Selective Image Segmentation and Tissue at Risk Assessment in Perfusion Weighted Imaging

Thesis submitted in accordance with the requirements of the University of Liverpool and National Tsing Hua University for the degree of Doctor in Philosophy by

Sukhdeep Singh Bal

3 April 2023

Contents

Acknowledgements

I'd like to begin by expressing my gratitude to my supervisors, Prof. Fan-pei Gloria Yang and Prof. Ke Chen for their encouragement, advice, and tolerance throughout my doctoral studies. I also want to extend my sincere thanks to Dr. Andreas Alpers, Dr. Ozgur Selsil and Dr. Rachel Bearon for their counsel and ongoing inspiration. Along with that, I want to thank Dr. Giia-Sheun Peng and Dr. Jiia Fu Lee for inspiring me to study medical imaging and giving me valuable insight.

I must also convey my appreciation for all of my CMIT colleagues who helped me over the past four years, especially Dr. Liam Burrows and Mr. Yiyao Zhang. I'd like to thank my family and friends, especially my parents and elder sister, for their unconditional support, reassurance, and encouragement throughout my life. Finally, I'd like to thank Miss Annie Lin, Miss Shirley Farrell, and the other support staff who deal with the paperwork for postgraduate research students at National Tsing Hua University and the University of Liverpool.

Abstract

Perfusion-weighted imaging (PWI) is a noninvasive Magnetic Resonance (MR)/Computed Tomography (CT) technique that assesses various hemodynamic parameters to examine blood flow in brain regions. These parameters are used in stroke patients to locate the penumbra or the tissue at risk, which can be salvaged with reperfusion therapies. Certain artefacts are associated with the imaging pipeline and segmentation methods used in scanner softwares to estimate perfusion parameters. The focus of this thesis is the development of an image analysis pipeline for perfusion parameter estimation and segmentation models. We concentrate on difficult problems such as arterial region segmentation on perfusion weighted images and tumor segmentation on images with low contrast, in-homogeneous intensity, and non-smooth edges.

We begin with developing a arterial region segmentation model in the variational framework. We propose a new model in which geometric constraints are incorporated into a distance function. The modified model employs discrete total variation in the distance term and locates arterial regions by minimizing the energy of a convex functional, outperforming previous selective segmentation works that typically employ either a cost function or a clustering-based approach. This enhancement enables our model to effectively select an arterial region that performs well in identifying tissue at risk. Another work investigates whether fitting a hemodynamic model to the Arterial input function (AIF) obtained from arterial segmentation and minimizing the partial volume effect during AIF selection improves volumetric estimation of core and penumbra in stroke patients.

In the second half of this thesis, we propose an efficient framework for selective segmentation using a new region force term and a geodesic distance penalty based on a discrete TV formulation. The proposed model is user-independent and allows for precise segmentation in tumor images and medical images with non-homogeneous, non-smooth, and scraggy boundary edges. A chapter is dedicated to integrate the variational segmentation method with deep learning. Despite being extremely popular recently, deep learning techniques are frequently constrained by the need for sizable sets of labelled data. We demonstrate how labels can be supplemented by using a variational method as a loss function in a unsupervised training algorithm for brain tumour segmentation.

List of Publications

Peer-reviewed Journal Publications

- Sukhdeep Singh Bal, Ke Chen, Fan-pei Gloria Yang, Giia-Sheun Peng. (2022). Arterial input function segmentation based on a contour geodesic model for tissue at risk identification in ischemic stroke. Medical Physics. 49:2475–2485. https://doi.org/10.1002/mp.15508.

- Sukhdeep Singh Bal, Fan-pei Gloria Yang, Yueh-Feng Sung, Ke Chen, Jiu-Haw Yin, and Giia-Sheun Peng. (2022). Optimal Scaling Approaches for Perfusion MRI with Distorted Arterial Input Function in Patients with Ischemic Stroke. Brain Sciences. 2022; 12(1):77. https://doi.org/10.3390/brainsci12010077.

- Fan Pei Gloria Yang, Sukhdeep Singh Bal, Jia-Fu Lee, and Chia-Chi Chen. (2021). White Matter Differences in Networks in Elders with Mild Cognitive Impairment and Alzheimer's Disease. Brain Connectivity. 11(3):180-188. doi: 10.1089/brain.2020.0767.

- Fan Pei Gloria Yang, Sukhdeep Singh Bal, Yueh-Feng Sung, Giia-Sheun Peng. (2020). Mathematical Framework of Deconvolution Algorithms for Quantification of Perfusion Parameters. Acta neurologica Taiwanica. Sep;29(3):79-85. PMID:32996115.

- Sukhdeep Singh Bal, Fan-pei Gloria Yang, Ke Chen. (2023). Selective Geodesic Variational model with new Region force for Medical image Segmentation. (*submitted*)

- Sukhdeep Singh Bal, Fan-pei Gloria Yang, Nai-Fang Chi, Jiu Haw Yin, Tao-Jung Wang, Giia Sheun Peng, Ke Chen, Cang - I - Chen. (2023). CNN based Curve Fitting leads to better estimates of Core and Penumbra in association with clinical measures in Stroke patients.(*submitted to insights into imaging*)

Conferences and presentations

- Sukhdeep Singh Bal, Fan-pei Gloria Yang, Nai-Fang Chi, Jiu Haw Yin, Tao-Jung Wang, Giia Sheun Peng, Ke Chen, Cang - I - Chen. Core and Penumbra estimation using Deep learning in association with clinical measures in Computed Tomography Perfusion. European Congress of Radiology (ECR 2023) (Poster). Place: Vienna, Austria. Date: 3 March 2023.

- Sukhdeep Singh Bal, Fan- Pei Gloria Yang, Jia -Fu Lee. White Matter Differences in Networks in Elders with MCI and AD. Brain Connects 2022 conference, 42-43. Place: Brain and Mind Research Center, Nagoya University, Nagoya, Japan. Date: 8 August 2022.

- Sukhdeep Singh Bal, Fan Pei Gloria Yang, Ke Chen. Talk on Functional Connectivity for cognitive control in elders. 9th Annual Bi-Lateral conference between National Tsing Hua University, Taiwan and University of Liverpool, UK. Place: National Tsing Hua University, Hsinchu, Taiwan. Date: 20/11/2019.

- Sukhdeep Singh Bal, Fan Pei Gloria Yang, Ke Chen. Talk on Artefacts in Perfusion weighted imaging in Ischemic stroke. 10th Annual Bi-Lateral conference between National Tsing Hua University, Taiwan, and University of Liverpool, UK. Place: University of Liverpool, Liverpool, United Kingdom. Date: 25 /05/2021.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Project Outline

This thesis describes the research findings of a four-year Dual PhD degree programme between National Tsing Hua University (NTHU), Taiwan and University of Liverpool (UoL), United Kingdom. Under the joint supervision of Prof. Fan-Pei Gloria Yang (NTHU) and Prof. Ke Chen (UoL), research was carried out for two years at the Center of Cognition Sciences, NTHU, and two years at the University of Liverpool in alternate years.

Conventional Computed Tomography (CT) and Magnetic Resonance (MR) imaging are not sufficiently sensitive to evaluate acute stroke [2]. Perfusion-weighted imaging (PWI) is a noninvasive MR/CT technique that infers how blood traverses the brain's vasculature by assessing various hemodynamic parameters such as cerebral blood volume, cerebral blood flow, mean transit time, and time to peak [3]. In stroke patients, these parameters are used to locate the penumbra and core [4, 5, 2, 3, 6]. Penumbra refers to brain regions that are on the verge of infarction but are still salvageable if reperfused [7, 3, 8]. The infarct core is the tissue that has already infarcted or will infarct regardless of reperfusion [8]. Risk of errors in core and penumbra volume estimations leads to misleading conclusions. Part of the risk stems from using established and more standard software, which employs simple mathematical models that are incapable of delivering the required accuracy and sensitivity [2].

The main aim of this project was to exploit novel variational approaches to create an image analysis pipeline for perfusion weighted imaging to increase our understanding of core and penumbra evaluation. This may aid in predicting clinical response to early reperfusion. Alongside, we purposed new segmentation models that could detect tumour boundaries in brain images when the boundary edges were non-homogeneous, non-smooth, and scraggy. Furthermore, although not initially part of the project, in the final section of the thesis we train a deep learning network to perform segmentation tasks on tumour images.

The main objectives / tasks were:

1. Familiarization with state of art techniques and software tools used in stroke imaging. This included both mathematical literature for image processing as well as patient data from acquisition in clinics. This was done in year one at NTHU, as discussed in chapters 1, 2, and 3.

2. Advanced segmentation model development for arterial segmentation on perfusion-weighted images. This was completed during year two at UoL and is discussed in chapter four.

3. Current perfusion theory algorithms were improved in order to calculate the core and penumbra in stroke patients. Proposed image analysis pipeline was validated for accurate patient data. This was completed in third year at NTHU and is discussed in chapter 5.

4. Advanced segmentation model development for multimodal images. This was done at UoL and is covered in detail in chapter 6 and chapter 7.

1.2 Perfusion Weighted Imaging

This section is based on authors published paper [9]. In the following section basic principles associated with Perfusion weighted imaging (PWI) are introduced. A brain perfusion image is an indicator of brain functioning as these images can be modelled to estimate the amount of blood taken up in different brain regions [2, 10, 8]. Perfusion data sets are used in the diagnosis of two types of stroke: ischemic (lack of blood flow) and hemorrhagic (bleeding in brain vessels) [8]. For general treatment and triaging decisions, Computed Tomography Perfusion (CTP) or Magnetic Resonance Perfusion images (MRP) are usually used as diagnostic tools [11].

The passage of blood to the capillary bed in brain tissue is referred to as cerebral perfusion. Perfusion MRI is a type of MRI sequence that assesses cerebral hemodynamic parameters (Figure 1.1) such as Cerebral Blood Flow (CBF), Mean Transit Time (MTT), Cerebral Blood Volume (CBV), T-MAX, and Time To Peak (TTP) [2]. Cerebral Blood Flow (CBF) is the rate of capillary blood flow in brain tissues. Cerebral Blood Volume (CBV) is the amount of blood in a given amount of brain tissue. Mean Transit Time (MTT) is the amount of time a given volume of blood spends in the cerebral vessel. T-MAX is the time interval at which the residue function attains its maximum or when the contrast shows its highest effect. T-MAX is calculated in terms of seconds. TTP is time interval (seconds) where a first highest peak is attained by concentration curves. To create perfusion maps of different hemodynamic parameters like CBF, CBV, and MTT, signal loss in cerebral vessels

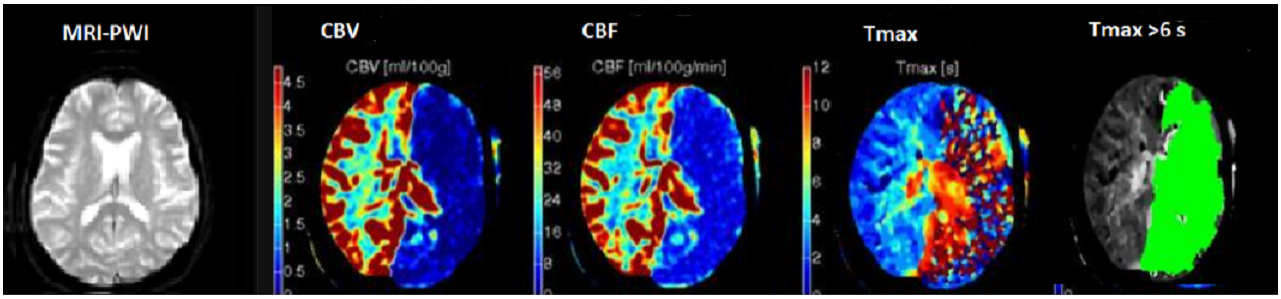


Figure 1.1: Perfusion Images: Input MRI PWI image, corresponding perfusion parameter maps of Cerebral Blood Volume (CBV), Cerebral Blood Flow (CBF) and T-MAX (Left to Right). Tmax> 6 is the region known as tissue at risk (Right).

is measured from PWI-MRI scans after the injection of a bolus of an MR paramagnetic contrast agent or tracer [5, 8]. These perfusion parameter maps or images are the final output that clinicals need to decide if the patients require reperfusion therapies. Accurate visual description of CBF, MTT, and CBV is crucial when evaluating tissue with restricted blood flow that may still be salvageable with proper treatment and reperfusion therapies [4].

A useful paradigm to select stroke patients for reperfusion therapies involves assessment of irreversibly damaged infarct core and the extent of hypoperfused tissue at risk or penumbra from CTP or MRP brain images. Accurate ischemic core and penumbra volumes are strong predictors of patient triage, as 87.0 % of patients have a favourable clinical outcome after endovascular thrombectomy (EVT) [12, 13]. Cerebral reperfusion following an accurate estimation of the core and penumbra results in the rescue of endangered brain tissues, with a significant correlation to the patients' clinical improvement. Tmax $>$ 6 seconds (Figure 1.1) for penumbra and relative CBF $\leq$ 30 % (a.u) for core are the commonly used thresholds to determine the optimal core and penumbra volumes [8, 6].

Next, we illustrate how the perfusion parameters are derieved from the perfusion images. Visually this is represented as a flowchart in Figure 1.2. In the perfusion model PWI scans are the initial input data. The image intensity function $S(t)$ in these images represents the signal loss information recorded during the administration of contrast agent injections to stroke patients [8, 5]. The following relationship is used to convert MR signal intensity functions $S(t)$ to concentration functions C(t) [5]-

$$C(t) = \frac{-k}{TE} ln(\frac{S(t)}{S_0}) \tag{1.1}$$

Here, TE is time to echo of the scanner, S_o is the baseline signal detected from initial scans, $S(t)$ is the signal time function and k is a proportionality constant. The baseline signal is less than the signal $S(t)$ and to prevent negative concentrations , k can be chosen as a negative constant. To estimate perfusion parameters, prior information in form of Arterial

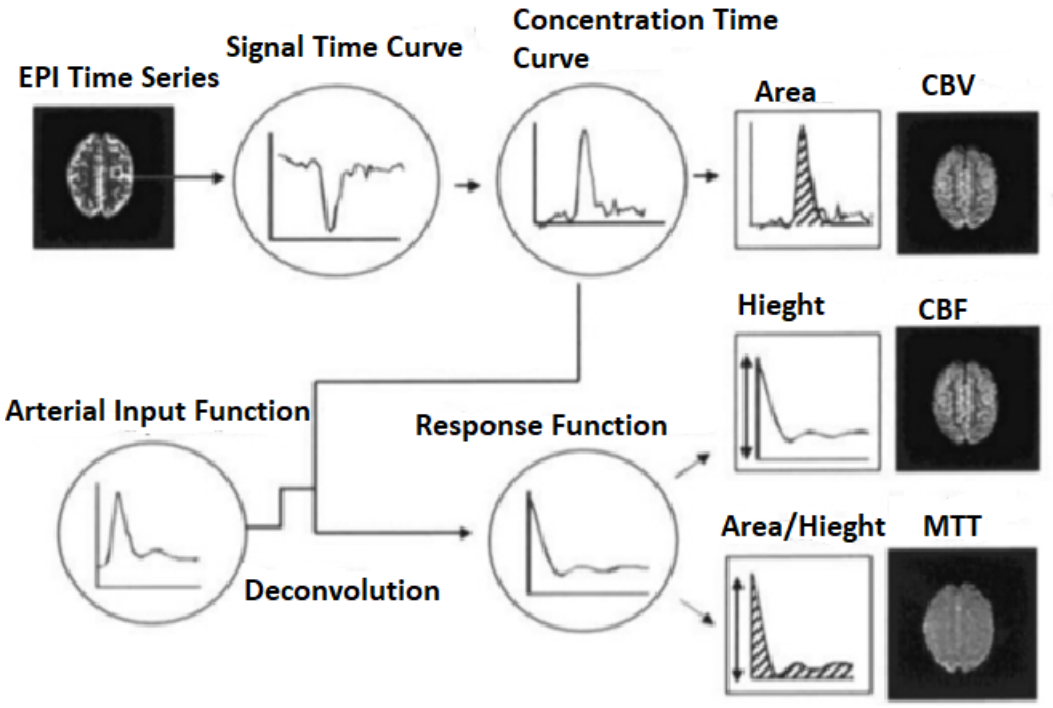


Figure 1.2: Flowchart of perfusion parameter estimations. EPI refers to Echo planar imaging.

Input function (AIF) from arterial region is essential [8]. The Arterial Input Function (AIF) is the time-dependent concentration of contrast agent supplied by the artery to the brain vessels [5, 8, 4, 2]. As AIF, a single or multiple pixels on the brain axial images can be chosen. The ideal AIF pixels should be in the region of large vessel, with early bolus arrival or early time-to-peak and also time dependent concentration curve of these AIF pixels should have maximum peak concentration with a large area [3]. The Tracer-kinetic perfusion theory suggests that tracer concentrations $C(t)$ in all brain regions are determined by the convolution of residue function $R(t)$ and the arterial input function C_a [5]. Once $R(t)$ is evaluated after deconvolving the equation 1.2, then the perfusion parameters are evaluated by the following relations.

$$C(t) = C_a(t) \otimes R(t) \tag{1.2}$$

$$\text{CBF} = 100.60[max(R(t))] \tag{1.3}$$

$$\text{CBV} = 100.\frac{\int C_t(t)dt}{\int C_a(t)dt} \tag{1.4}$$

$$\text{MTT} = \frac{60.CBV}{CBF} \tag{1.5}$$

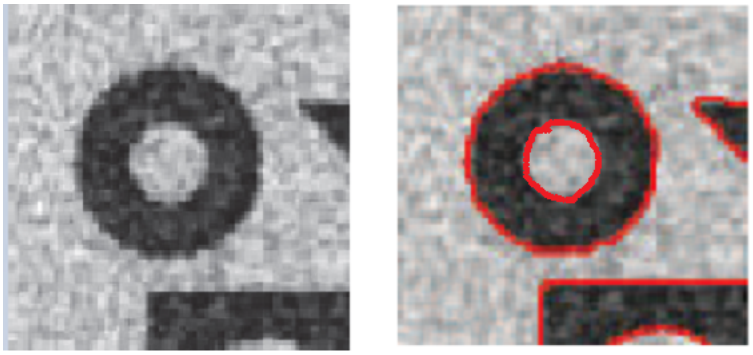


Figure 1.3: An example demonstrating a typical segmentation. Here, the image has been partitioned into segments. The right hand figure shows the segments, with the boundary separating the segments highlighted in Red.

The AIF concentration curve is the central quantity for calculating the perfusion parameters required by the clinicians. This implies that selection of AIF will affect the final output results which the clinical need in the form of brain images representing blood flow or volume in stroke patients. If an AIF with lower peak concentration and with delayed time to peak is used then the final results for perfusion parameters i.e blood flow or blood volume may not be accurate, sharp and specific [8]. This may lead to misleading decisions for the therapies or treatments to be administered to stroke patients. The clinicals are more concerned about the final output brain images representing blood flow or blood volume.

1.3 Image Processing and Image Segmentation

A digital image is a two-dimensional (2-D) data matrix $Z(x, y)$, with each element (x,y) representing a pixel [14, 15]. Each pixel or image element is associated with a finite, discrete numeric representation known as pixel intensity. Binary, gray-scale, or RGB (Red-Blue-Green) images are examples of digital images. Binary images stored as single bits are frequently referred to as bi-level or two-level, as each pixel has an intensity value of 1 or 0, representing black and white colour, respectively [14]. Each pixel in a gray-scale image represents intensity information, with intensity ranging from black (0) at the lowest intensity to white at the highest (255). Each pixel in a coloured image represents the intensity produced by combining three channels or colours, namely red, green, and blue [16]. In this thesis, we mainly work with gray-scale brain Magnetic Resonance (MR) and Computerised Tomography (CT) brain images.

Image processing is the digital manipulation of images using advanced algorithms to extract useful information from 2-D matrix image data. Its applications range from medicine to remote sensing, passing through geological processing and entertainment [14]. Image pro-

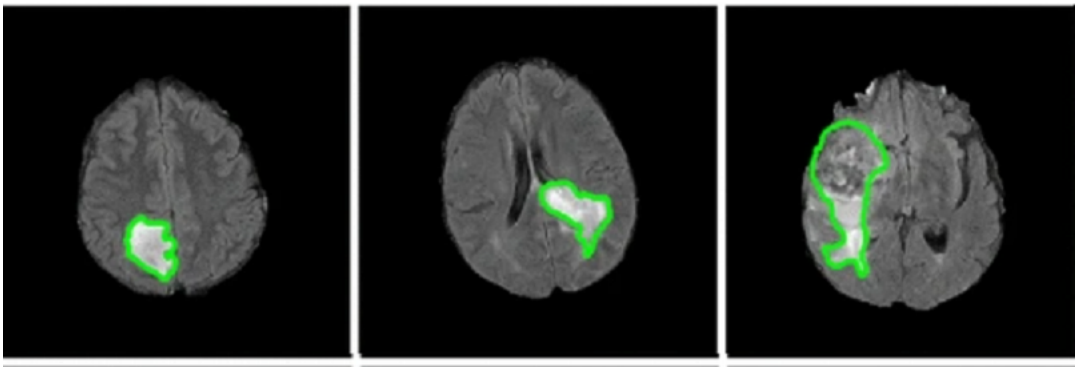


Figure 1.4: Example of Tumour Segmentation

cessing is further classified based on the type of information to be extracted from the input image. Detecting objects that are not visible in an image is part of image visualisation [16]. Image segmentation is the process of separating or segmenting a single or multiple objects in an image [17]. Image sharpening and restoration are processes that are used to create a clear, noise-free image from an original image [18, 19]. Pattern recognition involves determining various patterns surrounding the image's multiple objects [16]. Image registration involves transforming various scenes or slices of an image data into one coordinate system [18]. This thesis focuses on Image Segmentation and Image Processing applications to Perfusion weighted imaging (PWI), a noninvasive technique used to predict brain regions with decreased perfusion or blood flow in stroke patients.

In practise, it is possible that the viewer is not interested in all parts of the image, but only needs to visualise a specific region based on common characteristics. For example, a clinical might want to extract the tumour-containing region from a brain axial CT/MR image (Figure 1.4). Image segmentation is the process of dividing an input image into a number of different regions with similar pixel characteristics in order to identify objects or boundaries [20, 21, 22, 18]. The input in image segmentation is an image, and the output is a clear visualisation or image of one or more objects separated from the original input image (Figure 1.3). Image tumour segmentation is an important medical task [16]. Tumour segmentation facilitates in diagnosis, surgical planning, and also allows for the analysis of tumour shape and size prior to surgery, as well as comparison of tumour volume at various stages of treatment [15, 23].

The complexity of Brain CT/MR and PWI images, as well as the irregular shapes of the organs to be segmented, in-homogeneous intensities at tumour boundaries, noise levels, and variable contrasts at tumour edges, make segmentation a difficult task [14, 16]. No single image segmentation algorithm in the literature performs adequately for all types of images with variable noise levels. Furthermore, image segmentation performance varies from image to image. According to recent research, variational techniques are promising models for solving segmentation problems [22, 23, 24]. Finding the variational model solution involves

minimising nonlinear functionals. The minimiser, represented by a contour, locates the object's boundaries, separating it from other objects in the image [15].

The aims of the project include using state of the art variational segmentation models and developing new segmentation models for arterial region and tumour segmentation on brain images. The initial problem we consider here in chapter 4 is regarding application of variational segmentation model for the selection of arterial region or Arterial Input Function (AIF) to detect brain regions with reduced blood flow. Furthermore in chapter 5, we investigated whether using a deep learning approach for Arterial Input Function (AIF) selection improves the volumetric/visual estimation of penumbra and core regions. Along with optimal AIF, we studied the substantial influence of the Partial Volume Effect (PVE) induced by the MR scanner on PWI images in chapter 5. We investigated potential approaches for reducing volume averaging artefacts in order to achieve adequate results for identifying brain regions with reduced perfusion in stroke patients.

In chapter 6, we present novel selective segmentation methods for multi-modal image segmentation that are independent of user input. By using a region force term and a geodesic distance penalty based on a discrete TV formulation, we proposed an efficient framework for selective segmentation. We used empirical and quantitative evaluations on multi-modal (MRI, CT) images to compare the performance of our model to that of other state-of-the-art models. Following that, in chapter 7 we will discuss a deep learning architecture for unsupervised learning that uses information from the proposed variational segmentation model to detect tumours and lesions with high visual variance, shape difference, and ambiguous boundaries.

1.4 Chapters of Thesis

The chapters of thesis are organised as -

Chapter 2

This chapter introduces some basic mathematical tools that will be used throughout the thesis. We will briefly introduce useful preliminary definitions, theorems, and examples from normed linear spaces, calculus of variations, bounded space of variations, regularisation for image processing, and level set methods. A discussion of the notion of convexity, discretisation of partial differential equations (PDE) on regular domains using finite difference methods, as well as an overview of numerical methods such as additive operator splitting will be addressed.

Chapter 3

This chapter provides an overview of the literature's relevant segmentation models. We present global variational segmentation models first, followed by selective segmentation models. To segment the input image, these segmentation models employ variational methods based on calculus of variation. Each segmentation model includes an energy functional, and the goal is to find the functional's minimizer. The minimizer takes the form of a partial differential equation, and the solution corresponds to the image's segmented region. The notions and ideas presented in this chapter will be used in chapter 4 and chapter 6.

Chapter 4

This chapter is based on published paper [3]. The Arterial Input Function (AIF) is required as an input for core and penumbra estimation in stroke patients. As a result, the AIF segmentation paradigm has clinical significance. In this chapter, we propose a new technique for addressing the problem of AIF selection, which is based on a variational segmentation model that incorporates geometric constraints into a distance function. The modified model utilizes discrete total variation in the distance term and locates the arterial regions by minimising energy of a convex functional. The proposed model was compared to state of the art methods. The proposed segmentation model predicted AIF curves with a higher amplitude and earlier time to peak, as well as good performance in identifying the tissue at risk.

Chapter 5

This chapter is based on two papers, one that have been accepted in a conference proceedings and the other is published [6]. In the first part of this chapter, we looked into potential approaches for minimising the partial volume effect during AIF selection. The spatial resolution used in perfusion MRI and the average size of major arteries make a degree of partial volume unavoidable during AIF measurement. The proposed scaling method yields more reasonable absolute perfusion parameter values, as evidenced by higher mean CBF/Tmax values and CBF/Tmax images. This ensured that both the core and the infract region were not overlooked.

In second section, we investigate whether fitting a hemodynamic model to the AIF improves the volumetric estimation of core and penumbra in stroke patients in relation to clinical measures. With Institutional Review Board approval, the study included 160 acute stroke patients (male = 87, female = 73, median age = 73 years). CTP imaging, NIHSS, and ASPECTS grading had been performed on the patients. CNN AIF was created by training a Convolutional Neural Network (CNN) model to fit a raw AIF curve to a Gamma variate function. The core and penumbra volumes were estimated using CNN AIF.

Chapter 6

This chapter is based on the author's paper that has been submitted. In this chapter, we

propose an efficient framework for selective segmentation using a new region force term and a geodesic distance penalty based on a discrete TV formulation. The proposed segmentation model is robust and competitive with the current state-of-the-art, according to empirical evaluations. We compared the performance of our model to other segmentation methods by evaluating segmentation scores of multi-modal (MRI, CT) segmented images. In comparison to previous segmentation models, the proposed model is user-independent and allows for precise segmentation of objects in medical images with non-homogeneous, non-smooth, and scraggy boundary edges.

Chapter 7

In this chapter, we combine variational and deep learning approaches for segmentation. We demonstrate how a variational segmentation solution can be implemented as a loss function for a neural network. This allows unlabeled data to be integrated into the training set without the need for intensive ground truth labels. In this chapter, we combine the selective geodesic variational model with new region force from chapter 6 with a deep learning method to aid in the segmentation of challenging images.

Finally, In this last section, we propose possible future research directions derived from the work presented in the thesis.

Chapter 2

Mathematical Preliminaries

This chapter covers some fundamental mathematical theories, definitions, examples, and theorems that will help readers understand the mathematical concepts underlying the segmentation models in chapters 3, 4, and 6. This chapter's content is typically drawn from linear algebra or advanced calculus literature. Following basic definitions, this chapter includes a useful discussion of inverse problems, regularization methods, discretization of partial differential equations (PDEs) on regular domains, and methods for solving nonlinear systems of equations.

2.1 Linear Vector Spaces

Definition 2.1.1 (Field)**.** *F is a field with two operations addition($+$) and multiplication (.) which satisfy the following axioms -*

1. *If $x, y \in F$ then $x + y \in F$.*
2. *If $x + y = y + x$ for all $x, y \in F$*
3. *$x + (y + z) = (x + y) + z$ for all $x, y, z \in F$*
4. *Foe every $x \in F$ there exists $-x \in F$ such that $x + (-x) = 0$.*
5. *If $x, y \in F$ then $x \cdot y \in F$.*
6. *If $x.y = y.x$ for all $x, y \in F$.*
7. *$x.(y.z) = (x.y).z$ for all $x, y, z \in F$.*
8. *There exists two elements, $0, 1 \in F$ such that for $x \in F$, $x + 0 = x$ and $x \cdot 1 = x$.*
9. *For every $x \neq 0 \in F$ there exists an element $x^{-1} \in F$ such that $x \cdot x^{-1} = 1$.*

Example 2.1.1 (Field Examples)**.** *The set of complex numbers $\mathbb{C}$, the set of real numbers $\mathbb{R}$, set of rational numbers $\mathbb{Q}$, rational functions fields, algebraic function fields, algebraic number fields are some examples of fields.*

Definition 2.1.2 (Linear Vector Space)**.** *Let $\mathcal{F}$ be a field and V be a non-empty set with two defined operations of addition and scalar multiplication. For $u, v \in V$ sum is represented as $u+v$, and if $\alpha \in \mathcal{F}$, the multiplication of u by α is given by αu. V is a linear vector space if the following properties hold for any arbitrary $u, v, w \in V$ and for all scalars elements $\alpha, \beta \in \mathcal{F}$.*

1. *$u + v \in V$ (closure)*
2. *$u + v = v + u$ (commutativity)*
3. *$(u + v) + w = u + (v + w)$ (associativity of addition)*
4. *There exists an element $0 \in V$ such that $u + 0 = 0 + u = u$ (existence of identity element)*
5. *For all $u \in V$ there exists an element $-u \in V$ such that $u + (-u) = 0$ (existence of inverse under addition)*
6. *$\alpha u \in V$ (closure on scalar multiplication)*
7. *$\alpha(u + v) = \alpha u + \alpha v$ (distributivity)*
8. *$(\alpha + \beta)u = \alpha u + \beta u$ (distributivity)*
9. *$\alpha(\beta u) = (\alpha\beta)u$ (associativity under scalar multiplication)*
10. *There exists an element $1 \in V$ such that $u \cdot 1 = u$ (existence of identity of scalar multiplication)*

Example 2.1.2 (Linear Vector Space Examples)**.**

- *F^{mn} the set of mn matrices with entries in Field F, set of polynomials with coefficients in Field F, $\mathbb{R}^n$ and $\mathbb{C}^n$ for all $n \in \mathbb{N}$ are examples of linear vector space.*

Definition 2.1.3 (Norm)**.** *Let V be a vector space over a field $\mathcal{F}$. A real valued function $N : V \to \mathbb{R}$ is a norm if all $u, v \in V$ follow :*

1. *$N(u) = 0$ if $u = 0$.*
2. *$N(u) > 0$ for all $u \neq 0 \in V$.*
3. *$N(\alpha u) = |\alpha| N(u)$ for all scalars α.*
4. *$N(u + v) \leq N(u) + N(v)$.*

Definition 2.1.4 (Semi-norm)**.** *A semi-norm is defined similarly as a norm, with the difference that the above properties 1-2 are replaced with-*

$$N(u) \geq 0, \text{ for all } u \in V.$$

The norm of a vector u is usually represented by $||u||$.

Example 2.1.3 (Examples of norms)**.**

- *Euclidean norm of a vector, where $x \in \mathbb{R}^d$, defined as an example of norm that gives the ordinary distance from x to the origin.*

$$||x|| = \sqrt{x_1^2 + x_2^2 + \ldots + x_d^2}.$$

- *Another norm is Infinity norm, defined as:*

$$||x||_\infty = \max(|x_1|, |x_2|, \ldots, |x_d|).$$

- *Another widely used norm is ℓ_p-norm. $x \in \mathbb{R}^d$, $p \geq 1$, then the ℓ_p-norm of x is defined as:*

$$||x||_p = \left(\sum_{i=1}^{d} |x_i|^p \right)^{\frac{1}{p}}.$$

- *If for a continuous function f on Ω we have $\int_\Omega |f(x)|^p dx < \infty$, then the L^p norm is given by:*

$$||f(x)||_{L^p} = \Big(\int_\Omega |f(x)|^p dx \Big)^{\frac{1}{p}}.$$

  If $p = \infty$, then the norm takes the form:

$$||f(x)||_{L^\infty} = \sup_x |f(x)|.$$

- *Total Variation norm.*

  For $u \in \Omega \subset \mathbb{R}$, Total Variation (TV) norm of u is :

$$TV(u) = \int_\Omega |\nabla u| \, d\mathbf{\Omega}.$$

Definition 2.1.5 (Normed Linear Space)**.** *A vector space V with a norm is called a normed linear space.*

Definition 2.1.6 (Cauchy Sequence and Completeness)**.** *Let $\{v_i\}_{i=1}^{\infty}$ be a sequence in a normed linear space V. $\{v_i\}_{i=1}^{\infty}$ is Cauchy sequence if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that:*

$$||v_i - v_j|| < \epsilon, \qquad \text{for all } i, j > N.$$

if Cauchy sequence $\{v_i\}_{i=1}^{\infty} \subset V$ converges to an element $v \in V$ (i.e. $\lim_{i\to\infty} v_i = v$), then V is complete space.

Definition 2.1.7 (Banach Space)**.** *A complete normed linear space V is a Banach space.*

Definition 2.1.8 (Open set)**.** *For a normed space V a subset $S \subset V$ is said to be open if for each point $u \in S$ there exists $\delta > 0$ such that $|u - v| < \delta$ for all $v \in S$.*

Definition 2.1.9 (Complement of a set and Closed set)**.** *The complement of set A in V is set of all points $x \in V$ which does not belong to A. A subset $A \in V$ is closed if its complement is open.*

Definition 2.1.10 (Lipschitz Condition)**.** *if for any point $x, y \in S \subset R$ for some $M \in R$ the real function $f : S \to R$ satisfies*
$|f(x) - f(y)| \leq M|x - y|$
then f is said to satisfy the Lipschitz condition in S and f is known as Lipschitz function. R is set of set of real numbers.

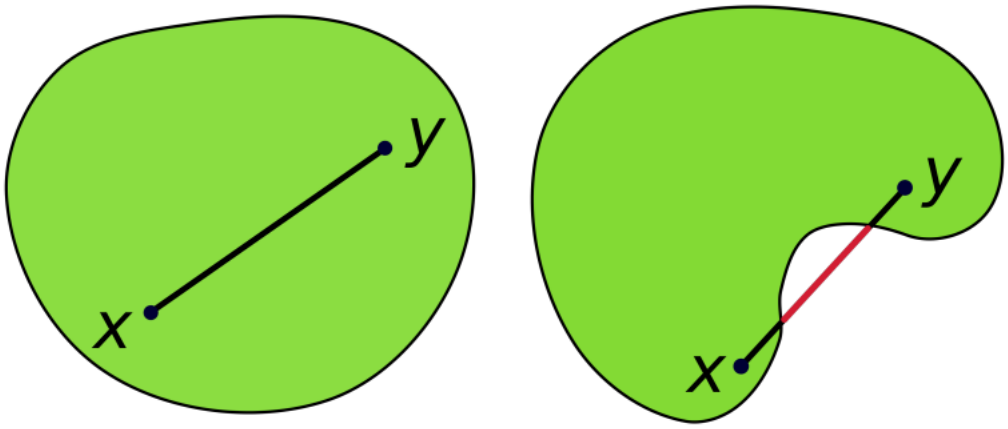
Definition 2.1.11 (Inner Product)**.** *V is a vector space defined over a scalar field $\mathcal{F}$. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathcal{F}$ is known as an inner product on V if all arbitrary $u, v, w \in V$ abide the following properties :*

1. $\langle u, u \rangle \geq 0$.
2. $\langle u, u \rangle = 0$ *if and only if* $u = 0$.
3. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$.
4. $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$, *for all* $\alpha \in \mathcal{F}$.
5. $\langle u, v \rangle = \overline{\langle v, u \rangle}$.

Example 2.1.4 (Examples of Inner Products)**.**

- *for $x, y \in \mathbb{R}^d$, the Euclidean inner product on $\mathbb{R}^d$ is :*

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_d y_d.$$



(a) A convex set (b) A non-convex set

Figure 2.1: Convex set and non convex set

- *$C[a,b]$ is the vector space of continuous real-valued functions defined on the interval $[a,b]$. With $g,h \in C[a,b]$, inner product on $C[a,b]$ is defined as:*

$$\langle g,h\rangle = \int_a^b g(x)h(x)\,dx.$$

Definition 2.1.12 (Hilbert Space)**.** *If every Cauchy sequence in a vector space V with inner product $\langle u,v\rangle$ converges to an element of V, then V is said to be a Hilbert space. A Hilbert space is always a Banach space, but the inverse is not necessarily true.*

Definition 2.1.13 (Convex sets)**.** *A set U in a vector space V is defined to be convex if, for all $u,v \in U$ and all $\theta \in [0,1]$, the point w defined by:*

$$w = (1-\theta)u + \theta v$$

is in U. This means that every point on the line segment connecting u and v should lie in U. We can see visually in figure 2.1 that the line segment joining two points in a convex set lies in the set.

Definition 2.1.14 (Convex function)**.** *A function $f : U \to \mathbb{R}$ defined on a convex set U is a convex function if*

$$f(\theta u + (1-\theta)v) \le f(u) + (1-\theta)f(v),$$

for all $u,v \in U$ and $\theta \in [0,1]$. f is strictly convex if this inequality is always strict for $u \neq v$.

