines. To the best of our knowledge, this is

the first work to directly learn biomechanically informed corrections of sparse-interpolation displacement fields.

- A large-scale dataset of patient-specific synthetic brain deformations generated using the meshless total Lagrangian explicit dynamics (MTLED) framework. These simulations model tissue relaxation following tumor resection and gravity-induced brain shift, capturing realistic deformations and providing ground-truth displacement fields that are otherwise unobtainable in real surgeries. This dataset effectively enables learning-based methods to develop an understanding of the biomechanical behaviors of brain deformations in neurosurgery.
- A computationally lightweight refinement strategy that introduces negligible latency relative to standard interpolation methods, thus meeting the practical constraints of real-time intra-operative integration. This efficiency at inference time, combined with sparse keypoint input requirements, makes the method particularly suitable for surgical environments with limited intra-operative imaging resources.
- Extensive ablation studies that reveal the contribution of each architectural component (residual learning, squeeze-and-excitation (SE) blocks, Jacobian regularization) to achieve precise, stable, and realistic displacement refinements.

Although the proposed framework shows strong performance and practical potential, several limitations remain and motivate future research. The model relies on sparse keypoints and the preMRI. In real surgeries, keypoints extracted intra-operatively, usually through ultrasound, may be noisy or mismatched due to artifacts and a limited field of view. While the refinement model showed robustness across a wide range of keypoint counts and distributions, extreme errors in the initial interpolation can still bias the refined field. This motivates future extensions incorporating uncertainty-aware keypoints, explicit outlier rejection, or probabilistic correspondence modeling. An end-to-end framework that jointly performs keypoint detection and matching, and refines displacement fields, may further improve the robustness of this registration task.

Additionally, while the Jacobian regularization promotes anatomical plausibility by reducing the number of noninvertible mappings, diffeomorphism is not strictly enforced. Complex deformations, such as those near the tumor cavity and skull boundaries, may still require stronger constraints. As diffeomorphic registration is preferred in clinical research because it guarantees both inverse consistency (IC) and topology preservation, it should be explicitly enforced into future iterations of this method for clinical robustness.

Regarding generalization, the model was trained on the University of Pennsylvania Glioblastoma (UPenn-GBM) dataset, which only includes high-grade gliomas. Other pathologies, such as low-grade gliomas and meningiomas, exhibit distinct deformation behaviors due to differences in tissue stiffness, mass effect, edema patterns, and resection dynamics, which are not represented in the current training distribution. In addition, training was performed using only contrast-enhanced T1-weighted (ceT1) magnetic resonance imaging (MRI) scans, chosen for their consistent availability in brain tumor cases. While this simplifies deployment, reliance on a single modality

limits exposure to complementary anatomical and textural cues present in other sequences (e.g., T2-weighted (T2) or fluid-attenuated inversion recovery (FLAIR)), likely hurting generalization. Moreover, the biomechanical simulations used for supervision rely on assumed material properties, simplified boundary conditions (BCs), and gravity-based modeling. These assumptions may not fully capture the diversity of real surgical scenarios involving cerebrospinal fluid (CSF) leakage, dynamic tissue collapse, or surgical tool interactions. As a result, both the pathology distribution and the simulation assumptions limit the extent to which the model can generalize to out-of-distribution cases. Broader training and validation across diverse tumor types, such as in the data found in the Brain Resection Multimodal Imaging Database (ReMIND), and the development of more expressive biomechanical simulations will be essential to ensure applicability across a wider range of neurosurgical cases.

Finally, a natural next step is to integrate the proposed interpolator into a complete intra-operative registration pipeline capable of continuous updates throughout surgery. Such integration would allow refined deformation estimates to directly inform surgical decision-making in real time, enhancing the reliability and interpretability of neuronavigation systems. Ultimately, this work contributes to a key building block toward enabling more precise, interpretable, and safer image-guided neurosurgery.

Bibliography

[1] R. L. Siegel, T. B. Kratzer, A. N. Giaquinto, H. Sung, and A. Jemal, "Cancer statistics, 2025," *CA: A Cancer Journal for Clinicians*, vol. 75, no. 1, pp. 10–45, 2025.

[2] M. Price, C. Ballard, J. Benedetti, C. Neff, G. Cioffi, K. A. Waite, C. Kruchko, J. S. Barnholtz-Sloan, and Q. T. Ostrom, "CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2017–2021," *Neuro-Oncology*, vol. 26, no. Suppl 6, p. vi1, 2024.

[3] Z. K. Asfaw, T. Young, C. Brown, and I. M. Germano, "Charting the success of neuronavigation in brain tumor surgery: from inception to adoption and evolution," *Journal of Neuro-Oncology*, vol. 170, no. 1, pp. 1–10, 2024.

[4] M. Lara-Velazquez, R. Al-Kharboosh, S. Jeanneret, C. Vazquez-Ramos, D. Mahato, D. Tavanaiepour, G. Rahmathulla, and A. Quinones-Hinojosa, "Advances in brain tumor surgery for glioblastoma in adults," *Brain Sciences*, vol. 7, no. 12, p. 166, 2017.

[5] L. Wang, B. Liang, Y. I. Li, X. Liu, J. Huang, and Y. M. Li, "What is the advance of extent of resection in glioblastoma surgical treatment—a systematic review," *Chinese Neurosurgical Journal*, vol. 5, no. 02, pp. 102–107, 2019.

[6] B. Sommer, C. Wimmer, R. Coras, I. Blumcke, B. Lorber, H. M. Hamer, H. Stefan, M. Buchfelder, and K. Roessler, "Resection of cerebral gangliogliomas causing drug-resistant epilepsy: short-and long-term outcomes using intraoperative MRI and neuronavigation," *Neurosurgical Focus*, vol. 38, no. 1, p. E5, 2015.

[7] M. Lacroix, D. Abi-Said, D. R. Fourney, Z. L. Gokaslan, W. Shi, F. DeMonte, F. F. Lang, I. E. McCutcheon, S. J. Hassenbusch, E. Holland, *et al.*, "A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival," *Journal of Neurosurgery*, vol. 95, no. 2, pp. 190–198, 2001.

[8] E. A. Spiegel, H. T. Wycis, M. Marks, and A. J. Lee, "Stereotaxic apparatus for operations on the human brain," *Science*, vol. 106, no. 2754, pp. 349–350, 1947.

[9] D. W. Roberts, J. W. Strohbehn, J. F. Hatch, W. Murray, and H. Kettenberger, "A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope," *Journal of Neurosurgery*, vol. 65, no. 4, pp. 545–549, 1986.

[10] D. Grönemeyer, R. Seibel, R. Erbel, A. Schmidt, A. Melzer, J. Plaßmann, M. Deli, A. Schmermund, D. Baumgart, R. Welsch, *et al.*, "Equipment configuration and procedures: preferences for interventional micro therapy," *Journal of Digital Imaging*, vol. 9, no. 2, pp. 81–96, 1996.

[11] D. A. Orringer, A. Golby, and F. Jolesz, "Neuronavigation in the surgical management of brain tumors: current and future trends," *Expert Review of Medical Devices*, vol. 9, no. 5, pp. 491–500, 2012.

[12] T. Noh, M. Mustroph, and A. J. Golby, "Intraoperative imaging for high-grade glioma surgery," *Neurosurgery Clinics of North America*, vol. 32, no. 1, p. 47, 2020.

[13] P. Juvekar, E. Torio, W. L. Bi, D. C. D. A. Bastos, A. J. Golby, and S. F. Frisken, "Mapping resection progress by tool-tip tracking during brain tumor surgery for real-time estimation of residual tumor," *Cancers*, vol. 15, no. 3, p. 825, 2023.

[14] I. J. Gerard, M. Kersten-Oertel, K. Petrecca, D. Sirhan, J. A. Hall, and D. L. Collins, "Brain shift in neuronavigation of brain tumors: A review," *Medical Image Analysis*, vol. 35, pp. 403–420, 2017.

[15] I. Reinertsen, F. Lindseth, C. Askeland, D. H. Iversen, and G. Unsgård, "Intra-operative correction of brain-shift," *Acta Neurochirurgica*, vol. 156, no. 7, pp. 1301–1310, 2014.

[16] L. H. Stieglitz, J. Fichtner, R. Andres, P. Schucht, A.-K. Krähenbühl, A. Raabe, and J. Beck, "The silent loss of neuronavigation accuracy: a systematic retrospective analysis of factors influencing the mismatch of frameless stereotactic systems in cranial neurosurgery," *Neurosurgery*, vol. 72, no. 5, pp. 796–807, 2013.

[17] I. J. Gerard, M. Kersten-Oertel, J. A. Hall, D. Sirhan, and D. L. Collins, "Brain shift in neuronavigation of brain tumors: an updated review of intra-operative ultrasound applications," *Frontiers in Oncology*, vol. 10, p. 618837, 2021.

[18] R. J. Benveniste and I. M. Germano, "Correlation of factors predicting intraoperative brain shift with successful resection of malignant brain tumors using image-guided techniques," *Surgical Neurology*, vol. 63, no. 6, pp. 542–548, 2005.

[19] A. Nabavi, P. M. Black, D. T. Gering, C.-F. Westin, V. Mehta, R. S. Pergolizzi Jr, M. Ferrant, S. K. Warfield, N. Hata, R. B. Schwartz, *et al.*, "Serial intraoperative magnetic resonance imaging of brain shift," *Neurosurgery*, vol. 48, no. 4, pp. 787–798, 2001.

[20] W. J. Elias, K.-M. Fu, and R. C. Frysinger, "Cortical and subcortical brain shift during stereotactic procedures," *Journal of Neurosurgery*, vol. 107, no. 5, pp. 983–988, 2007.

[21] S. Ohue, Y. Kumon, S. Nagato, S. Kohno, H. Harada, K. Nakagawa, K. Kikuchi, H. Miki, and T. Ohnishi, "Evaluation of intraoperative brain shift using an ultrasound-linked navigation system for brain tumor surgery," *Neurologia Medico-Chirurgica*, vol. 50, no. 4, pp. 291–300, 2010.

[22] Y. Peng, X. Liu, A. Wang, and R. Han, "The effect of mannitol on intraoperative brain relaxation in patients undergoing supratentorial tumor surgery: study protocol for a randomized controlled trial," *Trials*, vol. 15, no. 1, p. 165, 2014.

[23] H. A. Shenkin, B. Goluboff, and H. Haft, "The use of mannitol for the reduction of intracranial pressure in intracranial surgery," *Journal of Neurosurgery*, vol. 19, no. 10, pp. 897–901, 1962.

[24] M. Matsumae, J. Nishiyama, and K. Kuroda, "Intraoperative MR imaging during glioma resection," *Magnetic Resonance in Medical Sciences*, vol. 21, no. 1, pp. 148–167, 2022.

[25] P. Zaffino, S. Moccia, E. De Momi, and M. F. Spadea, "A review on advances in intraoperative imaging for surgery and therapy: imagining the operating room of the future," *Annals of Biomedical Engineering*, vol. 48, no. 8, pp. 2171–2191, 2020.

[26] D. C. D. A. Bastos, P. Juvekar, Y. Tie, N. Jowkar, S. Pieper, W. M. Wells, W. L. Bi, A. Golby, S. Frisken, and T. Kapur, "Challenges and opportunities of intraoperative 3D ultrasound with neuronavigation in relation to intraoperative MRI," *Frontiers in Oncology*, vol. 11, p. 656519, 2021.

[27] L. Dixon, A. Lim, M. Grech-Sollars, D. Nandi, and S. Camp, "Intraoperative ultrasound in brain tumor surgery: A review and implementation guide," *Neurosurgical Review*, vol. 45, no. 4, pp. 2503–2515, 2022.

[28] T. Selbekk, A. S. Jakola, O. Solheim, T. F. Johansen, F. Lindseth, I. Reinertsen, and G. Unsgård, "Ultrasound imaging in neurosurgery: approaches to minimize surgically induced image artefacts for improved resection control," *Acta Neurochirurgica*, vol. 155, no. 6, pp. 973–980, 2013.

[29] B. D. De Vos, F. F. Berendsen, M. A. Viergever, M. Staring, and I. Išgum, "End-to-end unsupervised deformable image registration with a convolutional neural network," in *International Workshop on Deep Learning in Medical Image Analysis*, pp. 204–212, Springer, 2017.

[30] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: a learning framework for deformable medical image registration," *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1788–1800, 2019.

[31] T. C. Mok and A. C. Chung, "Large deformation diffeomorphic image registration with laplacian pyramid networks," in *MICCAI 2020*, pp. 211–221, Springer, 2020.

[32] T. C. Mok and A. C. Chung, "Unsupervised deformable image registration with absent correspondences in pre-operative and post-recurrence brain tumor MRI scans," in *MICCAI 2022*, pp. 25–35, Springer, 2022.

[33] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain," *Medical Image Analysis*, vol. 12, no. 1, pp. 26–41, 2008.

[34] Y. Ou, A. Sotiras, N. Paragios, and C. Davatzikos, "DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting," *Medical Image Analysis*, vol. 15, no. 4, pp. 622–639, 2011.

[35] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, pp. S61–S72, 2009.

[36] B. Rister, M. A. Horowitz, and D. L. Rubin, "Volumetric image registration from invariant keypoints," *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4900–4910, 2017.

[37] M. P. Heinrich and L. Hansen, "Voxelmorph++: going beyond the cranial vault with keypoint supervision and multi-channel instance optimisation," in *International Workshop on Biomedical Image Registration*, pp. 85–95, Springer, 2022.

[38] A. Q. Wang, M. Y. Evan, A. V. Dalca, and M. R. Sabuncu, "A robust and interpretable deep learning framework for multi-modal registration via keypoints," *Medical Image Analysis*, vol. 90, p. 102962, 2023.

[39] H. Rasheed, R. Dorent, M. Fehrentz, T. Kapur, W. M. Wells III, A. Golby, S. Frisken, J. A. Schnabel, and N. Haouchine, "Learning to match 2D keypoints across preoperative MR and intraoperative ultrasoun," in *International Workshop on Advances in Simplifying Medical Ultrasound*, pp. 78–87, Springer, 2024.

[40] T. Assis, I. P. Machado, B. Zwick, N. C. Garcia, and R. Dorent, "Deep biomechanically-guided interpolation for keypoint-based brain shift registration," in *Collaborative Intelligence and Autonomy in Image-Guided Surgery* (Q. Dou, Y. Ban, Y. Jin, S. Bano, and M. Unberath, eds.), pp. 105–115, Springer Cham, 2026.

[41] J. Brunt, "Computed tomography–magnetic resonance image registration in radiotherapy treatment planning," *Clinical Oncology*, vol. 22, no. 8, pp. 688–697, 2010.

[42] K. K. Brock, S. Mutic, T. R. McNutt, H. Li, and M. L. Kessler, "Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132," *Medical Physics*, vol. 44, no. 7, pp. e43–e76, 2017.

[43] S. Mutic, J. F. Dempsey, W. R. Bosch, D. A. Low, R. E. Drzymala, K. C. Chao, S. M. Goddu, P. D. Cutler, and J. A. Purdy, "Multimodality image registration quality assurance for conformal three-dimensional treatment planning," *International Journal of Radiation Oncology, Biology, Physics*, vol. 51, no. 1, pp. 255–260, 2001.

[44] V. Fortunati, R. F. Verhaart, F. Angeloni, A. van der Lugt, W. J. Niessen, J. F. Veenland, M. M. Paulides, and T. van Walsum, "Feasibility of multimodal deformable registration for head and neck tumor treatment planning," *International Journal of Radiation Oncology, Biology, Physics*, vol. 90, no. 1, pp. 85–93, 2014.

[45] M. C. Yip, D. G. Lowe, S. E. Salcudean, R. N. Rohling, and C. Y. Nguan, "Tissue tracking and registration for image-guided surgery," *IEEE Transactions on Medical Imaging*, vol. 31, no. 11, pp. 2169–2182, 2012.

[46] N. Gerber, K. A. Gavaghan, B. J. Bell, T. M. Williamson, C. Weisstanner, M.-D. Caversaccio, and S. Weber, "High-accuracy patient-to-image registration for the facilitation of image-guided robotic microsurgery on the head," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 4, pp. 960–968, 2013.

[47] F. Alam, S. U. Rahman, S. Ullah, and K. Gulati, "Medical image registration in image guided surgery: Issues, challenges and research opportunities," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 1, pp. 71–89, 2018.

[48] E. Mazzucchi, G. La Rocca, P. Hiepe, F. Pignotti, G. Galieri, D. Policicchio, R. Boccaletti, P. Rinaldi, S. Gaudino, T. Ius, *et al.*, "Intraoperative integration of multimodal imaging to improve neuronavigation: a technical note," *World Neurosurgery*, vol. 164, pp. 330–340, 2022.

[49] S. K. Saha, D. Xiao, A. Bhuiyan, T. Y. Wong, and Y. Kanagasingam, "Color fundus image registration techniques and applications for automated analysis of diabetic retinopathy progression: A review," *Biomedical Signal Processing and Control*, vol. 47, pp. 288–302, 2019.

[50] T. De Silva, E. Y. Chew, N. Hotaling, and C. A. Cukras, "Deep-learning based multi-modal retinal image registration for the longitudinal analysis of patients with age-related macular degeneration," *Biomedical Optics Express*, vol. 12, no. 1, pp. 619–636, 2020.

[51] M. Tan, Z. Li, Y. Qiu, S. D. McMeekin, T. C. Thai, K. Ding, K. N. Moore, H. Liu, and B. Zheng, "A new approach to evaluate drug treatment response of ovarian cancer patients based on deformable image registration," *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 316–325, 2015.

[52] Y.-M. Lin, I. Paolucci, C. S. O'Connor, B. M. Anderson, B. Rigaud, B. M. Fellman, K. A. Jones, K. K. Brock, and B. C. Odisio, "Ablative margins of colorectal liver metastases

using deformable CT image registration and autosegmentation," *Radiology*, vol. 307, no. 2, p. e221373, 2023.

[53] I. Machado, M. Toews, E. George, P. Unadkat, W. Essayed, J. Luo, P. Teodoro, H. Carvalho, J. Martins, P. Golland, *et al.*, "Deformable MRI-ultrasound registration using correlation-based attribute matching for brain shift correction: Accuracy and generality in multi-site data," *NeuroImage*, vol. 202, p. 116094, 2019.

[54] A. Sedghi, T. Kapur, J. Luo, P. Mousavi, and W. M. Wells III, "Probabilistic image registration via deep multi-class classification: characterizing uncertainty," in *Workshop on Clinical Image-Based Procedures*, pp. 12–22, Springer, 2019.

[55] T. C. Mok and A. C. Chung, "Conditional deformable image registration with convolutional neural network," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 35–45, Springer, 2021.

[56] J. Chen, E. C. Frey, Y. He, W. P. Segars, Y. Li, and Y. Du, "TransMorph: Transformer for unsupervised medical image registration," *Medical Image Analysis*, vol. 82, p. 102615, 2022.

[57] M. P. Heinrich, "Closing the gap between deep and conventional image registration using probabilistic dense displacement networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 50–58, Springer, 2019.

[58] J. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[59] R. D. Hjelm, A. Fedorov, S. Lavoie-Marchildon, K. Grewal, P. Bachman, A. Trischler, and Y. Bengio, "Learning deep representations by mutual information estimation and maximization," in *International Conference on Learning Representations*, 2019.

[60] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 2002.

[61] S. Liu, B. Yang, Y. Wang, J. Tian, L. Yin, and W. Zheng, "2D/3D multimode medical image registration based on normalized cross-correlation," *Applied Sciences*, vol. 12, no. 6, p. 2828, 2022.

[62] J. N. Sarvaiya, S. Patnaik, and S. Bombaywala, "Image registration by template matching using normalized cross-correlation," in *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*, pp. 819–822, IEEE, 2009.

[63] A. Reithmeir, V. Spieker, V. Sideri-Lampretsa, D. Rueckert, J. A. Schnabel, and V. A. Zimmer, "From model-based to learned regularization in medical image registration: A comprehensive review," *Medical Image Analysis*, p. 103854, 2025.

[64] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.

[65] J. Chen, Y. Liu, S. Wei, Z. Bian, S. Subramanian, A. Carass, J. L. Prince, and Y. Du, "A survey on deep learning in medical image registration: New technologies, uncertainty, evaluation metrics, and beyond," *Medical Image Analysis*, vol. 100, p. 103385, 2025.

[66] J. L. Peng, D. Kahler, J. G. Li, S. Samant, G. Yan, R. Amdur, and C. Liu, "Characterization of a real-time surface image-guided stereotactic positioning system," *Medical Physics*, vol. 37, no. 10, pp. 5421–5433, 2010.

[67] S. Dhawan, Y. He, J. Bartek Jr, A. A. Alattar, and C. C. Chen, "Comparison of frame-based versus frameless intracranial stereotactic biopsy: systematic review and meta-analysis," *World Neurosurgery*, vol. 127, pp. 607–616, 2019.

[68] G. Eggers, J. Mühling, and R. Marmulla, "Image-to-patient registration techniques in head surgery," *International Journal of Oral and Maxillofacial Surgery*, vol. 35, no. 12, pp. 1081–1095, 2006.

[69] K. L. Holloway, S. E. Gaede, P. A. Starr, J. M. Rosenow, V. Ramakrishnan, and J. M. Henderson, "Frameless stereotaxy using bone fiducial markers for deep brain stimulation," *Journal of Neurosurgery*, vol. 103, no. 3, pp. 404–413, 2005.

[70] G. Burström, R. Nachabe, R. Homan, J. Hoppenbrouwers, R. Holthuizen, O. Persson, E. Edström, and A. Elmi-Terander, "Frameless patient tracking with adhesive optical skin markers for augmented reality surgical navigation in spine surgery," *Spine*, vol. 45, no. 22, pp. 1598–1604, 2020.