Example 2.1.5 (Examples of convex functions on $\mathbb{R}$ and $\mathbb{R}^n$)**.**

- *The function $f(x) = x^2$ defined on the domain $\mathbb{R}$, exponential function $f(x) = e^{ax}$ for $a \in \mathbb{R}$ and $a > 0$ or $a < 0$, defined on $\mathbb{R}$ are strictly convex.*

- *Another example for convex function are the norm $||x||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$, for $p \geq 1$, $||x||_\infty = \max_k(|x_k|)$.*

Definition 2.1.15 (Dual Space)**.** *let V be a normed linear space defined over a scalar field $\mathcal{F}$, equipped with a norm $||\cdot||$. Then the dual space V' of V, is the set of linear functions $f : V \to \mathcal{F}$. Dual norm $||\cdot||'$, on dual space V' is:*

$$||f||' = \sup\{|f(u)| : u \in V, ||u|| = 1\}.$$

Definition 2.1.16 (Reflexive Space)**.** *Let V be a normed linear space, V' be its dual space, and V'' be its bidual or dual of dual space. If the map, $h : V \to V''$ defined as*

$$h(f) = f(u),$$

for $f \in V', u \in V$, is an onto map then the dual space V' is said to be Reflexive space.

2.2 Inverse Problems

An inverse problem in science is the process of calculating from a set of observations the causal factors that produced them. Inverse problems are widely used in a variety of fields, and are of key importance in the field of imaging problems.

Inverse problems are either well-posed or ill-posed. Hadamard [25] defined the following criteria for a well-posed problem-

1. a solution exists;
2. solution should be unique;
3. solution's behaviour changes continuously with the initial conditions.

A solution can usually be calculated using a stable algorithm for well-posed problems. An inverse problem is said to be ill-posed if any of the three conditions listed above is not met. The third property is the most frequently violated by an ill-posed problem; imaging inverse problems frequently disregard the uniqueness property.

In this thesis, we will be concerned with discretizations of linear inverse problems of the form:

$$Ax = b$$

where the vector b represents measured data (typically with noise) and the matrix A represents the forward mapping. Given A and b, the aim is to compute an approximation of the unknown vector x. We are mainly interested in the problem which is ill-posed in the sense that the singular values of A gradually decay and cluster at zero. In chapter 1 section 1.2, we have the matrix system $C_a R(t) = C(t)$, so x is matrix $R(t)$, whereas A is C_a and b is $C(t)$. When the matrix is ill conditioned, i.e. has a high condition number, the computed solution is extremely sensitive to errors/noise in b. Regularization of the solution is thus required to produce stable solutions.

2.2.1 Regularisation

Regularisation is achieved via solving a penalised least squares problem of the form-

$$\arg\min_x ||Ax - b||_2^2 + \lambda^2 ||x||_2^2$$

where the penalty term $||x||_2^2$ is chosen to reflect the specific type of regularization that is suited for the problem. We obtain the classical Tikhonov regularization problem as [26].

A different way to achieve regularization is to apply an iterative method directly on the fit-to-data term and terminate the iterations when semi-convergence is achieved; that is, terminate when a desired approximation is obtained, but before noise starts to show up in the solution. Using an iterative method in this way is often referred to as iterative regularization.

Singular value decomposition (SVD) [27] is another way of regularising the solution system for a matrix system $Ax = b$. SVD decomposes A as -

$$A = U \sum V^T = \sum_{i=1}^{n} u_i \sigma_i v_i^T \tag{2.1}$$

Here, u_i and v_i are columns of U and V and $\sum$ is a matrix with diagonal elements as singular values $\sigma_1, \sigma_2 ... \sigma_n$. For the matrix system the solution x is computed as-

$$X = A^{-1} b = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i \tag{2.2}$$

The extremely large errors in the naive solution come from the noisy SVD components associated with the smaller singular values. The truncated SVD (TSVD) solution is obtained by retaining the first k components of the naive solution. The truncation parameter k ensures

all the noise-dominated SVD coefficients are discarded.

$$X = A^{-1}b = \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i \tag{2.3}$$

2.3 Calculus of Variations

In this section we discuss some well known Theorems and definitions from the literature of calculus of variations.

Definition 2.3.1 (Total Variation (TV)). *Let Ω be a bounded open subset of $\mathbb{R}^n$. The total variation of a function $u : \Omega \to \mathbb{R}$ is defined by*

$$TV(u) = \int_{\Omega} |\nabla u(\mathbf{x})| \, d\mathbf{x}, \tag{2.4}$$

for $\mathbf{x} \in \Omega$.

Total Variation is a widely used regulariser for many image processing problems [28]. In chapter 3, 4, and 6 we will discuss Total variation as a useful regulariser for image segmentation problems.

Definition 2.3.2 (Bounded Variation). *Let Ω be a bounded open subset of $\mathbb{R}^n$. Bounded variation is defined as:*

$$BV(\Omega) = \left\{ u \in L^1(\Omega) : TV(u) < \infty \right\}$$

Definition 2.3.3 (Gateaux derivative). *Let $J : V \to \mathbb{R}$ be a function defined on a Banach space V. The Gateaux derivative of J is defined as:*

$$\delta J(u; v) = \lim_{\epsilon \to 0} \frac{J(u + \epsilon v) - J(u)}{\epsilon} \tag{2.5}$$

for $u, v \in V$. If the limit exists for all $v \in V$ then we say that J is Gateaux differentiable at u.

2.3.1 Variation of a Functional

The Euler-Lagrange equation of a general functional is introduced here. Taking the first variation, this aims to identify the function for which the given functional is stationary. Consider a general functional on a normed linear space Ω , $\mathcal{J}(u) : \Omega \to \mathbb{R}$:

$$\mathcal{J}(u) = \int_{\Omega} L(\mathbf{x}, u(\mathbf{x}, \nabla u(\mathbf{x})) \, d\mathbf{x}, \tag{2.6}$$

The first variation means solving the following minimisation problem:

$$\min_u \mathcal{J}(u). \tag{2.7}$$

To solve the above minimisation problem, the necessary condition is that the for all test functions Gateaux derivative of $\mathcal{J}$ vanishes:

$$\delta\mathcal{J} = \delta J(u; v) = \lim_{\epsilon \to 0} \frac{J(u + \epsilon v) - J(u)}{\epsilon} = 0, \tag{2.8}$$

The quantity $\delta\mathcal{J}$ is the first variation of the functional $\mathcal{J}(u)$.

Theorem 2.3.1 (Gauss's Divergence Theorem)**.** *Let F be a continuously differentiable vector field in a bounded, closed domain $\Omega \subset \mathbb{R}^d$, whose boundary $\partial\Omega$ is smooth. Then, according to the divergence theorem :*

$$\int_\Omega (\nabla \cdot F)\, d\mathbf{x} = \int_{\partial\Omega} F \cdot \mathbf{n}\, ds, \tag{2.9}$$

here, $(\nabla \cdot F)$ is the divergence of the vector field F and $\mathbf{n}$ is the unit outward normal vector of $\partial\Omega$.

Definition 2.3.4 (Local minimiser)**.** *Let F: $\Omega \to \mathbb{R}$ be a real valued functional defined on normed space Ω. $F(u)$ has a local minimiser u^* if ϵ exist such that*

$$F(u^*) \leq F(u), \text{ for all } u \in \mathcal{N}_\epsilon(u^*), \tag{2.10}$$

here, $\mathcal{N}_\epsilon(u^*) = \{u \in \Omega : ||u - u^*|| < \epsilon\}$ is a small neighbourhood around u^*.

Definition 2.3.5 (Global minimiser)**.** *Let $J : \Omega \to \mathbb{R}$ be a real valued function defined on a normed space Ω. $J(u)$ has a global minimiser u^* if:*

$$J(u^*) \leq J(u), \text{ for all } u \in \Omega. \tag{2.11}$$

Definition 2.3.6 (Stationary point)**.** *Let $J : \Omega \to \mathbb{R}$ be a real valued functional on a normed space Ω. If J is Gateaux differentiable at $u \in \Omega$ for all test functions $v \in \Omega$, then u is a stationary point of J if $\delta J(u; v) = 0$.*

Definition 2.3.7 (Euler-Lagrange equation)**.** *The equation $\delta J(u; v) = 0$ is called the Euler-Lagrange equation for the minimisation problem $\min_u \mathcal{J}(u)$.*

Theorem 2.3.2 (Necessary condition for a local minimiser)**.** *For a Gateaux differentiable functional $J : \Omega \to \mathbb{R}$, if u is a local minimiser of $J(u)$, then u is a stationary point of $J(u)$.*

Example 2.3.1 (Euler-Lagrange equation for TV)**.** *Total Variation (TV) from Definition is:*

$$TV(u) = \int_\Omega |\nabla u(\mathbf{x})|\, d\mathbf{x},$$

defined on a domain $\Omega \subset \mathbb{R}^d$*. The Gateaux derivative of the Total Variation is:*

$$\begin{aligned}\delta TV(u;v) &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_\Omega \big(|\nabla(u+\epsilon v)| - |\nabla u|\big)\, d\mathbf{x}\\ &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_\Omega \big(|\nabla u| + \epsilon\frac{\nabla u\cdot\nabla v}{|\nabla u|} + \mathcal{O}(\epsilon^2) - |\nabla u|\big)\, d\mathbf{x}\\ &= \int_\Omega \frac{\nabla u\cdot\nabla v}{|\nabla u|}\, d\mathbf{x}\end{aligned}$$

Using integration and Gauss's theorem :

$$\int_\Omega \frac{\nabla u\cdot\nabla v}{|\nabla u|}\, d\mathbf{x} = \int_{\partial\Omega}\frac{\nabla u}{|\nabla u|}\cdot\mathbf{n}\, ds - \int_\Omega \nabla\cdot\Big(\frac{\nabla u}{|\nabla u|}\Big)v\, d\mathbf{x},$$

here, $\partial\Omega$ *is the boundary of* Ω *and* $\mathbf{n}$ *is the unit outward normal of* $\partial\Omega$*. For* $\delta TV(u;v) = 0$ *we are left with:*

$$\int_{\partial\Omega}\frac{\nabla u}{|\nabla u|}\cdot\mathbf{n}\, ds - \int_\Omega \nabla\cdot\Big(\frac{\nabla u}{|\nabla u|}\Big)v\, d\mathbf{x} = 0.$$

This should be true for all test functions $v \in \Omega$*. So, we can derive the following Euler-Lagrange equation:*

$$\nabla\cdot\Big(\frac{\nabla u(\mathbf{x})}{|\nabla u(\mathbf{x})|}\Big) = 0,\ \textit{for}\ \mathbf{x}\in\Omega,$$

with Neumann boundary conditions $\nabla u\cdot\mathbf{n} = \frac{\partial u}{\partial\mathbf{n}} = 0$*.*

2.4 Discretisation of Partial Differential Equations

In imaging problems, Euler-Lagrange equation is a partial differential equation (PDE). The solution of the PDE should be the minimiser. Analytic solutions are not possible in the vast majority of practical imaging problems, so we must solve the PDE by obtaining an approximation of the solution using numerical methods.

Definition 2.4.1 (Laplace operator)**.** *The Laplace operator is defined to be the divergence of the gradient of a function. For a scalar function f defined for $\mathbf{x} \in \mathbb{R}^d$, the Laplacian operator Δ of f is:*

$$\Delta f(\mathbf{x}) = \partial^2_{x_1} f + \ldots + \partial^2_{x_d} f$$

It is usually denoted as Δ, $\nabla^2, \nabla \cdot \nabla$ or $div \cdot \nabla$.

Many times in imaging problems, we must discretise and numerically solve a PDE. There are various methods available to discretise PDE problems, such as the finite element method, finite difference method, and finite volume method.

Considering two dimensional, $\Omega = [0,1] \times [0,1]$. In order to discretise on this domain we divide the domain into grid points by using $(n+1) \times (m+1)$ Cartesian grids with width $h_x = \frac{1}{n}$ and $h_y = \frac{1}{m}$ in the x and y directions respectively. This results in $n \times m$ grid points, and the grid point (i, j) is at the position :

$$(x_i, y_j) = \big(\frac{2i-1}{2} h_x, \frac{2j-1}{2} h_y\big), \text{ for } 1 \leq i \leq n \text{ and } 1 \leq j \leq m.$$

2.4.1 Finite Difference Schemes

If we consider the Taylor expansions on a discrete grid, then we can approximate continuous functions u of a PDE by

$$u(x_i + h_x, y_j) = u(x_i, y_j) + h_x \frac{\partial u}{\partial x}(x_i, y_j) + \mathcal{O}(h^2),$$

and

$$u(x_i - h_x, y_j) = u(x_i, y_j) - h_x \frac{\partial u}{\partial x}(x_i, y_j) + \mathcal{O}(h^2),$$

To approximate the partial derivative $\frac{\delta u}{\delta x}$, $\frac{\delta u}{\delta y}$at a particular point (x_i, y_j) we can use either of the following-

- **First order forward difference:**

$$\frac{\partial u}{\partial x}(x_i, y_j) \approx \nabla^+_x(u_{i,j}) = \frac{u(x_i + h_x, y_j) - u(x_i, y_j)}{h_x}.$$

- **First order backward difference:**

$$\frac{\partial u}{\partial x}(x_i, y_j) \approx \nabla^-_x(u_{i,j}) = \frac{u(x_i, y_j) - u(x_i - h_x, y_j)}{h_x}.$$

- **First order central difference:**

$$\frac{\partial u}{\partial x}(x_i, y_j) \approx \nabla_x^c(u_{i,j}) = \frac{u(x_i + h_x, y_j) - u(x_i - h_x, y_j)}{2h_x}.$$

Similarly, the second order approximation to $\frac{\partial^2 u}{\partial x^2}(x_i, y_j)$ can be done by:

$$\Delta_x^c(u_{i,j}) = \nabla_x^-(\nabla_x^+(u_{i,j})) = \frac{u(x_i + h_x, y_j) - 2u(x_i, y_j) + u(x_i - h_x, y_j)}{h_x^2}.$$

2.4.2 Boundary Conditions

Approximations of the points that appear outside the discretised grid, in this case the point $u_{n+1,j}$ is done from the boundary conditions. The two most common types of boundary conditions are-

Dirichlet boundary conditions specify the value of the function at the boundary, i.e. $\mathcal{L}_{\partial\Omega}u(\mathbf{x}) = u(\mathbf{x}) = g(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$. To calculate partial derivative, the point $u_{n+1,j}$ is replaced by the value of g at the boundary i.e.:

$$\frac{\partial u}{\partial x}(x_n, y_j) \approx \nabla_x^+(u_{i,j}) = \frac{g_{n,j} - u_{n,j}}{h_x}.$$

Neumann boundary conditions specify the value of the derivative of a function at the boundary, i.e. $\mathcal{L}_{\partial\Omega}u(\mathbf{x}) = \nabla u(\mathbf{x}) \cdot \mathbf{n} = g(\mathbf{x})$ for $\mathbf{x} \in \partial\Omega$, where $\mathbf{n}$ is normal to the boundary $\partial\Omega$. For partial derivative again:

$$\frac{\partial u}{\partial x}(x_n, y_j) \approx \nabla_x^+(u_{i,j}) = \frac{u_{n+1,j} - u_{n,j}}{h_x} = g_{n,j},$$

Example 2.4.1 (Discrete Poisson equation)**.** *Consider the Poisson equation*

$$\Delta u(\mathbf{x}) = f(\mathbf{x}), \text{ for } \mathbf{x} \in \Omega, \quad \mathcal{L}_{\partial\Omega}u(\mathbf{x}) = g(\mathbf{x}), \text{ for } \mathbf{x} \in \partial\Omega,$$

for some operator $\mathcal{L}_{\partial\Omega}$ as the boundary conditions. For domain $\Omega = [0,1]\times[0,1]$ we discretise the grid into an $n \times n$ grid ($h_x = h_y = h = \frac{1}{n}$). We can discretise using finite differences as follows:

$$(\Delta^h u)_{i,j} = \frac{u_{i+1,j} + u_{i-1,j} - 4u_{i,j} + u_{i,j+1} + u_{i,j-1}}{h^2}.$$

Each point requires 5 values of u, so the discretisation is called a $5-$point difference operator scheme. For the right boundary ($i = n$), if we impose zero-Dirichlet boundary conditions $u(\mathbf{x}) = 0$ for $\mathbf{x} \in \partial\Omega$, then:

$$(\Delta^h u)_{n,j} = \frac{u_{n-1,j} - 4u_{n,j} + u_{n,j+1} + u_{n,j-1}}{h^2}.$$

2.5 Numerical Methods

This section provides an overview of some basic numerical iterative methods that are widely used in the literature to solve systems of linear and nonlinear equations. Once an initial value is set, the iterative methods provide more and more accurate solutions with each iteration. When some stopping condition is met, the sequence of approximate solutions is terminated. First, we discuss methods to solve linear system of equations.

2.5.1 Numerical methods to solve Linear Systems of Equations

The linear system of equations is represented as -

$$A\mathbf{x} = \mathbf{b}, \tag{2.12}$$

here, A is an $n \times n$ matrix, $\mathbf{x}$ is an $n \times 1$ unknown vector to be obtained as a solution, and $\mathbf{b}$ is an $n \times 1$ vector, i.e. :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}, \ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

If the matrix A is invertible (i.e. determinant of A is non-zero) then x can be calculated as:

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

For large systems, this is known as a direct method, and it is computationally expensive.

Jacobi Method

The Jacobi method [29] is one of the more simpler iterative methods to solve equations of the form $Ax = b$. If we discretize the matrix system into elements, $i = 1, \dots, n$, we get:

$$b_i = \sum_{j=1}^{n} a_{i,j}x_j = a_{i,i}x_i + \sum_{j=1, j\neq i}^{n} a_{i,j}x_j.$$

x_i can be known in the form:

$$x_i = \frac{1}{a_{i,i}}\Big(b_i - \sum_{j=1, j\neq i}^{n} a_{i,j}x_j\Big).$$

In Jacobi method, the new elements of the sequence for the k th iteration are produced from values of the $(k-1)$th iteration.

$$x_i^{(k)} = \frac{1}{a_{i,i}}\Big(b_i - \sum_{j=1, j\neq i}^{n} a_{i,j}x_j^{(k-1)}\Big). \tag{2.13}$$

Update of $x_i^{(k)}$ requires prior values of $\mathbf{x}^{(k-1)}$. In each subsequent iteration, each element is updated independently of the others. Jacobi method can also be written in matrix form. Matrix A is decomposed into its diagonal part D, lower triangular part L and upper triangular part U as a linear sum i.e. $A = D + L + U$ where

$$D = \begin{pmatrix} a_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix},$$

$$L = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ a_{2,1} & 0 & 0 & \dots & 0 & 0 & 0 \\ a_{3,1} & a_{3,2} & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots \\ a_{n-2,1} & a_{n-2,2} & a_{n-2,3} & & 0 & 0 & 0 \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \dots & a_{n-1,n-2} & 0 & 0 \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,n-2} & a_{n,n-1} & 0 \end{pmatrix},$$

and

$$U = \begin{pmatrix} 0 & a_{1,2} & a_{1,3} & \dots & a_{1,n-2} & a_{1,n-1} & a_{1,n} \\ 0 & 0 & a_{2,3} & \dots & a_{2,n-2} & a_{2,n-1} & a_{2,n} \\ 0 & 0 & 0 & & a_{3,n-2} & a_{3,n-1} & a_{3,n} \\ \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & & 0 & a_{n-2,n-1} & a_{n-2,n} \\ 0 & 0 & 0 & \dots & 0 & 0 & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{pmatrix}.$$

Then, system $Ax = b$ can be rewritten as -

$$\mathbf{b} = (D + L + U)\mathbf{x},$$

x can be achieved by-

$$\mathbf{x} = D^{-1}\big(\mathbf{b} - (L + U)\mathbf{x}\big),$$

Gauss-Seidel Method

The Jacobi method was further improved into a new method known as Gauss-Seidel method [30]. This method uses newer approximations $\{x_j^{(k)}\}_{j=1}^{i-1}$ to update $x_i^{(k)}$, whereas old approximations $\{x_j^{(k-1)}\}_{j=1}^{i-1}$, were used in the Jacobi method. Because these new approximations are more accurate than the old ones, an exact solution can be found faster.

The Gauss-Seidel iterations are written as:

$$x_i^{(k)} = \frac{1}{a_{i,i}}\Big(b_i - \sum_{j=1}^{i-1} a_{i,j}x_j^{(k)} - \sum_{j=i+1}^{n} a_{i,j}x_j^{(k-1)}\Big), \tag{2.14}$$

For matrix form, element-wise update is written as:

$$a_{i,i}x_i^{(k)} + \sum_{j=1}^{i-1} a_{i,j}x_j^{(k)} = b_i - \sum_{j=i+1}^{n} a_{i,j}x_j^{(k)}.$$

The above system in matrix form is:

$$(D+L)\mathbf{x}^{(k)} = \mathbf{b} - U\mathbf{x}^{(k-1)},$$

The k th update is given as:

$$\begin{aligned}\mathbf{x}^{(k)} &= -(D+L)^{-1}U\mathbf{x}^{(k-1)} + (D+L)^{-1}\mathbf{b} \\ &= T_{GS}x^{(k-1)} + \mathbf{c}_{GS},\end{aligned} \tag{2.15}$$

where $T_{GS} = -(D+L)^{-1}U$ and $\mathbf{c}_{GS} = (D+L)^{-1}\mathbf{b}$.

2.5.2 Numerical methods to solve Non-Linear Systems of Equations

Usually variational imaging problems lead to non-linear systems. In this subsection, we discuss some commonly widely used non-linear solvers in image processing problems, such as Gradient Descent method [31], Additive Operator Splitting (AOS) [32], Primal Dual Gradient Method (PGDM) [33], and the Alternating Direction Method of Multipliers [34].

Gradient Descent Method

If $F : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ is a continuously differentiable functional with a minimiser $\mathbf{u}^*$. Descent methods have an initial solution $\mathbf{u}^{(0)}$, which is further updated as a more accurate solution as:

$$\mathbf{u}^{(k)} = \mathbf{u}^{(k-1)} - \alpha^{(k-1)}\mathbf{s}^{(k-1)}\, k = 1, 2, \ldots,$$

here $\alpha^{(k-1)} > 0$ (a scalar), is the step length and $-\mathbf{s}^{(k-1)}$ is the search direction. Step length may change at each iteration. A widely used descent method is the gradient descent method [31] or the steepest descent method. In this method the search direction is opposite to $\nabla F(u^{(k-1)})$. The rationale behind this is that F decreases most quickly in this direction. Gradient descent scheme is:

$$\mathbf{u}^{(k)} = \mathbf{u}^{(k-1)} - \alpha^{(k-1)} \nabla F\big(\mathbf{u}^{(k-1)}\big)\big) \; k = 1, 2, \ldots.$$

The main feature of a descent method is that each iteration reduces the function value. i.e.

$$F(\mathbf{u}^{(k)}) \leq F(\mathbf{u}^{(k-1)}),$$

$\alpha^{(k-1)}$ is chosen sufficiently small. In time marching, step length $\alpha^{(k-1)}$ is usually fixed as some time-step $\alpha^{(k-1)} = \tau$. Descent scheme is :

$$\mathbf{u}^{(k)} = \mathbf{u}^{(k-1)} - \tau \nabla F\big(\mathbf{u}^{(k-1)}\big)\big) \; k = 1, 2, \ldots.$$

Even though they are more straightforward, time marching schemes are constrained as they are only stable for low values of τ, possibly requiring a greater number of iterations to reach convergence to a consistent solution. A semi-implicit scheme, rather than an explicit scheme, could be used to reduce the stability constraints on τ.

Additive Operator Splitting

The semi-implicit, Additive Operator Splitting (AOS) scheme [32] was created to solve PDEs with the following m-dimensional anisotropic diffusion terms:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \nabla \cdot \big(G(u(\mathbf{x}))\nabla u(\mathbf{x})\big) + f(\mathbf{x}), \\ &= \partial_{x_1}\big(G(u^{(k-1)})\partial_{x_1} u^{(k)}\big) + \ldots + \partial_{x_m}\big(G(u^{(k-1)})\partial_{x_m} u^{(k)}\big) + f(\mathbf{x}) \end{aligned} \tag{2.16}$$

in $[0, T] \times \Omega$, with $\Omega \subset \mathbb{R}^m$, and with initial and zero Neumann boundary conditions:

$$u(0, \cdot) = u_0, \;\; \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \partial\Omega,$$

here, $\mathbf{n}$ is the normal to the boundary $\partial\Omega$. Here, f is the reaction term and G is the diffusivity function. Explicit, implicit or semi-implicit schemes can be used to discretise the equation

(2.16). While semi-implicit and implicit schemes are invariably stable, explicit schemes heavily rely on selecting a small time step. Because the implicit scheme is more difficult to solve, semi-implicit schemes are preferred. Discretising (2.16) with a semi-implicit scheme leads to:

$$\frac{u^{(k)} - u^{(k-1)}}{\tau} = \sum_{\ell}^{m} A_\ell\big(u^{(k-1)}\big)u^{(k)} + f,\ k \geq 1$$

where τ is the time-step and A_ℓ is a discretised version of $\partial_{x_\ell}\big(G(u^{(k-1)})\partial_{x_\ell}u^{(k)}\big)$. I being the identity matrix, this scheme can be rewritten as :

$$u^{(k)} = \Big(I - \tau \sum_{\ell}^{m} A_\ell(u^{(k-1)})\Big)^{-1}(u^{(k-1)} + \tau f),$$

final AOS scheme can be modified as:

$$u^{(k)} = \frac{1}{m}\sum_{\ell}^{m} \Big(I - \tau m A_\ell(u^{(k-1)})\Big)^{-1}(u^{(k-1)} + \tau f) \tag{2.17}$$

The key feature of the AOS scheme is that it treats each direction independently, allowing us to solve a set of m one-dimensional problems and average the results to obtain the new iterate. If we consider the two-dimensional case ($m = 2$). Matrices A_ℓ denote the diffusivity across the x_ℓ axis. For $m = 2$, the matrices are:

$$\begin{aligned}
\Big(A_1(u^{(k-1)})u^{(k)}\Big)_{i,j} &= \Big(\partial_{x_1} G(u^{(k-1)})\partial_{x_1}u^{(k)}\Big)_{i,j} \\
&= \frac{1}{h_1}\Bigg(G(u^{(k-1)})_{i+\frac{1}{2},j}(\partial_{x_1}u^{(k)})_{i+\frac{1}{2},j} - G(u^{(k-1)})_{i-\frac{1}{2},j}(\partial_{x_1}u^{(k)})_{i-\frac{1}{2},j}\Bigg) \\
&= \frac{1}{h_1}\Bigg(\Big(\frac{G(u^{(k-1)})_{i+1,j} + G(u^{(k-1)})_{i,j}}{2}\Big)\Big(\frac{u^{(k)}_{i+1,j} - u^{(k)}_{i,j}}{h_1}\Big) \\
&\quad - \Big(\frac{G(u^{(k-1)})_{i,j} + G(u^{(k-1)})_{i-1,j}}{2}\Big)\Big(\frac{u^{(k)}_{i,j} - u^{(k)}_{i-1,j}}{h_1}\Big)\Bigg) \\
&= \frac{1}{2h_1^2}\Bigg(\Big(G(u^{(k-1)})_{i+1,j} + G(u^{(k-1)})_{i,j}\Big)u^{(k)}_{i+1,j} \\
&\quad + \Big(G(u^{(k-1)})_{i,j} + G(u^{(k-1)})_{i-1,j}\Big)u^{(k)}_{i-1,j} \\
&\quad - \Big(G(u^{(k-1)})_{i+1,j} + G(u^{(k-1)})_{i-1,j} + 2G(u^{(k-1)})_{i,j}\Big)u^{(k)}_{i,j}\Bigg),
\end{aligned}$$

and similarly:

$$\begin{aligned}\Big(A_2(u^{(k-1)})u^{(k)}\Big)_{i,j} = \frac{1}{2h_2^2}\Bigg(&\Big(G(u^{(k-1)})_{i,j+1} + G(u^{(k-1)})_{i,j}\Big)u^{(k)}_{i,j+1} \\ &+ \Big(G(u^{(k-1)})_{i,j} + G(u^{(k-1)})_{i,j-1}\Big)u^{(k)}_{i,j-1} \\ &- \Big(G(u^{(k-1)})_{i,j+1} + G(u^{(k-1)})_{i,j-1} + 2G(u^{(k-1)})_{i,j}\Big)u^{(k)}_{i,j}\Bigg).\end{aligned}$$

Here, h_1 and h_2 are step lengths. The matrices A_ℓ are tridiagonal matrices, so Thomas algorithm can be used to implement the system (2.17). Instead of the $\mathcal{O}(N^3)$ operations needed by conventional Gaussian elimination methods (where N is the total number of discretized grid points), the Thomas algorithm [32] enables the solution to be reached in $\mathcal{O}(N)$ operations.

Next, we outline the criteria that an iterative scheme should satisfy in order to be a discrete scale-space. The advantages of satisfying these conditions is a guarantee of convergence and stability. Without loss of generality, we consider a general discretisation of (2.16) as:

$$u^{(0)} = \omega, \qquad u^{(k)} = Q(u^{(k-1)})u^{(k-1)}, \qquad \forall k \in \mathbb{N}, \tag{2.18}$$

We treat a discretised image as a vector $u_0 \in \mathbb{R}^N$ and denote index set by $J = \{1, 2, \ldots, N\}$. The discrete scheme (2.18) forms a non-linear scale-space if $Q = (q_{i,j})$ satisfies the following criteria:

- Continuity in its argument:
$$Q \in C(\mathbb{R}^N, \mathbb{R}^{N\times N}).$$
- Symmetry:
$$q_{i,j} = q_{j,i}, \quad \forall i, j \in J.$$
- Unit row sum:
$$\sum_{j\in J} q_{i,j} = 1, \quad \forall i \in J.$$
- Non-negativity:
$$q_{i,j} \geq 0, \quad \forall i, j \in J.$$
- Strictly positive diagonal:
$$q_{i,i} > 0, \quad \forall i \in J.$$
- Irreducibility: For all $i, j \in J$, there exist $k_0, \ldots, k_r \in J$ with $k_0 = i$, and $k_r = j$ such that $q_{k_p,k_{p+1}} \neq 0$ for $p = 0, \ldots, r-1$.

As shown in a past study by Weickert et al. [32] a discrete scale-space satisfies the following key properties:

- Average grey level invariance: The intensity value $\mu = \frac{1}{N}\sum_{j\in J}\omega_j$ is not affected by the discrete diffusion filter, i.e.:

$$\frac{1}{N}\sum_{j\in J} u_j^{(k)} = \mu, \quad \forall k \in \mathbb{N}.$$

- Extremum principle:

$$\min_{j\in J}\omega_j \leq u_i^{(k)} \leq \max_{j\in J}\omega_j, \qquad \forall i \in J, \quad \forall k \in \mathbb{N}.$$

- Convergence to a constant steady-state:

$$\lim_{k\to\infty} u_j^{(k)} = \mu, \quad \forall j \in J.$$

Furthermore, because the scheme is unconditionally stable, there are no restrictions on the size of the time step, demonstrating a significant advantage of AOS over explicit and semi-implicit time marching schemes.

Theorem 2.5.1. *The AOS scheme* (2.17) *with* $m = 2$ *corresponding to the finite difference equation:*

$$\begin{aligned}(Fu)^{(k-1)} := \frac{1}{\tau}u^{(k)} &- \frac{1}{2\tau}\Big(I - 2\tau A_1\big(u^{(k-1)}\big)\Big)^{-1}\big(u^{(k-1)} + \tau f\big) \\ &- \frac{1}{2\tau}\Big(I - 2\tau A_2\big(u^{(k-1)}\big)\Big)^{-1}\big(u^{(k-1)} + \tau f\big),\end{aligned}$$

for $k = 1, 2, \ldots$ *is an* $\mathcal{O}(\tau + h^2)$ *approximation, where* $h_1 = h_2 = h$. *Therefore, the scheme is consistent with PDE* (2.16).

Alternating Direction Method of Multipliers

Popular iterative approach for solving convex optimisation problems is the Alternating Direction Method of Multipliers (ADMM) approach [34]. If we consider the problem:

$$\min_{\mathbf{u},\mathbf{v}} f(\mathbf{u}) + g(\mathbf{v}), \quad \text{subject to } A\mathbf{u} + B\mathbf{v} = \mathbf{c}, \tag{2.19}$$

where $\mathbf{u} \in \mathbb{R}^{d_1}, \mathbf{v} \in \mathbb{R}^{d_2}, A \in \mathbb{R}^{p\times d_1}, B \in \mathbb{R}^{p\times d_2}$ and $\mathbf{c} \in \mathbb{R}^p$. ADMM's main objective is to

use partial updates for each of the variables $\mathbf{u}$ and $\mathbf{v}$. We define the augmented Lagrangian to incorporate the constraint into the formulation:

$$\mathcal{L}_\rho(\mathbf{u}, \mathbf{v}, \mathbf{x}) = f(\mathbf{u}) + g(\mathbf{v}) + \mathbf{x}^T(A\mathbf{u} + B\mathbf{v} - \mathbf{c}) + \frac{\rho}{2}||A\mathbf{u} + B\mathbf{v} - c||_2^2, \tag{2.20}$$

here, $\mathbf{x}$ is the dual variable and $\rho > 0$ is a penalty parameter controlling the constraint. The ADMM update that solves (2.20) is given by :

$$\begin{cases} \mathbf{u}^{(k)} = \arg\min_{\mathbf{u}} \mathcal{L}_\rho(\mathbf{u}, \mathbf{v}^{(k-1)}, \mathbf{x}^{(k-1)}), \\ \mathbf{v}^{(k)} = \arg\min_{\mathbf{v}} \mathcal{L}_\rho(\mathbf{u}^{(k)}, \mathbf{v}, \mathbf{x}^{(k-1)}), \\ \mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + \rho(A\mathbf{u}^{(k)} + B\mathbf{v}^{(k)} - \mathbf{c}). \end{cases} \tag{2.21}$$

Instead of solving the $\mathbf{u}$ and $\mathbf{v}$ together, splitting them into two separate sub-problems can greatly reduce the implementation and computation demand required for each iteration. The alternating minimisation in ADMM can be thought of as a single Gauss-Seidel pass over $\mathbf{u}$ and $\mathbf{v}$. Another advantage of ADMM with simpler implementation is that it can have convergence guarantees under modest f and g assumptions, see for more details [34, 35].

Primal Dual Gradient methods

This approach usually does not require expensive minimization sub-steps like in the ADMM. The Primal Dual Gradient method (PDHG) [33, 36] is a popular iterative method to solve convex optimisation problems. Consider the saddle point problem:

$$\min_{x \in X} \max_{y \in Y} f(x) + y^t A x - g(y) \tag{2.22}$$

Here, X, Y, are convex sets $A \in R^{M \times N}$ is a matrix, and $X \subset R^N$, $Y \subset R^M$ are convex sets.

Primal Dual Hybrid Gradient method is listed in Algorithm 1. The forward-backward algorithm is used to update the primal x and dual parameter y, which are updated using a combination of forward and backward steps. In steps (2-3), the method updates x to lower the energy 2.22 by first performing a gradient descent step with respect to the inner product term in Equation 2.22, and then performing a "backward" or proximal step involving f. By first marching up the gradient of the inner product term with respect to y and then taking a step backward with respect to g, steps (5–6) increase the energy 2.22.

Algorithm 1 Basic PDHG

Require: : $x_o \in R^N, y_o \in R^M, \sigma_k, \tau_k > 0$

1. **While** not converged **do**.
2. $x^{k+1} = x_k - \tau_k A^T y_k$.
3. $x^{k+1} = \arg\ \min_{x \in X} f(x) + \frac{1}{2\tau_k}||x - \hat{x}_{k+1}||^2$
4. $\hat{x} = x_{k+1} + (x_{k+1} - x_k)$.
5. $\hat{y}_{k+1} = y_k + \sigma_{.k} A\hat{y}_{k+1}$.
6. $y_{k+1} = argmin_{y \in Y} g(y) + \frac{1}{2\sigma_k}||y - \hat{y}_{k+1}||^2$
7. **EndWhile** $k = k + 1$

Chapter 3

Review of Segmentation Models

3.1 Introduction

Image segmentation partitions an image into segments in which each pixel is assigned to an object [16]. It is used in a wide range of practical applications, such as medical image analysis, face recognition and detection, satellite image analysis, video surveillance, computer vision for autonomous vehicles and many more [14, 15].

In the past decades, variational methods based on variational calculus have proven to be effective [37, 38]. Depending on the application, models are devised to use both region (intensity) and edge-based information from the image [38, 39, 37]. Variational models can be formulated to achieve either multiple object segmentation's or selective single object segmentation. The following sections introduce segmentation models that segment the input image using variational methods based on calculus of variation. Each segmentation model includes an energy functional, and the goal is to find the functional's minimizer. The minimizer of the energy functional is achieved using calculus of variations. The minimizer takes the form of a partial differential equation, and the solution corresponds to the image's segmented region.

Potts model, which is defined in a discrete framework for multi-phase image segmentation [40], is the origin of many variational methods. Many previous segmentation models adopted the Potts model in some way [41]. Blake-Zisserman model [42] proposed that smooth images can be reconstructed on discrete domains . Both the Potts model and the Blake-Zisserman model have undergone numerous approximations due to the difficult task of solving the energy functional of these models [40, 42]. The Mumford and Shah model [37] was an early segmentation model that found the piecewise-constant approximation of the image and the edge that represents discontinuity. Following a discussion of the Mumford and Shah model, we will discuss the Chan-Vese model [38], which is a two-dimensional piecewise constant variant of the Mumford and Shah model, in the following section of global segmentation models. Prior to these segmentation models, we will discuss about the energy functionals that these segmentation models use.