[71] M. Y. Wang, C. R. Maurer, J. M. Fitzpatrick, and R. J. Maciunas, "An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head," *IEEE Transactions on Biomedical Engineering*, vol. 43, no. 6, pp. 627–637, 2002.

[72] S. Wolfsberger, K. Rössler, R. Regatschnig, and K. Ungersböck, "Anatomical landmarks for image registration in frameless stereotactic neuronavigation," *Neurosurgical Review*, vol. 25, no. 1, pp. 68–72, 2002.

[73] A. Alansary, O. Oktay, Y. Li, L. Le Folgoc, B. Hou, G. Vaillant, K. Kamnitsas, A. Vlontzos, B. Glocker, B. Kainz, *et al.*, "Evaluating reinforcement learning agents for anatomical landmark detection," *Medical Image Analysis*, vol. 53, pp. 156–164, 2019.

[74] Y. Zheng, D. Liu, B. Georgescu, H. Nguyen, and D. Comaniciu, "3D deep learning for efficient and robust landmark detection in volumetric data," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 565–572, Springer, 2015.

[75] H. Chen, C. Shen, J. Qin, D. Ni, L. Shi, J. C. Cheng, and P.-A. Heng, "Automatic localization and identification of vertebrae in spine CT via a joint learning model with deep neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 515–522, Springer, 2015.

[76] S. Joutard, R. Dorent, S. Ourselin, T. Vercauteren, and M. Modat, "Driving points prediction for abdominal probabilistic registration," in *International Workshop on Machine Learning in Medical Imaging*, pp. 288–297, Springer, 2022.

[77] K. Rohr, M. Fornefett, and H. S. Stiehl, "Spline-based elastic image registration: integration of landmark errors and orientation attributes," *Computer Vision and Image Understanding*, vol. 90, no. 2, pp. 153–168, 2003.

[78] G. Allasia, R. Cavoretto, and A. De Rossi, "Local interpolation schemes for landmark-based image registration: a comparison," *Mathematics and Computers in Simulation*, vol. 106, pp. 1–25, 2014.

[79] J. Woo, M. Stone, and J. L. Prince, "Multimodal registration via mutual information incorporating geometric and spatial context," *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 757–769, 2014.

[80] F. Maes, D. Vandermeulen, and P. Suetens, "Medical image registration using mutual information," *Proceedings of the IEEE*, vol. 91, no. 10, pp. 1699–1722, 2003.

[81] X. Cao, J. Yang, J. Zhang, D. Nie, M. Kim, Q. Wang, and D. Shen, "Deformable image registration based on similarity-steered CNN regression," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 300–308, Springer, 2017.

[82] J. D. Foley, J. F. Hughes, A. van Dam, and S. K. Feiner, *Geometrical Transformations*, pp. 201–227. Addison-Wesley, 2nd ed., 1996.

[83] D. Rueckert and J. A. Schnabel, "Medical image registration," in *Biomedical Image Processing*, pp. 131–154, Springer, 2010.

[84] M. D. Buhmann, "Radial basis functions," *Acta numerica*, vol. 9, pp. 1–38, 2000.

[85] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 567–585, 1989.

[86] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Non-rigid registration using free-form deformations: application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[87] C. E. Rasmussen, "Gaussian processes in machine learning," in *Summer School on Machine Learning*, pp. 63–71, Springer, 2003.

[88] R. Bajcsy and S. Kovačič, "Multiresolution elastic matching," *Computer Vision, Graphics, and Image Processing*, vol. 46, no. 1, pp. 1–21, 1989.

[89] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Partial Differential Equations*, pp. 863–871. Cambridge University Press, 2nd ed., 1989.

[90] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.

[91] G. E. Christensen, S. C. Joshi, and M. I. Miller, "Volumetric transformation of brain anatomy," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 864–877, 1997.

[92] J. Ashburner, "A fast diffeomorphic image registration algorithm," *NeuroImage*, vol. 38, no. 1, pp. 95–113, 2007.

[93] V. Arsigny, O. Commowick, X. Pennec, and N. Ayache, "A log-euclidean framework for statistics on diffeomorphisms," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 924–931, Springer, 2006.

[94] J.-P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[95] J. Shackleford, N. Kandasamy, and G. Sharp, *Deformable Registration Using Optical-Flow Methods*, pp. 95–106. Elsevier, 2013.

[96] M. Modat, T. Vercauteren, G. R. Ridgway, D. J. Hawkes, N. C. Fox, and S. Ourselin, "Diffeomorphic demons using normalized mutual information, evaluation on multimodal brain MR images," in *Medical Imaging 2010: Image Processing*, vol. 7623, pp. 800–807, SPIE, 2010.

[97] A. Tristán-Vega, G. Vegas-Sánchez-Ferrero, and S. Aja-Fernández, "Local similarity measures for demons-like registration algorithms," in *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1087–1090, IEEE, 2008.

[98] F. P. Oliveira and J. M. R. Tavares, "Medical image registration: a review," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 17, no. 2, pp. 73–93, 2014.

[99] A. Joshi and Y. Hong, "MetaRegNet: Metamorphic image registration using flow-driven residual networks," in *International Workshop on Computational Mathematics Modeling in Cancer Analysis*, pp. 160–170, Springer, 2023.

[100] M. Meng, L. Bi, D. Feng, and J. Kim, "Brain tumor sequence registration with non-iterative coarse-to-fine networks and dual deep supervision," in *International MICCAI Brainlesion Workshop*, pp. 273–282, Springer, 2022.

[101] L. Tian, H. Greer, F.-X. Vialard, R. Kwitt, R. S. J. Estépar, R. J. Rushmore, N. Makris, S. Bouix, and M. Niethammer, "GradICON: Approximate diffeomorphisms via gradient inverse consistency," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18084–18094, 2023.

[102] M. Wodzinski, A. Jurgas, N. Marini, M. Atzori, and H. Müller, "Unsupervised method for intra-patient registration of brain magnetic resonance images based on objective function weighting by inverse consistency: Contribution to the BraTS-Reg challenge," in *International MICCAI Brainlesion Workshop*, pp. 241–251, Springer, 2022.

[103] A. Z. B. Aziz, M. S. T. Karanam, T. Kataria, and S. Y. Elhabian, "EFFICIENTMORPH: Parameter-efficient Transformer-based architecture for 3D image registration," in *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 1330–1341, IEEE, 2025.

[104] M. Rahmani, H. Moghaddasi, A. Pour-Rashidi, A. Ahmadian, E. Najafzadeh, and P. Farnia, "D2BGAN: Dual discriminator bayesian generative adversarial network for deformable MR–ultrasound registration applied to brain shift compensation," *Diagnostics*, vol. 14, no. 13, p. 1319, 2024.

[105] S. Almahfouz Nasser, N. C. Kurian, M. Meena, S. Shamsi, and A. Sethi, "WSSAMNet: Weakly supervised semantic attentive medical image registration network," in *International MICCAI Brainlesion Workshop*, pp. 15–24, Springer, 2022.

[106] R. Han, C. K. Jones, J. Lee, P. Wu, P. Vagdargi, A. Uneri, P. A. Helm, M. Luciano, W. S. Anderson, and J. H. Siewerdsen, "Deformable MR-CT image registration using an unsupervised, dual-channel network for neurosurgical guidance," *Medical Image Analysis*, vol. 75, p. 102292, 2022.

[107] A. Pirhadi, S. Salari, M. O. Ahmad, H. Rivaz, and Y. Xiao, "Robust landmark-based brain shift correction with a siamese neural network in ultrasound-guided brain tumor resection," *International Journal of Computer Assisted Radiology and Surgery*, vol. 18, no. 3, pp. 501–508, 2023.

[108] M. Hoffmann, B. Billot, D. N. Greve, J. E. Iglesias, B. Fischl, and A. V. Dalca, "SynthMorph: learning contrast-invariant registration without acquired images," *IEEE Transactions on Medical Imaging*, vol. 41, no. 3, pp. 543–558, 2021.

[109] K. Briechle and U. D. Hanebeck, "Template matching using fast normalized cross correlation," in *Optical Pattern Recognition XII*, vol. 4387, pp. 95–102, SPIE, 2001.

[110] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.

[111] A. Collignon, "Automated multi-modality image registration based on information theory," in *Information Processing in Medical Imaging*, pp. 263–274, 1995.

[112] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[113] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 69, no. 6, p. 066138, 2004.

[114] M. I. Belghazi, A. Baratin, S. Rajeshwar, S. Ozair, Y. Bengio, A. Courville, and D. Hjelm, "Mutual information neural estimation," in *International Conference on Machine Learning*, pp. 531–540, PMLR, 2018.

[115] C. Studholme, D. L. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.

[116] G. Hermosillo, C. Chefd'Hotel, and O. Faugeras, "Variational methods for multimodal image matching," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 329–343, 2002.

[117] A. Kelman, M. Sofka, and C. V. Stewart, "Keypoint descriptors for matching across multiple image modalities and non-linear intensity variations," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–7, IEEE, 2007.

[118] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, pp. 1150–1157, Ieee, 1999.

[119] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[120] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *European Conference on Computer Vision*, pp. 404–417, Springer, 2006.

[121] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, vol. 2, pp. 1508–1515, Ieee, 2005.

[122] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *European Conference on Computer Vision*, pp. 430–443, Springer, 2006.

[123] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *2011 International Conference on Computer Vision*, pp. 2548–2555, Ieee, 2011.

[124] M. P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. V. Gleeson, M. Brady, and J. A. Schnabel, "MIND: Modality Independent Neighbourhood Descriptor for multi-modal deformable registration," *Medical Image Analysis*, vol. 16, no. 7, pp. 1423–1435, 2012.

[125] D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperPoint: Self-supervised interest point detection and description," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition workshops*, pp. 224–236, 2018.

[126] K. M. Yi, E. Trulls, V. Lepetit, and P. Fua, "Lift: Learned invariant feature transform," in *European Conference on Computer Vision*, pp. 467–483, Springer, 2016.

[127] C. Liu, J. Xu, and F. Wang, "A review of keypoints' detection and feature description in image registration," *Scientific Programming*, vol. 2021, no. 1, p. 8509164, 2021.

[128] T. Eelbode, J. Bertels, M. Berman, D. Vandermeulen, F. Maes, R. Bisschops, and M. B. Blaschko, "Optimization for medical image segmentation: theory and practice when evaluating with dice score or jaccard index," *IEEE Transactions on Medical Imaging*, vol. 39, no. 11, pp. 3679–3690, 2020.

[129] A. Hering, L. Hansen, T. C. Mok, A. C. Chung, H. Siebert, S. Häger, A. Lange, S. Kuckertz, S. Heldmann, W. Shao, *et al.*, "Learn2reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning," *IEEE Transactions on Medical Imaging*, vol. 42, no. 3, pp. 697–712, 2022.

[130] T. C. Mok and A. Chung, "Fast symmetric diffeomorphic image registration with convolutional neural networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4644–4653, 2020.

[131] Z. Min, Z. M. Baum, S. U. Saeed, M. Emberton, D. C. Barratt, Z. A. Taylor, and Y. Hu, "Biomechanics-informed non-rigid medical image registration and its inverse material property estimation with linear and nonlinear elasticity," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 564–574, Springer, 2024.

[132] J. Yang, Y. Wu, D. Zhang, W. Cui, X. Yue, S. Du, and H. Zhang, "LDVoxelMorph: A precise loss function and cascaded architecture for unsupervised diffeomorphic large displacement registration," *Medical Physics*, vol. 49, no. 4, pp. 2427–2441, 2022.

[133] P. Tzitzimpasis, M. Ries, B. W. Raaymakers, and C. Zachiu, "Generalized div-curl based regularization for physically constrained deformable image registration," *Scientific Reports*, vol. 14, no. 1, p. 15002, 2024.

[134] M. Wodzinski and A. Skalski, "Adversarial affine registration for real-time intraoperative registration of 3-D US-US for brain shift correction," in *International Workshop on Advances in Simplifying Medical Ultrasound*, pp. 75–84, Springer, 2021.

[135] K. Tang, L. Wang, M. Yang, J. Xu, X. Cheng, J. Zhang, Y. Zhu, and H. Wei, "Deformable registration framework for glioma images with absent correspondence based on auxiliary-image-aided intensity-consistency constraint," in *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 3752–3757, IEEE, 2024.

[136] C. Wu, X. Zeng, H. Wang, X. Zhang, W. Fang, Q. Li, and Z. Wang, "Noise removed inconsistency activation map for unsupervised registration of brain tumor MRI between pre-operative and follow-up phases," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 683–693, Springer, 2024.

[137] C. Chefd'Hotel, G. Hermosillo, and O. Faugeras, "Flows of diffeomorphisms for multi-modal image registration," in *Proceedings IEEE International Symposium on Biomedical Imaging*, pp. 753–756, IEEE, 2002.

[138] L. Song, M. Ma, and G. Liu, "TS-Net: two-stage deformable medical image registration network based on new smooth constraints," *Magnetic Resonance Imaging*, vol. 99, pp. 26–33, 2023.

[139] Y. Zhu and S. Lu, "Swin-VoxelMorph: A symmetric unsupervised learning model for deformable medical image registration using swin transformer," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 78–87, Springer, 2022.

[140] L. Zhang, G. Ning, L. Zhou, and H. Liao, "Symmetric pyramid network for medical image inverse consistent diffeomorphic registration," *Computerized Medical Imaging and Graphics*, vol. 104, p. 102184, 2023.

[141] A. Joshi and Y. Hong, "R2Net: Efficient and flexible diffeomorphic image registration using lipschitz continuous residual networks," *Medical Image Analysis*, vol. 89, p. 102917, 2023.

[142] Y. Hu, E. Gibson, N. Ghavami, E. Bonmati, C. M. Moore, M. Emberton, T. Vercauteren, J. A. Noble, and D. C. Barratt, "Adversarial deformation regularization for training image registration neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 774–782, Springer, 2018.

[143] R. Bhalodia, S. Y. Elhabian, L. Kavan, and R. T. Whitaker, "A cooperative autoencoder for population-based regularization of CNN image registration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 391–400, Springer, 2019.

[144] A. Bône, P. Vernhet, O. Colliot, and S. Durrleman, "Learning joint shape and appearance representations with metamorphic auto-encoders," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 202–211, Springer, 2020.

[145] B. Fischer and J. Modersitzki, "Curvature-based image registration," *Journal of Mathematical Imaging and Vision*, vol. 18, no. 1, pp. 81–85, 2003.

[146] E. Haber and J. Modersitzki, "Image registration with guaranteed displacement regularity," *International Journal of Computer Vision*, vol. 71, no. 3, pp. 361–372, 2007.

[147] T. Rohlfing, C. R. Maurer, D. A. Bluemke, and M. A. Jacobs, "Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint," *IEEE Transactions on Medical Imaging*, vol. 22, no. 6, pp. 730–741, 2003.

[148] G. E. Christensen and H. J. Johnson, "Consistent image registration," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 568–582, 2002.

[149] H. Greer, R. Kwitt, F.-X. Vialard, and M. Niethammer, "ICON: Learning regular maps through inverse consistency," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3396–3405, 2021.

[150] J. Hadamard, "Sur les problèmes aux dérivées partielles et leur signification physique," *Princeton University Bulletin*, pp. 49–52, 1902.

[151] V. Mani and S. Arivazhagan, "Survey of medical image registration," *Journal of Biomedical Engineering and Technology*, vol. 1, no. 2, pp. 8–25, 2013.

[152] G. Song, J. Han, Y. Zhao, Z. Wang, and H. Du, "A review on medical image registration as an optimization problem," *Current Medical Imaging Reviews*, vol. 13, no. 3, pp. 274–283, 2017.

[153] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 2006.

[154] S. Klein, M. Staring, and J. P. Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2879–2890, 2007.

[155] I. Branislav, M. Haifeng, M. Dijana, *et al.*, "A survey of gradient methods for solving nonlinear optimization," *Electronic Research Archive*, vol. 28, no. 4, pp. 1573–1624, 2020.

[156] M. P. Wachowiak, R. Smolíková, Y. Zheng, J. M. Zurada, and A. S. Elmaghraby, "An approach to multimodal biomedical image registration utilizing particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 289–301, 2004.

[157] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, "Dense image registration through MRFs and efficient linear programming," *Medical Image Analysis*, vol. 12, no. 6, pp. 731–741, 2008.

[158] R. W. So, T. W. Tang, and A. C. Chung, "Non-rigid image registration of brain magnetic resonance images using graph-cuts," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2450–2467, 2011.

[159] M. P. Heinrich, M. Jenkinson, S. M. Brady, and J. A. Schnabel, "Globally optimal deformable registration on a minimum spanning tree using dense displacement sampling," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 115–122, Springer, 2012.

[160] Y. Nesterov, "A method for solving the convex programming problem with convergence rate O (1/k2)," in *Dokl Akad Nauk SSSR*, vol. 269, p. 543, 1983.

[161] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning," in *International Conference on Machine Learning*, pp. 1139–1147, pmlr, 2013.

[162] K. D. B. J. Adam *et al.*, "A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, vol. 1412, no. 6, 2014.

[163] Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang, "Deep learning in medical image registration: a review," *Physics in Medicine & Biology*, vol. 65, no. 20, p. 20TR01, 2020.

[164] M. Kim, J. Yun, Y. Cho, K. Shin, R. Jang, H.-j. Bae, and N. Kim, "Deep learning in medical imaging," *Neurospine*, vol. 16, no. 4, p. 657, 2019.

[165] H.-M. Zhang and B. Dong, "A review on deep learning in medical image reconstruction," *Journal of the Operations Research Society of China*, vol. 8, no. 2, pp. 311–340, 2020.

[166] M. Radak, H. Y. Lafta, and H. Fallahi, "Machine learning and deep learning techniques for breast cancer diagnosis and classification: a comprehensive review of medical imaging studies," *Journal of Cancer Research and Clinical Oncology*, vol. 149, no. 12, pp. 10473–10491, 2023.

[167] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, and D. Shen, "Unsupervised deep feature learning for deformable registration of MR brain images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 649–656, Springer, 2013.

[168] S. Miao, Z. J. Wang, and R. Liao, "A CNN regression approach for real-time 2D/3D registration," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1352–1363, 2016.

[169] X. Chen, Y. Meng, Y. Zhao, R. Williams, S. R. Vallabhaneni, and Y. Zheng, "Learning unsupervised parameter-specific affine transformation for medical images registration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 24–34, Springer, 2021.

[170] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, 2015.

[171] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 424–432, Springer, 2016.

[172] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021.

[173] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. Pluim, "Elastix: a toolbox for intensity-based medical image registration," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 196–205, 2009.

[174] M. Jaderberg, K. Simonyan, A. Zisserman, *et al.*, "Spatial Transformer networks," *Advances in Neural Information Processing Systems*, vol. 28, 2015.

[175] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[176] C. Qin, W. Bai, J. Schlemper, S. E. Petersen, S. K. Piechnik, S. Neubauer, and D. Rueckert, "Joint learning of motion estimation and segmentation for cardiac MR image sequences," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 472–480, Springer, 2018.

[177] T. Estienne, M. Vakalopoulou, S. Christodoulidis, E. Battistela, M. Lerousseau, A. Carre, G. Klausner, R. Sun, C. Robert, S. Mougiakakou, *et al.*, "U-ReSNet: Ultimate coupling of registration and segmentation with deep nets," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 310–319, Springer, 2019.

[178] Z. Tan, H. Zhang, F. Tian, L. Zhang, W. Sun, and H. Lu, "Progressively coupling network for brain MRI registration in few-shot situation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 623–633, Springer, 2023.

[179] S.-i. Amari, "Backpropagation and stochastic gradient descent method," *Neurocomputing*, vol. 5, no. 4-5, pp. 185–196, 1993.

[180] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[181] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," *Advances in Neural Information Processing Systems*, vol. 33, pp. 18661–18673, 2020.

[182] J. Hu, S. Sun, X. Yang, S. Zhou, X. Wang, Y. Fu, J. Zhou, Y. Yin, K. Cao, Q. Song, *et al.*, "Towards accurate and robust multi-modal medical image registration using contrastive metric learning," *IEEE Access*, vol. 7, pp. 132816–132827, 2019.

[183] N. Pielawski, E. Wetzer, J. Öfverstedt, J. Lu, C. Wählby, J. Lindblad, and N. Sladoje, "CoMIR: Contrastive multimodal image representation for registration," *Advances in Neural Information Processing Systems*, vol. 33, pp. 18433–18444, 2020.

[184] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, "Signature verification using a "Siamese" time delay neural network," *Advances in Neural Information Processing Systems*, vol. 6, 1993.

[185] L. Liu, X. Hu, L. Zhu, and P.-A. Heng, "Probabilistic multilayer regularization network for unsupervised 3D brain image registration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 346–354, Springer, 2019.

[186] M. Kang, X. Hu, W. Huang, M. R. Scott, and M. Reyes, "Dual-stream pyramid registration network," *Medical Image Analysis*, vol. 78, p. 102379, 2022.

[187] Y. Hu, M. Modat, E. Gibson, W. Li, N. Ghavami, E. Bonmati, G. Wang, S. Bandula, C. M. Moore, M. Emberton, *et al.*, "Weakly-supervised convolutional neural networks for multimodal image registration," *Medical Image Analysis*, vol. 49, pp. 1–13, 2018.

[188] B. D. De Vos, F. F. Berendsen, M. A. Viergever, H. Sokooti, M. Staring, and I. Išgum, "A deep learning framework for unsupervised affine and deformable image registration," *Medical Image Analysis*, vol. 52, pp. 128–143, 2019.