In this thesis, we are concerned mostly with selective segmentation problems. After focusing on global segmentation formulations, we move on to selective segmentation models. Edge detector function in selective models constructs an active contour to locate the boundary of an object using intensity information [43]. The first studies to use this function were snake models proposed by Kass et al. [44] and the Geodesic Active Contours model proposed by Caselles et al [43]. We will also review the formulations of recent selective models like - Chen Badshah model [21], Rada Chen model [45], Spencer Chen model [46], and convex selective models [24]. Then, we review a selective segmentation model known as the convex geodesic model [15], which employs additional regularisation terms to achieve better selective segmentation.

3.1.1 Variational Segmentation approach

The functional $F(u)$ is made up of two parts. The first is the fitting term, and the second is the regularisation term. A fitting function is f in the following functional. The fitting term's function is to define the correspondence between the objective function and the data. J is the regularisation term or a term that imposes regularity. The regularisation term penalises the objective function for over-fitting or for making the optimal solution unique.

$$F(u) = \int_{\Omega} J(u)dx + \alpha \int_{\Omega} f(u)dx \tag{3.1}$$

Such energy functionals are minimised using variational methods. By increasing the parameter α the fitting term becomes more dominant, resulting in a solution that closely represents the average data. Reducing the parameter α would make the regularisation term more prominent, resulting in a smoother or regular solution. The solution is generally in the form-

$$u* = arg\, min_{u \epsilon S} F(u) \tag{3.2}$$

where $u*$ is an optimiser of the functional $F(u)$, defined on an appropriate space S. If F is continuous and differentiable, the first variation or the minimizer of the energy functional is computed using the Euler-Lagrange equation.

$$\frac{\delta F}{\delta u} = 0 \tag{3.3}$$

3.2 Global Segmentation

Here, in this section we will detail some previous global segmentation and selective segmentation models. However, Roberts convex geodesic model is more relevant as the models we introduce in chapter 4 and chapter 6 have roots in the Roberts convex model.

3.2.1 Mumford and Shah Model

The Mumford-Shah segmentation model [37] involves a functional whose minimizer serves as a criterion for segmenting an input image into sub-regions. It is concerned with the piecewise smooth approximation of an input image $u_0(x)$ by a pair (u,C), where u represents the desired segmented region with sharp boundary C. Ω is a bounded domain, and $u_0(x)$ is a bounded measurable function defined on Ω. Then the Mumford-Shah functional is defined as-

$$F^{(MS)}(u, C) = \mu.Length(C) + \lambda \int_{\Omega} |u_0(x) - u(x)|^2 dx + \int_{\Omega/C} ||\nabla u(x)||^2 dx \tag{3.4}$$

A fitting term ($|u_0(x) - u(x)|^2$) and two regularity terms ($\mu.Length(C), \nabla u(x)$) are included in the above functional. The fitting term imposes smoothness on u, while the other terms impose regularity on C. The solution image u is in form of smooth regions with sharp boundaries C and is obtained by minimising the functional. The functional penalises the distance between the model and the input image, the lack of smoothness of the model within the sub-regions, and the length of the sub-region boundaries, as represented by the three terms. A solution for (u, C) can be found by minimising the functional 3.4. Because of the undefined domain caused by non-regularity of C, computing the minimiser of this functional is difficult.

Many attempts have been made to approximate the Mumford- Shah functional in order to find the minimiser. Ambrosio and Tortorelli [47] proposed that the functional can be approximated by a sequence of elliptic variational problems with a phase field energy term. Level set methods [22], as described in the following subsection, have also been widely used. Pock et al. [48] proposed a primal-dual scheme for using the dual domain to find the minimiser of Mumford-Shah model. Cai et al. [49] proposed another convex approximation to the Mumford-Shah model.

Mumford and Shah also discuss the restriction of F to piecewise-constant functions u. In other words if we consider the input image to be made up of piecewise constant regions k with $u = c_k$ on each open set k defined on domain Ω_k, where the values c_k are simply the average values of u_0 in each region k. The piecewise-constant functional takes the form -

$$F^{(MS)}(u, C) = \mu.Length(C) + \lambda \int_{\Omega_k} |u_0(x) - c_k|^2 dx \tag{3.5}$$

If images are to be partitioned into two components, foreground and background, then the curve C has two components, c_1 and c_2. This is known as the two-phase constant Mumford-Shah functional, and the functional takes the following form -

$$F^{(MS)}(C, c_1, c_2) = \mu.Length(C) + \lambda \int_{\Omega_1} |u_0(x) - c_1|^2 dx + \lambda \int_{\Omega_2} |u_0(x) - c_2|^2 dx \tag{3.6}$$

Next, we need to find the minimizer for this functional. We will discuss details of computing minimiser of this functional by using level set methods in relation to the Chan-Vese model in next subsection.

3.2.2 Chan-Vese Model

The Chan vese Functional [38] is the two dimensional piece wise constant variant of the Mumford and Shah model and takes the form-

$$\begin{aligned} F^{(CV)}(C, c_1, c_2) = \mu.Length(C) + \nu.Area(insideC) + \lambda_1 \int_{\Omega_1} |u_0(x, y) - c_1|^2 dxdy \\ + \lambda_2 \int_{\Omega_2} |u_0(x, y) - c_2|^2 dxdy \end{aligned} \tag{3.7}$$

Here the foreground Ω_1 is the required region to be segmented and $\Omega_2 = \frac{\Omega}{\Omega_1}$is background. λ_1,λ_2, μ, ν are fixed non negative parameters and c_1 and c_2 are average intensities of input image u_o inside Ω_1 and Ω_2 . Level set method is utilised for minimization of an energy based-functional F. To use level set functions, it is assumed that the image u_0 is formed by two regions of approximately piece-wise-constant intensities, of distinct values u_0^i and u_0^o. The object to be detected is represented by the region with the value u_0^i. Within a closed curve C, boundary of object is denoted by C_0, c_1 and c_2 are average intensities inside and outside the curve C. The following is the fitting term:

$$F_1(C) + F_2(C) = \int_{insideC} |u_0(x, y) - c_1|^2 dxdy + \int_{outsideC} |u_0(x, y) - c_2|^2 dxdy \tag{3.8}$$

In this simple case, it is obvious that C_0, which is the boundary of the object is the minimizer of the fitting term.

$$inf_C F_1(C) + F_2(C) = 0 = F_1(C_0) + F_2(C_0) \tag{3.9}$$

It is to be noted that, if the curve is outside the object then $F_1(C) > 0$ and $F_2(C) \approx 0$. If the curve C is inside the object, then $F_1(C) \approx 0$ but $F_2(C) > 0$. If the curve is both inside and outside the object, then $F_1(C) > 0$ and $F_2(C) > 0$. Finally, we can see here that the fitting energy is minimized when $C = C_0$. The above fitting term along with two regularization terms i.e. the length of curve and area inside C are formulated as Chan- Vase functional [38]-

$$\begin{aligned} F(c_1, c_2, C) = \mu.Length(C) + \nu Area(inside(C)) + \lambda_1 \int_{insideC} |u_0(x,y) - c_1|^2 dxdy \\ +\lambda_2 \int_{outsideC} |u_0(x,y) - c_2|^2 dxdy \end{aligned} \tag{3.10}$$

here $\mu \geq 0$, $\nu \geq 0$, $\lambda_1 > 0$, $\lambda_2 > 0$
The minimization problem is the following

$$inf_{c_1,c_2,C} F(c_1, c_2, C) \tag{3.11}$$

This C can be initialised with parameters and then updated with each iteration. However, changes in boundary topology (splitting or merging of the boundary) are far too complex to be encoded by parametrisation. A straightforward solution is to embed the boundary in higher dimensions. This is accomplished using the level set method, and instead of tracking a parametrized C, we now track the zero level set of a function ϕ. We are now concentrating on determining the function ϕ, and the object boundary C is obtained indirectly from the information of ϕ.

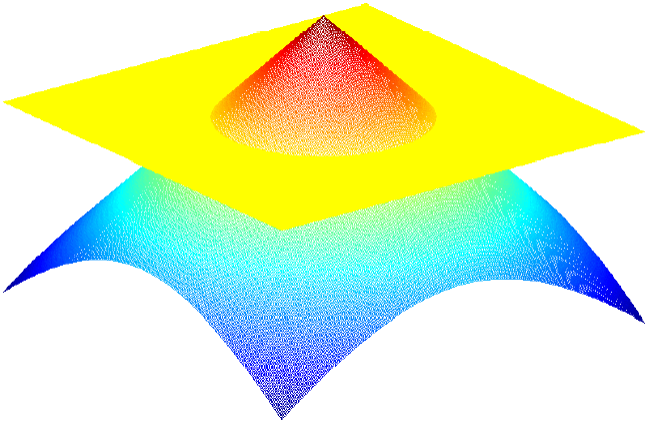
3.2.3 Level set Method

Here, C is represented in terms of the zero level set of Lipschitz function $\phi : \Omega \to R$ such that-

$$\begin{cases} C = (x,y) \in \Omega : \phi(x,y) = 0 \\ inside(C) = w = (x,y) \in \Omega : \phi(x,y) > 0 \\ outside(C) = \frac{\Omega}{w} = (x,y) \in \Omega : \phi(x,y) < 0 \end{cases} \tag{3.12}$$

By using level set formulation the unknown variable C is replaced by the unknown variable ϕ. Heaviside function H and Dirac measure δ_0 are defined as-

$$H(z) = \begin{cases} 1, if z \geq 0 \\ 0, if z < 0 \end{cases} \tag{3.13}$$



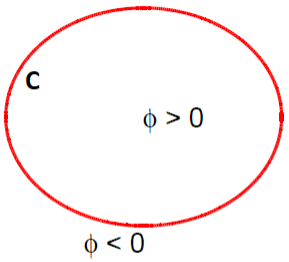


Figure 3.1: Plot of the function $\phi(x)$ and the interface C $=\{(x,y)\ \epsilon\ \Omega : \phi(x,y) = 0\}$.

$$\delta_0(z) = \frac{d}{dz}H(z) \tag{3.14}$$

By utilising the Heaviside and the Dirac function, the length term and area term in F as well as the fitting terms take the following form-

$$Length(\phi = 0) = \int_\Omega |\nabla H(\phi(x,y))|dxdy = \int_\Omega \delta_0(\phi(x,y))|\nabla\phi(x,y)|dxdy \tag{3.15}$$

$$Area(\phi \geq 0) = \int_\Omega H(\phi(x,y))dxdy \tag{3.16}$$

$$\int_{\phi>0} |u_0(x,y) - c_1|^2 dxdy = \int_\Omega |u_0(x,y) - c_1|^2 H(\phi(x,y))dxdy \tag{3.17}$$

$$\int_{\phi<0} |u_0(x,y) - c_2|^2 dxdy = \int_\Omega |u_0(x,y) - c_2|^2 (1 - H(\phi(x,y)))dxdy \tag{3.18}$$

Then energy can be rewritten as -

$$F(c_1,c_2,\phi) = \mu.\int_\Omega \delta_0(\phi|\nabla\phi|dxdy + \nu.\int_\Omega H(\phi)dxdy + \lambda_1\int_\Omega |u_0(x,y) - c_1|^2 H(\phi)dxdy + \lambda_2\int_\Omega |u_0(x,y) - c_2|^2(1 - H(\phi)dxdy \tag{3.19}$$

The objective is to find the minimizer of the above functional. Keeping ϕ fixed and minimizing F with respect to C_1 gives us -

$$\frac{\partial F}{\partial c_1} = 0 \tag{3.20}$$

we get

$$2\lambda_1 \int_\Omega |u_0(x,y) - c_1| H(\phi(x,y))dxdy = 0 \tag{3.21}$$

$$\int_\Omega u_0(x,y)H(\phi(x,y))dxdy = \int_\Omega c_1 H(\phi(x,y))dxdy \tag{3.22}$$

$$c_1(\phi) = \frac{\int_\Omega u_0(x,y)H(\phi(x,y))dxdy}{\int_\Omega H(\phi(x,y))dxdy} \tag{3.23}$$

similarly minimizing with respect to c_2

$$c_2(\phi) = \frac{\int_\Omega u_0(x,y)(1 - H(\phi(x,y)))dxdy}{\int_\Omega (1 - H(\phi(x,y)))dxdy} \tag{3.24}$$

For Euler–Lagrange equation of the unknown function , slightly regularized versions of the functions H and δ_0 are used and are denoted by H_ϵ and δ_ϵ. Regularised functions are visually demonstrated in Figure 3.2. Associated regularized functional is formulated as -

$$F_\epsilon(c_1,c_2,\phi) = \mu.\int_\Omega \delta_\epsilon(\phi)|\nabla\phi|dxdy + \nu.\int_\Omega H_\epsilon(\phi)dxdy + \lambda_1\int_\Omega |u_0(x,y) - c_1|^2 H_\epsilon(\phi)dxdy + \lambda_2\int_\Omega |u_0(x,y) - c_2|^2(1 - H_\epsilon(\phi))dxdy \tag{3.25}$$

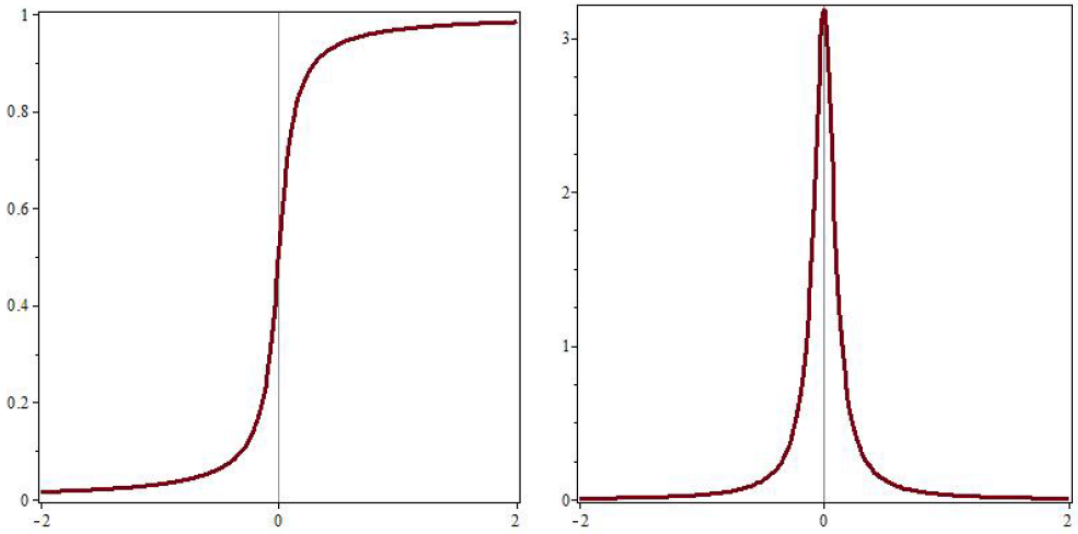


Figure 3.2: Regularised Heaviside and Delta function.(a) $H_\epsilon(\phi)$ (b) $\delta_\epsilon\ (\phi)$, $\epsilon = 0.1$

Keeping c_1 and c_2 fixed and minimizing with respect to ϕ i.e. $\frac{\delta F}{\delta \phi}$, we deduce the associated Euler–Lagrange equation for ϕ. To minimize F we need to find Eulers equation, so

minimizing with respect to ϕ-

$$lim_{h\to o}\frac{dF_{\epsilon}^{CV}(\phi+h\psi,c_1,c_2)}{dh}=0 \tag{3.26}$$

Here ψ is any test function.

$$\begin{aligned}\frac{d}{dh}\int_{\Omega}\mu(\delta_{\epsilon}(\phi+h\psi)|\nabla(\phi+h\psi)|dxdy|_{h=0}+\frac{d}{dh}\nu\int_{\Omega}H_{\epsilon}(\phi+h\psi)dxdy|_{h=0}+\\ \frac{d}{dh}\int_{\Omega}(\lambda_1|u_0(x,y)-c_1|^2H_{\epsilon}(\phi+h\psi)+\lambda_2|u_0(x,y)-c_2|^2(1-H_{\epsilon}(\phi+h\psi)))|_{h=0}=0\end{aligned} \tag{3.27}$$

Computing derivatives we have -

$$\begin{aligned}\mu\int_{\Omega}\delta_{\epsilon}^{'}(\phi)|\nabla(\phi)|\psi dxdy+\mu\int_{\Omega}\delta_{\epsilon}(\phi)\frac{\nabla\phi}{|\nabla(\phi)|}.\nabla\psi dxdy+\int_{\Omega}\delta_{\epsilon}(\phi)(\nu+\lambda_1|u_0(x,y)-c_1|^2\\ -\lambda_2|u_0(x,y)-c_2|^2)\psi dxdy=0\end{aligned} \tag{3.28}$$

By using Greens theorem we have -

$$\int_{\Omega}v\nabla.\bar{w}dx=-\int_{\Omega}\nabla v.\bar{w}dx+\int_{d\Omega}v\bar{w}.\bar{n}ds \tag{3.29}$$

Taking $\psi=v$ and $\frac{\delta_{\epsilon}(\phi)\nabla\phi}{|\nabla\phi|}=\bar{w}$. The above equation becomes-

$$\int_{\Omega}\delta_{\epsilon}(\phi)\frac{\nabla\phi\nabla\psi}{|\nabla\phi|}dxdy=-\int_{\Omega}\psi\nabla.(\frac{\delta_{\epsilon}(\phi)\nabla\phi}{|\nabla\phi|})dxdy+\int_{d\Omega}\psi\frac{\delta_{\epsilon}(\phi)}{|\nabla\phi|}\frac{\delta\phi}{\delta\bar{n}}ds \tag{3.30}$$

using above and we have $\nabla\phi.\vec{n}=\frac{\delta\phi}{\delta n}$

$$\begin{aligned}\int_{\Omega}\mu\delta_{\epsilon}^{'}(\phi)|\nabla\phi|\psi dxdy-\int_{\Omega}\mu\delta_{\epsilon}(\phi)\nabla.(\frac{\nabla\phi}{|\nabla\phi|})\psi dxdy-\int_{\Omega}\mu\delta_{\epsilon}^{'}(\phi)\nabla\phi.\frac{\nabla\phi}{|\nabla\phi|}\psi dxdy\\ +\int_{d\Omega}\mu\psi\frac{\delta_{\epsilon}(\phi)}{|\nabla\phi|}\frac{\delta\phi}{\delta\bar{n}}ds+\int_{\Omega}\delta_{\epsilon}(\phi)(\nu+\lambda_1(u_0(x,y)-c_1)^2-\lambda_2(u_0(x,y)-c_2)^2)\psi dxdy=0\end{aligned} \tag{3.31}$$

Now we have the following Eular-Lagrange equation for ϕ.

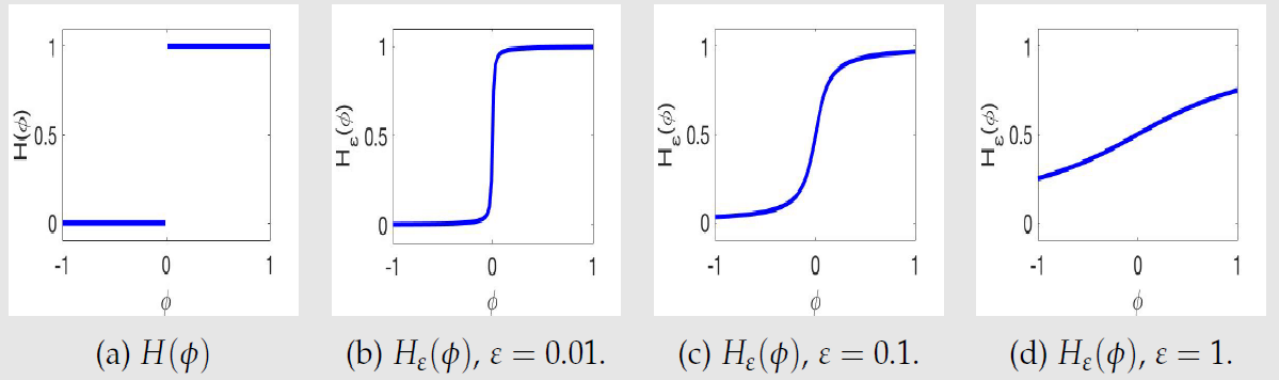
$$\delta_{\epsilon}(\phi)\big[\mu\nabla.(\frac{\nabla\phi}{|\nabla\phi|})-\nu-\lambda_1(u_0-c_1)^2+\lambda_2(u_0-c_2)^2\big]=0,\ \ \frac{\delta_{\epsilon}(\phi)}{|\nabla\phi|}\frac{\delta\phi}{\delta\bar{n}}=0\text{ on }\delta\Omega \tag{3.32}$$

The approximation of above equation can be done by introducing an artificial time step t and using the gradient descent method. In this way we get the following evolution equation which represents the minimizer of the functional in terms of partial differential equation.

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi)\big[\mu div(\frac{\nabla \phi}{|\nabla \phi|}) - \nu - \lambda_1(u_0 - c_1)^2 + \lambda_2(u_0 - c_2)^2\big] = 0 \text{ in } \Omega$$
$$\phi(0, x, y) = \phi_0(x, y) \text{in } \Omega \tag{3.33}$$
$$\frac{\delta_\epsilon(\phi)}{|\nabla \phi|}\frac{\delta \phi}{\delta \bar{n}} = 0 \text{ on } \delta\Omega$$

In the next section we discuss the scheme used to solve the above equation to get the solution ϕ. However, as we must find $|\nabla H(\phi)|$, but $H(\phi)$ is discontinuous at zero, so we regularise $H(\phi)$ at the discontinuity by setting it as-

$$H_{2,\epsilon}(z) = \frac{1}{2}(1 + \frac{2}{\pi} arctan(\frac{z}{\epsilon})) \tag{3.34}$$



(a) $H(\phi)$ (b) $H_\varepsilon(\phi)$, $\varepsilon = 0.01$. (c) $H_\varepsilon(\phi)$, $\varepsilon = 0.1$. (d) $H_\varepsilon(\phi)$, $\varepsilon = 1$.

Figure 3.3: Regularised Heaviside Function

Here with $H_{2,\epsilon}$ the algorithm has tendency to compute a global minimizer. To discretize the equation in ϕ, we use finite difference implicit scheme. h is the space step, Δt is the time step and $(x_i, y_i) = (ih, jh)$ are the grid points for $1 \leq i$, and $j \leq M$. Let $\phi^n_{i,j} = \phi(n\Delta t, x_i, y_i)$ be an approximation of $\phi(t, x, y)$ with $\phi^0 = \phi_0$. We now apply the finite differences method -

$$\begin{aligned} \Delta^x_- \phi_{i,j} &= \phi_{i,j} - \phi_{i-1,j} \\ \Delta^x_+ \phi_{i,j} &= \phi_{i+1,j} - \phi_{i,j} \\ \Delta^y_- \phi_{i,j} &= \phi_{i,j} - \phi_{i,j-1} \\ \Delta^y_+ \phi_{i,j} &= \phi_{i,j+1} - \phi_{i,j} \end{aligned} \tag{3.35}$$

With information of ϕ^0, we first compute $c_1(\phi^n)$ and $c_2(\phi^n)$ by (3.26) and (3.27). By discretization in ϕ and using definitions of forward and central difference we obtain solution

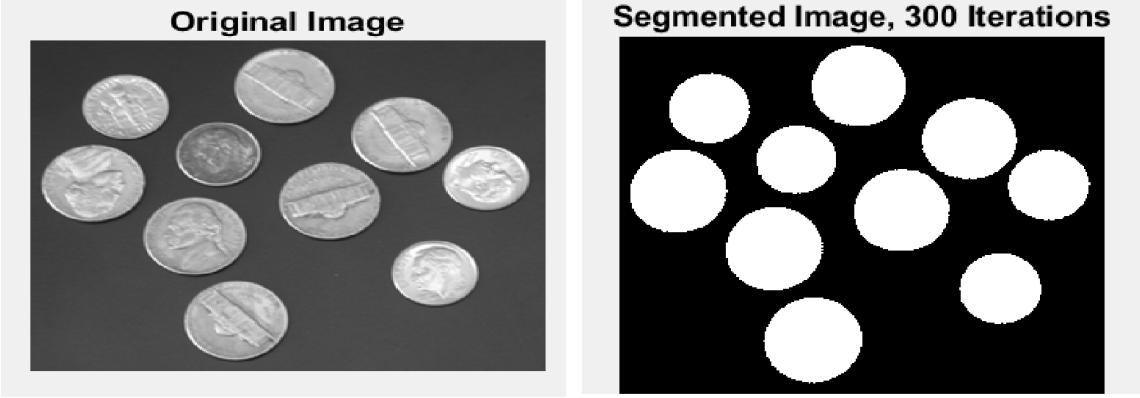


Figure 3.4: Example of multiple object segmentation done by chan-vase model.(a) input (Left) (b) segmented image (right)

ϕ^{n+1} from-

$$\frac{\phi_{i,i}^{n+1}-\phi_{i,i}^{n}}{\Delta t}=\delta_h(\phi_{i,j}^n)\big[\frac{u}{h^2}\Delta_-^x.(\frac{\Delta_+^x\phi_{i,j}^n}{\sqrt{(\frac{(\Delta_+^x\phi_{i,j}^n)^2)}{(h^2)}+\frac{(\phi_{i,j+1}^n-\phi_{i,j-1}^n)^2}{(2h)^2}}})$$
$$+\frac{u}{h^2}\Delta_-^y.(\frac{\Delta_+^y\phi_{i,j}^n}{\sqrt{(\frac{(\Delta_+^y\phi_{i,j}^n)^2)}{(h^2)}+\frac{(\phi_{i+1,j}^n-\phi_{i-1,j}^n)^2}{(2h)^2}}})-v-\lambda_1(u_{0,i,j}-c_1(\phi^n))^2+\lambda_2(u_{0,i,j}-c_2(\phi^n))^2\big] \tag{3.36}$$

Figure 3.4 is an experimental result using the Chan Vase model. The input coins image is segmented with the objects being recognized as coins.

3.2.4 Convex version of Chan Vese Model

The Chan-Vese method is non-convex with respect to ϕ as the optimization problem is non-convex, which makes it prone to locate local minima. That is, a different initialisation of ϕ may yield a significantly different result, as shown in Figure 3.6 (a)-(d). As a result, some research has been conducted into convex relaxation methods. The global minimiser of such convex models, which is preferred in practise, is independent of initialization.

The non-convex Chan-Vese model was reformulated to an equivalent convex model by Chan, Esdoglu, and Nikolova in order to find a global minimiser [50]. The authors' key observation is that the Chan-Vese algorithm employs a non compactly supported, smooth approximation H_ϵ for H. The authors proposed that the stationary solution of Chan-Vase functional [38, 51] is similar to the stationary solution ϕ of the following equation:

$$\frac{\delta\phi}{\delta t}=\nabla.\frac{\nabla\phi}{|\nabla\phi|}-(\lambda_1(u_0(x,y)-c_1)^2-\lambda_2(u_0(x,y)-c_2)^2) \tag{3.37}$$

The above scheme is the gradient descent of the following energy functional-

$$\int_\Omega |\nabla\phi|dx + \int_\Omega (\lambda_1(u_0(x,y)-c_1)^2 - \lambda_2(u_0(x,y)-c_2)^2)\phi dx \tag{3.38}$$

To find the minimiser of the above energy, Chan, Esdoglu, and Nikolova [50] proposed a simple way by restricting the minimization of ϕ to $[0,1]$.

Theorem 3.2.1. *(Theorem 2 in [50]) For any given $c_1, c_2 \in R$, a global minimiser of the Chan-Vese model (3.33) can be found by solving out the following convex minimisation problem:*

$$min_{0\le u(x)\le 1} \int_\Omega |\nabla u(x)|dx + \int_\Omega \lambda((u_0(x,y)-c_1)^2 - (u_0(x,y)-c_2)^2)u(x)dxdy \tag{3.39}$$

and then setting $\Omega_1 = \{x : u(x) \ge 0.5\}$

The authors proposed the following unconstrained minimization to balance the drawback of imposing the constraint $u \in [0,1]$ in Theorem 3.2.1.

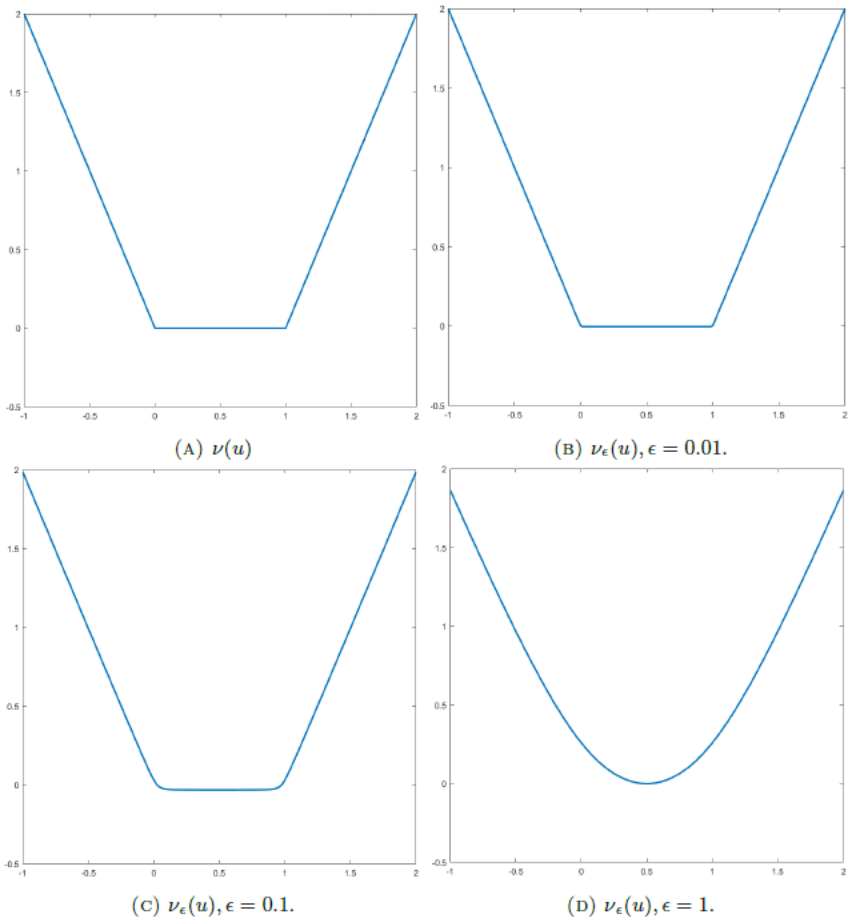
Theorem 3.2.2. *(Claim 1 from [50]) Let $r(x) \in L^\infty(\Omega)$. Then the convex, constrained minimisation problem in Theorem 3.2.1 has the same set of minimisers as the following convex, unconstrained minimisation problem:*

$$min_{u(x)} \int_\Omega |\nabla u(x)|dx + \int_\Omega \lambda((u_0-c_1)^2 - (u_0-c_2)^2)u(x)dx + \alpha\int_\Omega \nu dx \tag{3.40}$$

where $\nu(u) := max\{0, 2|u-\frac{1}{2}|-1\}$ is an exact penalty function, provided that, $\alpha > \frac{\lambda}{2}||(u_0-c_1)^2 - (u_0-c_2)^2||_{L^\infty(\Omega)}$

The addition of exact penalty term $\nu(u)$ forces u in the range [0, 1]. It would be more practical to use the regularised version as in Figure 3.5 we can observe that $\nu(u)$ is non-differentiable at 0 and 1. In Figure 3.5 we show different regularised versions of $\nu_\epsilon(u)$.

$$\nu_\epsilon(u) = H_\epsilon(\sqrt{(2u-1)^2+\epsilon}-1)[\sqrt{(2u-1)^2+\epsilon}-1] \tag{3.41}$$



(A) $\nu(u)$ (B) $\nu_\epsilon(u), \epsilon = 0.01$.

(C) $\nu_\epsilon(u), \epsilon = 0.1$. (D) $\nu_\epsilon(u), \epsilon = 1$.

Figure 3.5: Penalty function $\nu(u)$ and the regularised version $\nu_\epsilon(u)$ with different values for ϵ.

by using the above convex energy functional we can see in figure 3.6, initialisation of solution is not a concern for convex framework as initialisations provide the same segmentation output if we utilise Theorem 3.2.2.

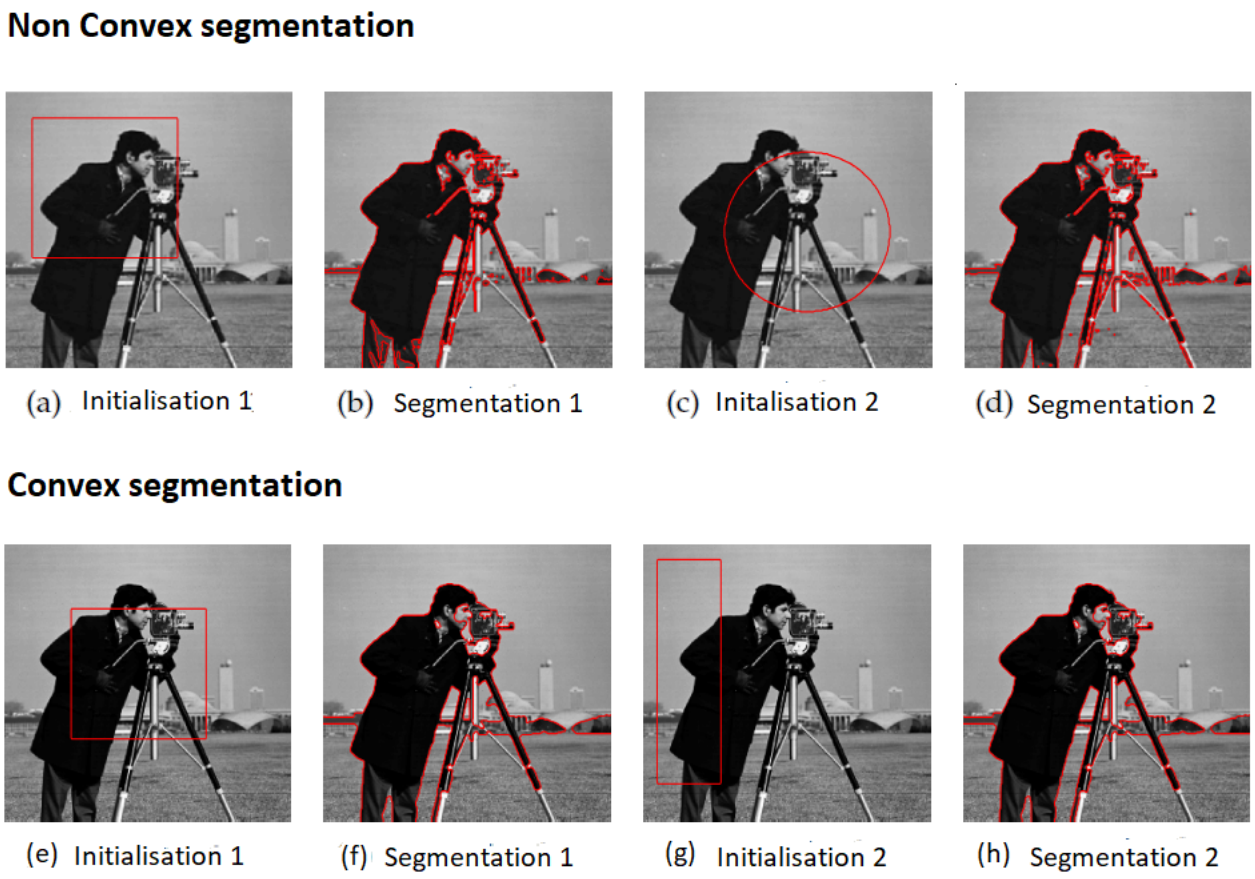


Figure 3.6: Segmentation comparison of convex and non convex Chan-Vese model. In the non convex relaxed setting, different initialisations (a), (c) provide the different segmentation results (b), (d). In the convex relaxed setting, different initialisations (e), (g) provide the same segmentation results (f), (h) .

3.3 Selective Segmentation

Selective segmentation models are used when a user needs to segment only one object from an image that contains multiple objects or regions. Selective segmentation has numerous applications in medical imaging since clinicals require the segmentation of a tumour area or a specific organ from MR or CT scans [14] . We have a domain Ω and an input image $z(x, y)$. The selective segmentation models require user input in the form of a marker set M, with these points being close to the object to be segmented. The Energy functional locates the τ contour that is close to the points in M. Here, we review methods of selective segmentation in the literature, and discuss current challenges.

3.3.1 Geodesic Active Contours

Geodesic Active contour (GAC) model is one of the early selective segmentation model purposed by Caselles et al. [43] where the energy functional is -

$$F_{GAC} = \int_{\Omega} g(|\nabla z(x,y)|)dxdy \tag{3.42}$$

The term g is edge detector function and Ω is image domain. The commonly used edge detector function is -

$$g(s) = \frac{1}{1+\beta s^2} \tag{3.43}$$

The idea behind the edge detector function was that g is small near object boundaries and is controlled by the parameter β. While providing segmentation results, this model faces a number of challenges. First, the model is based on the gradient of the image, images with noise are unsuitable for this method. Secondly, if the edge detector function is not negligible at an edge, the contour will transcend the object's boundary, generating an incorrect segmentation result. Generally, $s = |\nabla(G_\sigma \otimes z)|$, where G_σ is a Gaussian filter. Gaussian filter is convolved with image z, which blurs it. This convolution blurs the image's noisy pixels and ensures that noise does not distract from the image's edges.