[189] J. Shi, Y. He, Y. Kong, J.-L. Coatrieux, H. Shu, G. Yang, and S. Li, "XMorpher: Full transformer for deformable medical image registration via cross attention," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 217–226, Springer, 2022.

[190] X. Yang, R. Kwitt, M. Styner, and M. Niethammer, "Quicksilver: Fast predictive image registration–a deep learning approach," *NeuroImage*, vol. 158, pp. 378–396, 2017.

[191] H. Wang, D. Ni, and Y. Wang, "Recursive deformable pyramid network for unsupervised medical image registration," *IEEE Transactions on Medical Imaging*, vol. 43, no. 6, pp. 2229–2240, 2024.

[192] S. Zhao, Y. Dong, E. I. Chang, Y. Xu, *et al.*, "Recursive cascaded networks for unsupervised medical image registration," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10600–10610, 2019.

[193] H. Siebert, L. Hansen, and M. P. Heinrich, "Fast 3D registration with accurate optimisation and little learning for Learn2Reg 2021," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 174–179, Springer, 2021.

[194] I. Y. Ha and M. P. Heinrich, "Modality-agnostic self-supervised deep feature learning and fast instance optimisation for multimodal fusion in ultrasound-guided interventions," *Computer Methods and Programs in Biomedicine*, vol. 211, p. 106374, 2021.

[195] D. Waldmannstetter, B. Wiestler, J. Schwarting, I. Ezhov, M. Metz, S. Bakas, B. Baheti, S. Chakrabarty, D. Rueckert, J. S. Kirschke, *et al.*, "Primitive simultaneous optimization of similarity metrics for image registration," in *International MICCAI Brainlesion Workshop*, pp. 57–68, Springer, 2023.

[196] K. Miller, G. R. Joldes, G. Bourantas, S. K. Warfield, D. E. Hyde, R. Kikinis, and A. Wittek, "Biomechanical modeling and computer simulation of the brain during neurosurgery," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 35, no. 10, p. e3250, 2019.

[197] F. Morin, M. Chabanas, H. Courtecuisse, and Y. Payan, "Biomechanical modeling of brain soft tissues for medical applications," in *Biomechanics of Living Organs*, pp. 127–146, Elsevier, 2017.

[198] Y. Zhuang, *Real-time simulation of physically realistic global deformations*. University of California, Berkeley, 2000.

[199] G. Székely, C. Brechbühler, R. Hutter, A. Rhomberg, N. Ironmonger, and P. Schmid, "Modelling of soft tissue deformation for laparoscopic surgery simulation," *Medical Image Analysis*, vol. 4, no. 1, pp. 57–66, 2000.

[200] G. Picinbono, H. Delingette, and N. Ayache, "Non-linear anisotropic elasticity for real-time surgery simulation," *Graphical Models*, vol. 65, no. 5, pp. 305–321, 2003.

[201] K. Miller, G. Joldes, D. Lance, and A. Wittek, "Total Lagrangian explicit dynamics finite element algorithm for computing soft tissue deformation," *Communications in Numerical Methods in Engineering*, vol. 23, no. 2, pp. 121–134, 2007.

[202] R. Kikinis, M. E. Shenton, D. V. Iosifescu, R. W. McCarley, P. Saiviroonporn, H. H. Hokama, A. Robatino, D. Metcalf, C. G. Wible, C. M. Portas, *et al.*, "A digital brain atlas for surgical planning, model-driven segmentation, and teaching," *IEEE Transactions on Visualization and Computer Graphics*, vol. 2, no. 3, pp. 232–241, 1996.

[203] W. L. Nowinski, A. Fang, B. T. Nguyen, J. K. Raphel, L. Jagannathan, R. Raghavan, R. N. Bryan, and G. A. Miller, "Multiple brain atlas database and atlas-based neuroimaging system," *Computer Aided Surgery*, vol. 2, no. 1, pp. 42–66, 1997.

[204] M. M. Chakravarty, G. Bertrand, C. P. Hodge, A. F. Sadikot, and D. L. Collins, "The creation of a brain atlas for image-guided neurosurgery using serial histological data," *NeuroImage*, vol. 30, no. 2, pp. 359–376, 2006.

[205] F. Isensee, M. Schell, I. Pflueger, G. Brugnara, D. Bonekamp, U. Neuberger, A. Wick, H.-P. Schlemmer, S. Heiland, W. Wick, *et al.*, "Automated brain extraction of multisequence MRI using artificial neural networks," *Human Brain Mapping*, vol. 40, no. 17, pp. 4952–4964, 2019.

[206] F. Ségonne, A. M. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn, and B. Fischl, "A hybrid approach to the skull stripping problem in MRI," *NeuroImage*, vol. 22, no. 3, pp. 1060–1075, 2004.

[207] A. V. Faria, A. Hoon, E. Stashinko, X. Li, H. Jiang, A. Mashayekh, K. Akhter, J. Hsu, K. Oishi, J. Zhang, *et al.*, "Quantitative analysis of brain pathology based on MRI and brain atlases—applications for cerebral palsy," *NeuroImage*, vol. 54, no. 3, pp. 1854–1861, 2011.

[208] C. Velasco-Annis, A. Akhondi-Asl, A. Stamm, and S. K. Warfield, "Reproducibility of brain MRI segmentation algorithms: Empirical comparison of local MAP PSTAPLE, FreeSurfer, and FSL-FIRST," *Journal of Neuroimaging*, vol. 28, no. 2, pp. 162–172, 2018.

[209] H. K. Bø, O. Solheim, A. S. Jakola, K.-A. Kvistad, I. Reinertsen, and E. M. Berntsen, "Intra-rater variability in low-grade glioma segmentation," *Journal of Neuro-Oncology*, vol. 131, no. 2, pp. 393–402, 2017.

[210] K. Miller, A. Horton, G. Joldes, and A. Wittek, "Beyond finite elements: a comprehensive, patient-specific neurosurgical simulation utilizing a meshless method," *Journal of Biomechanics*, vol. 45, no. 15, pp. 2698–2701, 2012.

[211] J. Y. Zhang, G. R. Joldes, A. Wittek, and K. Miller, "Patient-specific computational biomechanics of the brain without segmentation and meshing," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 29, no. 2, pp. 293–308, 2013.

[212] M. Mazumder, K. Miller, S. Bunt, A. Mostayed, G. Joldes, R. Day, R. Hart, and A. Wittek, "Mechanical properties of the brain-skull interface," *Acta of Bioengineering and Biomechanics*, vol. 15, no. 2, pp. 3–11, 2013.

[213] S. Agrawal, A. Wittek, G. Joldes, S. Bunt, and K. Miller, "Mechanical properties of brain–skull interface in compression," in *Computational Biomechanics for Medicine: New Approaches and New Applications*, pp. 83–91, Springer, 2015.

[214] A. Hagemann, K. Rohr, H. S. Stiehl, U. Spetzger, and J. M. Gilsbach, "Biomechanical modeling of the human head for physically based, nonrigid image registration," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 875–884, 2002.

[215] M. I. Miga, K. D. Paulsen, P. J. Hoopes, F. E. Kennedy, A. Hartov, and D. W. Roberts, "In vivo quantification of a homogeneous brain deformation model for updating preoperative images during surgery," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 2, pp. 266–273, 2002.

[216] J. Hu, X. Jin, J. B. Lee, L. Zhang, V. Chaudhary, M. Guthikonda, K. H. Yang, and A. I. King, "Intraoperative brain shift prediction using a 3D inhomogeneous patient-specific finite element model," *Journal of Neurosurgery*, vol. 106, no. 1, pp. 164–169, 2007.

[217] A. Wittek and K. Omori, "Parametric study of effects of brain-skull boundary conditions and brain material properties on responses of simplified finite element brain model under angular acceleration impulse in sagittal plane," *JSME International Journal Series C Mechanical Systems, Machine Elements and Manufacturing*, vol. 46, no. 4, pp. 1388–1399, 2003.

[218] F. Wang, Y. Han, B. Wang, Q. Peng, X. Huang, K. Miller, and A. Wittek, "Prediction of brain deformations and risk of traumatic brain injury due to closed-head impact: quantitative analysis of the effects of boundary conditions and brain tissue constitutive model," *Biomechanics and Modeling in Mechanobiology*, vol. 17, no. 4, pp. 1165–1185, 2018.

[219] A. Wittek, K. Miller, R. Kikinis, and S. K. Warfield, "Patient-specific model of brain deformation: Application to medical image registration," *Journal of Biomechanics*, vol. 40, no. 4, pp. 919–929, 2007.

[220] A. Mostayed, R. R. Garlapati, G. R. Joldes, A. Wittek, A. Roy, R. Kikinis, S. K. Warfield, and K. Miller, "Biomechanical model as a registration tool for image-guided neurosurgery: evaluation against Bspline registration," *Annals of Biomedical Engineering*, vol. 41, pp. 2409–2425, 2013.

[221] S. Frisken, P. Unadkat, X. Yang, M. Miga, and A. Golby, *Intraoperative Measurement of Brain Deformation*, pp. 303–319. Springer Cham, 8 2019.

[222] H. J. Kim, C. Figueroa, T. Hughes, K. Jansen, and C. Taylor, "Augmented Lagrangian method for constraining the shape of velocity profiles at outlet boundaries for three-dimensional finite element simulations of blood flow," *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 45-46, pp. 3551–3566, 2009.

[223] H. J. Kim, I. E. Vignon-Clementel, C. A. Figueroa, J. F. LaDisa, K. E. Jansen, J. A. Feinstein, and C. A. Taylor, "On coupling a lumped parameter heart model and a three-dimensional finite element aorta model," *Annals of Biomedical Engineering*, vol. 37, no. 11, pp. 2153–2169, 2009.

[224] F. Drakopoulos, C. Tsolakis, A. Angelopoulos, Y. Liu, C. Yao, K. R. Kavazidi, N. Foroglou, A. Fedorov, S. Frisken, R. Kikinis, *et al.*, "Adaptive physics-based non-rigid registration for immersive image-guided neuronavigation systems," *Frontiers in Digital Health*, vol. 2, p. 613608, 2021.

[225] Y. Yu, S. Safdar, G. Bourantas, B. Zwick, G. Joldes, T. Kapur, S. Frisken, R. Kikinis, A. Nabavi, A. Golby, *et al.*, "Automatic framework for patient-specific modelling of tumour

resection-induced brain shift," *Computers in Biology and Medicine*, vol. 143, p. 105271, 2022.

[226] L. E. Bilston, "Brain tissue mechanical properties," in *Biomechanics of the Brain* (K. Miller, ed.), pp. 71–95, Springer Cham, 2 ed., 8 2019.

[227] K. Sahay, R. Mehrotra, U. Sachdeva, and A. Banerji, "Elastomechanical characterization of brain tissues," *Journal of Biomechanics*, vol. 25, no. 3, pp. 319–326, 1992.