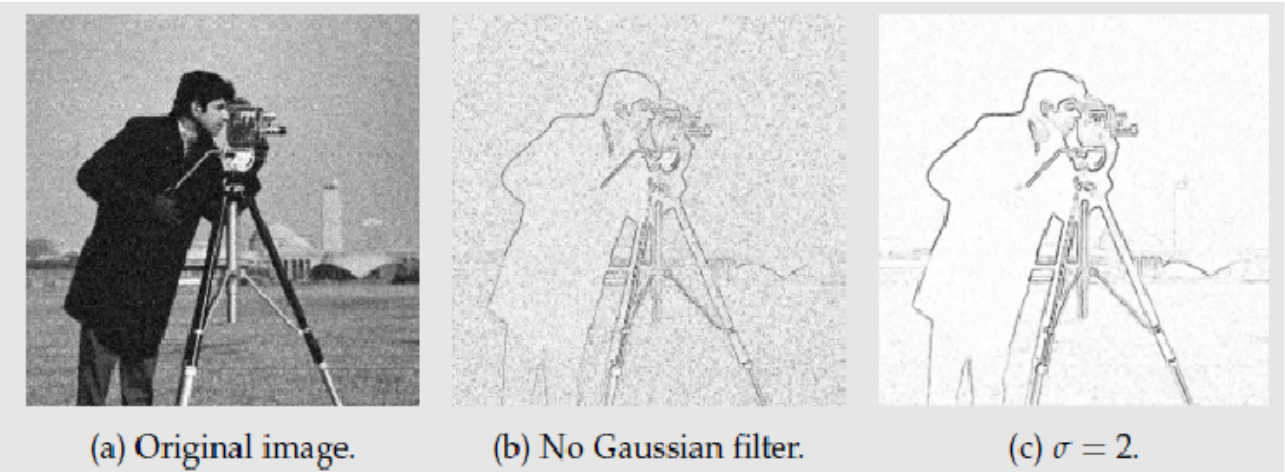


Figure 3.7: (a) original image (b) segmentation result for no Gaussian filtering of input image. (b) Edge detector segmentation result for varying σ.

3.3.2 Gout et al.

Certain improvements were done to the Geodesic contour model. Gout et al. [39] used an additional distance term D in the integral, where the integrand is $Dg(|\nabla z|)$. The distance

term being the penalty on distance from set M (with k points) allows the contour to be near M. However, the model struggles with when object edges are blurred in case of blurred images.

$$F_{Gout} = \int_{\Omega} Dg(|\nabla z(x,y)|)dxdy \tag{3.44}$$

We can rewrite in terms of level set function as -

$$F_{Gout}(\phi) = \mu \int_{\Omega} D(x,y)g(|\nabla z(x,y)|)|\nabla H_{\epsilon}(\phi)|dx \tag{3.45}$$

The commonly used Distance term is the following-

$$D(x) = \Pi_{i=1}^{k}(1 - exp(\frac{|x - x_i|^2}{2\sigma^2}))\forall x_i \in M \tag{3.46}$$

Here, σ is the tuning parameter which tunes the distance map. However, even with the additional distance penalty the model struggles with when object edges are blurred in case of blurred images.

3.3.3 Badshah and Chen

The Gout et al. model [39] can generate a effective selective segmentation performance on images with no noise, whereas in the presence of background noise, the edge detector g is susceptible to picking false edges. To improve Gout model, Badshah and Chen [20] added intensity fitting terms from the Chan-Vese model.

$$F_{BC}(\phi, c_1, c_2) = \mu \int_{\tau} Dg(|\nabla z|)|\nabla H_{\epsilon}(\phi)|dx + \lambda_1 \int_{\Omega} (z - c_1)^2 H_{\epsilon}(\phi)dx + \lambda_2 \int_{\Omega} (z - c_2)^2 (1 - H_{\epsilon}(\phi))dx \tag{3.47}$$

When compared to Gout et al. model, the addition of the Chan-Vese fitting terms allows the Badshah-Chen model to more robustly segment noisy images, since it is less dependent on edge detection.

3.3.4 Rada and Chen

When the intended object is close to or connected with another object of similar intensity strength, the Badshah-Chen model is vulnerable to failure. Rada and Chen [52, 45]

proposed the following model to address this flaw by imposing size constraints on the object.

$$F_{RC}(\phi, c_1, c_2) = \mu \int_\tau Dg(|\nabla z|)|\nabla H_\epsilon(\phi)|dx + \lambda_1 \int_\Omega (z - c_1)^2 H_\epsilon(\phi)dx$$
$$+\lambda_2 \int_\Omega (z - c_2)^2 (1 - H_\epsilon(\phi))dx + \gamma[\int_\Omega (H_\epsilon(\phi)dx - A_1)^2 + \int_\Omega ((1 - H_\epsilon(\phi))dx - A_2)^2] \tag{3.48}$$

where A_1 and A_2 are the areas of the target and the background respectively. A_1 is the polygon area formed by joining the points of M and $A_2 = |\Omega| - A_1$. The selective fitting term uses no local information from set M, as a result segmentation result can be separated over the domain in small fragments, whose sum area shall be identical to the area fitting term used in the model.

3.3.5 Spencer and Chen Model

Spencer and Chen [24] separated the distance function from the edge detector term and used it as a single penalty term. The distance term D_E used was normalised euclidean distance from marker set M.

$$min_{\phi,c_1,c_2}[F_(u, c_1, c_2)] = \mu \int_\Omega g(|\nabla z|)|\nabla H_\epsilon(\phi)|dx + \lambda_1 \int_\Omega [(z - c_1)^2 H_\epsilon(\phi)]dx$$
$$+\lambda_2 \int_\Omega [(z - c_2)^2 (1 - H_\epsilon(\phi))]dx + \theta \int_\Omega D_E(x, y) H_\epsilon(\phi)dx \tag{3.49}$$

The Idea to use of D_E is to stop the contour from evolving far away from M by restricting $H(\phi) \in \frac{\Omega}{M}$to be close to zero. The selection of parameter θ is important for appropriate segmentation result. If θ too large, the segmentation result will be identical to marker set M. If θ is small, multiple objects are segmented. This model might get stuck between local minimisers. Spencer and Chen used the ideas discussed in 3.2.4 to reformulate a segmentation solution into a convex minimisation problem. The convex relaxed model [24] uses indicator function u as a replacement for the level set function ϕ.
Given, a two dimensional gray scale image $z(x, y) : \Omega \leftarrow R$, the Spencer-Chen Convex Geodesic model is-

$$min_{u,c_1,c_2}[F_(u, c_1, c_2)] = \mu \int_\Omega g(|\nabla z(x, y)|)|\nabla u|d\Omega + \int_\Omega [\lambda_1 (z(x, y) - c_1)^2 - \lambda_2 (z(x, y) - c_2)^2]ud\Omega$$
$$+\theta \int_\Omega D_E(x, y)ud\Omega + \alpha \int_\Omega \nu_\epsilon(u)d\Omega \tag{3.50}$$

Where μ, λ_1, λ_2, θ, and α are positive constants and u is the selective segmentation solution. c_1 and c_2 are the average image intensities of the foreground and background, respectively. The edge detector function is denoted by g, and the Euclidean distance term is denoted by

D_E. The additional penalty term $\nu_\epsilon = \max\{0, 2|u - \frac{1}{2}| - 1\}$ confines the minimizer to be in the range $[0, 1]$. Moreover, additional requirement for $u \in [0, 1]$ is to select parameter α as:

$$\alpha > \frac{1}{2}||[\lambda_1(z - c_1)^2 - \lambda_2(z - c_2)^2] + \theta D_E|| \tag{3.51}$$

The Euclidean distance term is extremely sensitive to the marker set chosen. The marker points must be evenly distributed in the region of interest to be segmented for satisfactory segmentation. Even in the segmentation region, the Euclidean penalty is always present, and it can be very high if there are few marker points. In the following subsection, we discuss a more robust geodesic distance penalty, which allows us to have small or zero distance penalties at the edges and within the segmentation region.

3.3.6 Convex Liu et.al

Recent convex models [23] also considered applying weighting to the data fitting terms of Mumford-Shah convex variant models. The convex Liu et. al [23] functional is -

$$F_{Liu}(u) = \mu \int_\Omega |\nabla u| dx + \mu_2 \int_\Omega |\nabla u|^2 dx + \lambda \int_\Omega w^2(x, y)|z - u|^2 dx \tag{3.52}$$

here, μ, μ_2, λ are non negative and w requires information of Distance term i.e. $w(x, y) = 1 - D(x, y)g|\nabla z|$. The effect of this weight means that, close to edges and marker points (where ω^2 is large) the fidelity term plays an important role and important features are preserved, whereas away from edges and marker points (where ω^2 is small) smoothing plays a more prominent role, smoothing out unwanted objects.

3.3.7 Roberts Chen Convex Geodesic Selective Model

In chapter 6 of this thesis, we will detail a selective segmentation model that has its roots in Roberts Chen convex geodesic model [15]. Here, we discuss the geodesic model in detail [15].

We discuss a more robust geodesic distance penalty, which allows us to have small or zero distance penalties at the edges and within the segmentation region as replacement for non zero euclidean distance. The Energy functional of convex geodesic selective model differs from initial segmentation models as it includes intensity fitting terms from the Chan-Vese model and a distance penalty term which uses geodesic distance from the marker set M in the distance penalty term rather than the Euclidean distance. The model involves a convex functional, which is to be minimized to achieve segmentation. The minimizer of this functional specifies the criteria to segment selective objects. The minimizer of the following

functional is in the form of partial differential equation. The next subsection reviews the methods used to solve the obtained partial differential equation in order to get the final segmentation.

Let $z(x, y)$ represent the input image, defined on a image domain $\Omega \subset R^2$. u represents the level set of contour . c_1, c_2 are average intensities of z inside and outside u. The functional is in the following form-

$$F(u, c_1, c_2) = \mu \int_\Omega g(|\nabla z|)|\nabla u| d\Omega + \int_\Omega [\lambda_1(z - c_1)^2 - \lambda_2(z - c_2)^2]u d\Omega + \theta \int_\Omega D_M(x, y)u d\Omega + \alpha \int_\Omega \nu_\epsilon(u) d\Omega \tag{3.53}$$

θ,μ,λ_1, λ_2 are non negative parameters. The term $g(|\nabla z|)$ is the edge detector which is $g(s) = 1/1 + \beta s^2$ where β is tuning parameter. The last term is an exact penalty term due to convex formulation of the functional, where $v(u)= max\{0, 2|u - \frac{1}{2}| - 1\}$. This is done to achieve unconstrained minimization as this encourages the minimizer to be in range [0,1]. Next we review the calculation of Geodesic term D_M.

The geodesic distance from the marker set M is given by $D_M(x, y) = 0$ for $(x, y) \in M$ and $D_M(x, y) = \frac{D^0_M(x,y)}{||D^0_M||_{L^\infty}}$ for $(x, y) \notin M$, where $D^0_M(x, y)$ is the solution of the following PDE:

$$|\nabla D^0_M(x, y)| = f(x, y), D^0_M(x0, y0) = 0, (x0, y0) \in M. \tag{3.54}$$

If $f(x, y) = 1$ then the distance penalty $D_M(x, y)$ is simply the normalised Euclidean distance. For selective image segmentation, we want small gradients in homogeneous areas of the image and large gradients at edges. So the f is set as -

$$f(x, y) = \epsilon_D + \beta_G |\nabla z(x, y)|^2 \tag{3.55}$$

Setting f in the above form ensures that in areas where $|\nabla z(x, y)| = 0$, the distance function increases by some small amount ϵ_D. Here, image z(x, y) is scaled to [0, 1] and at edges, geodesic distance increases as $|\nabla z(x, y)|$ is large. The constants of f are used as β_G = 1000 and $\epsilon_D = 10^{-3}$.

To improve noise robustness and qualitative nature of segmentation results, author considered an-isotropic TV denoising. The new f is formulated as -

$$f_1(x, y) = \epsilon_D + \beta_G |\nabla S^k(z(x, y))|^2 \tag{3.56}$$

Here S^k represents the gauss-siedel iterative scheme done to update z in order to do anisotropic

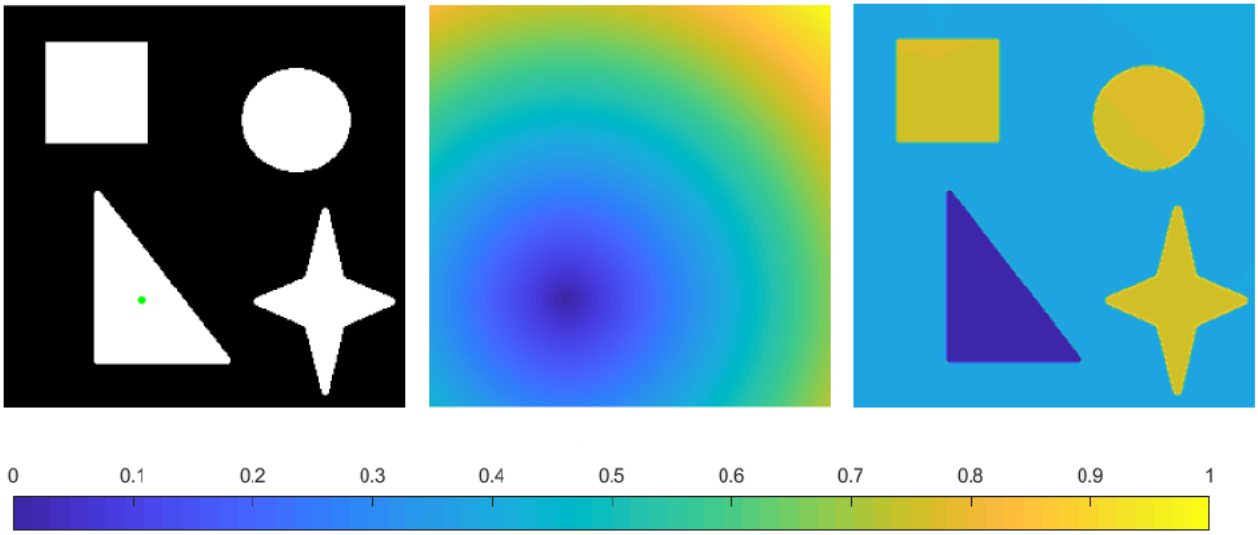


Figure 3.8: Comparison of distance constraints. (a) Input image with a marker point (b) Euclidean distance constraint (c) Geodesic distance constraint

denoising. To compensate the distance penalty for objects that are far from marker set with low penalty the author modified f as-

$$f_2(x,y) = \epsilon_D + \beta_G |\nabla S^k(z(x,y))|^2 + \nu D_E(x,y) \tag{3.57}$$

If ν increases, the distance function resembles the Euclidean distance D_E more. Authors used $\nu = 10^{-1}$ [15] as it adds a reasonable penalty to pixels far from the marker set. If there are blurred edges between objects in an image, the geodesic distance will not increase significantly at this edge. Therefore, the final segmentation result will include unwanted objects and will not result in the required accurate segmentation results. To improvise the model uses anti markers. Anti-markers are the markers or set of points which indicate the objects that we do not want to segment, i.e. the opposite of marker points, denoted by AM. Geodesic distance map from the set AM is denoted by $D_{AM}(x,y)$. Pixels near to the set AM are penalized as the following-

$$D_{AM}(x,y) = \frac{exp(-\bar{\alpha}\bar{D}_{GAM}(x,y)) - exp(-\bar{\alpha})}{1 - exp(-\bar{\alpha})} \tag{3.58}$$

Here $\bar{\alpha} = 200$ and $D_{GAM}(x,y)$ is normalised geodesic distance from set AM. The reformulated model is-

$$min_{u,c_1,c_2} F_(u,c_1,c_2) = \mu \int_\Omega g(|\nabla z(x,y)|)|\nabla u| d\Omega + \int_\Omega [\lambda_1(z(x,y)-c_1)^2 - \lambda_2(z(x,y)-c_2)^2] u d\Omega$$
$$+\theta \int_\Omega D_G(x,y) u d\Omega + \alpha \int_\Omega \nu_\epsilon(u) d\Omega$$

(3.59)

where $D_G(x,y) = \frac{(D_M(x,y)+D_{AM}(x,y))}{2}$ is the geodesic distance from marker set M. This is average of distance map of marker and antimarker set. D_M is to be calculated using $f = f_2$. Like the Chan Vase model we use calculus of variation and solve above equation with respect to c_1 and c_2 with u fixed. This leads to-

$$c_1(u) = \frac{\int_\Omega u.z(x,y))d\Omega}{\int_\Omega u d\Omega} \tag{3.60}$$

$$c_2(u) = \frac{\int_\Omega (1-u).z(x,y))d\Omega}{\int_\Omega (1-u) d\Omega} \tag{3.61}$$

Using calculus of variation and solving above equation with respect to u with fixed c_1 and c_2 leads to Euler's equation -

$$\mu\nabla(g(|\nabla z(x,y)|)\frac{\nabla u}{|\nabla u|_{\epsilon_2}}) - [\lambda_1(z(x,y)-c_1)^2 - \lambda_2(z(x,y)-c_2)^2] - \theta D_G(x,y) - \alpha\nu_\epsilon^{'}(u) = 0 \tag{3.62}$$

We also have Neumann boundary conditions $\frac{\delta u}{\delta n} = 0$ on $\delta\Omega$ where n is the outward unit normal vector. Now we discuss the numerical scheme to solve the above PDE.

3.3.8 Additive Operator Splitting

The AOS scheme [32, 53] allows equal treatment of all coordinate axes and is stable for big time steps. The scheme presents the semi-implicit algorithm [32, 53] based on a discrete non-linear diffusion scale-space framework. This scheme is applied to the m-dimensional diffusion equation and it takes the following form:

$$\frac{\delta u}{\delta t} = \mu\nabla(G(u)\nabla u) - f \tag{3.63}$$

$$\frac{\delta u}{\delta t} = \sum \frac{\delta}{\delta x_j}(G_j(u)\frac{\delta u}{\delta x_j}) - f \tag{3.64}$$

initial and boundary conditions are - $u(0,.) = u_0$ and $\frac{\delta u}{\delta n} = 0$ on $\delta\Omega$. Here, g is diffusivity function and f is reaction term. We consider discrete times $t_k = k\tau$, and m =2. After discretisation -

$$u^{k+1} = \frac{1}{2}\sum_{l=1}^{2}(I - 2\mu\tau A_l(u^k))^{-1}(u^k + \tau f) \tag{3.65}$$

τ is the time step. Matrices $A_1(u)$ and $A_2(u)$ are obtained as following-

$$(A_1(u^k)u^{k+1})_{i,j} = \frac{(G_{i+\frac{1}{2},j})}{h_x^2}u_{i+1,j}^{k+1} + \frac{(G_{i-\frac{1}{2},j})}{h_x^2}u_{i-1,j}^{k+1} + \frac{(G_{i+\frac{1}{2},j} + G_{i-\frac{1}{2},j})}{h_x^2}u_{i,j}^{k+1} \tag{3.66}$$

$$(A_2(u^k)u^{k+1})_{i,j} = \frac{(G_{i,j+\frac{1}{2}})}{h_y^2}u_{i,j+1}^{k+1} + \frac{(G_{i,j-\frac{1}{2},})}{h_y^2}u_{i,j-1}^{k+1} + \frac{(G_{i,j+\frac{1}{2}} + G_{i,j-\frac{1}{2}})}{h_y^2}u_{i,j}^{k+1} \tag{3.67}$$

For the half points in G, average of the surrounding pixels is considered. The AOS method described here assumes f does not depend on u; however, in the case of convex geodesic model, it depends on ν which has jumps around 0 and 1, so the algorithm has stability issues. The major part is to extract a linear part out of the non linearity in $f = f(u)$. Using Taylor expansion of $\nu_\epsilon(u)$ around u = 0 and u = 1 and grouping the terms into the constant and linear components in u, respectively, $\nu_\epsilon(u) = a_0(\epsilon) + b_0(\epsilon)u + O(u^2)$ and $\nu_\epsilon(u) = a_1(\epsilon) + b_1(\epsilon)u + O(u^2)$. It actually turns out that $b_0 = b_1$ and denote the linear term as b. This allows to approximate a change in $v_\epsilon^{'}(u)$ as $b.\delta u$. Next, interval is defined where $v_\epsilon^{'}(u)$ jumps as -

$$I_\xi = [0 - \xi, 0 + \xi] \cup [1 - \xi, 1 + \xi] \tag{3.68}$$

and linear function is -

$$\bar{b}_{i,j}^k = \{b, u_{i,j}^k \in I_\xi\} \tag{3.69}$$

Using these, we can now offset the change in $\nu_\epsilon^{'}(u^k)$ by changing the formulation to-

$$\frac{\delta u}{\delta t} = \mu\nabla(G(u)\nabla u) - \alpha\bar{b}^k u + [\alpha\bar{b}^k u - f] \tag{3.70}$$

this can be written in AOS form as-

$$u^{k+1} = u^k + \tau\mu\nabla(G(u^k)\nabla u^{k+1}) - \tau\alpha\bar{b}^k u^{k+1} + [\tau\alpha\bar{b}^k u^k - f^k] \tag{3.71}$$

This can be reformulated as following-

$$u^{k+1} = \frac{1}{2}\sum_{l=1}^{2}(I + \bar{B}^k - 2\mu\tau A_l(u^k))^{-1}((I + \bar{B}^k)u^k + \tau f^k) \tag{3.72}$$

here $\bar{B}^k = diag(\tau\alpha\bar{b}^k)$. This scheme does not satisfy the discrete scale-space conditions of Weickert [53] (which guarantee convergence of the scheme). It does not satisfy all of them.

In particular, the matrix $I + \bar{B}^k - 2\mu\tau A_l(u^k)$ does not have unit row sum and is not symmetrical. The authors [15] adapted the scheme to the equivalent which is -

$$u^{k+1} = \frac{1}{2}\sum_{l=1}^{2}(I - 2\mu\tau(I + \bar{B}^k)^{-1}A_l(u^k))^{-1}(u^k + \tau(I + \bar{B}^k)^{-1}f^k) \tag{3.73}$$

where the matrix $Q2 = (I - 2\mu\tau(I + \bar{B}^k)^{-1}A_l(u^k))$ does have unit row sum. Convergence is achieved in past literature for any small value of ξ [53]. Figure 3.9 demonstrates a selective segmentation solution achieved on abdominal image by utilising AOS in Robert Chen Convex Geodesic model.

Figure 3.9: (a) Input abdominal image with marker set (b) segmented image (c) zoomed segmentation result on a abdominal image.

Chapter 4

Arterial Input Function Segmentation based on a Contour Geodesic model for Tissue at Risk identification in Ischemic Stroke

Deconvolution of concentration -time curves with the Arterial Input Function (AIF) is necessary to identify regions with decreased blood flow in Ischemic stroke patients. This has been discussed in chapter 1 section 1.3 (1.2). The AIF is a key reference input curve used in the deconvolution model to obtain quantitative CBF, CBV and perfusion-diffusion mismatch estimation. Selection of the AIF curve influences the result of the deconvolution operation and this makes optimal AIF segmentation very important. The optimal AIF is segmented in form of multiple pixels (blue dots) in arterial region as in Figure 4.1. The corresponding AIF curve of the selected AIF pixels should have a baseline followed by a peak along with a regular recirculation part (Figure 4.1).

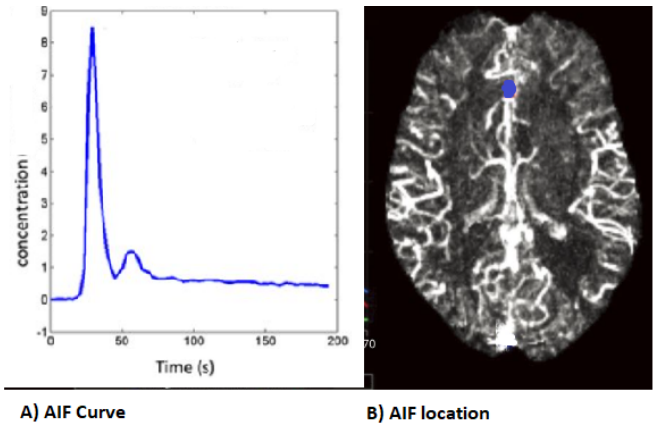


Figure 4.1: A)The AIF curve. B) selected arterial region on brain image.

In this chapter, we propose a segmentation model which via minimizing an energy, locates the arterial regions more accurately. After segmenting the arterial region, matrix analysis is utilised to find the voxel with maximum peak height within the contour to overcome the problem associated with shallow or low peak height AIF selection. This chapter is based on the author's published paper [3].

4.1 Introduction

There has been plenty of progress in recent years regarding how and where to measure AIF [54, 55]. Although the AIF should, in principle, be measured from inside an artery (or at least from a voxel that contains primarily arterial contributions), but many studies in past often considered measuring the AIF from the region outside or from a region in the vicinity of an artery [54, 56, 57, 55]. Also, from a practical point of view, due to the coarse spatial resolution of DSC- (the typical voxel size is 2x2x5 mm^3) it is difficult to measure the signal from inside a small artery[58]. Usually in MR-PWI studies suitable AIF voxels are chosen by inspecting the peak shape characteristics (e.g., arrival time, height, width, etc.) in a region in and around arteries. The name given to this input function is generally Arterial Input Function (AIF) [54, 59].

To improve reliability, quality, and reproducibility of the AIF selection several automatic and semiautomatic methods have been proposed [60, 58, 61, 62]. The majority of the toolboxes preinstalled in MR scanners use either manual, clustering or arterial likelihood methods for AIF estimation. For manual AIF selection, a trained clinical operator based on his or her experience and judgement selects a small number of pixels containing region of one of the principal arterial vessels [63]. Manual location of AIF is not preferred as this reduces the procedure reproducibility [58]. Low spatial resolution of MR-PWI data also makes manual selection difficult on contrast-MRI-PWI images [61, 60].

Automatic methods were developed to overcome the shortcomings of the manual AIF selection procedure [60, 63, 62]. The clustering based method uses the middle cerebral artery (MCA) as a elliptical region of interest (ROI) and then utilizes a recursive cluster analysis to select the arterial voxels [60]. Inefficient AIF selection usually occurs in the cases where the elliptical ROI does not segment the MCA precisely and some of the arterial voxels are left on the boundary or in the vicinity of the elliptical marker.

Some softwares use arterial likelihood methods so as to select the potential AIF to match the arterial features [61]. This includes minimizing the bolus arrival time, peak width and maximize the peak height. AIF detection algorithm searches for locations or voxels with signals of above-average amplitude or height along with below-average width and early bolus-arrival time using a cost function. Final AIF locations are selected in a region with the highest sum of the clustered values of cost function. Incorrect or flawed selection in this

method arises from the weighting factors used as the penalty factor used for peak height is much lower than the other penalty factors used in calculation of cost function. This results in selection of an AIF voxel with a shallow or low peak height. Apart from these methods several studies use different approaches, like a local AIF extraction method was introduced to replace the global AIF [59, 64, 65]. Despite of the presence of multiple studies to select AIF, in this study we mainly focus to use to a model to select a AIF with higher amplitude and early time to peak.

To overcome the limitation of past methods, we propose an improved convex segmentation model. The PWI images are usually of low contrast which makes detection of edges difficult [54]. To solve this problem we use a new idea of discrete Total Variation (TV) in a convex geodesic model. This TV helps in locating the boundary of arterial regions to separate homogeneous regions or intensity jumps [66]. The modified segmentation model via minimizing an energy can locate the arterial regions more accurately. After segmenting the arterial region, we use matrix analysis to find the voxel with maximum peak height within the contour to overcome the problem associated with shallow or low peak height AIF selection. Furthermore, to demonstrate better accuracy and arterial features obtained by the proposed model, a statistical comparison based on PWI dataset of 15 patients is made between the present method and the previous methods.

4.2 Methods

In the proposed method, we focus on the selective segmentation or specifying the location of the potential voxels which could be used as AIF in the vicinity of an artery. Initially a contour representing a region of interest (ROI) is drawn in the surrounding of the arterial location (Figure 4.2). For this purpose, convex based geodesic selective model is used to draw the contour on the middle slice of the brain axial images [67]. The advantage with a contour-based selective segmentation is exclusion of the CSF region as the contour model segments the ROI region based on homogeneous intensity values. After the segmentation of the ROI, the matrix analysis is used to find the potential voxel with maximum peak height within the contour (Figure 4.2). This ensures that the location or pixel with maximum height within the contour is selected as the potential AIF.

4.2.1 Proposed contour based AIF Segmentation method

The energy functional of convex geodesic selective model differs from initial segmentation models as it includes intensity fitting terms as well as distance penalty term which uses geodesic distance from the marker set rather than the Euclidean distance [67, 68, 69]. Here, we utilise a Total Variation function in the distance term of the model for denoising the image (cf. [66] for more information on the Discrete TV utilised in the model). The contour model involves a convex function and is to be minimized to achieve segmentation. The

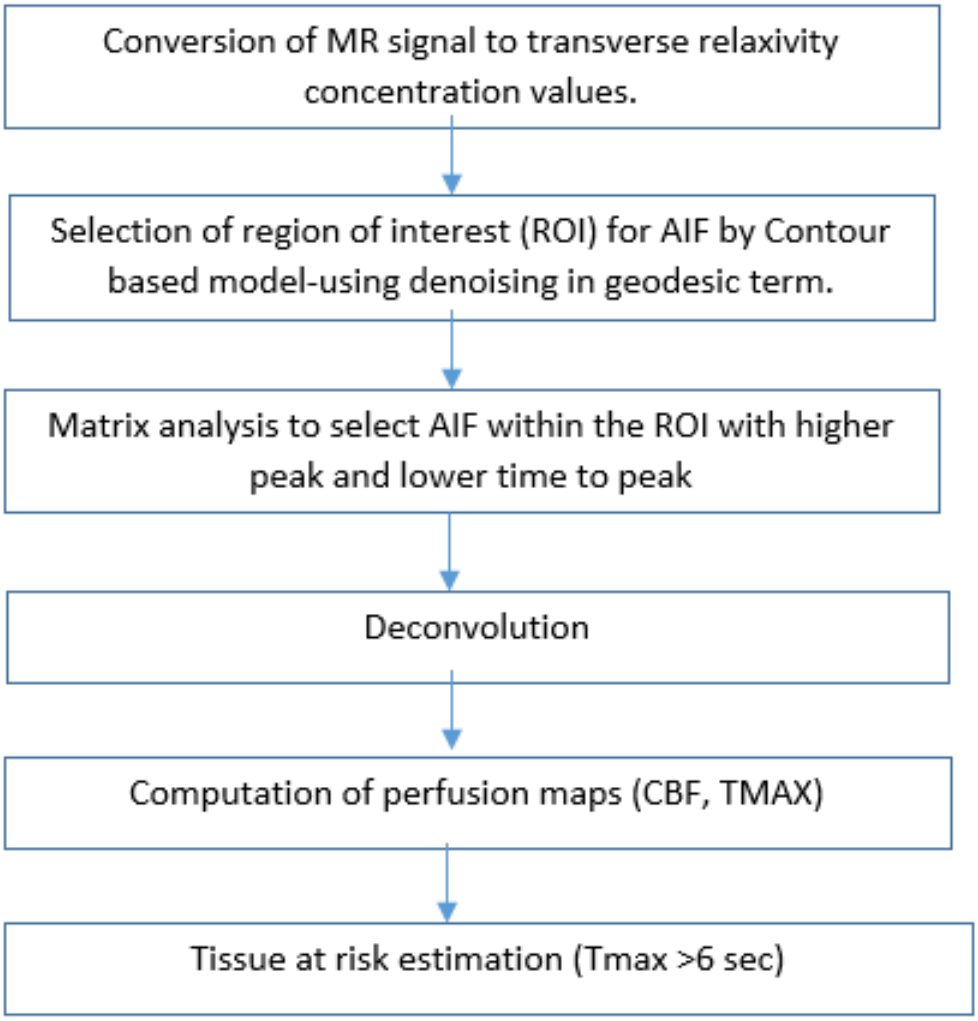


Figure 4.2: Model Pipeline used to estimate perfusion parameters after extracting AIF by a contour based geodesic model.

minimizer of this function specifies the criteria to segment selective objects. The minimizer of the function is in the form of partial differential equation. The definition of the function is -

Let $z(x, y)$ represent the input PWI image, defined on a image domain $\Omega \subset R^2$. u represents the level set of initial contour. c_1, c_2 are average intensities of z inside and outside u. The functional is in the following form-

$$F(u, c_1, c_2) = \mu \int_\Omega g(|\nabla z(x, y)|)|\nabla u| d\Omega + \int_\Omega [\lambda_1(z(x, y) - c_1)^2 - \lambda_2(z(x, y) - c_2)^2]u d\Omega + \theta \int_\Omega D_M(x, y)u d\Omega + \alpha \int_\Omega \nu_\epsilon(u) d\Omega \tag{4.1}$$

θ,μ,λ_1, λ_2 are non negative parameters. The term $g(|\nabla z|)$ is the edge detector which is $g(s) = 1/1 + \beta s^2$ where β is tuning parameter. The last term is an exact penalty term due to convex formulation of the functional, where $v(u)= max\{0, 2|u - \frac{1}{2}| - 1\}$. This is done to achieve unconstrained minimization as this encourages the minimizer to be in range [0,1]. We refer the reader to [68, 67, 70] for more information on the model. Next we illustrate the calculation of Geodesic term D_M [67].

The geodesic distance from the marker set M is given by $D_M(x, y) = 0$ for $(x, y) \in M$ and $D_M(x, y) = \frac{D^0_M(x,y)}{||D^0_M||_{L^\infty}}$ for $(x, y) \notin M$, where $D^0_M(x, y)$ is the solution of the following PDE:

$$|\nabla D^0_M(x, y)| = f(x, y), D^0_M(x_0, y_0) = 0, (x_0, y_0) \in M. \tag{4.2}$$

To improve noise robustness and qualitative nature of segmentation results, we considered TV denoising by utilising the new definition of TV. The formulation of the discrete TV to be used in the geodesic term is [66] -

$$TV(x) = min\{||v_\updownarrow||_{1,2} + ||v_\leftrightarrow||_{1,2} + ||v_\cdot||_{1,2} : L^*_\updownarrow v_\updownarrow + L^*_\leftrightarrow v_\leftrightarrow + L^*_\cdot v_\cdot = Dx\} \tag{4.3}$$

Here, v is the whole gradient field, which is the concatenation of $v_\updownarrow, v_\leftrightarrow, v_\cdot$ vector fields solution to above equation. Its elements $v_\updownarrow(n_1, n_2), v_\leftrightarrow(n_1, n_2), v_\cdot(n_1, n_2)$ are vectors located at positions $(n_1 + \frac{1}{2}, n_2), (n_1, n_2 + \frac{1}{2}), (n_1, n_2)$. The proposed TV is the $l_{1,2}$ norm of the gradient field v associated to the image x, defined on a grid three times more dense than the one of x [66]. Defining it on a three times finer grid allow this TV to detect edges in low contrast regions, when used in segmentation model (cf.[66] for more information). The new f is formulated as -

$$f(x, y) = \epsilon_D + \beta_G |TV_p(z(x, y))|^2 + \nu D_E(x, y) \tag{4.4}$$

Here TV_p $(z(x, y))$ represents the gradient field achieved after denoising is done with the new purposed TV and D_E is the euclidean distance. We use calculus of variation and solve

above equation (1) with respect to c_1 and c_2 with u fixed (cf.[68] for more information on solving the equation). This leads to-

$$c_1(u) = \frac{\int_\Omega u.z(x,y))d\Omega}{\int_\Omega u d\Omega} \tag{4.5}$$

$$c_2(u) = \frac{\int_\Omega (1-u).z(x,y))d\Omega}{\int_\Omega (1-u) d\Omega} \tag{4.6}$$

Using calculus of variation and solving above equation with respect to u with fixed c_1 and c_2 leads to Euler's equation [68, 70]

$$\mu\nabla(g(|\nabla z(x,y)|)\frac{\nabla u}{|\nabla u|_{\epsilon_2}}) - [\lambda_1(z(x,y)-c_1)^2 - \lambda_2(z(x,y)-c_2)^2] - \theta D_G(x,y) - \alpha\nu_\epsilon^{'}(u) = 0 \tag{4.7}$$

We also have Neumann boundary conditions $\frac{\delta u}{\delta n} = 0$ on $\delta\Omega$ where n is the outward unit normal vector. The Numerical solution of the above equation decides the contour that segments the arterial region (cf.[67, 53] for information on the numerical solution and the scheme used).

4.2.2 Purposed Matrix analysis to find the potential AIF within the contour

The steps used to select appropriate AIF voxels inside the segmented region were as following-

1) The coordinates (i, j) of segmented region inside contour represented by u were formed into an array A.

$$A = \begin{pmatrix}(i_1, j_1) & (i_2 & j_2) & ...(i_n, j_n)\end{pmatrix}^T \tag{4.8}$$

2) Matrix C had the information of concentration of contrast agent at each (x, y, z, t) of the brain images, where x, y were location of coordinates in brain image, z was the slice selected for AIF extraction and t represented time points.

$$\text{Conc} = \begin{pmatrix}x, y, z, t\end{pmatrix} \tag{4.9}$$

3) For the the n segmented (i,j) coordinates in the A array, we form following $1 \times n$ row vector C_1,C_2,..C_k at different time points k.

$$C_k = \begin{pmatrix}Conc(i_1, j_1, z, k) & Conc(i_2, j_2, z, k) & .....Conc(i_n, j_n, z, k)\end{pmatrix} \tag{4.10}$$

This is done to form a final $k \times n$ matrix F which represents information of concentration of contrast agent in all the selected voxels inside contour at different time point in a row wise manner.

$$F = \begin{pmatrix} C_1 & C_2 & .....C_k \end{pmatrix}^T \tag{4.11}$$

4) Maximum of F matrix will be the highest amplitude of concentration curves among all time points and all the voxels. This purposed analysis is used to trace back the spatial location of the best potential AIF voxels. Finally, global AIF for perfusion analysis is represented by the contrast agent concentration of the selected AIF voxel.