[228] J. Ruan, T. Khalil, and A. I. King, "Dynamic response of the human head to impact by three-dimensional finite element analysis," *Journal of Biomechanical Engineering*, vol. 116, no. 1, pp. 44–50, 1994.

[229] K. Miller, K. Chinzei, G. Orssengo, and P. Bednarz, "Mechanical properties of brain tissue in-vivo: experiment and computer simulation," *Journal of Biomechanics*, vol. 33, no. 11, pp. 1369–1376, 2000.

[230] L. E. Bilston, Z. Liu, and N. Phan-Thien, "Large strain behaviour of brain tissue in shear: some experimental data and differential constitutive model," *Biorheology*, vol. 38, no. 4, pp. 335–345, 2001.

[231] M. T. Prange and S. S. Margulies, "Regional, directional, and age-dependent properties of the brain undergoing large deformation," *Journal of Biomechanical Engineering*, vol. 124, no. 2, pp. 244–252, 2002.

[232] A. Wittek, T. Hawkins, and K. Miller, "On the unimportance of constitutive models in computing brain deformation for image-guided surgery," *Biomechanics and Modeling in Mechanobiology*, vol. 8, no. 1, pp. 77–84, 2009.

[233] K. Miller and J. Lu, "On the prospect of patient-specific biomechanics without patient-specific properties of tissues," *Journal of the Mechanical Behavior of Biomedical Materials*, vol. 27, pp. 154–166, 2013.

[234] R. R. Garlapati, A. Roy, G. R. Joldes, A. Wittek, A. Mostayed, B. Doyle, S. K. Warfield, R. Kikinis, N. Knuckey, S. Bunt, *et al.*, "Biomechanical modeling provides more accurate data for neuronavigation than rigid registration," *Journal of Neurosurgery*, vol. 120, no. 6, p. 1477, 2014.

[235] G. Zhang, A. Wittek, G. Joldes, X. Jin, and K. Miller, "A three-dimensional nonlinear meshfree algorithm for simulating mechanical responses of soft tissue," *Engineering Analysis with Boundary Elements*, vol. 42, pp. 60–66, 2014.

[236] Y. Yu, G. Bourantas, B. Zwick, G. Joldes, T. Kapur, S. Frisken, R. Kikinis, A. Nabavi, A. Golby, A. Wittek, *et al.*, "Computer simulation of tumour resection-induced brain deformation by a meshless approach," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 38, no. 1, p. e3539, 2022.

[237] K.-J. Bathe, *Solution of Equilibrium Equations in Dynamics Analysis*, pp. 768–837. Klaus-Jürgen Bathe, 2nd ed., 2014.

[238] B. Zwick, G. R. Joldes, A. Wittek, and K. Miller, "Numerical algorithm for simulation of soft tissue swelling and shrinking in a total Lagrangian explicit dynamics framework," in *Computational Biomechanics for Medicine: New Approaches and New Applications*, pp. 37–46, Springer, 2015.

[239] G. R. Joldes, A. Wittek, and K. Miller, "Computation of intra-operative brain shift using dynamic relaxation," *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 41-44, pp. 3313–3320, 2009.

[240] G. R. Joldes, A. Wittek, and K. Miller, "An adaptive dynamic relaxation method for solving nonlinear finite element problems. application to brain shift estimation," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 27, no. 2, pp. 173–185, 2011.

[241] M. T. Townsend and N. Sarigul-Klijn, "Updated Lagrangian finite element formulations of various biological soft tissue non-linear material models: a comprehensive procedure and review," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 19, no. 11, pp. 1137–1142, 2016.

[242] A. Horton, A. Wittek, G. R. Joldes, and K. Miller, "A meshless total Lagrangian explicit dynamics algorithm for surgical simulation," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 26, no. 8, pp. 977–998, 2010.

[243] F. Alonso, J. Cuadrado, U. Lugrís, and P. Pintado, "A compact smoothing-differentiation and projection approach for the kinematic data consistency of biomechanical systems," *Multibody System Dynamics*, vol. 24, no. 1, pp. 67–80, 2010.

[244] M. Li, K. Miller, G. R. Joldes, R. Kikinis, and A. Wittek, "Biomechanical model for computing deformations for whole-body image registration: A meshless approach," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 32, no. 12, p. e02771, 2016.

[245] T. Belytschko, "An overview of semidiscretization and time integration procedures," *Computational Methods for Rransient Analysis(A 84-29160 12-64). Amsterdam, North-Holland, 1983,*, pp. 1–65, 1983.

[246] K.-J. Bathe, "Finite element method," *Wiley Encyclopedia of Computer Science and Engineering*, pp. 1–12, 2007.

[247] G. R. Joldes, A. Wittek, and K. Miller, "Suite of finite element algorithms for accurate computation of soft tissue deformation for surgical simulation," *Medical Image Analysis*, vol. 13, no. 6, pp. 912–919, 2009.

[248] M. Ferrant, B. Macq, A. Nabavi, and S. K. Warfield, "Deformable modeling for characterizing biomedical shape changes," in *International Conference on Discrete Geometry for Computer Imagery*, pp. 235–248, Springer, 2000.

[249] F. Morin, H. Courtecuisse, I. Reinertsen, F. Le Lann, O. Palombi, Y. Payan, and M. Chabanas, "Brain-shift compensation using intraoperative ultrasound and constraint-based biomechanical simulation," *Medical Image Analysis*, vol. 40, pp. 133–153, 2017.

[250] M. Luo, S. F. Frisken, J. A. Weis, L. W. Clements, P. Unadkat, R. C. Thompson, A. J. Golby, and M. I. Miga, "Retrospective study comparing model-based deformation correction to intraoperative magnetic resonance imaging for image-guided neurosurgery," *Journal of Medical Imaging*, vol. 4, no. 3, pp. 035003–035003, 2017.

[251] L. GR, "1013 mesh free methods: Moving beyond the finite element method," in *The Proceedings of The Computational Mechanics Conference 2003.16*, pp. 937–938, The Japan Society of Mechanical Engineers, 2003.

[252] M. Luo, P. S. Larson, A. J. Martin, and M. I. Miga, "Accounting for deformation in deep brain stimulation surgery with models: comparison to interventional magnetic resonance imaging," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 10, pp. 2934–2944, 2020.

[253] S. Frisken, J. Luo, N. Haouchine, S. Pieper, Y. Wang, W. M. Wells, and A. J. Golby, "Incorporating uncertainty into path planning for minimally invasive robotic neurosurgery," *IEEE Transactions on Medical Robotics and Bionics*, vol. 4, no. 1, pp. 5–16, 2021.

[254] N. Haouchine, P. Juvekar, A. Golby, and S. Frisken, "Predicted microscopic cortical brain images for optimal craniotomy positioning and visualisation," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 9, no. 4, pp. 407–413, 2021.

[255] I. Chen, R. E. Ong, A. L. Simpson, K. Sun, R. C. Thompson, and M. I. Miga, "Integrating retraction modeling into an atlas-based framework for brain shift prediction," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 12, pp. 3494–3504, 2013.

[256] P. Dumpuri, R. C. Thompson, A. Cao, S. Ding, I. Garg, B. M. Dawant, and M. I. Miga, "A fast and efficient method to compensate for brain shift for tumor resection therapies measured between preoperative and postoperative tomograms," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp. 1285–1296, 2010.

[257] C. Li, X. Fan, J. Hong, D. W. Roberts, J. P. Aronson, and K. D. Paulsen, "Model-based image updating for brain shift in deep brain stimulation electrode placement surgery," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 12, pp. 3542–3552, 2020.

[258] D.-X. Zhuang, Y.-X. Liu, J.-S. Wu, C.-J. Yao, Y. Mao, C.-X. Zhang, M.-N. Wang, W. Wang, and L.-F. Zhou, "A sparse intraoperative data-driven biomechanical model to compensate for brain shift during neuronavigation," *American Journal of Neuroradiology*, vol. 32, no. 2, pp. 395–402, 2011.

[259] K. Sun, T. S. Pheiffer, A. L. Simpson, J. A. Weis, R. C. Thompson, and M. I. Miga, "Near real-time computer assisted surgery for brain shift correction using biomechanical models," *IEEE journal of translational engineering in health and medicine*, vol. 2, pp. 1–13, 2014.

[260] N. Haouchine, P. Juvekar, W. M. Wells III, S. Cotin, A. Golby, and S. Frisken, "Deformation aware augmented reality for craniotomy using 3D/2D non-rigid registration of cortical vessels," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV 23*, pp. 735–744, Springer, 2020.

[261] N. Haouchine, P. Juvekar, M. Nercessian, W. M. Wells III, A. Golby, and S. Frisken, "Pose estimation and non-rigid registration for augmented reality during neurosurgery," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 4, pp. 1310–1317, 2021.

[262] S. Lee, G. Wolberg, and S. Y. Shin, "Scattered data interpolation with multilevel B-splines," *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, no. 3, pp. 228–244, 2002.

[263] G. R. Joldes, A. Wittek, S. K. Warfield, and K. Miller, "Performing brain image warping using the deformation field predicted by a biomechanical model," in *Computational Biomechanics for Medicine: Deformation and Flow*, pp. 89–96, Springer, 2012.

[264] R. R. Garlapati, A. Roy, G. R. Joldes, A. Wittek, A. Mostayed, B. Doyle, S. K. Warfield, R. Kikinis, N. Knuckey, S. Bunt, *et al.*, "More accurate neuronavigation data provided by biomechanical modeling instead of rigid registration," *Journal of Neurosurgery*, vol. 120, no. 6, pp. 1477–1483, 2014.

[265] S. Frisken, M. Luo, P. Juvekar, A. Bunevicius, I. Machado, P. Unadkat, M. M. Bertotti, M. Toews, W. M. Wells, M. I. Miga, *et al.*, "A comparison of thin-plate spline deformation and finite element modeling to compensate for brain shift during tumor resection," *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, pp. 75–85, 2020.

[266] P. A. López, H. Mella, S. Uribe, D. E. Hurtado, and F. Sahli Costabal, "WarpPINN: Cine-MR image registration with physics-informed neural networks," *Medical Image Analysis*, vol. 89, p. 102925, 2023.

[267] A. Reithmeir, L. Felsner, R. Braren, J. A. Schnabel, and V. A. Zimmer, "Data-driven tissue- and subject-specific elastic regularization for medical image registration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 575–585, Springer, 2024.

[268] Y. Salehi and D. Giannacopoulos, "PhysGNN: A physics–driven graph neural network based model for predicting soft tissue deformation in image–guided neurosurgery," *Advances in Neural Information Processing Systems*, vol. 35, pp. 37282–37296, 2022.

[269] H. Rivaz and D. L. Collins, "Near real-time robust non-rigid registration of volumetric ultrasound images for neurosurgery," *Ultrasound in Medicine & Biology*, vol. 41, no. 2, pp. 574–587, 2015.

[270] M. P. Heinrich, "Intra-operative ultrasound to MRI fusion with a public multimodal discrete registration tool," in *International Workshop on Point-of-Care Ultrasound*, pp. 159–164, Springer, 2018.

[271] W. Wein, "Brain-shift correction with image-based registration and landmark accuracy evaluation," in *International Workshop on Point-of-Care Ultrasound*, pp. 146–151, Springer, 2018.

[272] I. Machado, M. Toews, J. Luo, P. Unadkat, W. Essayed, E. George, P. Teodoro, H. Carvalho, J. Martins, P. Golland, *et al.*, "Deformable MRI-ultrasound registration via attribute matching and mutual-saliency weighting for image-guided neurosurgery," in *International Workshop on Point-of-Care Ultrasound*, pp. 165–171, Springer, 2018.

[273] H. Sokooti, B. De Vos, F. Berendsen, B. P. Lelieveldt, I. Išgum, and M. Staring, "Nonrigid image registration using multi-scale 3D convolutional neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 232–239, Springer, 2017.

[274] R. A. Zeineldin, M. E. Karar, Z. Elshaer, M. Schmidhammer, J. Coburger, C. R. Wirtz, O. Burgert, and F. Mathis-Ullrich, "iRegNet: Non-rigid registration of MRI to interventional US for brain-shift compensation using convolutional neural networks," *IEEE Access*, vol. 9, pp. 147579–147590, 2021.

[275] Y. Feng, Y. Zheng, D. Huang, J. Wei, T. Liu, Y. Wang, and Y. Liu, "Stepwise corrected attention registration network for preoperative and follow-up magnetic resonance imaging of glioma patients," *Bioengineering*, vol. 11, no. 9, p. 951, 2024.

[276] J. Zhang, X. Xie, X. Cheng, T. Li, J. Zhong, X. Hu, L. Sun, and H. Yan, "Deep learning-based deformable image registration with bilateral pyramid to align pre-operative and follow-up magnetic resonance imaging (MRI) scans," *Quantitative Imaging in Medicine and Surgery*, vol. 14, no. 7, p. 4779, 2024.

[277] L. Siegert, P. Fischer, M. P. Heinrich, and C. F. Baumgartner, "PULPo: Probabilistic unsupervised laplacian pyramid registration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 717–727, Springer, 2024.

[278] Y. Liu, X. Li, R. Li, S. Huang, and X. Yang, "A multi-view assisted registration network for MRI registration pre-and post-therapy," *Medical & Biological Engineering & Computing*, vol. 61, no. 12, pp. 3181–3191, 2023.

[279] J. Abderezaei, A. Pionteck, A. Chopra, and M. Kurt, "3D Inception-based TransMorph: Pre-and post-operative multi-contrast MRI registration in brain tumors," in *International MICCAI Brainlesion Workshop*, pp. 35–45, Springer, 2022.

[280] R. A. Zeineldin, M. E. Karar, F. Mathis-Ullrich, and O. Burgert, "Self-supervised iRegNet for the registration of longitudinal brain MRI of diffuse glioma patients," in *International MICCAI Brainlesion Workshop*, pp. 25–34, Springer, 2022.

[281] B. Duan, M. Zhong, and Y. Yan, "Towards saner deep image registration," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 12459–12468, 2023.

[282] L. Mercier, R. F. Del Maestro, K. Petrecca, D. Araujo, C. Haegelen, and D. L. Collins, "Online database of clinical MR and ultrasound images of brain tumors," *Medical Physics*, vol. 39, no. 6Part1, pp. 3253–3261, 2012.

[283] Y. Xiao, M. Fortin, G. Unsgård, H. Rivaz, and I. Reinertsen, "REtroSpective Evaluation of Cerebral Tumors (RESECT): A clinical database of pre-operative mri and intra-operative ultrasound in low-grade glioma surgeries," *Medical Physics*, vol. 44, no. 7, pp. 3875–3882, 2017.

[284] T. Shimamoto, Y. Sano, K. Yoshimitsu, K. Masamune, and Y. Muragaki, "Precise brain-shift prediction by new combination of w-net deep learning for neurosurgical navigation," *Neurologia Medico-Chirurgica*, vol. 63, no. 7, pp. 295–303, 2023.

[285] P. Juvekar, R. Dorent, F. Kögl, E. Torio, C. Barr, L. Rigolo, C. Galvin, N. Jowkar, A. Kazi, N. Haouchine, *et al.*, "ReMIND: The brain resection multimodal imaging database," *Scientific Data*, vol. 11, no. 1, p. 494, 2024.

[286] B. Baheti, S. Chakrabarty, H. Akbari, M. Bilello, B. Wiestler, J. Schwarting, E. Calabrese, J. Rudie, S. Abidi, M. Mousa, *et al.*, "The brain tumor sequence registration (BraTS-Reg) challenge: Establishing correspondence between pre-operative and follow-up mri scans of diffuse glioma patients," 2024.

[287] L. Mercier, R. F. Del Maestro, K. Petrecca, A. Kochanowska, S. Drouin, C. X. Yan, A. L. Janke, S. J.-S. Chen, and D. L. Collins, "New prototype neuronavigation system based on preoperative imaging and intraoperative freehand ultrasound: system description and validation," *International Journal of Computer Assisted Radiology and Surgery*, vol. 6, no. 4, pp. 507–522, 2011.

[288] B. Behboodi, F.-x. Carton, M. Chabanas, S. de Ribaupierre, O. Solheim, B. K. Munkvold, H. Rivaz, Y. Xiao, and I. Reinertsen, "Open access segmentations of intraoperative brain tumor ultrasound images," *Medical Physics*, vol. 51, no. 9, pp. 6525–6532, 2024.

[289] Y. Xiao, H. Rivaz, M. Chabanas, M. Fortin, I. Machado, Y. Ou, M. P. Heinrich, J. A. Schnabel, X. Zhong, A. Maier, *et al.*, "Evaluation of MRI to ultrasound registration methods for brain shift correction: the CuRIOUS2018 challenge," *IEEE Transactions on Medical Imaging*, vol. 39, no. 3, pp. 777–786, 2019.

[290] L. Hansen, W. Heyer, C. Großbröhmer, F. Madesta, T. Sentker, W. Jiazheng, Y. Zhang, H. Zhang, M. Liu, J. Wang, *et al.*, "Learn2Reg 2024: New benchmark datasets driving progress on new challenges," *arXiv preprint arXiv:2509.01217*, 2025.

[291] A. Q. Wang, R. Saluja, H. Kim, X. He, A. Dalca, M. R. Sabuncu, *et al.*, "BrainMorph: A foundational keypoint model for robust and flexible brain mri registration," *Machine Learning for Biomedical Imaging*, vol. 3, no. May 2025 issue, pp. 181–203, 2025.

[292] P. Bergström and O. Edlund, "Robust registration of point sets using iteratively reweighted least squares," *Computational Optimization and Applications*, vol. 58, no. 3, pp. 543–561, 2014.

[293] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," in *International Conference on Learning Representations*, 2021.

[294] R. Dorent, L. Rigolo, C. P. Galvin, J. Chen, M. P. Heinrich, A. Carass, O. Colliot, D. Wassermann, A. Golby, T. Kapur, *et al.*, "The Brain Resection Multimodal Image Registration (ReMIND2Reg) 2025 Challenge," *arXiv preprint arXiv:2508.09649*, 2025.

[295] R. Dorent, N. Haouchine, A. Golby, S. Frisken, T. Kapur, and W. Wells, "Unified cross-modal medical image synthesis with hierarchical mixture of product-of-experts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

[296] D. Morozov, R. Dorent, and N. Haouchine, "A 3D cross-modal keypoint descriptor for MR-US matching and registration," *arXiv preprint arXiv:2507.18551*, 2025.

[297] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.

[298] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin Transformer: Hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10012–10022, 2021.

[299] A. Joshi and Y. Hong, "Diffeomorphic image registration using Lipschitz continuous residual networks," in *International Conference on Medical Imaging with Deep Learning*, pp. 605–617, PMLR, 2022.

[300] S. Salari, A. Rasoulian, H. Rivaz, and Y. Xiao, "Towards multi-modal anatomical landmark detection for ultrasound-guided brain tumor resection with contrastive learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 668–678, Springer, 2023.

[301] G. R. Joldes, A. Wittek, and K. Miller, "Real-time nonlinear finite element computations on GPU–application to neurosurgical simulation," *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 49-52, pp. 3305–3314, 2010.

[302] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," in *Sensor Fusion IV: Control Paradigms and Data Structures*, vol. 1611, pp. 586–606, Spie, 1992.

[303] X. Fan, S. Ji, A. Hartov, D. W. Roberts, and K. D. Paulsen, "Stereovision to MR image registration for cortical surface displacement mapping to enhance image-guided neurosurgery," *Medical Physics*, vol. 41, no. 10, p. 102302, 2014.

[304] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, *et al.*, "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012.

[305] S. Bakas, C. Sako, H. Akbari, M. Bilello, A. Sotiras, G. Shukla, J. D. Rudie, N. F. Santamaría, A. F. Kazerooni, S. Pati, *et al.*, "The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: advanced MRI, clinical, genomics, & radiomics," *Scientific Data*, vol. 9, no. 1, p. 453, 2022.

[306] T. Rohlfing, N. M. Zahr, E. V. Sullivan, and A. Pfefferbaum, "The SRI24 multichannel atlas of normal adult human brain structure," *Human Brain Mapping*, vol. 31, no. 5, pp. 798–819, 2010.

[307] R. Dorent, T. Booth, W. Li, C. H. Sudre, S. Kafiabadi, J. Cardoso, S. Ourselin, and T. Vercauteren, "Learning joint segmentation of tissues and brain lesions from task-specific hetero-modal domain-shifted datasets," *Medical Image Analysis*, vol. 67, p. 101862, 2021.

[308] M. Himmetoglu, I. Ciernik, and E. Konukoglu, "Learning to segment anatomy and lesions from disparately labeled sources in brain MRI," *arXiv preprint arXiv:2503.18840*, 2025.

[309] B. Billot, D. N. Greve, O. Puonti, A. Thielscher, K. Van Leemput, B. Fischl, A. V. Dalca, J. E. Iglesias, *et al.*, "SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining," *Medical Image Analysis*, vol. 86, p. 102789, 2023.

[310] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.

[311] T. Lewiner, H. Lopes, A. W. Vieira, and G. Tavares, "Efficient implementation of marching cubes' cases with topological guarantees," *Journal of Graphics Tools*, vol. 8, no. 2, pp. 1–15, 2003.

[312] L. Chauvin, K. Kumar, C. Wachinger, M. Vangel, J. de Guise, C. Desrosiers, W. Wells, M. Toews, A. D. N. Initiative, *et al.*, "Neuroimage signature from salient keypoints is highly specific to individuals and shared by close relatives," *NeuroImage*, vol. 204, p. 116208, 2020.

[313] T. Assis, R. Dorent, B. F. Zwick, N. C. Garcia, and I. P. Machado, "NeuriPhy - Neuroimaging Dataset for Physics-Informed Learning." Zenodo, 2025.

[314] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in Neural Information Processing Systems*, vol. 33, pp. 6840–6851, 2020.