4.2.3 Perfusion Data acquisition

During the diagnostic MR procedure, fifteen stroke patients underwent perfusion imaging. A clinical 1.5 T MR scanner at the Tri-service General Hospital, Taipei (Signa; General Electric) was used to acquire contrast-enhanced T2*-weighted images. Single-shot gradient-echo EPI sequence was utilised (TR : 1800 ms, TE : 40 ms). During Perfusion imaging, bolus injection (Magnevist; gadopentetate dimeglumine, Bayer Health Care pharmaceuticals Inc.) was injected with the speed of 5 ml/sec and quantity was 20 ml. After the contrast agent passes through the tissues, the decrease in signal intensity depends on the contrast agent concentration, which is considered as a proxy for perfusion. The acquired time series data are then postprocessed to obtain perfusion maps with different parameters. The additional benefit of using this type of dataset is to accentuate local magnetic homogeneity effects to aid in the detection of hemorrhage, core and better segmentation [71]. This study was granted IRB approval from the Tri-Service General Hospital, Taipei, Taiwan.

4.2.4 Statistical analysis and Perfusion parameter estimation

Statistical analysis

AIF location on the brain axial slices was decided by utilising different methods: clustering method, arterial likelihood method and contour based AIF segmentation method. Due to the different patient conditions, physical condition, severity of the disease, the contrast injection time, and due to variable time to start the scan, statistical comparisons are only made by using the differences of the curve parameters [58]. These curve parameters are amplitude (peak), the center position of the peak of concentration curve or time to peak, the differences are represented by Δ amplitude (a.u), Δ center (sec). One-way ANOVA statistical test was used to establish whether there is a significant difference in terms of amplitude of AIF selected by the different three methods.

Perfusion Parameter estimation

Perfusion DSC model was used to compute the perfusion parameters (CBF and Tmax) with the global AIFs deduced from different methods: clustering method, arterial likelihood

method and contour based AIF segmentation method. Perfusion analysis was done once global AIF was decided by the AIF selection methods. On the lines on past perfusion studies in ischemic stroke, perfusion analysis was done by deconvolution of the tracer kinetic equation [55, 72, 9, 73].

$$C_t = C_a \otimes R(t) \tag{4.12}$$

All the image analysis were implemented in MATLAB (Mathworks, Natick, MA). Here $C_t(t)$ denotes the tissue concentration curve at each pixel, $C_a(t)$ is the AIF either using one of the three AIF selection methods described above, symbol $\otimes$ represents the convolution operator and $R(t)$ represents the residue impulse response function. Deconvolution of Eq. 4.12 to estimate CBF, Tmax was done using the singular value decomposition method (SVD) [74, 55, 75]. Deconvolution of Eq. 4.12 for known values of $C_a(t)$, $C_t(t)$ at each pixel of axial slices leads to evaluation of the residue function $R(t)$. CBF is measured as the maximum of $R(t)$ [61]. Tmax is the time t for which $R(t)$ attains maximum value [61, 73]. After estimating CBF and Tmax for all brain tissues, CBF and Tmax are represented visually on axial slices. Tissue at risk was identified by thresholding the Tmax values by Tmax >6 seconds.

4.3 Results

4.3.1 Statistical analysis of Curve Characteristics

Subjectively, the concentration curve of AIF extracted by contour based AIF segmentation method confirmed to the arterial characteristics, such as large amplitude, small width, fast attenuation, and gamma-like shape (Figure 4.3). In terms of AIF location, it is visible that the location selected by contour based AIF segmentation method is quite close or in proximity of the AIF location selected by arterial likelihood method (Fig. 4.3). In terms of curve characterstics comparison, contour based AIF segmentation method selects AIF curve with larger amplitude or higher peak position and with fast attenuation represented by early time to peak or positive Δ center (Figure 4.3). We also calculated the similarity of the AIF concentration curves. The similarity was calculated by Correlation Coefficient, which indicates that the curves are positively correlated (Table 4.1, Table 4.2, Table 4.3).

The AIF curve characterstics of all other subjects are represented in the form of statistical tables (Table 4.1, Table 4.2). We also show the Group mean differences between the contour based peak height AIF method and the previous AIF selection methods to get a group overview. The group mean differences indicate that overall the AIF selected by contour based AIF segmentation method has better arterial features of higher peak position (Figure 4.3), and fast attenuation as compared to the other AIF selection methods (Table 4.3). A one-way ANOVA (Figure 4.5) revealed that there was a statistically significant difference in AIF amplitude (peak) between the three AIF selection methods ($F(2, 42) = 5.66$, $p = .0067$).

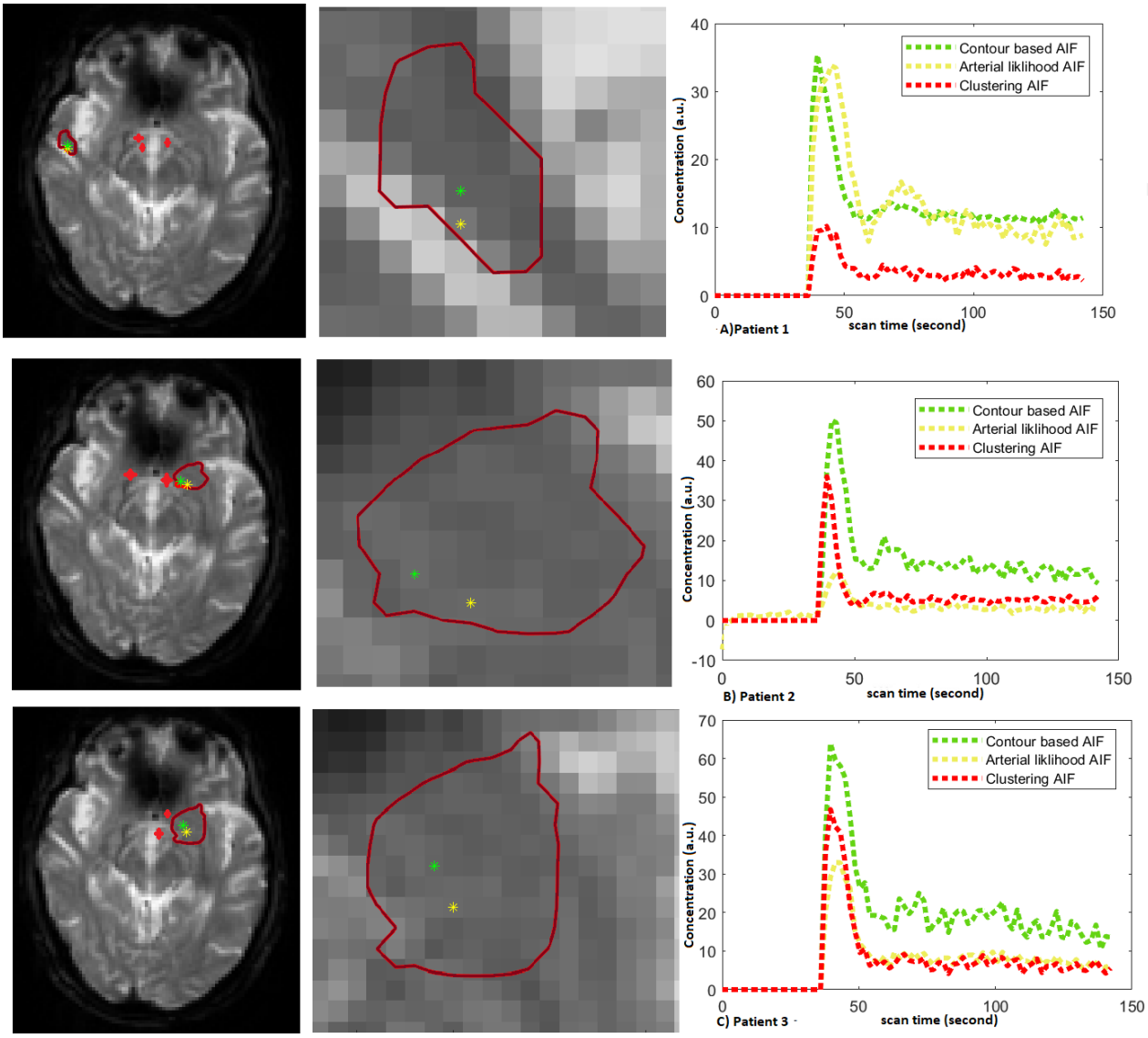


Figure 4.3: (a)AIF voxels selected by the contour based AIF (green), arterial likelihood selection method (yellow) and clustering method (red) (left column) (b) Zoomed in images of contour (dark red) used for AIF selection demonstrates that the voxels selected for contour based AIF (green), arterial likelihood selection method (yellow) were very close to each other. (central column) (c) Concentration curves of the selected AIF voxels (Right column). contour based AIF segmentation method (green curve) selects AIF curve with larger amplitude or higher peak position, and fast attenuation represented by early time to peak or positive Δ center than the latter two methods. Each row demonstrates different patient.

Table 4.1: The difference of curve characteristics between the arterial liklihood method and contour based peak height AIF selection method.

Sample	amplitude (a.u)			center (s)			Correlation
	Contour based AIF	Arterial likelihood	Δamplitude (a.u)	Arterial likelihood	Contour based AIF	Δcenter (s)	
1	29	21	8	41.4	39.6	1.8	0.8
2	41.4	41.4	0	17.7	17.7	0	1
3	90.8	27	63.8	27	27	0	0.8
4	46	20	26	43.2	41.4	1.8	0.6
5	80.6	64.3	16.3	33.9	37.5	-3.6	0.8
6	12	4.3	7.7	50.4	45	5.4	0.7
7	36	19.6	16.4	48.6	46.8	1.8	0.9
8	59.7	13.2	46.5	43.2	45	-1.8	0.7
9	39.5	33.7	5.8	45	43.2	1.8	0.9
10	50.1	12.2	37.9	43.2	41.4	1.8	0.9
11	53.8	33.7	20.1	41.4	41.4	0	0.9
12	69.7	64.1	5.6	39.6	37.8	1.8	0.8
13	42.5	16.1	26.4	34.2	34.2	0	0.9
14	109.5	42.8	66.7	36	37.8	-1.8	0.7
15	56.5	10.8	45.7	41.4	36	5.4	0.8

Table 4.2: The difference of curve characteristics between the clustering method and contour based peak height AIF selection method.

Sample	amplitude (a.u)			center (s)			Correlation
	Contour based AIF	Clustering method	Δ amplitude (a.u)	Clustering method	Contour based AIF	Δcenter (s)	
1	29	3.4	25.6	41.4	39.6	1.8	0.8
2	41.4	29.8	11.6	19.5	17.7	1.8	0.8
3	90.8	45.7	45.1	30.6	27	3.6	0.2
4	46	36	10	41.4	41.4	0	0.8
5	80.6	65.2	15.4	37.8	37.8	0	0.8
6	12	60.1	-48.1	50.4	45	5.4	0.7
7	36	61.7	-25.7	50.4	46.8	3.6	0.9
8	59.7	48.9	10.8	39.6	45	-5.4	0.7
9	39.5	10.2	29.3	43.2	43.2	0	0.9
10	50.11	36	14.11	39.6	41.4	-1.8	0.9
11	53.8	46.7	7.1	39.6	41.4	-1.8	0.9
12	69.7	64.1	5.6	41.4	37.8	3.6	0.8
13	42.5	16.1	26.4	36	34.2	1.8	0.8
14	109.5	42.8	66.7	37.8	37.8	0	0.7
15	56.5	10.8	45.7	39.6	36	3.6	0.8

The clustering method and arterial likelihood methods have the same peak for patients 12-15 (Tables 4.1 ,4.2). The contour-based method uses a matrix based approach to ensure that out of all the potential AIF voxels in the marked contour the selected AIF voxel has the maximum peak concentration. In subjects 12-15 the other two methods miss out AIF with maximum peak which is a feature of utilizing the matrix analysis post selection of ROI for AIF by contour-based model.

In Figure 4.3 we have shown the cases where the contour method selects AIF with better arterial features i.e., high peak and early time to peak than the latter two methods. In Figure 4.4 , both the methods select a similar AIF voxel and this represents that in some cases both arterial likelihood method and contour-based method may yield the same result for AIF i.e. in this case arterial likelihood method may not miss out the peak AIF voxel. In contrast, for all other subjects both the AIF locations are quite close but the arterial likelihood method misses out the location with highest peak (Figure 4.3). This could be due to the varying physical conditions, severity of the disease, noise associated with signals and the variability of contrast injection time among different patients. Although we processed all the samples, considering the number of samples, we only showed selected AIF location and corresponding concentration curve of three of them. For a patient, the contour method yields an AIF curve after 14 s (seconds). Time taken by the clustering method and arterial likelihood method for the AIF estimation was 9 s and 13 s. (Intel I5/Ram :8gb/ MATLAB 2020(a)).

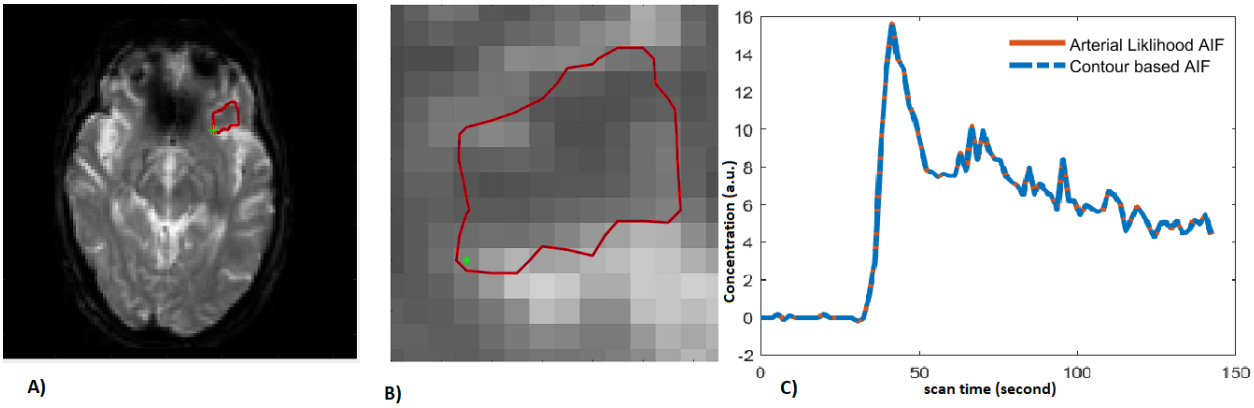


Figure 4.4: (A)Similar AIF voxel (Green) selected by the contour based AIF and arterial likelihood selection method (left column) (B) Zoomed in images of contour (dark red) used for AIF selection demonstrates similar voxel selected for contour based AIF (green) and arterial likelihood selection method (central column). (C) AIF Concentration curves of the selected AIF voxel (right column).

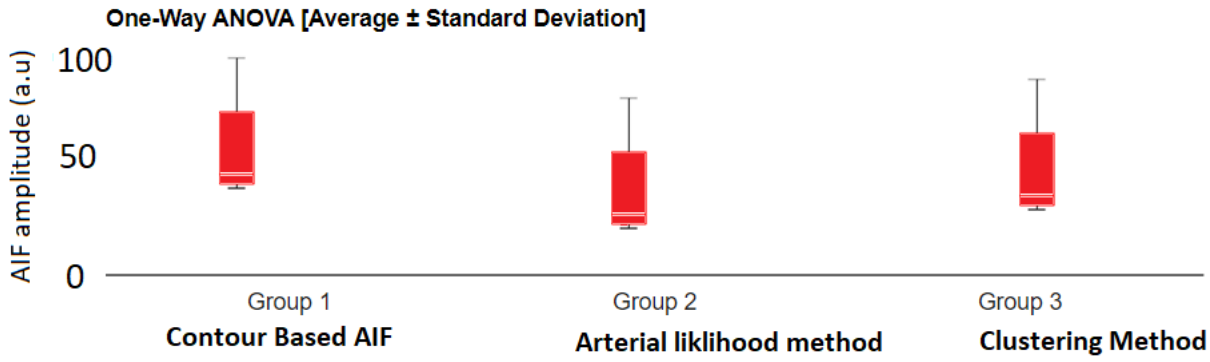


Figure 4.5: Comparison of AIF amplitude for the AIF selected by the three methods in the patient cohort. A one-way ANOVA revealed that there was a statistically significant difference in AIF amplitude (peak) between the three AIF selection methods (F value = 5.66 , P value = .0067).

Table 4.3: Group mean difference between the Contour based peak height AIF method and the previous AIF selection methods.

Method	amplitude (a.u)	center (s)	Correlation coefficient
Clustering method	16.7	1.08	0.7
Arterial likelihood method	26.1	0.9	0.8

Table 4.4: Perfusion parameters (Tmax and CBF) for different AIF selection methods.

Method		Tmax (s)	CBF (a.u)
Contour Based AIF	Mean	1.87	81.9
	SD	2.09	64.7
Arterial Likelihood method	Mean	1.6	178.3
	SD	3.01	139.9
Clustering method	Mean	1.13	296.5
	SD	2.8	229.6

4.3.2 Perfusion maps

We derived the perfusion parameters (CBF, Tmax) corresponding to the AIF given by all three methods in each pixel in each sample. For comparison we used a similar SVD deconvolution method with the optimal standard threshold [74]. The perfusion maps were accessed by an experienced clinical from veterans hospital. Based on the feedback investigators concluded that perfusion parameter maps could be utilised for diagnosis.

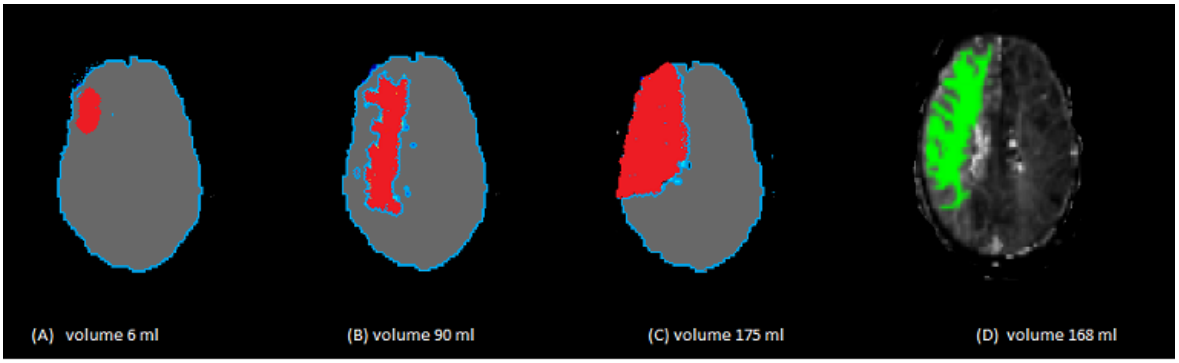


Figure 4.6: Tissue at risk (Red) maps estimated by AIF from (A) clustering (B) arterial likelihood method and (C) contour based AIF from left to right. (D) Tissue at risk identified by the commercial software. Tissue at risk is identified by Tmax> 6 sec and is overlaid on brain masks. Among the three methods, contour based AIF method has the closest prediction of Tissue at risk (168 mL) with the tissue at risk identified by the commercial software (175 mL) considered as golden standard for perfusion processing outcome.

The distribution of CBF and Tmax maps based on the AIF selected by all three methods is basically the same, however ischemic regions or tissues at risk can be clearly located through the perfusion maps given by contour based segmentation method (Figure 4.6, Figure 4.7). The mean and standard deviation of perfusion parameters (CBF and Tmax) over a cohort of all patients are shown in Table 4.4. The other two AIF selection methods estimate

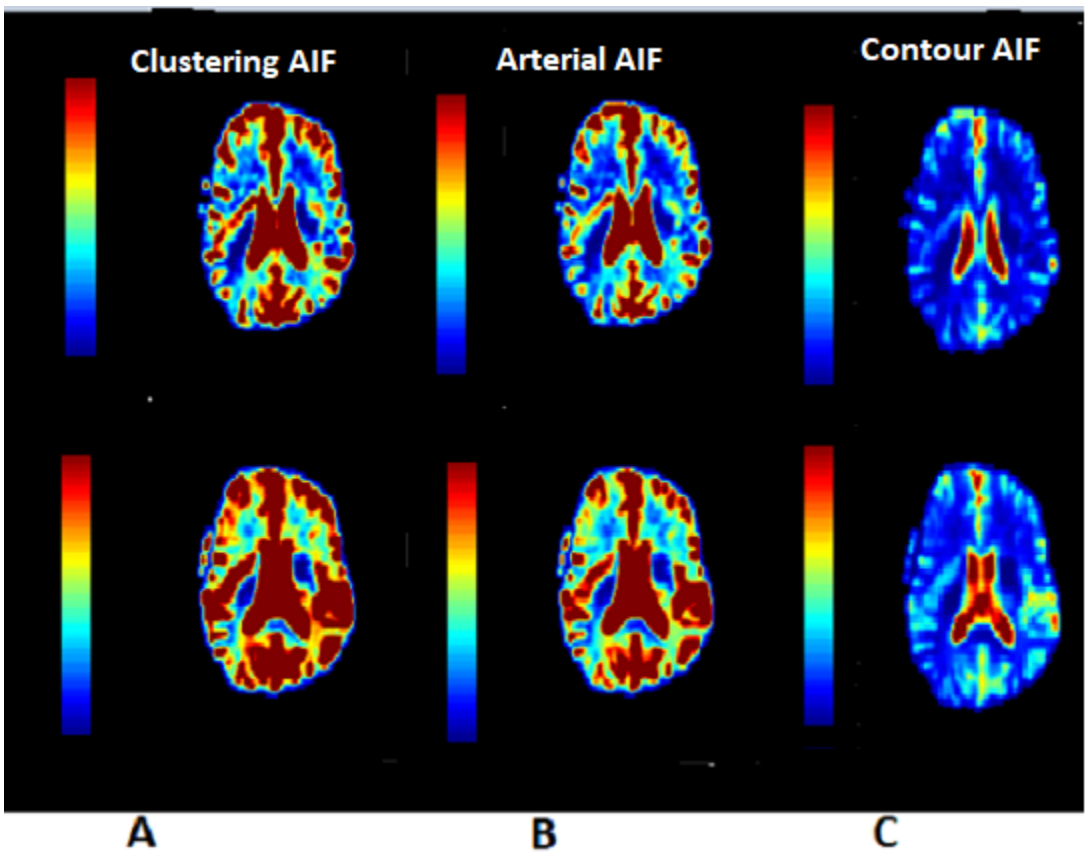


Figure 4.7: CBF maps estimated by AIF from (A) clustering (B) arterial likelihood method and (C) contour based AIF from left to right. CBF values obtained from the contour based AIF method are lower than the latter two methods due to the larger peak and lower time to peak of the AIF. This allows to locate the core regions with decreased blood flow more precisely and accurately.

lower Tmax and higher CBF values, which misleads us in terms of severity of Ischemia and the size of tissue at risk. The mean CBF values obtained by our method are in general lower than those obtained by the other two methods (Table 4.4). The mean Tmax values obtained by Contour based AIF selection method are higher than those obtained by the other two methods (Table 4.4). Higher Tmax and lower CBF values reported are due to the early time to peak and the larger peak value of the AIF curve.

With the help of a sample example of a stroke patient we illustrate that tissue at risk was clearly located with improved visual specificity (Figure 4.6). The clustering based AIF method failed to estimate the tissue at risk in this case (Figure 4.6). Contour based method estimated the size and volume of tissue at risk similar to the size estimated by a widely used commercial software considered as golden standard for perfusion processing outcome.

4.4 Discussion

4.4.1 AIF and contour based models

Contour based segmentation models have been widely used for object segmentation in images with noise and inhomogeneous intensities [67, 76, 46]. They are usually based on a functional and the minimizer of this functional decides the accuracy of segmentation [46, 68, 70]. However, up to date this has not yet been applied for AIF estimation in PWI studies. Experimental and comparison results suggest that the discussed method performs better in terms of AIF estimation as compared to earlier methods. The present method has been proved to be robust to detect voxels with large amplitudes, small width, fast attenuation, and gamma-like shape as potential AIF. The utilization of discrete TV allows the contour model to deal with noisy data sets as well as with in homogeneous intensities.

Recent studies utilized a deep convolutional neural network (CNN) to automatically identify AIFs in computed tomography perfusion (CTP) and perfusion-weighted MRI (PWI) datasets [77, 58]. These studies concluded that CNN network models could be potentially viable as the cross-correlation values of manual AIFs with CNN AIFs were observed higher than the AIF decided by the traditional methods [77]. The CNN-derived AIFs for the PWI data-sets showed marginally greater peak heights and early time to peak. CNN models require the choice of ground truth as an input and this ground truth is mainly a manual selected AIF. Here, to provide ground truth user has to inspect each voxel which is inside the arterial ROI. This might be time consuming and there are high chances of missing a voxel which could represent a AIF with better arterial features. Comparatively, the purposed method is selective and just requires a single click to set a marker point or to localize the arterial region as region of interest (ROI) and find a suitable AIF voxel.

4.4.2 Tissue at risk and limitations

Tissue at risk or Ischemic penumbra denotes the stroke region that is at risk of progressing to infarction but is still salvageable [61, 78]. Ischemic penumbra is usually located around an infarct core which represents the infarcted or blocked necrotic tissue [79]. Cerebral perfusion in terms of parameters is the key information that helps to locate the penumbra around the infract core [58, 79, 80]. AIF plays a central role in cerebral perfusion estimation. PWI studies proved that AIF measured with a lower amplitude, large width and slow attenuation could produce a four times blood flow overestimation along with inaccurate ischemic penumbra [81]. We could observe that (Figure 4.6) difference in AIF selection makes a substantial bias in the estimation of ischemic penumbra.

With the help of an example (Figure 4.6) we demonstrate that the accuracy and visual reliability to identify tissue at risk with our model is promising. Among the three methods, contour-based AIF method has the closest prediction of the tissue at risk in comparison to a commercial software. Detection of ischemic infraction is important because of narrow window of therapeutic efficacy. Inclusion of this fast and efficient AIF selection algorithm presented in this study in clinical settings may optimise the delivery of stroke care. The proposed method could potentially be considered as part of the calculation for perfusion imaging in general.

This study has several limitations. In this study, we used MR-PWI data set of 15 patients. In clinical settings, collecting the data-sets for a broader patient cohort is challenging due to the restricted access to urgent MR-PWI and the contraindications (e.g., uncharacterized metallic foreign bodies) related to MR-PWI acquisition [82]. Recent studies have demonstrated that Computed Tomography Perfusion (CTP) can provide information related to treatment decision making at a level similar to MR-PWI [82]. Due to the greater accessibility of CTP, further CTP studies on a large data set with variability of onset of stroke are required to demonstrate the consistency of purposed AIF selection method. Also, it would be worth to see if the proposed model can deal with the noise and in-homogeneity in the CT perfusion images.

4.5 Conclusion

This chapter proposes a contour-based segmentation model for estimating AIF curves in brain perfusion images. This segmentation framework worked on perfusion images at levels superior to the current clinical state of the art. The model estimated AIF curves with higher amplitude and early time to peak along with a good performance in identifying the tissue at risk. Inclusion of this improved AIF selection methodology discussed in the study will facilitate prediction and localization of the ischemic penumbra, which in turn may optimise the delivery of stroke care and surgical pharmaceutical treatments.

Chapter 5

Partial Volume Effects correction and Curve fitted Arterial Input Function for Core and Penumbra estimation

After optimal Arterial Input Function (AIF) segmentation as purposed in chapter 4, there are still some artefacts that affect the penumbra and core estimation in stroke patients. In this chapter, we discuss two key problems associated with the perfusion parameter estimations.

In first section of this chapter, we focus on the partial volume effect correction (PVE) for AIF. Firstly, we introduce basic notion of PVE. During AIF segmentation or AIF pixel selection on brain axial slices, most of the times the AIF represents only some fraction of the arterial blood, known as partial volume effect. In light of the substantial influence of the PVE on AIF as well as perfusion parameters, we investigate potential approaches for minimization of the volume averaging artifacts associated with the PVE. If not corrected, these artifacts generate misleading CBF and Tmax brain images that fail to identify infract regions as shown by an example (Figure 5.1). This section is based on authors published paper [6] .

Segmented AIF curves with a distorted baseline, shape-amplitude errors in the first passage, and non-identical recirculation portions as shown in Figure 5.2 (B), predict incorrect perfusion parameters [2]. In the second section of this chapter, we use deep learning to generate a distortion free AIF. An example of distortion free AIF or CNN AIF is given in Figure 5.2 (C). In this example, predicted AIF combines features of consistent baseline, peak amplitude, and identical recirculation portions. The utilisation of this CNN based AIF in perfusion parameter pipeline 1.2 improves volumetric assessment of core and penumbra is association with clinical scores. This section is based on a published conference paper (European Congress of Radiology 2023, https://www.myesr.org/congress).

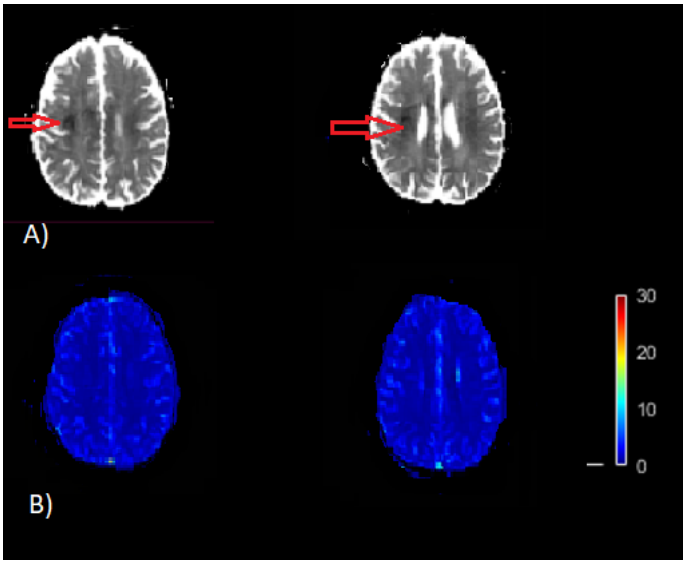


Figure 5.1: Example of misleading perfusion parameter maps. (A) Apparent diffusion coefficient (ADC) image is an indicator of infract region. The dark region on ADC image thresholded by ADC $\leq 620 \times 10^{-6} mm^2/s$ is the infracted core (red arrow). (B) CBF map (bottom) [mL/100 g/min]. This CBF map does not indicates the infract region as represented on the ADC map, which is a result of inaccurate quantification of CBF.

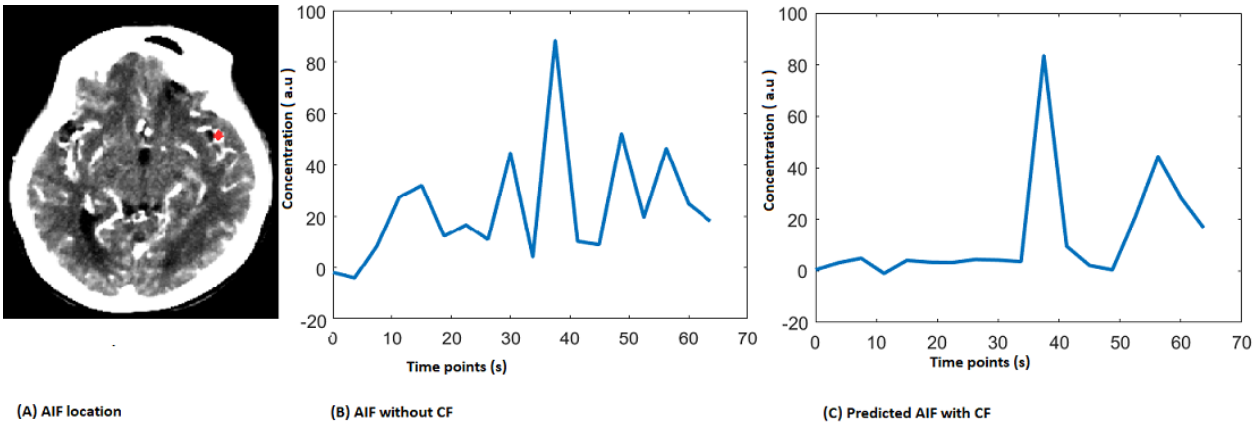


Figure 5.2: Curve Fitted AIF for a single dataset (A) AIF location is marked as a red dot. (B) AIF curve with distorted baseline and shape amplitude errors (C) Predicted curve fitted AIF curve by CNN model as an output.

5.1 Partial volume affect correction for Arterial Input Function (AIF)

As discussed in chapter 4, in perfusion quantification's, one of the most important function which is required as input from the perfusion data, is the arterial input function (AIF) [2]. When the voxel or region that has been selected for quantification of AIF has only some fraction of the arterial blood, the partial volume effect (PVE) arises. During measurement of AIF, spatial resolution used in perfusion MRI and the average size of major arteries make a degree of partial volume unavoidable [83, 2, 10]. The arterial and tissue contributions are complex numbers with amplitude and phase. The magnitude and phase of tissue components in the selected AIF voxel may decide whether there will either destructive or constructive contributions by the tissue components [2].

The perfusion model 1.2 uses AIF as an initial input to calculate perfusion parameters as final output [2]. In the process of calculation of perfusion imaging parameters, the PVE seriously affects the estimation of Arterial Input Function (AIF) [10, 84, 85, 86, 87, 88] . As voxels with signals from both artery and surrounding tissues may result in distortion of the signal loss of the contrastive agent during the passage of blood, this may lead to erroneous estimation of AIF, which consequently yields misleading brain maps of CBF and Tmax. In current clinical practices, volumes on brain image with Tmax greater than 6 s are considered to be the critically hypoperfused region, which is also known as the penumbra, and tissues with relative CBF < 30 % are considered to be the infarct core [89]. Early and correct assessment of the hypoperfused as well as infract regions are critical for appropriate diagnosis and treatment decisions in acute stroke [10, 88, 89].

We hypothesize that correct estimation and reasonable perfusion parameters can be achieved by several rescaling methods of AIF. This enhancement could be evidently seen from the Tmax and CBF images before and after rescaling. This proposed enhancement method will use multiplicative rescaling on the multiple AIF voxels to minimize the underestimation or overestimation of AIF and CBF values. By increasing the size of the region from which the AIF is sampled, we will demonstrate the increase in the PVE and the increased distortion of AIF estimation. Three different multiplicative rescaling approaches are used to rescale the AIF, as follows: (a) scaling using AIF curve; (b) scaling using VOF curve; (c) scaling by matching peaks. The rescaling factor is decided according to different rescaling approach and is applied to the AIF concentration curves. Finally, the variation in CBF value estimated from the reference AIF with minimal PVE and the AIF concentration curve after rescaling is evaluated. Based on these comparisons, an optimal scaling method to minimize the PVE is determined and the perfusion parameter maps are generated. It is anticipated that the scaling approach will generate rational parameters, as it takes into factor the conservation of the time integral of the tracer concentration curve, C(t), through the vasculature, which might affect the AIF calculation most.

5.1.1 Materials and Methods

Data Acquisition

Fifteen patients suffering from acute ischemic stroke underwent perfusion imaging as a part of their diagnostic MR procedure. A single-shot gradient–echo EPI sequence (TR/TE/flip angle/number of slices/voxel size: 1800 ms/40 ms/60/23/1 × 1 × 5 mm) on a clinical 1.5 T MR scanner (Signa; General Electric) was used to acquire contrast-enhanced T2*-weighted images from the Tri-Service General Hospital, Taipei. During perfusion imaging, with the speed of 5 mL s^{-1}, a dose of 20 mL of bolus injection (Magnevist; gadopentetate dimeglumine, Bayer Health Care pharmaceuticals Inc, Berlin, Germany) was injected. The present study was granted the review board approval by the Tri-Service General Hospital, Taipei, Taiwan.

Approaches for correction of the Artefacts that arise from the PVE

Voxels containing signal contributions from both the artery and the surrounding tissues are referred as voxels affected by the PVE [84, 85]. Since the signals come from the artery and the neighboring tissues, we first define the signals from these regions. Suppose we have PVE-affected voxels, which are selected for AIF estimation; they are composed of k and t fractions of arterial blood signal (S_a) and tissue signal (S_t). The MR signal from the entire voxel (S_v) then reflects the weighted average of signals S_a and S_t, as follows:

$$S_v = kS_a + tS_t \tag{5.1}$$

If it is assumed that the tissue contribution is much smaller than the arterial contribution ($kS_a > tS_t$), then multiplicative rescaling can be used to estimate correct arterial signal from the measured voxel signal by multiplication with the inverse volume fraction of arterial blood [85] i.e.,

$$S_a = \frac{S_v}{k} \tag{5.2}$$

A direct evaluation of S_a is not possible since k is unknown. The present study implements three different criteria to determine the rescaling factor k. The MR signals are converted to concentration curves based on a traditional nonlinear relationship provided by earlier studies [5, 2].

$$S(t) = S_0 e^{\frac{-TE}{p} C(t)} \tag{5.3}$$

S_0 is the baseline (pre-bolus or pre-contrast agent) MR signal intensity, TE is the echo time, p is the proportionality constant taken as p = 1 [90], and $C(t)$ represents the concentration time function. A direct expression for concentration values based on MRI signal data is derived by inverting Equation 5.3, as follows:

$$C(t) = -(\frac{p}{TE})ln\frac{S(t)}{S_0} \tag{5.4}$$

A reference arterial curve derived from a voxel with minimal PVE is manually selected. The MR concentration from the entire voxel (C_v) reflects the weighted average of concentration in artery (C_a) and concentration in tissue (C_t), which can be expressed as $C_v = kC_a + tC_t$. Assuming that the concentration in tissue is much smaller than the concentration in the artery, this can be simplified as $C_v = k \times C_a$. Following this argument, k is estimated from concentration values, as follows:

$$\int_0^\infty C_{ref}(t)dt = \int_0^\infty C_a(t)dt = \int_0^\infty \frac{C_v(t)}{k}dt \tag{5.5}$$

The first two parts of the equation imply that AIF concentration time curve (C_a) has the same area under the curve (AUC) as any other manually selected reference concentration time curve (C_{ref}) [91]. The relationship of conversion of MRI signal to concentration values is nonlinear (Equations 5.3). In this study, we intend to make use of AUC of concentration curves rather than the signal curves to derive scaling factor k. So, we assume that during the calculation of AUC of arterial curve the non-linearity of signal to concentration conversion will have minimal effect. On the similar pattern of Equation (5.2), AUC of arterial curve would be ratio of AUC of the concentration curve from multiple voxels to the scaling factor, this is represented analytically in last two parts of Equation (5.5).

The first rescaling approach referred as scaling by AIF uses concentration curve of a reference AIF as (C_{ref}) and concentration curve of selected multivoxel AIF region as (C_v) in Equation 5.5 to estimate the scaling factor k. The second rescaling approach referred as scaling by VOF uses concentration curve of a venous output function (VOF) as C_{ref} and concentration curve of selected multivoxel AIF region as (C_v) in Equation 5.5 to estimate the scaling factor k. The first two rescaling approaches are based on the principle of conservation of time integral of tracer concentration curve $C(t)$ through the vasculature. The third rescaling approach, referred to as scaling by peak, estimates the scaling factor k by matching the peak height of the multi-voxel AIF concentration time curves with the reference concentration time curve. This follows that the multivoxel AIF will have similar characteristics to any other concentration curve in terms of peak height [85].

Reference AIF Curve

During PWI-MRI, the internal carotid artery (ICA) is nearly perpendicular to the axial plane and offers the advantage of easy and reliable manual identification with minimal errors from volume-averaging artifacts [2]. As demonstrated in Figure 5.3 (a), internal carotid artery (ICA) is used as C_{ref} in the model Equation 5.5, as it is associated with minimal errors from volume-averaging artifacts. The increased size of the AIF sampling region represents the increased degree of the PVE. The AIF concentration was measured from 3, 5, 7, 9, and 11 voxels centered around the reference ICA voxel (Figure 5.3b).

Perfusion Analysis

The reference AIF curve (C_{ref}) for the rescaling approach, was measured from an ICA voxel where the concentration curve had the features of large amplitude (peak), small width, and fast attenuation. The venous output function (VOF) is the concentration–time curve measured in a vein that drains the organ of interest. Based on practical and theoretical considerations, manual VOF is often chosen from the sagittal or transverse sinus [2, 85, 89]. To obtain a VOF with peak from the first-pass bolus passage followed by a recirculation peak, voxel with the maximum signal in the sagittal sinus is chosen as the reference VOF curve. To reproduce AIFs with an increasing degree of the PVE, we used concentration curves measured from ROIs of widths 3, 5, 7, 9, and 11 voxels, centered on the reference AIF voxel for 15 patients. A region of interest (ROI) tissue was manually selected in the grey matter [92] to evaluate the CBF percentage change (ΔCBF). ΔCBF is the percentage change of CBF estimated from rescaled AIFs relative to the CBF estimated from reference AIF curve. When using the VOF approach, the reference AIF concentration curve (width of 1 voxel) was also rescaled with the reference VOF curve.

Perfusion quantification for CBF was performed by deconvolution of the tracer kinetic equation (Equation 5.6) [5, 83, 8], implemented using MATLAB scripts (MathWorks, Natick, MA, USA).

$$C_t(t) = \mathrm{CBF}(C_a(t) \otimes R(t)) \tag{5.6}$$

$$\mathrm{CBF} = \frac{1 - H_{sv}}{1 - H_{lv}} \frac{1}{\rho} \max\, R(t) \tag{5.7}$$

where $C_t(t)$ denotes the tissue concentration curve of the ROI located in gray matter, $C_a(t)$ is the AIF either corrected using one of the three rescaling criteria described above or without rescaling, $\otimes$ symbol represents the convolution operator, and $R(t)$ represents the residue impulse response function. Deconvolution of Equation 5.6 to estimate CBF was carried out using the block circulant singular value decomposition method (cSVD) [5, 8, 93]. The block circulant decomposition method has an advantage of being less sensitive to tracer arrival timing differences. Deconvolution of Equation 5.6 for known values of $C_a(t)$,$C_t(t)$ leads to

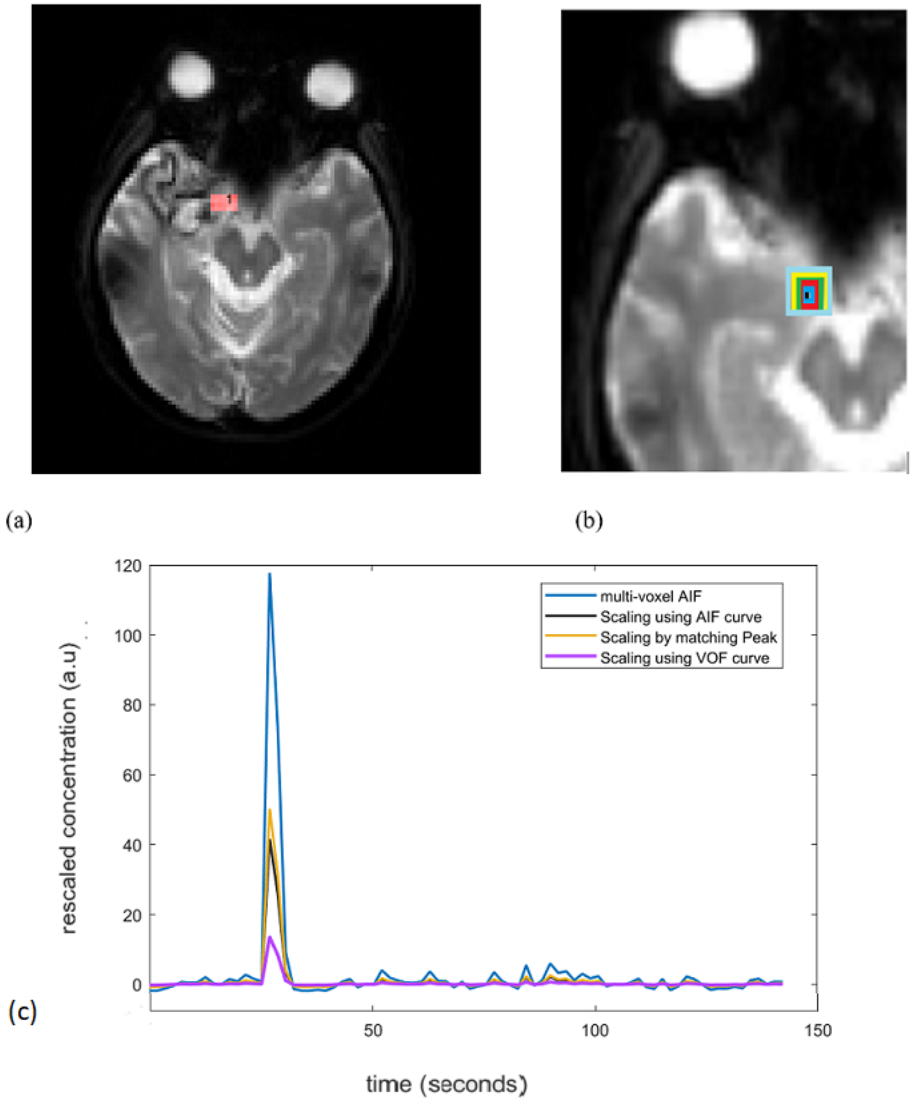


Figure 5.3: (a) Red colored square represents ICA used for reference AIF determination. (b) Increased size of the AIF, i.e., AIF concentrations were measured from 3, 5, 7, 9, and 11 voxels centered around the reference ICA voxel shown by black, blue, red, green, yellow, light blue colored squares, respectively. (c) An example of non-corrected AIFs (3 voxel AIF; blue curve) and corrected AIFs by all 3 scaling approaches. For a single subject, the unscaled AIF was derived from a 3-voxel-wide region to include the effect of the PVE. ICA: internal carotid artery; PVE: partial volume effect.

evaluation of the residue function $R(t)$. CBF is measured as the maximum of $R(t)$. Furthermore, H_{lv} and H_{sv} represent a correction for different levels of hematocrit in large vessels and small vessels. Here, the values used are $H_{sv} = 0.25$, $H_{lv} = 0.45$ and $\rho = 1.04$ g/mL (density of the brain) [92, 94]. After estimating CBF and Tmax for all brain tissues, CBF and Tmax are represented visually on axial brain maps.

5.1.2 Results

Figure 5.3 (c) is an example to see the effect of the AIF correction by all the three scaling approaches. For a single subject, the unscaled AIF was derived from a 3-voxel-wide region to include the effect of the PVE. After correction, the peak of unscaled multi voxel AIF (3 voxels wide) reduces for all the approaches and there is slight change in the tail (recirculation) part. The VOF approach rescales the AIF to large extent.

Figure 5.4 shows curves of rescaled AIFs with increasing degree of the PVE plotted as function of time using different scaling approaches. Overall, the deviation of the rescaled AIFs from the reference AIF increases with increasing PVE (i.e., increased number of voxels used for measuring AIF). Rescaling of AIFs by using scaling by AIF approach leads to small deviations at the tail, but large deviations at the peak (Figure 5.4 (A)). Rescaling of AIFs by matching peak reproduces peak similar to reference AIF for rescaled AIFs. However, this approach fails to accurately reproduce the tail similar to the tail of reference AIF (Figure 5.4 (B)).

Rescaling of measured AIFs using scaling by VOF approach gives rise to results similar to the first approach apart from the decrease in the overall peak height estimates of rescaled AIFs (Figure 5.4(C)). The tail accounts for the recirculation of tracer in the brain vasculature after an initial bolus passage, whereas the peak represents the maximum amplitude bolus rush through the brain vasculature [2]. The ideal AIF concentration curve has to represent correct tail and peak in order to reproduce more reasonable perfusion parameter maps.

Rescaled AIFs curves do not coincide with the exact reference AIF curve (Figure 5.4). Rescaled AIFs match the exact reference AIF curve either at peak height or at tail. The least percentage change of CBF values estimated using rescaled AIFs from CBF values estimated using reference AIF may decide the most appropriate scaling approach. The least percentage change of CBF indicates the approach that will be least affected by the volume-averaging artifacts.

Deviation of CBF, based on all 15 patients, which is represented as Δ CBF, was estimated as percentage difference in CBF estimated using rescaled AIFs and reference AIF. Rescaling of the AIFs was conducted using the following four approaches: no scaling, scaling by AIF, scaling by VOF, and scaling by peak height. The increasing degree of the PVE and its

association with ΔCBF in 15 patients is shown in Figure 5.5. As shown in the Figure 5.5, ΔCBF values estimated from non-rescaled AIFs without any modification increases as the number of voxels used for measuring AIF increases. This shows that the increased PVE resulting from the increased number of voxels might seriously affect the estimation of AIF and consequently the calculation of perfusion parameters. Rescaling of the AIF using either of the three approach results in the reduced ΔCBF values. Overall, scaling by AIF and scaling by VOF seemed to achieve the best and similar results as they yield the least ΔCBF values when the PVE increases maximally among all scaling approaches.

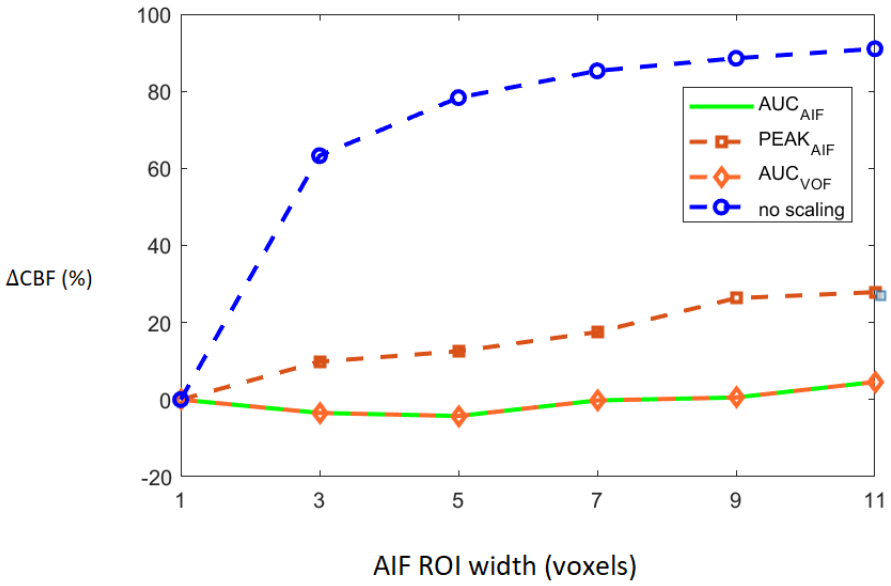


Figure 5.5: Average CBF divergence from reference CBF plotted against increasing number of voxels. Average CBF divergence for group of 15 patients is plotted according to increased partial volume effect (PVE) for all four scaling approaches indicated by the legend on right.

The CBF brain map was generated in the absence of scaling of AIF as well as by using the VOF rescaling approaches (Figure 5.6). In some cases, relative to the CBF map generated by using non-rescaled AIF, the CBF map generated by rescaled AIF approaches showed increased CBF values on the slices in the left and right hemispheres (red color) (visible in Figure 5.6). From Equations (5.5) and (5.6), it follows that the ratio of scaled and unscaled CBF values should be the scaling factor k. The images of the ratio of scaled and unscaled maps are expected to show the factor k for every voxel (Figure 5.6 (c)). The mean CBF values, based on all 15 subjects using non-rescaled (AIF ROI width = 5 voxels) and VOF approach, were 43.98 and 61.16 mL/100g/min, respectively. The mean CBF values for AIF-rescaled and peak scaled approach were 57.10 and 47.10 mL/100g/min, respectively. At individual level, all the fifteen patients in this study did not follow similar pattern of

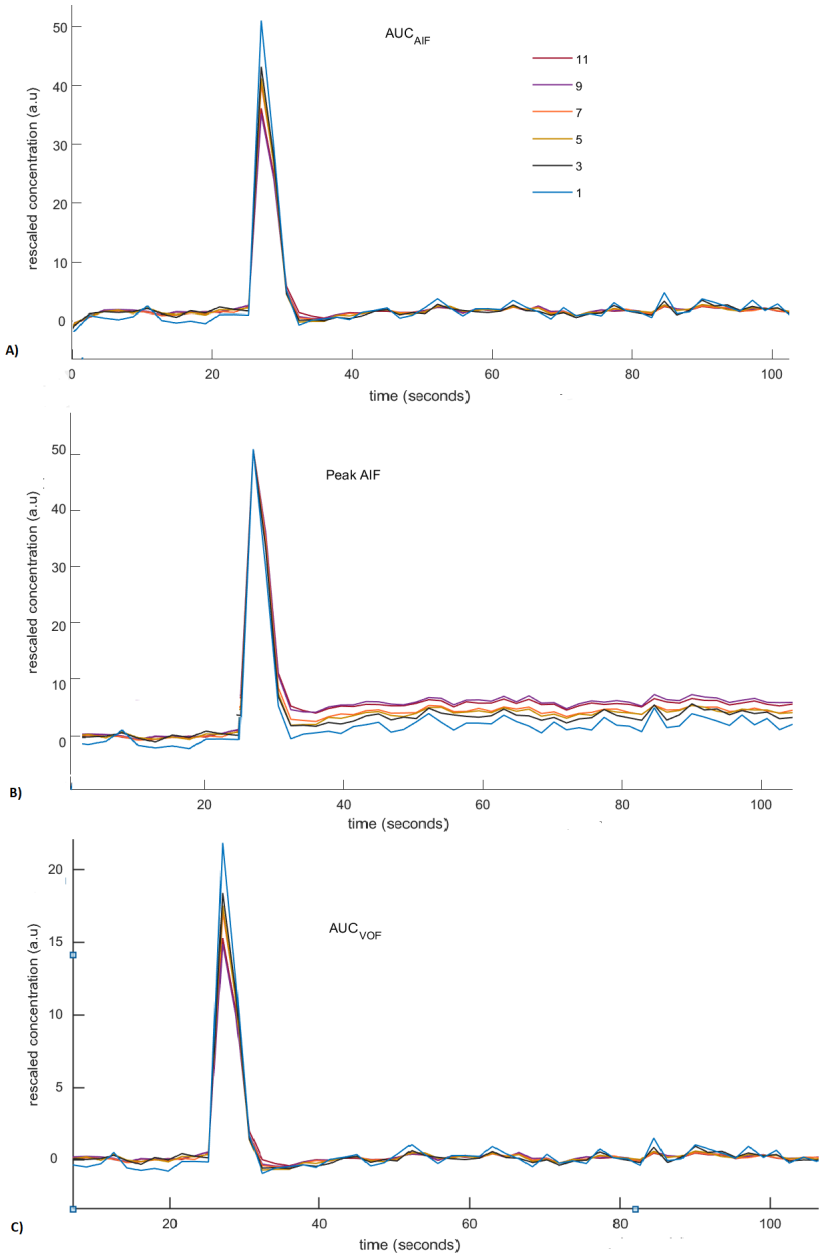


Figure 5.4: Rescaled concentration curves of Arterial Input Functions (AIFs) generated using different scaling methods. The legend in (A) indicates the width (in voxels) of ROI used for measuring the AIF. (A) Rescaled AIFs generated using scaling by AIF approach. (B) Rescaled AIFs generated using scaling by matching peak height approach. (C) Rescaled AIFs generated using scaling by VOF approach. VOF : Venous Output Function.

underestimated CBF values due to the PVE. To generalize, we need a larger dataset to conclude whether the scaling corrects the underestimation of CBF, as this could vary patient to patient in a small cohort.

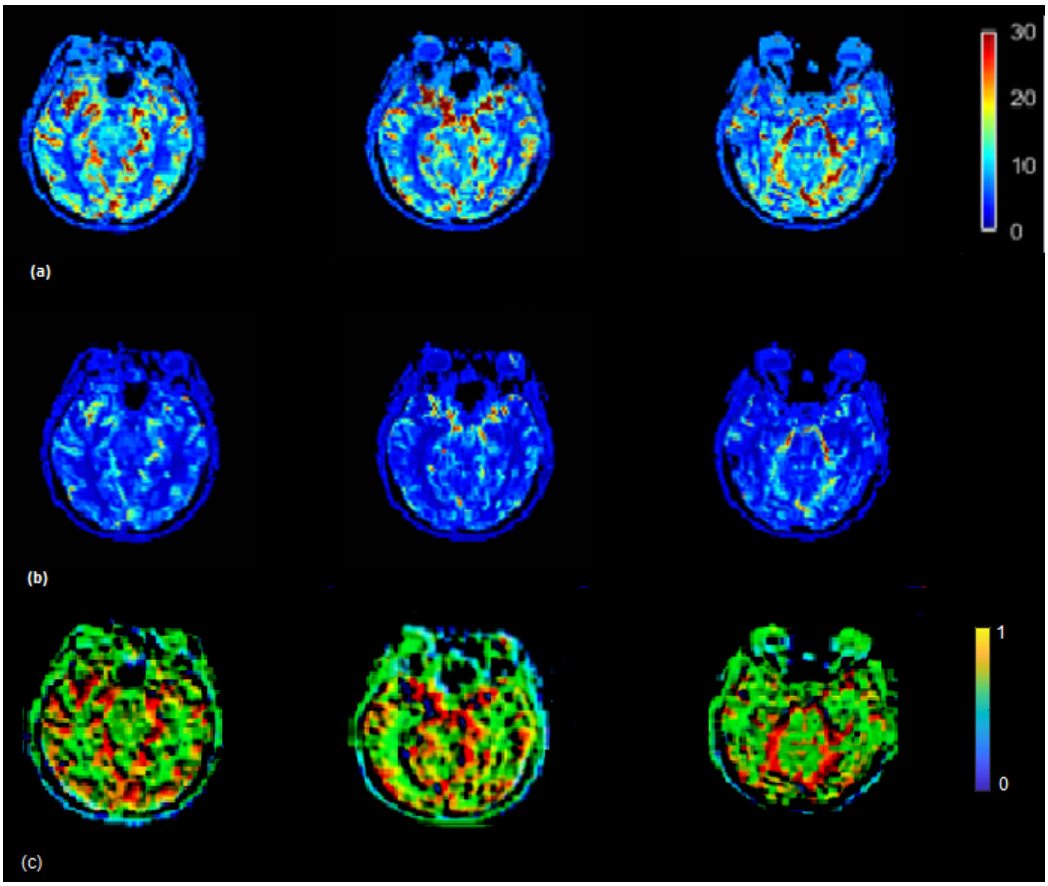


Figure 5.6: (a) CBF (mL/100 g/min) map generated by using rescaled AIF (b) and non-rescaled AIF. CBF map generated using rescaled AIF represents increased CBF values in the shown axial brain slices. CBF maps from non-rescaled AIF display mostly all the ROIs with decreased blood flow which makes it difficult to locate the regions which actually have a decreased flow. CBF images derived using rescaled AIF display ROIs with increased flow (red color) which helps to segregate the regions with decreased blood flow. This may help clinicals to identify the infract regions as well as regions with decreased blood flow on visual brain CBF images. (c) Maps illustrating the ratio between CBF values derived from the re-scaled and the non-scaled AIF.

The Tmax (seconds) map was generated using the rescaled AIF (VOF approach) (Figure 5.7) (top) and non-rescaled AIF (Figure 5.7) (bottom). The Tmax map generated using the rescaled AIF showed increased values in the axial brain slices in the left and right hemispheres. The mean Tmax values (range of 0–12 s) based on 15 subjects using non rescaled

(AIF ROI width = 5 voxels) and VOF approach were 4 s and 7 s, respectively. The derivation of Tmax is performed from the residue function ($R(t)$), which is achieved by deconvolution of Equation 5.6. Tmax is the argument, i.e., t of the maximum of $R(t)$. The deconvolution utilizes a matrix method called the circulant singular deconvolution, which is sensitive to the peak amplitudes of AIF (cSVD) [93]. The changed AIF amplitude used in the cSVD algorithm shifts the maximum of $R(t)$ to a higher time points t which accounts for higher Tmax values. The change in Tmax is consistent with a previous study where different AIFs with changed amplitudes and similar shape selected by different algorithms resulted in change of Tmax values [4]. The increased Tmax maps generated by the rescaled AIF may allow clinicals to visualize the critically hypoperfused regions which are likely to be salvageable.

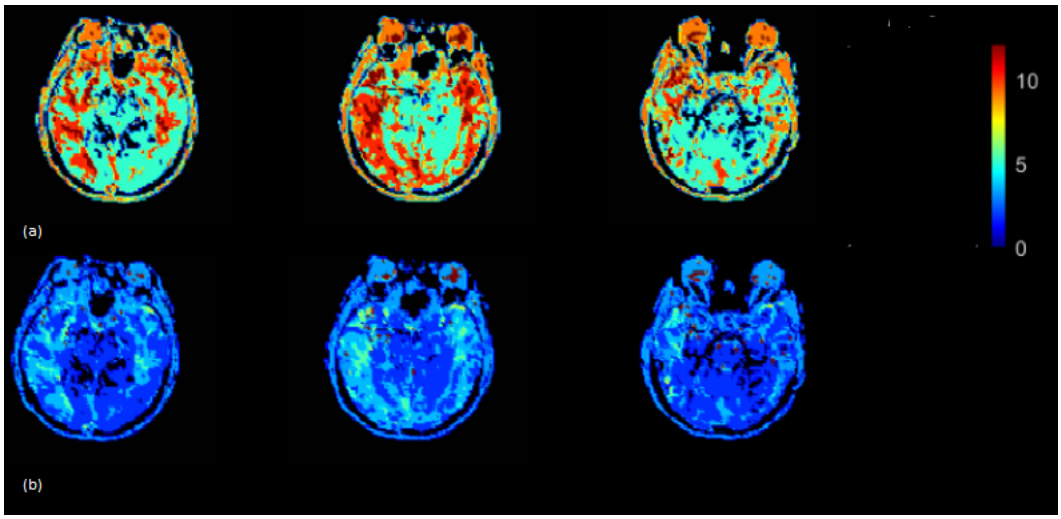


Figure 5.7: Tmax (seconds) map generated by using (a) rescaled AIF (b)non-rescaled AIF for one subject. Tmax map generated using rescaled AIF represents increased values in the shown axial brain slices.

5.1.3 Discussion

In the present chapter, we used multiple AIF rescaling approaches using perfusion imaging data so as to correct the amplitude falsification of the multi-voxel AIF. This, thorough investigation, has allowed us to study the effect of the PVE on a multi-voxel AIF, which is a prerequisite for obtaining accurate CBF measurements using MR bolus tracking [84]. The significant findings of the study revealed that rescaling AIF using scaling by VOF approach leads to more reasonable absolute perfusion parameter values, represented by the increased mean CBF/Tmax values and CBF/Tmax images. This may assure that the core or brain regions with decreased blood flow will not be overlooked. The present study has shown that the absence of multi-voxel AIF scaling results in inaccurate and untrue CBF values.

The spatial resolution typically used in the cerebral DSC-PWI, $1.9 \times 1.9 \times 5mm^3$ makes it difficult to identify vessels and only voxels placed in the ICA could be selected as reference AIF, free from partial volume of the tissue [95]. Selecting more voxels or a large region for AIF estimation can lead to significant PVE. However, PVE-corrected, multi-voxel AIF is necessary as AIF obtained from a single voxel or a small region is not reliable enough due to noise in spatial measurements and motion in temporal measurements [4]. In this study, we intend to make use of scaling as a way out to calibrate a multiple voxel AIF which would further lead to reduce the effect of the PVE on the quantification of absolute CBF and Tmax values. The great benefit associated with this type of linear scaling is that predefined thresholds could be used for evaluation/comparison of perfusion images and parameter values obtained from different scanners examined at different time points.

For the typical spatial resolution used in DSC-MRI studies, average size of major arteries and considering that the true AIF as signal is saturated at peak concentrations for a voxel with 100% blood, a degree of partial volume is in practice unavoidable when measuring the AIF [2]. The signals from the arterial and tissue contributions are complex numbers (with amplitude and phase), which makes selection of PVE-free AIF more complicated [2]. The scaling method used in this study might be a pragmatic way of using a multi-voxel AIF. Considering the difficulty involved in the selection of PVE-free reference AIF, we can consider that reference AIF in the study is an approximation of the true AIF with minimal PVE.

Clinically, AIF selection depends on the expertise, experience, and skill of experts accompanied by time consumption, low reproducibility, and often including tissue signals in AIF. In the past, perfusion studies utilized AIF selection approaches such as slice-specific AIF selection, clustering methods that require ROI to be marked manually prior to AIF extraction [11, 96, 97, 98], and multi stream 3D CNN method [4, 99]. No matter what selection strategy is taken, the PVE is always present as the voxels selected may exhibit partial signals. Therefore, a proper approach must be taken to solve the problem. Rescaling of AIF discussed in the present study can be carried out for PVE correction even if the AIF selection procedure is slice specific.

The extent of the influence of PVE on the output perfusion parameters has been observed in previous research. In vivo studies have reported large variations in perfusion parameters due to the PVE [84, 85, 95, 100]. Past simulation results in DSC MRI proved that uncorrected AIF measured with a partial volume fraction of about 50% could produce a four times CBF overestimation along with distortions of AIF frequency characteristics [84]. Investigation of the impacts of the PVE on quantitative perfusion metrics at 1.5 T and 3.0 T has reported broaden tissue contribution, resulting in fluctuations in the AIF which further compromises quantitative perfusion estimates in a nonlinear fashion [95].

Inaccurate AIF estimation can be minimized by correction for partial volume effects by utilizing specific post processing approaches or data acquisition techniques. Past MRI

research has reported that correction for the PVE was appropriate for arteries that were parallel to the main magnetic field by estimation and subtraction of the static signal of the surrounding tissue [100]. However, to measure quantitative input function of vessels that were not parallel to the main magnetic field was still a challenge [100]. CTP studies suggested that AIF measurements should be done with smaller section thickness, i.e., small location or reduced voxel volume, as AIF and VOF measurements from thicker sections would cause an overestimation of CBV and CBF [101, 102]. In vivo MRI studies have suggested that minimal impact of the PVE in AIF measurements may be achieved with reduced contrast dosage and minor adjustments to the pulse sequence [84]. Previous research used linear scaling methods for the correction of PVE-affected AIF on T1-weighted perfusion imaging datasets as they allow easy identification of the large arteries [85]. Rescaling using VOF as a reference was reported as most feasible approach [85]. The present study used T2-weighted perfusion imaging datasets and used rescaling on AIF concentration curves rather than the MR signal curves and proved to be consistent with past results of PVE correction of T1-weighted datasets [85].

Recent studies in DCE-MRI demonstrate that inter-frame realignment have a huge effect on parameter mapping [103]. In this study, reference AIF (red colored square represented on ICA (Figure 5.3a) and ROI used as multivoxel AIF (i.e., AIF concentrations measured from 3, 5, 7, 9, and 11 voxels centered around the reference ICA voxel shown by black, blue, red, green, yellow, light blue squares, respectively, in Figure 5.3b) were measured on a same brain MRI axial slice for all subjects. For the geometrical alignment of axial brain slice selected for AIF determination, co-registration to a common template for all subjects after the acquisition was considered. However, in case if multiple axial slices are utilized for AIF selection across the subjects, then inter-frame realignment has to be considered due to its impact on parameters. The limitation of the present study is that multiplicative rescaling can only be used if tissue contribution is much smaller than the arterial contribution. This condition restricts AIF correction in case when there is ample amount of surrounding signal contributions to the AIF. The use of scaling factor assumes that any PVE in the AIF can be represented by linear scaling; however, in the case of large signal contribution from surrounding tissue, the PVE can be very complex and can distort the AIF shape.

5.1.4 Conclusions

In summary, the present study demonstrates that utilizing scaling approach provides more reasonable absolute perfusion parameter values, represented by the increased mean CBF/Tmax values and CBF/Tmax images. This will assure that the core, as well as the infract region, will not be overlooked. Distortions due to the PVE might be still present in AIF after the scaling as it does not affect the shape of the curve to a large extent. Rescaling AIF by VOF approach seems to be the robust and adaptable approach for correction of the PVE-affected multivoxel AIF. Absence of multi-voxel AIF scaling during deconvolution of the tracer kinetic equation may lead to inaccurate CBF values.

5.2 Core and Penumbra estimation using Deep learning based AIF in association with clinical measures in Computed Tomography Perfusion (CTP)

5.2.1 Introduction

Deconvolution of two time-dependent functions, Concentration time function $C_t(t)$ and Arterial Input function (AIF) $C_a(t)$, is required for CTP parameter estimation [83]. AIF curves, in general, have a baseline period, a first passage, and a recirculation part. AIF with a distorted baseline, shape-amplitude errors in the first passage, and non-identical recirculation portions predicts incorrect perfusion parameters [2]. These inherent errors may lead to the identification of incorrect core and penumbra volumes [2].

The estimation of the CTP parameters can be improved by curve fitting (CF) of hemodynamic models to the AIF, according to recent studies [104]. Arterial concentration time curves plotted against time points signify that contrast agent is injected into a blood vessel upstream and dilutes downstream [2]. Curve fitting can assure an AIF with a constant baseline, a first passage with a peak amplitude, and identical recirculation portions [90, 104]. To achieve this, previous studies modelled arterial concentration time curves by using the gamma variate function, which has the value [104] :

$$C(t)_{GVM} = K(t - AT)^{\alpha} exp(-\frac{(t - AT)}{\beta}) \tag{5.8}$$

Here t represents the time points, AT is the bolus arrival time, K the constant scale factor, and to describe the shape of the curve and are used as arbitrary parameters.

Apart from the thresholded perfusion parameters, medical professionals also refer to the NIHSS and ASPECT scores to make thrombectomy decisions [105, 106]. Clinical trials require a severity assessment, NIHSS is considered as the gold standard for stroke severity rating as it has been shown to be a predictor of both short and long term outcome of stroke patients [107]. In the NIHSS scale (0-42), score of 1-4 represents minor stroke, 5-15 indicates moderate stroke, 16- 20 characterizes moderate to severe stroke and a score of 21-42 indicates that the patient has severe stroke [108].

ASPECTS measures early ischemic changes in anterior circulation hyperacute ischemic stroke [7]. For ASPECTS, a normal brain has a score of 10 and the score falls as more brain regions are affected [109]. Patients with ASPECT score 0–5 benefit from mechanical thrombectomy without increasing the risk of symptomatic intracerebral hemorrhage

[13]. The American Stroke Association recently updated their stroke management guidelines, and one of the key selection criteria now includes ASPECTS. Patients with baseline ASPECTS$\geq$ 6 are advised to receive endovascular therapy [110].

The current study investigates whether curve fitted AIF estimation based on CNN, which combines features of consistent baseline, peak amplitude, and identical recirculation portions, improves volumetric assessment of core and penumbra. It is hypothesized that core and penumbra assessment with CNN AIF could be validated with clinical scores such as NIHSS and ASPECTS. We expect CNN-based AIF to identify core and penumbra in stroke patients where traditional AIF methods fail. This may aid clinicians in making thrombectomy and reperfusion therapy decisions.

5.2.2 Methods

Patient Population

The current study used CTP datasets of 160 stroke patients with large vessel occlusions (male = 87, female = 73, median age = 73 years). These datasets were obtained from the Veterans General Hospital in Taipei and its branch in Hsinchu, Taiwan. The Institutional Review Boards for human studies gave this study their ethical approval (IRB-TPEVGH 2021-06-016 BC, IRB-2020-02-006B).

Imaging Protocol

A dose of 70 ml of contrast agent (iodine) was injected at a rate of 4 ml per second. Images were acquired on a clinical CT scanner (Phillips: Ingenuity CT) using a sequenced acquisition (KVP/X-Ray Tube Current/slice thickness/slices 80 kv/190 mA/5/16) with a 24-hour onset-to-imaging time. Three experienced neuroradiologists scored the NIHSS and ASPECTS, with the median NIHSS being 10(4-19) and the median ASPECTS being 8 (6-10).

Core and Penumbra estimation from CNN based AIF

CTP datasets were randomly divided into training (128 datasets) and validation datasets (22 datasets). A widely used clustering algorithm estimated three AIFs for each dataset in the training and validation sets to generate training and validation AIF curves. To estimate the AIF, this clustering algorithm employs recursive cluster analysis in the Middle Cerebral Artery (MCA) region. Training dataset density was improved for model training performance using feature augmentation and spline interpolation. Mirroring, rotation, and both mirroring and rotation were used for data augmentation. Initially, AIFs obtained from the clustering algorithm on 128 training datasets contain intrinsic errors associated with shallow

peak, distorted baseline, and non-identical recirculation phase (Figure 5.8).

Training AIF curves are curve fitted (CF) to Gamma Variate hemodynamic function (C_{GVM}) (Equation 5.8) by well-known past algorithms in MATLAB and used as labels to adjust these shape-based errors. Following augmentation, 1152 distorted arterial curves estimated by the clustering method were used as input sources, and 1152 curve fitted arterial curves were used as labelled data for the CNN's supervised training.

The CNN model architecture was built with Python's Keras library. The CNN architecture consists of a single input layer that accepts input in the form of an interpolated AIF curve (500 points), two Convolutional layers (kernel length =2), a pooling layer, a flattening layer, and two dense layers that are fully connected to the output layer. To connect the convolutional layer with the average pooling layers, the ReLu activation function was used. To connect the 36-neuron dense layer to the output nodes, the softmax function was used. With a batch size of 32 and 17,238 total trainable parameters, the network was trained for 300 epochs. The optimizer was root means square propagation with an initial learning rate of 0.001. The CNN model training was done on a workstation with hardware: Intel I5/Ram :16 GB / GPU: 1070 (8 GB).

Validation was performed on 22 datasets following model training. For the testing phase, distorted AIFs were chosen from a location other than the AIF location of the training dataset. The CNN model takes as input a distorted AIF curve and outputs a probability map of gamma variate fitted AIF curve (Figure 5.8). This Gamma variate fitted AIF curve obtained from trained CNN is referred as CNN AIF. Deconvolution of Concentration time curves (c_t) with the CNN AIF (c_a) estimates the residue function R(t) (Equation 4.12), which further estimates core by thresholding CBF$<$ 30% and penumbra volumes by thresholding Tmax $>$ 6s. The workflow to estimate perfusion parameters is described in Figure 5.8. The core and penumbra volumes estimated by using CNN AIF are compared with the volumes obtained from the AIF selected by the clustering algorithm.

Statistical Analysis

Spearman's correlation coefficient was used to assess the agreement of the NIHSS and ASPECTS with core and penumbra volume. Mean volumetric differences were demonstrated using Bland-Altman plots. Wilcoxon signed rank test was used at the group level to compare perfusion parameters. MedCalc software (https://www.medcalc.org/) was used for all statistical analysis.

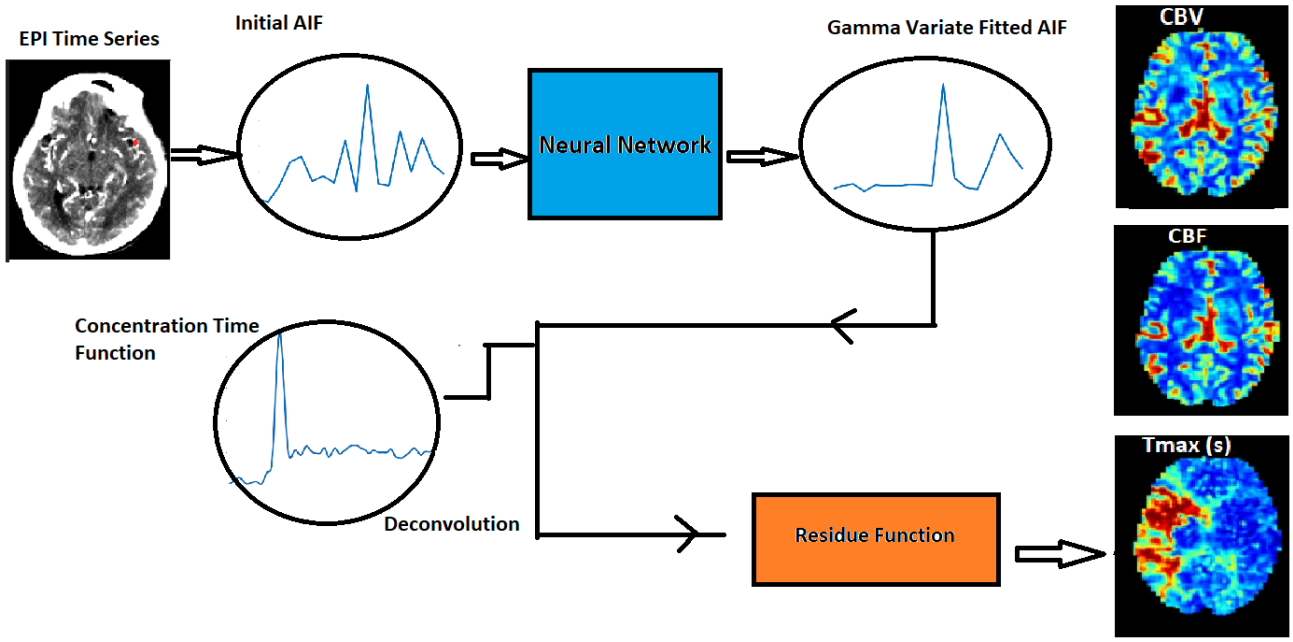


Figure 5.8: Workflow to estimate perfusion parameters.