[315] K. Lee, J. Zung, P. Li, V. Jain, and H. S. Seung, "Superhuman accuracy on the SNEMI3D connectomics challenge," *arXiv preprint arXiv:1706.00120*, 2017.

[316] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

[317] A. G. Roy, N. Navab, and C. Wachinger, "Concurrent spatial and channel 'squeeze & excitation'in fully convolutional networks," in *MICCAI 2018*, pp. 421–429, Springer, 2018.

[318] W. C. of Tumours Editorial Board, *Gliomas, glioneuronal tumours, and neuronal tumours*, vol. 6, pp. 15–55. World Health Organization, 5 ed., 2021.

[319] D. L. Hill, C. R. Maurer Jr, R. J. Maciunas, J. A. Barwise, M. J. Fitzpatrick, and M. Y. Wang, "Measurement of intraoperative brain surface deformation under a craniotomy," *Neurosurgery*, vol. 43, no. 3, pp. 514–526, 1998.

[320] C. Nimsky, O. Ganslandt, S. Cerny, P. Hastreiter, G. Greiner, and R. Fahlbusch, "Quantification of, visualization of, and compensation for brain shift using intraoperative magnetic resonance imaging," *Neurosurgery*, vol. 47, no. 5, pp. 1070–1080, 2000.

[321] S. K. Warfield, S. J. Haker, I.-F. Talos, C. A. Kemper, N. Weisenfeld, A. U. Mewes, D. Goldberg-Zimring, K. H. Zou, C.-F. Westin, W. M. Wells, *et al.*, "Capturing intraoperative deformations: research experience at Brigham and Women's Hospital," *Medical Image Analysis*, vol. 9, no. 2, pp. 145–162, 2005.

[322] S. Frisken, P. Unadkat, X. Yang, M. I. Miga, and A. J. Golby, "Intra-operative measurement of brain deformation," *Biomechanics of the Brain*, pp. 303–319, 8 2019.

[323] N. P. de Arriba, A. A. Ramos, V. M. Fernandez, M. d. C. R. Sanchez, J. R. G. Alarcon, M. A. A. Vega, N. Pérez, and J. R. G. Alarcon Sr, "Risk factors associated with inadequate brain relaxation in craniotomy for surgery of supratentorial tumors," *Cureus*, vol. 14, no. 5, 2022.

[324] S. Schnaudigel, C. Preul, T. Ugur, H. J. Mentzel, O. W. Witte, M. Tittgemeyer, and G. Hagemann, "Positional brain deformation visualized with Magnetic Resonance Morphometry," *Neurosurgery*, vol. 66, pp. 376–384, 2 2010.

[325] J. G. Rosenman, E. P. Miller, and T. J. Cullip, "Image registration: an essential part of radiation therapy treatment planning," *International Journal of Radiation Oncology, Biology, Physics*, vol. 40, no. 1, pp. 197–205, 1998.

[326] V. Gorbunova, P. Lo, H. Ashraf, A. Dirksen, M. Nielsen, and M. de Bruijne, "Weight preserving image registration for monitoring disease progression in lung CT," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 863–870, Springer, 2008.

[327] J. Blumenfeld, J. Carballido-Gamio, R. Krug, D. J. Blezek, I. Hancu, and S. Majumdar, "Automatic prospective registration of high-resolution trabecular bone images of the tibia," *Annals of Biomedical Engineering*, vol. 35, no. 11, pp. 1924–1931, 2007.

[328] M. M. Chakravarty, A. F. Sadikot, J. Germann, G. Bertrand, and D. L. Collins, "Towards a validation of atlas warping techniques," *Medical Image Analysis*, vol. 12, no. 6, pp. 713–726, 2008.

[329] D. W. Shattuck, M. Mirza, V. Adisetiyo, C. Hojatkashani, G. Salamon, K. L. Narr, R. A. Poldrack, R. M. Bilder, and A. W. Toga, "Construction of a 3D probabilistic atlas of human cortical structures," *NeuroImage*, vol. 39, no. 3, pp. 1064–1080, 2008.

[330] V. Gravemeier, A. Comerford, L. Yoshihara, M. Ismail, and W. A. Wall, "A novel formulation for Neumann inflow boundary conditions in biomechanics," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 28, no. 5, pp. 560–573, 2012.

[331] D. Brands, G. Peters, and P. Bovendeerd, "Design and numerical implementation of a 3-D non-linear viscoelastic constitutive model for brain tissue during impact," *Journal of Biomechanics*, vol. 37, no. 1, pp. 127–134, 2004.

[332] E. Walsh and A. Schettini, "Calculation of brain elastic parameters in vivo," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 247, no. 4, pp. R693–R700, 1984.

[333] K. Miller, "Constitutive model of brain tissue suitable for finite element analysis of surgical procedures," *Journal of Biomechanics*, vol. 32, no. 5, pp. 531–537, 1999.

[334] N. M. Newmark, "A method of computation for structural dynamics," *Journal of the Engineering Mechanics Division*, vol. 85, no. 3, pp. 67–94, 1959.

[335] A. Wittek, N. M. Grosland, G. R. Joldes, V. Magnotta, and K. Miller, "From finite element meshes to clouds of points: a review of methods for generation of computational biomechanics models for patient-specific applications," *Annals of Biomedical Engineering*, vol. 44, no. 1, pp. 3–15, 2016.

[336] B. Couteau, Y. Payan, and S. Lavallée, "The mesh-matching algorithm: an automatic 3D mesh generator for finite element structures," *Journal of Biomechanics*, vol. 33, no. 8, pp. 1005–1009, 2000.

[337] A. D. Castellano-Smith, T. Hartkens, J. Schnabel, D. R. Hose, H. Liu, W. A. Hall, C. L. Truwit, D. J. Hawkes, and D. L. Hill, "Constructing patient-specific models for correcting intraoperative brain deformation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 1091–1098, Springer, 2001.

[338] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[339] M. Wu and N. Goodman, "Multimodal generative models for scalable weakly-supervised learning," *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[340] R. Ranganath, D. Tran, and D. Blei, "Hierarchical variational models," in *International Conference on Machine Learning*, pp. 324–333, PMLR, 2016.

[341] L. Scarpace, T. Mikkelsen, S. Cha, S. Rao, S. Tekchandani, D. Gutman, J. H. Saltz, B. J. Erickson, N. Pedano, A. E. Flanders, *et al.*, "The Cancer Genome Atlas Glioblastoma Multiforme collection (TCGA-GBM)," *The Cancer Imaging Archive*, 2016.

[342] N. Pedano, A. E. Flanders, L. Scarpace, T. Mikkelsen, J. M. Eschbacher, B. Hermes, V. Sisneros, J. Barnholtz-Sloan, and Q. Ostrom, "The Cancer Genome Atlas Low Grade Glioma collection (tcga-lgg)," *(No Title)*, 2016.

[343] R. B. Puchalski, N. Shah, J. Miller, R. Dalley, S. R. Nomura, J.-G. Yoon, K. A. Smith, M. Lankerovich, D. Bertagnolli, K. Bickley, *et al.*, "An anatomic transcriptional atlas of human glioblastoma," *Science*, vol. 360, no. 6389, pp. 660–663, 2018.

[344] J. Liu, S. Cao, K. J. Imbach, M. A. Gritsenko, T.-S. M. Lih, J. E. Kyle, T. M. Yaron-Barir, Z. A. Binder, Y. Li, I. Strunilin, *et al.*, "Multi-scale signaling and tumor evolution in high-grade gliomas," *Cancer Cell*, vol. 42, no. 7, pp. 1217–1238, 2024.

[345] L.-B. Wang, A. Karpova, M. A. Gritsenko, J. E. Kyle, S. Cao, Y. Li, D. Rykunov, A. Colaprico, J. H. Rothstein, R. Hong, *et al.*, "Proteogenomic and metabolomic characterization of human glioblastoma," *Cancer Cell*, vol. 39, no. 4, pp. 509–528, 2021.

[346] A. Bilger, J. Dequidt, C. Duriez, and S. Cotin, "Biomechanical simulation of electrode migration for deep brain stimulation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 339–346, Springer, 2011.

[347] S. Ghose, J. Mitra, D. Mills, D. T.-B. Yeo, T. K. Foo, A. Golby, and S. Frisken, "Automatic non-rigid registration of preoperative MRI and intraoperative US for US-guided neurosurgery-a preliminary study," in *2021 IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, IEEE, 2021.

[348] J. Wang, X. Chen, Y. Zhang, M. Liu, Y. Wang, and H. Zhang, "Unsupervised multimodal 3d medical image registration with multilevel correlation balanced optimization," *arXiv preprint arXiv:2409.05040*, 2024.

Appendix A

Data augmentation

To synthetically increase the variability of the contrast-enhanced T1-weighted (ceT1) pre-operative magnetic resonance imaging (preMRI) used for the training of the network, several intensity-based data augmentation strategies were applied. Augmentation parameters were randomly sampled from predefined ranges, following procedures adapted from the work in nnU-Net [172]. In this context, $x \sim U(a, b)$ denotes that the variable x is sampled from a uniform distribution over the interval $[a, b]$, subject to a stochastically defined probability.

To prevent interpolation artifacts, no geometric transformations were applied to the displacement fields. Specifically, rotations, translations, scaling, shearing, and mirroring transformations present in the original data augmentation procedures cited above were intentionally excluded. All augmentations were applied exclusively to the MRI intensity volumes and executed in the following order on-the-fly during training:

1. **Gaussian noise.** Since all input volumes are normalized using z-score normalization, voxel intensities have approximately zero mean and unit variance. Zero-mean Gaussian noise was therefore added to each voxel with a probability of 0.15. The noise variance was sampled from $U(0, 0.1)$.

2. **Gaussian blur.** Spatial smoothing was applied with a probability of 0.20 per sample. The standard deviation of the Gaussian kernel σ was sampled from $U(0.5, 1.5)$.

3. **Brightness scaling.** Voxel intensities were multiplicatively scaled by a factor $x \sim U(0.7, 1.3)$ with a probability of 0.15.

4. **Contrast adjustment.** Voxel intensities were scaled by a factor $x \sim U(0.65, 1.5)$ with a probability of 0.15. Following scaling, intensity values were clipped to the original value range.

5. **Simulated low resolution.** With a probability of 0.25, samples were downsampled by a factor drawn from $U(1, 2)$ using nearest-neighbor interpolation and subsequently upsampled back to the original resolution using cubic interpolation.

6. **Gamma augmentation.** Gamma augmentation was applied with an overall probability of 0.15, following two equally likely variants. In the first variant, voxel intensities were normalized to the $[0, 1]$ range, transformed according to $i_{\text{new}} = i_{\text{old}}^{\gamma}$ with $\gamma \sim U(0.7, 1.5)$, and then rescaled to the original intensity range. In the second variant, intensities were inverted prior to the transformation, i.e., $1 - i_{\text{new}} = (1 - i_{\text{old}})^{\gamma}$, before being rescaled back.