5.2.3 Results

Figure 5.2 depicts the curve fitting performed by the CNN model. AIF location is marked as a red dot on an axial CTP slice for a single dataset (Figure 5.2 (A)). The raw AIF is the concentration curve of this arterial location plotted against time (Figure 5.2 (B)). The CNN model accepts the curve as input after interpolation. As an output, the CNN model predicts gamma curve fitted AIF (Figure 5.2 (C)).

Volumetric correlation of penumbra and infarct core with NIHSS and ASPECTS

Penumbra volumes calculated with the CNN AIF were positively related to the NIHSS score (r=0.69; P 0.001) and negatively related to the ASPECTS (r=-0.43; P 0.001) (Figure 5.9(A), Figure 5.9(B)). Table 5.1 shows that when penumbra volume is estimated using the CNN AIF, it has a stronger positive correlation with the NIHSS score. When Tmax>4s, 8s, and 10s volumes are calculated using the CNN AIF, they have a higher positive NIHSS correlation (Table 5.1). According to the Bland-Altman plots, the mean volumetric difference between the tissue at risk estimated by the CNN AIF and without CNN AIF was 12.1 mL. (Limits of agreement, -311.5 to 335.8 mL; Figure 5.9C).

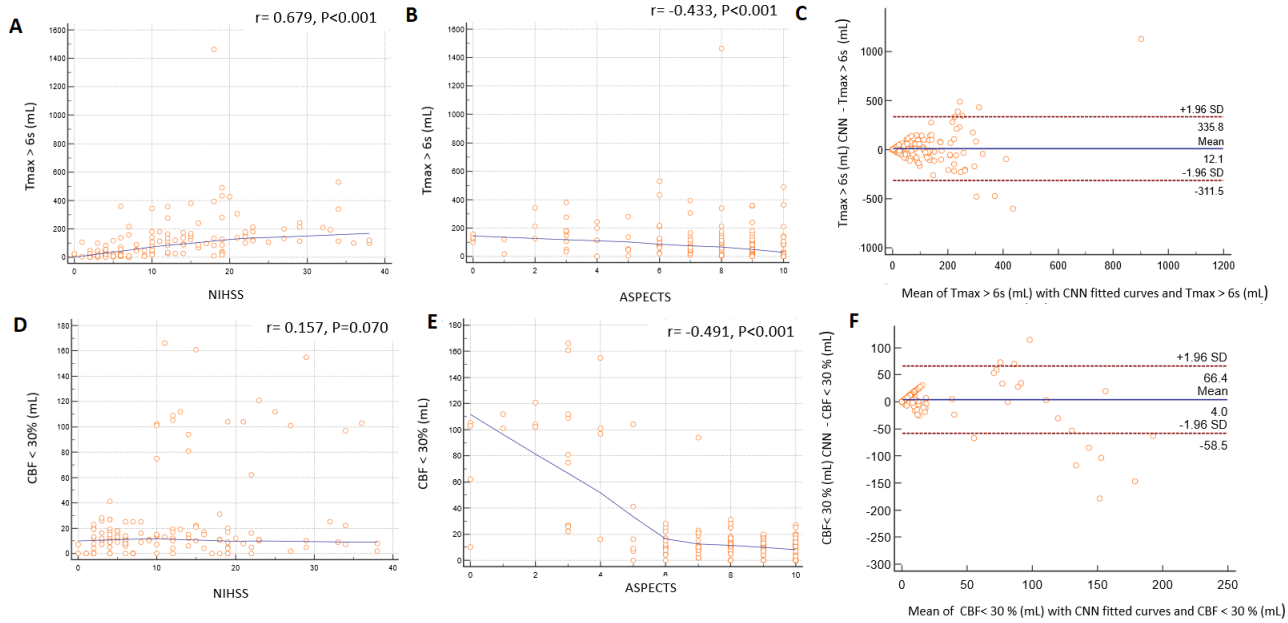


Figure 5.9: Volumetric agreement of the penumbra (Tmax >6s) and infract volume (CBF < 30%) with NIHSS and ASPECTS. (A) Association of Tmax > 6s volume with NIHSS (B) Association of Tmax > 6s volume with ASPECTS. (C)Bland -Altman plot for penumbra estimated by CNN AIF and without CNN AIF. (D)Association of CBF < 30% volume with NIHSS (E) Association of CBF < 30% volume with ASPECTS. (F) Bland -Altman plot for the Infract core estimated by CNN AIF and without CNN AIF. NIHSS : National Institutes of Health Stroke Scale, ASPECTS : Alberta Stroke Program Early CT Score.

Tmax volumes	CNN AIF		Without CNN AIF	
	r	P value	r	P value
Tmax >10s	0.70	<0.001	0.50	<0.001
Tmax >8s	0.71	<0.001	0.54	<0.001
Tmax >6s	0.68	<0.001	0.54	<0.001
Tmax >4s	0.58	<0.001	0.46	<0.001

Table 5.1: Association between CT perfusion parameters and NIHSS scores (Spearman's correlation).

Infarct core estimated by CNN AIF correlated negatively with the ASPECTS (r=-0.49; P <0.001) (Figure 5.9). Table 5.2 shows that Tmax>6s volumes estimated from CNN AIF have higher negative correlation with the ASPECT score. Infarct core prediction using CNN AIF or without it both demonstrate negative and similar correlation to ASPECTS (Table 5.2). Mean volumetric difference for the Infarct region estimated from CNN AIF without

CNN AIF was 4.0 mL (Limits of agreement, -58.5 to 66.4 Ml; Figure 5.9F).

Tmax volumes	CNN AIF		Without CNN AIF	
	r	P value	r	P value
Tmax >8s	-0.40	<0.001	-0.45	<0.001
Tmax >6s	-0.43	<0.001	-0.39	<0.001
Tmax >4s	-0.35	<0.001	-0.25	0.003
CBF< 20 %	-0.46	<0.001	-0.47	<0.001
CBF < 30 %	-0.49	<0.001	-0.49	<0.001

Table 5.2: Association between CT perfusion parameters and ASPECTS scores (Spearman's correlation).

Penumbra and core regions estimation

In some cases, penumbra and core regions estimated without CNN AIF were not rational for patient symptoms suffering from severe stroke. These patients were typically scored with higher NIHSS (> 20) and lower ASPECTS (<5). We demonstrate this with examples of two patients in Figure 5.10 and Figure 5.11. Patient 1 had a NIHSS score of 23 (severe stroke), ASPECTS of 0 and acute occlusion of the left main coronary artery (LMCA occlusion). Without CNN AIF, penumbra volume was 99 mL along with absence of infarct core on left side representing flawed estimates. With CNN AIF, penumbra volume was 123 mL and volume of core region was 10 mL (Figure 5.10).

Patient 2 had a NIHSS score of 23 (severe stroke) and right internal carotid artery (RICA) occlusion. This patient reported for ischemic region at bilateral brain without CF, which was not reasonable for his symptoms as ischemia should be on the right side only. Volume of tissue at risk estimated without CNN AIF was 736 mL along with absence of infarct core (Figure 5.11). With CNN AIF, volume of tissue at risk was 137 mL and volume of infarct core was 6 mL along with presence of ischemic and core region mostly on right side (Figure 5.11). With CNN AIF mean core volume for these patients was 5 mL with median of 2 mL. The ischemic region location could be visually more precisely located through the perfusion maps derived with CF CNN model.

Group comparison of penumbra and core volume for stroke patients with CNN AIF and without CNN AIF are demonstrated in Table 5.3. Wilcoxon signed rank test indicated that the median rCBF< 20%, rCBF< 30%, rCBF< 38%, Tmax >10s, volumes estimated with CNN AIF were statistically significantly higher. The median infarct core (rCBF< 30%) with CNN AIF is 12 mL whereas without CNN AIF the median core volume is 0 (Table 5.3). As these people have stroke, so the median of infarct core volume as zero might not reflect the hypoperfusion in patients.

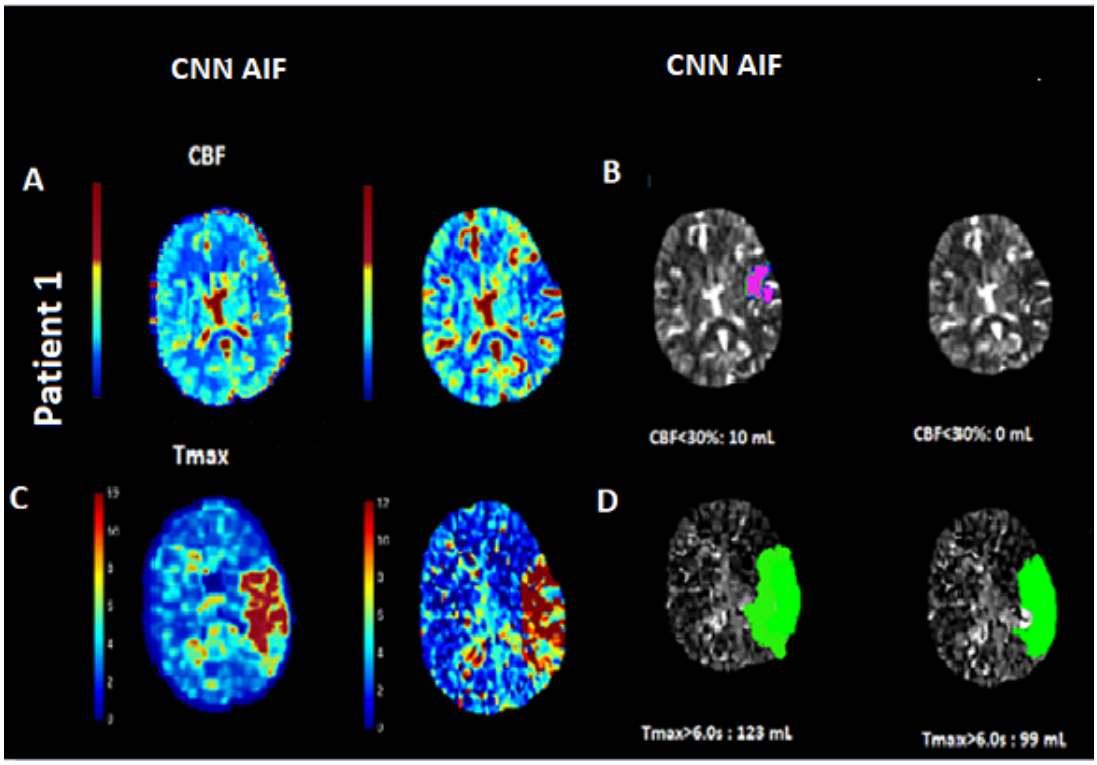


Figure 5.10: Comparison of infarct core and penumbra for stroke patients. A) CBF maps derived with CNN AIF and without CNN AIF. B) Infarct Core estimated with CNN AIF and without CNN AIF. (C) Tmax maps derived with CNN AIF and without CNN AIF. D) Penumbra volumes estimated with CNN AIF and without CNN AIF.

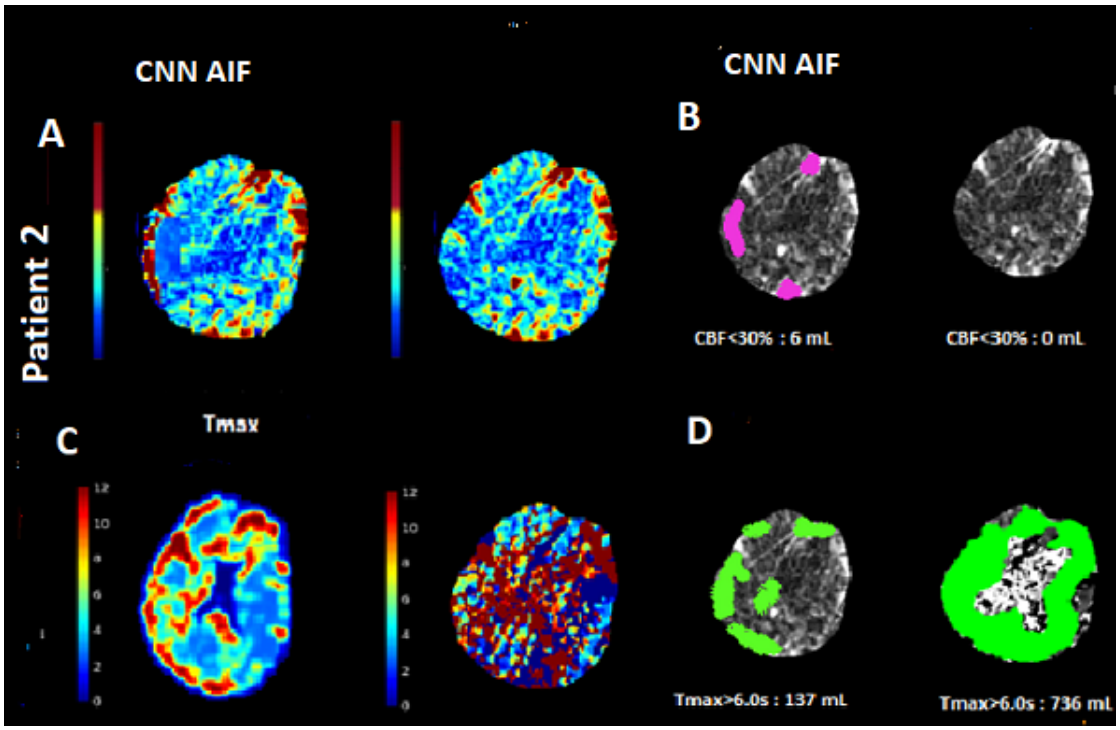


Figure 5.11: Comparison of infarct core and penumbra for stroke patients. A) CBF maps derived with CNN AIF and without CNN AIF. B) Infarct Core estimated with CNN AIF and without CNN AIF. (C) Tmax maps derived with CNN AIF and without CNN AIF. D) Penumbra volumes estimated with CNN AIF and without CNN AIF.

	Mean (CNN AIF)	Mean	Median (CNN AIF)	Median	P value
CBF $< 20\%$	13	10	3	0	< 0.001
CBF $< 30\%$	23	19	12	0	< 0.001
CBF $< 38\%$	32	25	18	0	< 0.001
Tmax >10s	70	50	27	6	0.002
Tmax >8s	86	67	47	27	0.012
Tmax >6s	113	101	64	55	0.386
Tmax >4s	243	213	140	163	0.416

Table 5.3: Mean and Median volume comparison of CT Perfusion parameters (Wilcoxon signed rank test). CBF $< 20\%$ indicates blood flow reduction of 20% as compared to the normal side.

The deconvolution-based model without CF used in recent stroke studies to estimate ischemic regions is linear, despite the fact that estimation of core and penumbra depends on a variety of factors [5, 2, 87]. Certain risk factors for ischemia include distorted CTP signals, collateral status, and gray/white matter content [109]. Using perfusion thresholds without correcting for distorted AIF signals may result in negligible core estimates. Fitting distorted AIF to hemodynamic models greatly reduces distortion in the recirculation and baseline parts of the AIF curve [90, 104]. The advantage of the CNN-based AIF algorithm is that it solves the AIF distortion problem and allows for the selection of the best AIF curve corresponding to distorted data points while ignoring noise or errors.

Conclusion

CNN-based AIF improves the estimation of penumbra and infarct core volumes. Better correlation of penumbra and infarct volume with NIHSS and ASPECT scores were obtained using the CNN-AIF. This serves as a motivation as well as evidence to include CF prior to Tmax/CBF estimations. CF AIF could identify patients with core regions likely to be ignored by conventional approaches. The inclusion of CF AIF can provide physicians with reasonably accurate and precise perfusion parameter brain images that may aid them to determine suitable triage, transport, and treatment decisions for stroke patients.

Chapter 6

Selective Geodesic Variational Segmentation Model with new Region force

In this chapter, we focus on developing a segmentation model that can do selective segmentation on challenging medical images. The used medical images include T2 MRI brain tumour images (Figure 6.1), where the focus is to segregate/segment the tumour region. Along with the tumour images, we also test the purposed model on other medical images (Figure 6.1 (D), (F)) to do selective organ segmentation.

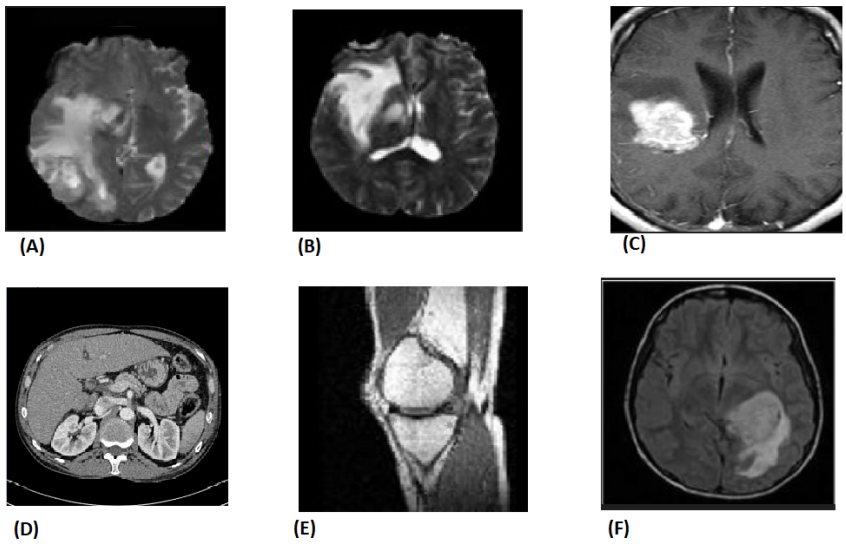


Figure 6.1: Medical images for selective segmentation. (A), (B), (C), (F) Brain Tumour T2 MRI images with the bright part being the Tumour region. (D) Abdominal CT image (E) Joint image.

We propose a effective framework for selective segmentation by adopting a region force term along with a geodesic distance penalty based on a discrete TV formulation. The proposed segmentation model is robust and competitive with the current state-of-the-art, according to empirical evaluations. We compared the performance of our model to other testing methods by evaluating segmentation scores of multimodal (MRI, CT) segmented images. In comparison to previous segmentation models, the proposed model is independent of user input and allows for segmentation in medical images with high precision scores. This chapter is based on the author's submitted paper.

6.1 Introduction

Selective segmentation models extract object of interest from image z after some points inside or near the target area are available as marker points [20, 21, 45]. The marker set is represented as $M = \{(x_i, y_i) \in \Omega, 1 \leq i \leq k\}$, Ω is the image domain. With an initialisation obtained from marker set, segmentation solution ϕ involves obtaining minimiser of the functional of the following form -

$$F = F_{Regulariser} + F_{Fitting} + F_{Distance} \tag{6.1}$$

$F_{Regulariser}$ is a regulariser such as Total Variation (TV) or weighted-TV [28, 19]. The regulariser considers the geometric properties to ensures that solution has a smooth boundary. The regulariser widely used is the sum of weighted length of boundaries assembled with a edge detector [20, 15, 44, 41, 23, 111]. The widely used formulation is - $\int_\Omega \alpha(x,y) d\Omega = \int_\Omega \alpha(x,y)|\nabla\phi| d\Omega$. Here, α is used as edge detector. The popular choice for edge detector is $\alpha(x,y) = \frac{\beta}{1+|\nabla z(x,y)|^2}$, with α, β being some properly chosen positive numbers [41]. Additionally, Gaussian smoothing of image z before computation of edge detector has also been considered [41, 23].

Except in the vicinity of edges, the fitting term $F_{Fitting}$ ensures that the object to be segmented has homogeneous intensity inside it [38, 41]. The fitting term is defined in most variational segmentation approaches as the euclidean distance between the data point intensity and the average data point intensity of the marker set [41, 38, 112]. This quadratic region force penalises the object of interest's heterogeneity or lack of smoothness in the foreground and background. This works well when the data points are homogeneous and spread out over a large smooth region. When image data points have complex geometry and it is difficult to separate background and foreground, penalising heterogeneity becomes difficult [41].

Recent studies use geodesic distance penalty or Euclidean distance penalty as $F_{Distance}$ [15]. The distance term penalises the distance between the markers in the set [15, 113, 114, 115]. This is done to encourage the segmented result to be close to some of the marker set's

points [116]. A recent study confirmed that the edge weighted geodesic distance increases at edges and is intuitive for images with blurry edges [15]. Prior Total Variation (TV) denoising is required for the calculation of geodesic distance. Recent image denoising research has shown that TV denoising using a more explicit discrete formulation is more robust and produces sharp edges [19] . This motivates the use of the geodesic distance penalty involving discrete TV formulation as $F_{Distance}$ in the intended model.

The proposed model is a different approach to the previous convex relaxed models [15, 41], including a new region force as a fitting term and geodesic distance with discrete TV formulation and as a standalone distance penalty. We use a primal-dual hybrid gradient method to solve the functional model [33]. This is accomplished by solving a dual problem derived from the functional of the primal model. It should be noted that the region force was previously used in the reformulated Potts model for multi-phase segmentation and data clustering without a geodesic distance penalty [41]. The prior denoising required to compute the geodesic term has also been evaluated using the discrete TV formulation 4.3 [19]. To the best of our knowledge, no other segmentation study has attempted to combine both region force and geodesic term in a single segmentation framework.

The contributions of this chapter can be summarised as follows:

- We incorporate (a) new region force term (b) geodesic distance penalty in selective segmentation framework.
- We improve the geodesic penalty term, with a new TV formulation for noise robustness.
- We use a primal-dual hybrid gradient numerical algorithm to achieve solution from the dual formulation
- We propose a convex selective segmentation model and compare it with other segmentation models.

The chapter is structured as follows; in section 2, we review some related studies. In section 3, we discuss the formulation of the purposed model and also address computation of the region force term and the geodesic distance term. Here, we also discussed the formulation of a dual problem from the primary problem. Next, we present the numerical algorithm used to achieve the segmented solution. Section 4 contains numerical results. In this section we show that along with segmentation superiority, the proposed model is robust to noise and user input. Finally, we conclude this chapter in section 5.

6.2 Review of Selective Segmentation Model

It would be beneficial to first examine and review some selective segmentation models. Throughout, the original image is denoted by $z(x, y)$ with image domain $\Omega \subset R^2$. We

present here some quick thoughts on two previous selective segmentation models [15, 24]. In next sections, along with our purposed model we will also present reformulated segmentation models that incorporate region force and discrete total variation in the following models. We will compare the models discussed here to our purposed model.

Spencer- Chen Selective Model

Spencer and Chen used the ideas from previous segmentation literature to reformulate a segmentation solution into a convex minimisation problem [24]. Given, a two dimensional gray scale image $z(x, y) : \Omega \to R$, the Spencer-Chen model performs selective segmentation by solving the following minimization problem

$$min_{u,c_1,c_2}[F_(u, c_1, c_2)] = \mu \int_\Omega g(|\nabla z(x, y)|)|\nabla u| d\Omega + \int_\Omega [\lambda_1(z(x, y) - c_1)^2 - \lambda_2(z(x, y) - c_2)^2]u d\Omega$$
$$+\theta \int_\Omega D_E(x, y)u d\Omega + \alpha \int_\Omega \nu_\epsilon(u) d\Omega \tag{6.2}$$

Where μ, λ_1, λ_2, θ, and α are positive constants and u is the selective segmentation solution. c_1 and c_2 are the average image intensities of the foreground and background, respectively. The edge detector function is denoted by g, and the Euclidean distance term is denoted by D_E. The additional penalty term $\nu_\epsilon = \{max0, 2|u - \frac{1}{2}| - 1\}$ encourages the minimizer to be in the range [0,1]. The Euclidean distance term is extremely sensitive to the marker set chosen. The marker points must be evenly distributed in the region of interest to be segmented for satisfactory segmentation. Even in the segmentation region, the Euclidean penalty is always present, and it can be very high if there are few marker points. In the following section, we employ a more robust geodesic distance penalty, which allows us to have small or zero distance penalties at the edges and within the segmentation region.

Roberts Convex Geodesic Model

Given, a two dimensional gray scale image $z(x, y) : \Omega \to R$, the Roberts Convex Geodesic model [15] performs selective segmentation by solving the following minimization problem -

$$min_{u,c_1,c_2}[F_(u, c_1, c_2)] = \mu \int_\Omega g(|\nabla z(x, y)|)|\nabla u| d\Omega + \int_\Omega [\lambda_1(z(x, y) - c_1)^2 - \lambda_2(z(x, y) - c_2)^2]u d\Omega$$
$$+\theta \int_\Omega D_g(x, y)u d\Omega + \alpha \int_\Omega \nu_\epsilon(u) d\Omega \tag{6.3}$$

Where μ, λ_1, λ_2, θ, α are positive constants and u is the solution for selective segmentation. c_1 and c_2 are average image intensities of the foreground and the background. g is the edge detector function. This model was an improvement on the Convex Spencer - Chen model. In this model, the distance penalty term is edge weighted geodesic distance rather than the Euclidean distance penalty. For noise robustness, the geodesic distance term is computed using the isotropic total variation of image pixels. The geodesic distance increases closer to the edges, providing a more accurate reflection of the true similarity of pixels in an image from the marker set. One disadvantage of this model is that the fitting term does not incorporate any local or statistical information from the marker set. As a result, segmentation results in images with non-homogeneous, non-smooth, and scraggy boundary edges are unsatisfactory. In next section, we present the concept of region force term to use local information from the marker set.

6.3 Purposed Selective Segmentation Model

This work extends Roberts' model by enforcing a new region force term and a discrete TV-based geodesic distance penalty term as discussed in 4.3. We consider a problem with two partitions, one for the region of interest to be segmented and one for the background.

$$\text{Purposed model} = \text{New Region force} + \text{Edge detector term} + \text{Geodesic term} \tag{6.4}$$

Let $\phi_k(x)$ be an indicator function for $1 \leq k \leq K$. The general representation of indicator function associated with k-th sub domain can be defined as -

$$\phi_k(x) = \begin{cases} 1 & x \in \Omega_k \\ 0 & x \notin \Omega_k \end{cases} \tag{6.5}$$

We consider the collection of all indicator function represented by a set $S = \{[\phi_k] : \sum_{k=1}^{K} \phi_k = 1, 0 \leq \phi_k \leq 1\}$. For the implementation of model, We shall consider two partitions ϕ_1 and ϕ_2 as two indicator functions to from a set S_1. Ω_1 represents foreground and Ω_2 is the background. The average intensity of Ω_1 and Ω_2, calculated as mean average intensity is represented by c_1 and c_2 respectively.

$$S_1 = \{\phi_1, \phi_2 : \phi_1 + \phi_2 = 1, 0 \leq \phi_1 \leq 1, 0 \leq \phi_1 \leq 1\} \tag{6.6}$$

$$c_1 = \frac{\int \phi_1 . z d\Omega}{\int \phi_1 d\Omega} \tag{6.7}$$

$$c_2 = \frac{\int (1 - \phi_1) z d\Omega}{\int (1 - \phi_1) d\Omega} \tag{6.8}$$

Region Force

We present a region force term based on the Wei-Yin segmentation model [41]. Under the Bernoulli model, two region force functions f_1 and f_2 for the foreground and background were obtained as negative log-likelihood functions. The probability functions p_1 and p_2 were used to define the region force functions. The probability of a pixel (x, y) being part of a segmented area is denoted by p_1, and the probability of a pixel being part of the background is denoted by p_2. The region force functions were created with this purpose in mind.

$$\begin{aligned} f_1(x,y) &= 1 - 2p_1(x,y) \\ f_2(x,y) &= 1 - 2p_2(x,y) \end{aligned} \tag{6.9}$$

The proximity of the image intensity $z(x, y)$ to the average intensity of z(x,y) within Ω_1, denoted by c_1, is used to calculate p_1. A similar expression was used with the average intensity of z(x,y) inside Ω_2 for p_2. σ represents the image variance. We assume that each subdomain's image intensity follows the Gaussian random model given by $z(x) \approx N(c_1, c_2, \sigma^2)$. p_1 and p_2 probability functions were computed as -

$$\begin{aligned} p_1(x,y) &= \frac{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} \\ p_2(x,y) &= \frac{e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} \end{aligned} \tag{6.10}$$

The new region force term is -

$$\text{Region force term} = \frac{f_1 - f_2}{2} = p_2 - p_1 = \frac{e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} - \frac{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} \tag{6.11}$$

Edge Detector

The edge detector or regularisation term measures the geometry properties of the boundaries of the region to be segmented out of the image [18, 41]. The sum of the weighted lengths of each boundary was used as the regularizer in this work i.e $\int_{\delta\Omega_1} \alpha(x)dx$. A popular choice for the edge detector involves gaussian smoothing of image function $z(x)$ as -

$$\alpha(x) = \frac{\beta}{1 + \gamma|\nabla z(x)^2|} \tag{6.12}$$

$$\text{Edge detector} = \int_{\Omega} \alpha(x)|\nabla\phi_1|dx = \int_{\delta\Omega_1} \alpha(x)dx \tag{6.13}$$

Geodesic Term with Discrete TV

To calculate D_g, we firstly select a marker set M. We use *roipoly* in MATLAB to choose small number of points as marker points (Marker points are the vertices of the polygon drawn) . Based on the Roberts convex geodesic model [15], geodesic distance (D_g) from the marker set M is given by-

$$D_g(x) = \begin{cases} 0 & (x,y) \in M \\ \frac{D_g^0(x,y)}{||D_g^0(x,y)||_{L_\infty}} & (x,y) \notin M \end{cases} \tag{6.14}$$

We want (a) small gradients in homogeneous pixels and large gradients at edges for selective image segmentation, (b) noise robustness, and (c) enhanced qualitative nature of segmentation results. Based on recent works, the authors considered discrete TV denoising [19] as discussed in 4.3 and a penalty factor to include these aspects [19]. We use a penalty factor ϵ_D to compensate for the distance penalty for objects that are far from a marker set with a low penalty. $D_g^0(x,y)$ is derived from:

$$|\nabla D_g^0(x,y)| = \epsilon_D + \beta_g |\nabla z|^2 + \gamma_g D_E \tag{6.15}$$

Here, ϵ_D is penalty, $|\nabla z|$ is total variation term calculated by the discrete TV formulation and D_E is the euclidean distance term.

Proposed model and Dual formulation

With the inclusion of region force, edge detector and regularisation term the purposed model takes the form-

$$F = inf_{\phi_1 \epsilon \{0,1\}, \phi_1 \in S} \int_\Omega \mu[p2 - p1]\phi_1 dx + \int_\Omega \alpha(x)|\nabla \phi_1| dx + \theta \int_\Omega D_g(x,y)\phi_1 dx \tag{6.16}$$

In a past study , Chan, Esedoglu and Nicolova [50] transformed the above functional to a convex functional by relaxing the binary constraint between 0 and 1. Then Functional F can be written as -

$$F = inf_{0 \leq \phi_1 \leq 1, \phi_1 \in S} \int_\Omega \mu[p2 - p1]\phi_1 dx + \int_\Omega \alpha(x)|\nabla \phi_1| dx + \theta \int_\Omega D_g(x,y)\phi_1 dx \tag{6.17}$$

Using p_1 and p_2,

$$F(\phi_1, c_1, c_2) = min_{0\leq\phi_1\leq 1, \phi_1\in S} \int_\Omega \mu \left[\frac{e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} - \frac{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} \right] \phi_1 dx + \int_\Omega \alpha(x)|\nabla\phi_1| dx + \theta \int_\Omega D_g(x,y)\phi_1 dx \tag{6.18}$$

Next, we derive the dual formulation [117] with inclusion of dual variable q_1. We utilise min-max theorem [118] and following proven equality -

$$\int_\Omega \alpha(x,y)|\nabla\phi_1| dx = max_{|q_1|\leq\alpha(x,y)} \int_\Omega \phi_1 \text{div} q_1 dx \tag{6.19}$$

With the above equality the formulation of F reduces to -

$$min_{0\leq\phi_1\leq 1, \phi_1\in S} \quad max_{|q_1\leq\alpha(x,y)|} \int_\Omega \left[\frac{\mu e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} - \frac{\mu e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} + \theta D_g \right] \phi_1 dx + \int_\Omega \phi_1 \text{div} q_1 dx \tag{6.20}$$

using min-max theorem the formulation reduces to -

$$max_{|q_1\leq\alpha(x,y)|} \quad min_{0\leq\phi_1\leq 1, \phi_1\in S} \int_\Omega \left[\frac{\mu e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} - \frac{\mu e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} + \theta D_g + \text{div} q_1 \right] \phi_1 dx \tag{6.21}$$

We can also write the above equation as a dual problem -

$$max_{|q_1\leq\alpha(x,y)|} \int_\Omega \min \left[\frac{\mu e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} - \frac{\mu e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} + \theta D_g + \text{div} q_1 \right] dx \tag{6.22}$$

In this work, we use recent primal-dual algorithms [48, 36] to solve the aforementioned dual problem. For the above min-max problem, we use the primal-dual hybrid gradient (PDHG) algorithm [33] (12). A projected gradient descent step updates the primal and dual variables after each iteration. This is a subset of the general primal-dual algorithms. The convergence analysis of primal-dual algorithms has been extensively researched in the literature and can be found in [119, 120, 121, 122, 123].

Algorithm 2 PDHG algorithm

Set β, γ and θ
Compute $\alpha(x) = \frac{\beta}{1+\gamma|\nabla z(x)^2|}$
Compute $D_g(x) = \begin{cases} 0 & (x,y) \in M \\ \frac{D_g^0(x,y)}{||D_g^0(x,y)||_{L_\infty}} & (x,y) \notin M \end{cases}$ from 6.15
initialize ϕ_1^0 arbitrarily
for l =1 to maximum iterations **do**
Calculate c_1 and c_2 using 6.7 and 6.8
calculate and update for q_1 andϕ_1
$q_1^l = \Pi_{|q_k \leq \alpha(x,y)|}\ (q_1 - \beta_l \nabla\ \phi_1^l)$
$\phi_1^l = \Pi_S\ (\phi_1^l - \gamma_l\ (div q_1^l + (\frac{e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} - \frac{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}}) + \theta D_g))$
$\phi_1^{l+1} = \theta_1\ [\ \phi_1^l\ ] + (1-\theta_1)[\phi_1^{l+1}\ ]$
end for
$\phi_1^* \leftarrow \phi_k$

6.3.1 Mathematical analysis

Proposition 6.3.1. *The model (??) is convex and there exists a minimiser $\phi \in W^{1,2}(\Omega)$ for F.*

Proof. Using the concept of [124], we first show that $0 \leq \inf_\phi F(\phi) < \infty$. For the lower bound it can be easily shown by $\phi_0 = 0$. For the upper bound, we consider $\phi_0 = 1$ and obtain the inequality-

$$inf_\phi F(\phi) \leq F(\phi_0) = \int_\Omega p_2 - p_1 dx + \theta \int_\Omega D_g dx < \infty$$

We then demonstrate that functional $F(\phi)$ is convex by checking the convex inequality [125] for each term of the functional, for $\psi \epsilon [0,1]$:

$$\int_\Omega [p_2 - p_1](\psi\phi_1 + (1-\psi)\phi_2)dx = \psi \int_\Omega [p_2 - p_1]\phi_1 dx + (1-\psi)\int [p_2 - p_1]\phi_2 dx$$

$$\theta \int_\Omega D_g(\psi\phi_1 + (1-\psi)\phi_2)dx = \theta\psi \int_\Omega D_g(\phi_1)dx + (1-\psi)\theta \int_\Omega D_g(\phi_2)dx$$

$$\begin{aligned}\int_\Omega \alpha|\nabla(\psi\phi_1 + (1-\psi)\phi_2)|dx &\leq \int_\Omega \alpha|\nabla(\psi\phi_1)|dx + \int_\Omega \alpha|\nabla(1-\psi)\phi_2|dx \\ &= \psi \int_\Omega \alpha|\nabla(\phi_1)|dx + (1-\psi)\int_\Omega \alpha|\nabla\phi_2|dx\end{aligned}$$

So, F is convex. Also, recent study [124] has proved that the second term of $F(\phi)$ is lower semi-continuous. The other terms are also lower semi-continuous as $\forall \phi_1, \phi_0 \epsilon W^{1,2}(\Omega)$-

$$\lim_{\phi_1 \to \phi_o} \inf \int_\Omega D_g \phi_1 dx = \int_\Omega D_g \phi_0 dx$$

$$\lim_{\phi_1 \to \phi_o} \inf \int_\Omega [p_2 - p_1]\phi_1 dx = \int_\Omega [p_2 - p_1]\phi_0 dx$$

Now, we can say that our model $F(\phi)$ is lower semi continuous. $F(\phi)$ is coercive also as $\phi_1 \to \infty$ implies $F \to \infty$. With the information that $W^{1,2}(\Omega)$ is a reflexive banach space, $F(\phi)$ is convex, lower semi-continuous as well as coercive, it can be concluded that a minimiser exists in $W^{1,2}(\Omega)$.

□

6.4 Numerical Results

In this section we will provide the serval results in MATLAB to demonstrate the advantages, efficiency, and stability of the intended segmentation model over previous and related selective image segmentation models. We compared the following models -

- M1- the convex Spencer - Chen model
- M2- the Roberts convex geodesic model
- M3- the reformulated Roberts model with inclusion of discrete TV in geodesic distance penalty
- M4- the reformulated Spencer - Chen model with region force term
- M5- the purposed selective geodesic variational model with new region force term and isotropic TV based geodesic distance penalty term
- M6- the purposed selective geodesic variational model with new region force term and discrete TV based geodesic distance penalty

We extend M2-M3 to see if including discrete TV in the geodesic penalty term improves M2. Specifically, the model M3 [15] is -

$$min_{u,c_1,c_2}[F_(u,c_1,c_2)] = \mu \int_\Omega g(|\nabla z(x,y)|)|\nabla u| d\Omega + \int_\Omega [\lambda_1 (z(x,y) - c_1)^2 - \lambda_2 (z(x,y) - c_2)^2] u d\Omega$$

$$+\theta \int_\Omega D_g(x,y) u d\Omega + \alpha \int_\Omega \nu_\epsilon(u) d\Omega \tag{6.23}$$

Where μ, λ_1, λ_2, θ, α are positive constants and u is the solution for selective segmentation. c_1 and c_2 are average image intensities of the foreground and the background. g is the edge detector function. D_g is the geodesic distance penalty which is zero for the marker set. For all other image pixels it is computed as -

$$|\nabla D_g(x,y)| = \epsilon_D + \beta_g |\nabla z|^2 + \gamma_g D_E \tag{6.24}$$

Here, ϵ_D is penalty, $|\nabla z|$ is total variation term calculated by the discrete TV formulation [19] and D_E is the euclidean distance term. For more information on discrete TV, we refer readers [19]. Similarly, we extend M1-M4 to see if including new region force as a fitting term in a previous variational segmentation model is preferable. Specifically, M4 [24] is -

$$min_{u,c_1,c_2}[F_(u,c_1,c_2)] = \mu \int_\Omega g(|\nabla z(x,y)|)|\nabla u| d\Omega + \int_\Omega \lambda \left[\frac{e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}}} - \frac{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}}}{e^{\frac{(-|z(x)-c_1|)}{2\sigma^2}}} + e^{\frac{(-|z(x)-c_2|)}{2\sigma^2}} \right] u d\Omega + \theta \int_\Omega D_E(x,y) u d\Omega + \alpha \int_\Omega \nu_\epsilon(u) d\Omega \tag{6.25}$$

Where μ, λ, θ, α are positive constants and u is the solution for selective segmentation. c_1 and c_2 are average image intensities of the foreground and the background. g is the edge detector function. D_E is the Euclidean distance penalty term.

As test images, multi-modal medical images (MRI/CT) are used. After scaling medical images in the range [0,1], numerical algorithms are applied. The linear -stretch scaling formula is

$$z_s = \frac{z - z_{min}}{z - z_{max}} \tag{6.26}$$

Here, z represents the input image, and z_{min} and z_{max} represent the minimum and maximum pixel values of z. Three different sets of test results are showcased. Models M1-M4 are compared to the two versions of purposed model M5 and M6 in Test 1. The tested images have (a) average target intensity greater than surrounding intensity and (b) non-homogeneous, non-smooth, and scraggy boundary edges for the objects to be segmented. In Test 2, we look at how the position and number of marker points affect the results. In Test 3, we will investigate the impact of noise on our purposed model. Finally, we use widely used segmentation scores as the gold standard to assess segmentation quality. Experts labelled the ground truths. The parameters μ and θ used for image segmentation comparison tests are given in Table 6.1. We set $\beta = 1$, $\gamma = 10^2$, $\epsilon_D = 10^3$,$\beta_g = 10^3$, $\gamma_g = 0.1$ and $\theta_1 = -0.5$. Identical sets of parameters were used across M1-M6 for the comparison results.

	1	2	3	4	5(a)	5(d)	5(g)
μ	.225	.35	.25	.23	.23	.20	.25
θ	10	10	100	10	10	20	5
l	100	100	100	100	100	100	100

Table 6.1: Parameters used for the image segmentation comparison tests in Figures 1, 2, 3, 4, 5(a), 5(d), and 5(g).

6.4.1 Test 1

In this subsection, we present the segmentation results for models M1 - M6 for challenging test images with average target intensity greater than surrounding intensity. The boundary edges of the objects to be segmented are non-homogeneous, non-smooth, and scraggy. Images are displayed alongside marker sets. Figure 6.2 and Figure 6.3 are $T2$ MRI brain scans with tumor region to be segmented. Figure 6.4 illustrates the results of a bone image. Figure 6.5 demonstrates the segmentation of the liver from a CT abdominal scan. Figure 6.6 shows the segmentation result of a purposed model on $T2$ MRI tumour scans from different scanners and a CT image to generalise our model to images from different scanners. In figure 6.7, we display the convergence curve of the used numerical method.

The task of segmentation is difficult due to the high variance in shape and complicated edges, as shown in figure 6.2 and figure 6.3. We can see that M1, M2, M3 and M4 are unable to capture the boundaries of the object of interest. In particular, the tumour can be segmented using the M5 and M6 models, which enforce the new region term. In the case of M1 and M4, the regularisation term includes a euclidean distance term, which, because it is negligible, captures fewer details in order to reach the boundaries. In other words, the distance term is insufficient to offset the effect of fitting term.

The geodesic distance term is used as an additional regularisation term to weight the fitting terms in M2 and M3. The fitting term encourages segmentation of bright objects while ignoring non-homogeneous edges with low intensity, resulting in poor segmentation. M5 and M6 utilise region force term along with geodesic distance term with the only difference being the TV formulation used. Visually, M5 and M6 achieve comparable segmentation quality, with M6 being able to segment sharp edges. The region force models M5 and M6 outperform the previous models. The similar performance of M5 and M6 in comparison results suggests that the region force term has a much significant impact on the segmented solution than the TV formulation used in the computation of the geodesic distance term.

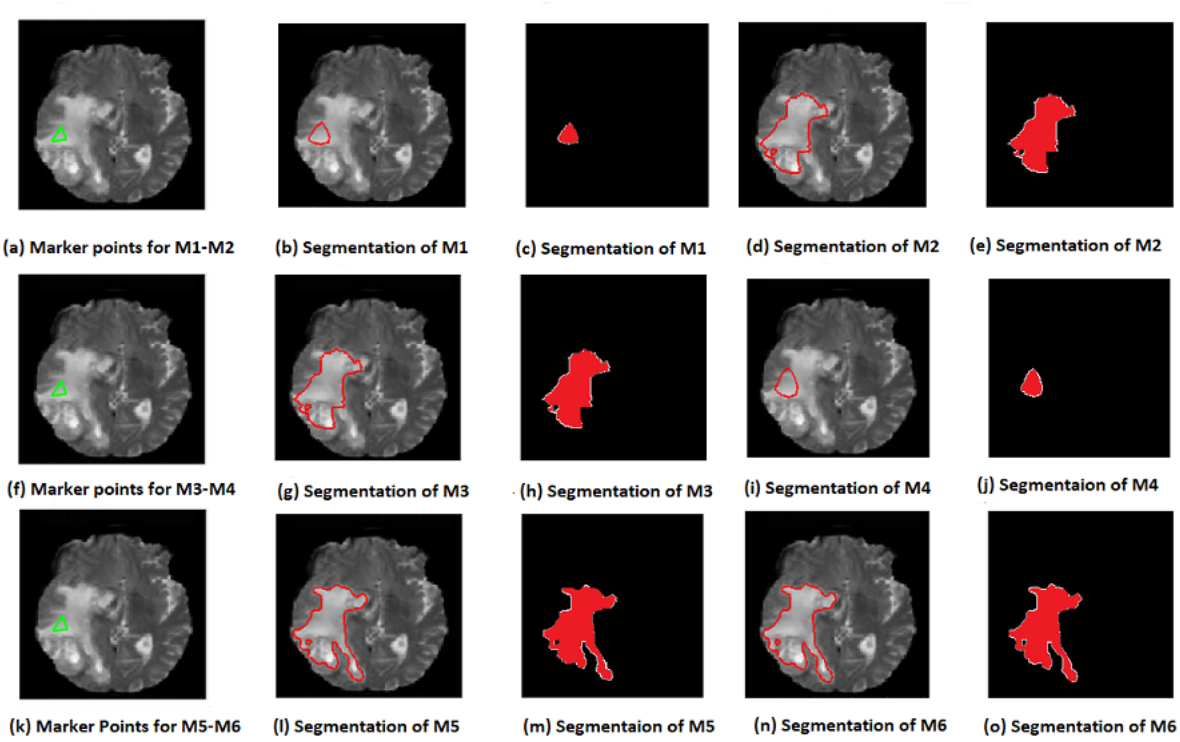


Figure 6.2: Segmentation results of Models M1-M6 on brain tumor image

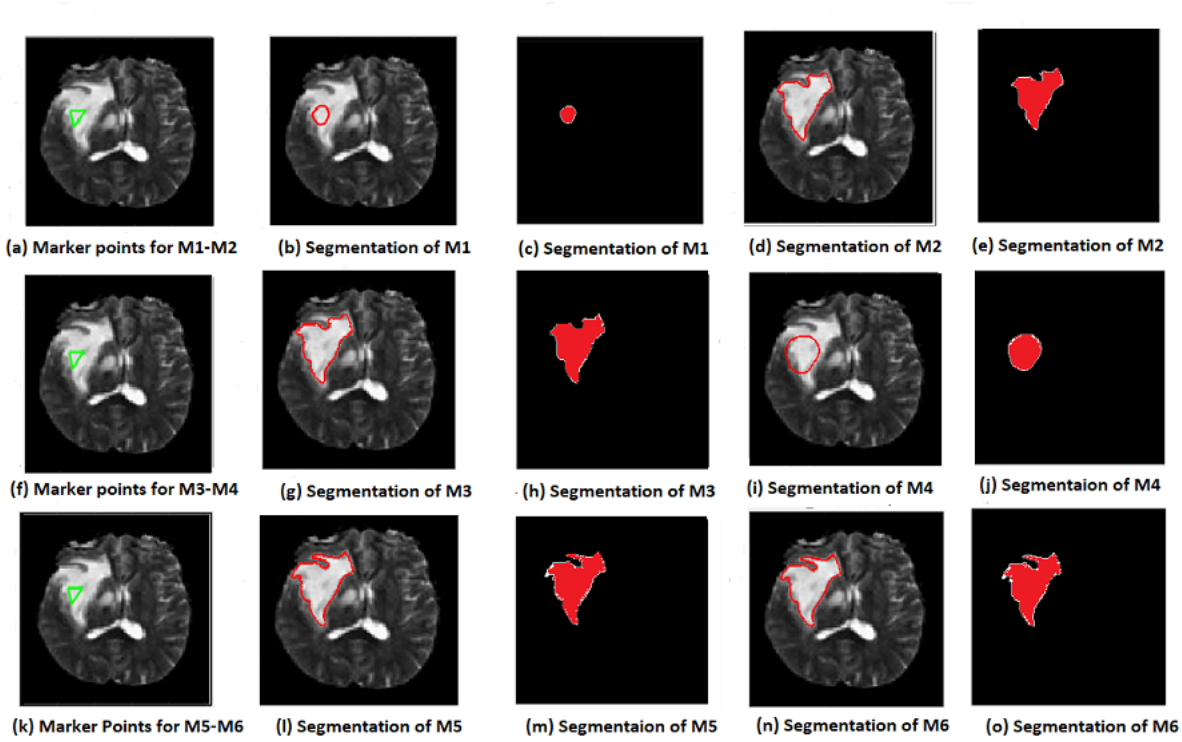


Figure 6.3: Segmentation results of Models M1-M6 on brain tumor image.

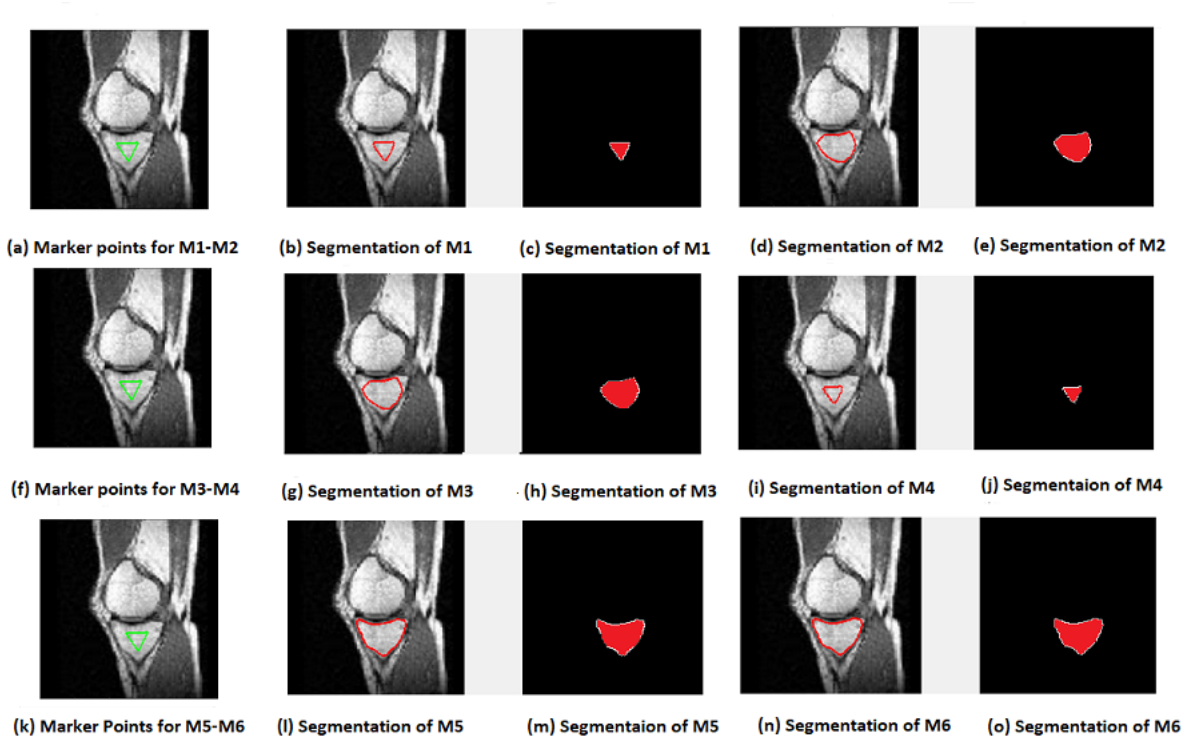


Figure 6.4: Segmentation results of Models M1-M6 on bone image.

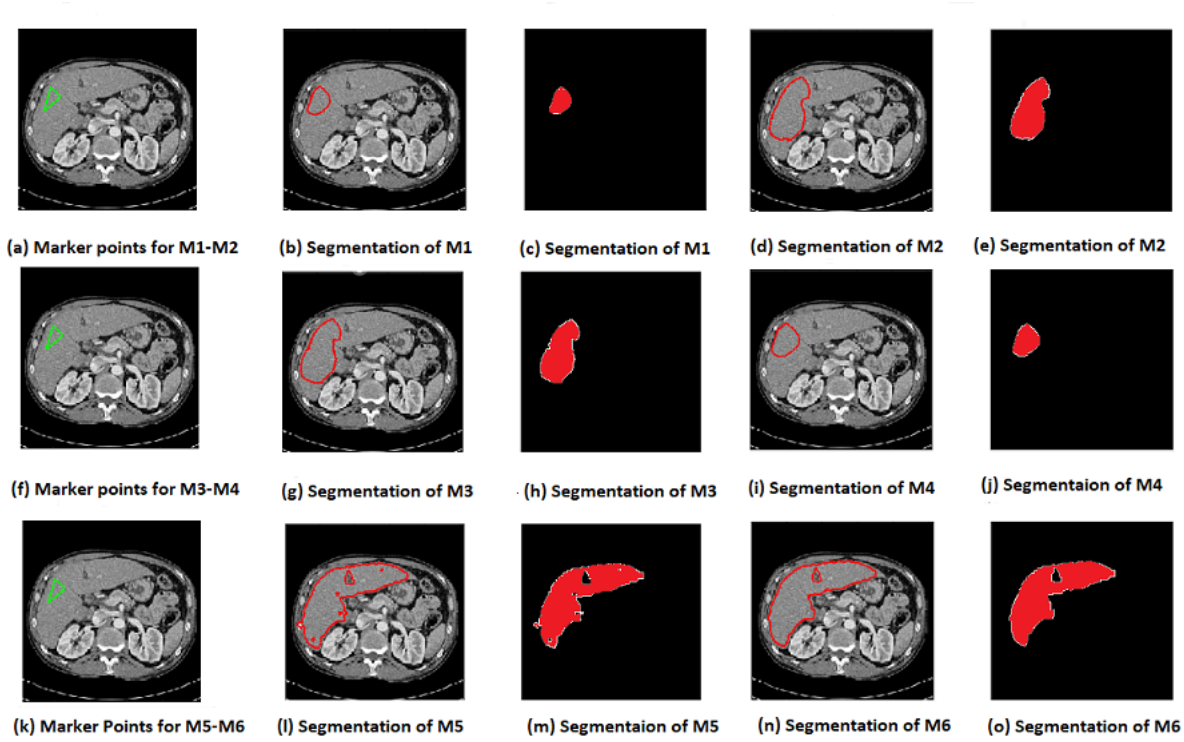


Figure 6.5: Segmentation results of Models M1-M6 on abdominal CT image.

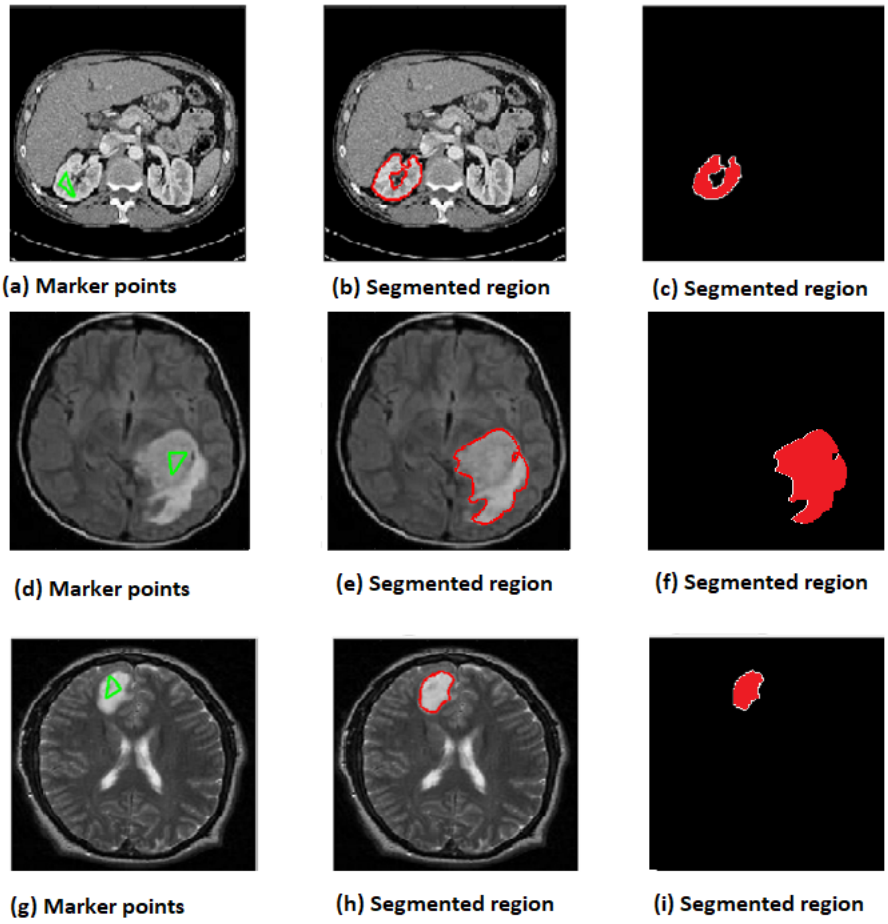


Figure 6.6: Segmentation results of M6.

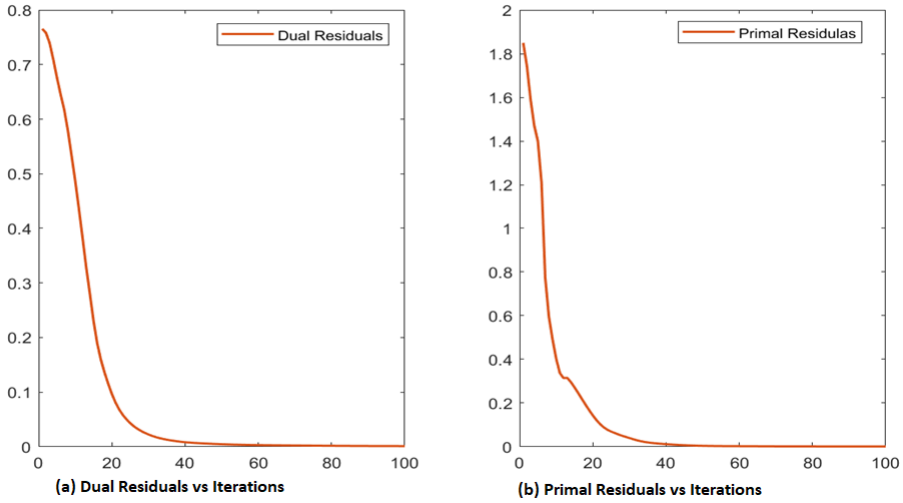


Figure 6.7: Convergence curve of the used numerical method

6.4.2 Quantitative comparisons

We use the Dice coefficient and the Jaccard coefficient as quantitative measures to evaluate the performance of segmentation results. The Jaccard score starts at one for a segmentation that matches the ground label and decreases to zero as the quality of the segmentation decreases. The dice score, a popular statistic score, assesses the similarity of the segmented mask and the ground label. The DICE score and Jaccard Score are calculated using a model segmentation result Σ and a ground truth label GT.

$$\text{Dice score} = \frac{\Sigma \cap GT}{\Sigma + GT} \tag{6.27}$$

$$\text{Jaccard score} = \frac{\Sigma \cup GT}{\Sigma + GT - (\Sigma \cap GT)} \tag{6.28}$$

Alongside the above scores we calculate sensitivity of segmentation results by -

$$\text{Sensitivity score} = \frac{\text{correctly classified pixels}}{\text{correctly classified pixels} + \text{incorrect or falsely classified pixels}} \times 100 \tag{6.29}$$

We refer sensitivity score as segmentation accuracy (SA) score. The mean and standard deviation values of the DICE and JACCARD scores are shown in Table 6.2. For further visualisation, in Figure 6.8 we present box plots of segmentation scores on a dataset of 50 medical images. These higher mean segmentation scores clearly demonstrate that our model M6 is more effective than the models M1-M5.

Model	Dice		Jaccard		Sensitivity	
	μ	σ	μ	σ	μ	σ
M1	0.2121	0.072	0.12	0.0467	0.126	0.0467
M2	0.7517	0.1603	0.6204	0.1894	0.6299	0.196
M3	0.7587	0.1574	0.6287	0.1869	0.6382	0.1929
M4	0.3077	0.1434	0.1888	0.1096	0.1942	0.1183
M5	0.9107	0.0174	0.8364	0.0293	0.8632	0.0406
M6	0.9143	0.0163	0.8424	0.0276	0.8688	0.04

Table 6.2: Segmentation scores (A) Dice score and (B) Jaccard score.

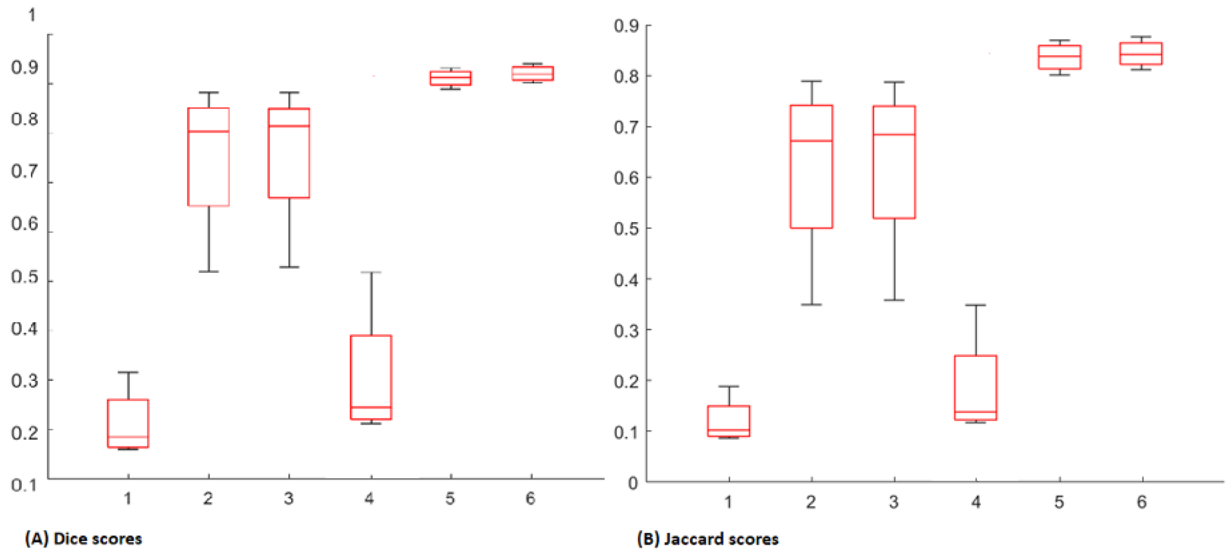


Figure 6.8: Box plot of Segmentation scores for M1-M6.

6.4.3 Test 2

We examine the impact of the number of marker points in this subsection. Different numbers of marker points (3, 6, and 9 points) are used to segment the same MRI tumour scan in figure 6.9. The proposed Model M6 is unaffected by the number of markers within the segmented region of interest. M6 generates similar segmentation accuracy in all three cases. According to recent studies, the number of marker points influences the Euclidean distance term [15, 24]. The geodesic distance map and region force term remains nearly identical regardless of the number of markers chosen, allowing our selective model to be independent of marker set size.

Second, we investigate the effect of marker point position. To validate position independence, we choose four marker points close to the target object from various positions, namely

all four points inside the boundary, two points inside and two points on the boundary, two points outside and two points inside the boundary. When two marker points are on the boundary in Figure 6.10 (d), the model M6 produces the same strict segmentation accuracy as when all marker points are inside the boundary in Figure 6.10 (a).This is not the case when the marker points are outside as shown in Figure 6.10 (g). To get better segmentation results from the purposed algorithm, we conclude that the marker points should be inside the object or near the boundary.

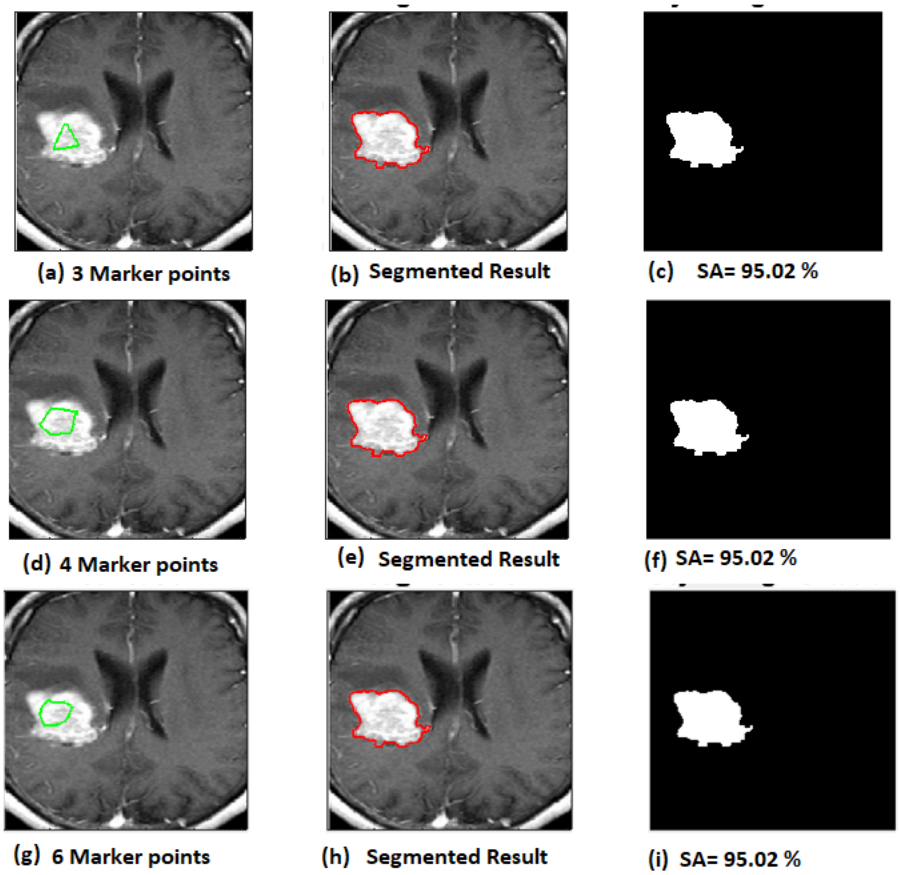


Figure 6.9: Segmentation results of M6 for different Marker points.

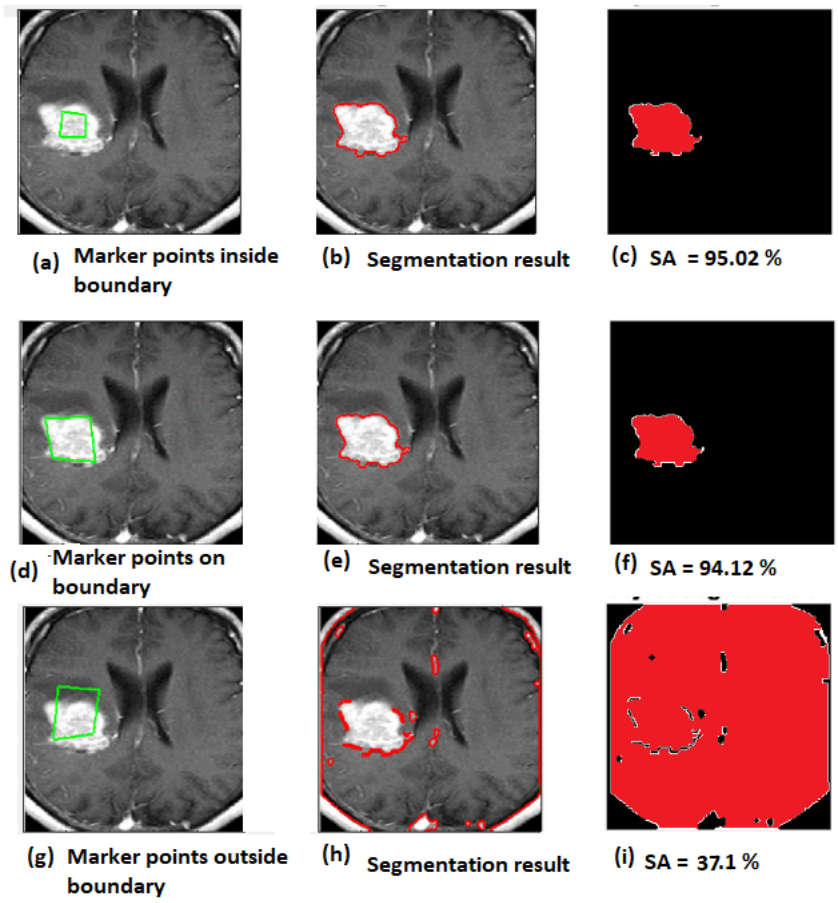


Figure 6.10: Segmentation results of M6 for different Marker point positions.

6.4.4 Test 3

Finally, in figure 6.11, we examine the impact of noise on our model. The original image is contaminated by Gaussian noise in MATLAB at various noise levels. A zero mean noise with a variance of .05, .07, and .09 is added to the tumour image. Our model can achieve adequate segmentation sensitivity of 92% when the variance is .05. However, in the presence of excessive noise, the model produces an insufficient level of segmentation with segmentation accuracy as low as 77%.

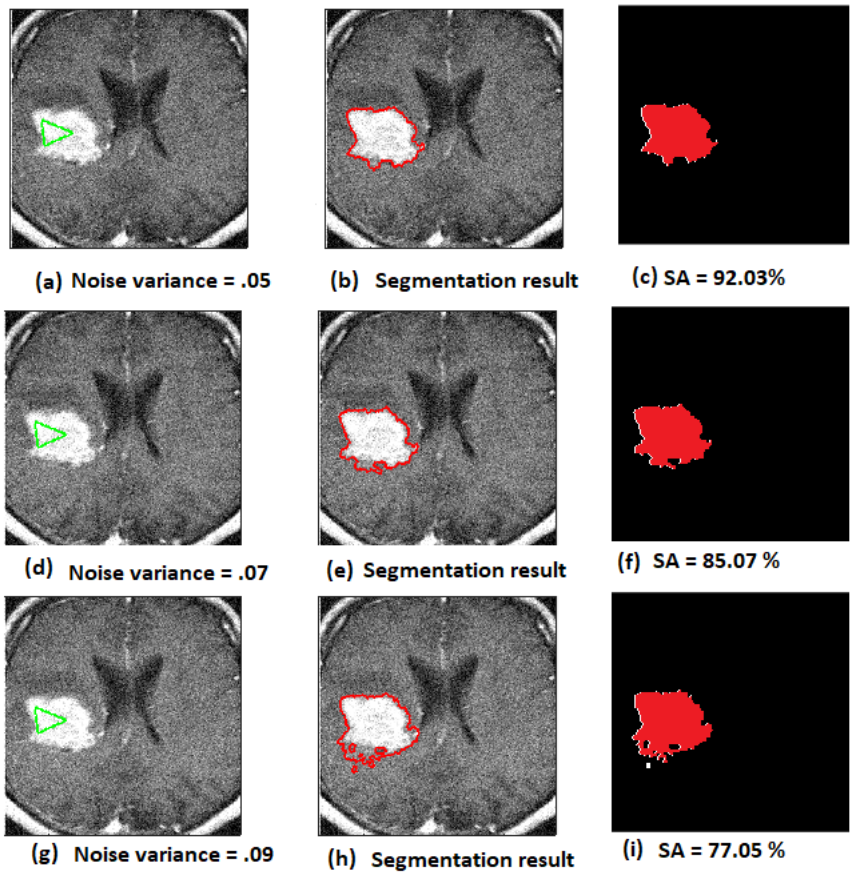


Figure 6.11: Segmentation results for M6 with different noise levels.

6.5 Conclusion

We propose a convex selective segmentation model with a novel region force and geodesic distance penalty term. We tested the model on some challenging medical images to compare the segmentation efficiency and effectiveness with the other segmentation models. The model outperformed competing models in terms of segmentation results and segmentation scores. The proposed model's segmentation results are sufficiently robust to user input and noise levels. To improve future work, we advise readers to incorporate higher-order regularizers and data-driven new region force terms into the existing model. Application to 3D segmentation data sets is also worth considering.

Chapter 7

Deep Learning Segmentation and Future Work

Although not initially part of the project, but we were curious to see the application of variational models in unsupervised segmentation. In this chapter, we incorporate a variational model into a deep learning algorithm in an unsupervised framework. We will use purposed variational segmentation method in chapter 6 as a loss function in a deep learning architecture. We will also discuss some extensions to the work done in this thesis.

7.1 Unsupervised Segmentation

7.1.1 Introduction

Since the last ten years, deep learning-based segmentation has gained popularity [126, 127]. Deep learning segmentation is frequently regarded as the best method for automatic segmentation [126]. In addition, variational segmentation methods have been extensively researched for several decades. In this thesis, in chapter 4 and chapter 6 we have seen that variational methods with objective of minimizing a functional as a segmentation solution are effective on challenging images with low contrast variability, in-homogeneous intensity and scraggy, uneven edges.

Deep learning is data driven and require manually marked ground truths [126, 128]. Variational methods often suffer with local minimums and have not been generalised on a larger multimodal datasets [24]. This implies that deep learning and variational methods both have limitations. Using both variational segmentation methods and deep learning based segmentation in a single framework can help to address some of the previous weaknesses and provide better segmentation results [128, 129, 130]. The standard learning approach in deep learning networks is based on adjusting network weights and biases. Typically, a loss function is minimised via backpropagation, which helps in iteratively updating the weights

and biases of the convolutional networks [126]. Minimising the loss function encourages the network's outputs to be similar to the ground truth [126, 127]. Deep learning networks are data driven and to achieve satisfactory results, usually thousands of labelled images are used in the training stage. For tumour segmentation, labelled tumour images are usually marked manually by multiple clinicians. As a result, labelled images are costly and time-consuming to obtain.

Usually in variational segmentation methods, we have an energy functional and segmentation results are achieved by minimising this energy functional. This inspires to use the energy functionals as a loss function in a convolutional neural network [128, 129]. This will eliminate the need for a substantial amount of labelled data during the training phase and enable a network to use unlabelled data.

7.1.2 Deep Learning Implementation

U-Net

U-Net [1] is a popular CNN architecture that was first used for semantic image segmentation in the biomedical field, but it has since shown to be successful in a variety of other image processing tasks, including denoising, super-resolution, registration, and deblurring. There are two main parts to the architecture as shown in Figure 7.1. Convolutions, non-linear activation's (ReLUs), and max-pooling layers make up the one path, also known as the contracting path, which minimises the spatial size. The second path, involves more convolutions, non-linear activation's (ReLUs), and up-sampling to enlarge the spatial domain. Skip connections help in concatenating early layers with later layers, making it easier to update parameters in the early layers.

Purposed Network

In order to produce a precise and accurate prediction, UNet is a popular architecture for semantic segmentation because it can extract both low level information from the initial layers and high level information from the final layers [1]. For Tumour segmentation task, we use UNet like architecture as shown in Figure 7.2. The inputs to this network are in form of image z and geodesic distance penalty D_g as discussed in 6.3. The image and geodesic distance are subjected to convolutional layers downwards separately. We use a dot product to combine the two paths at the bottleneck. This creates a single upward path that produces segmentation results similar to input image size.

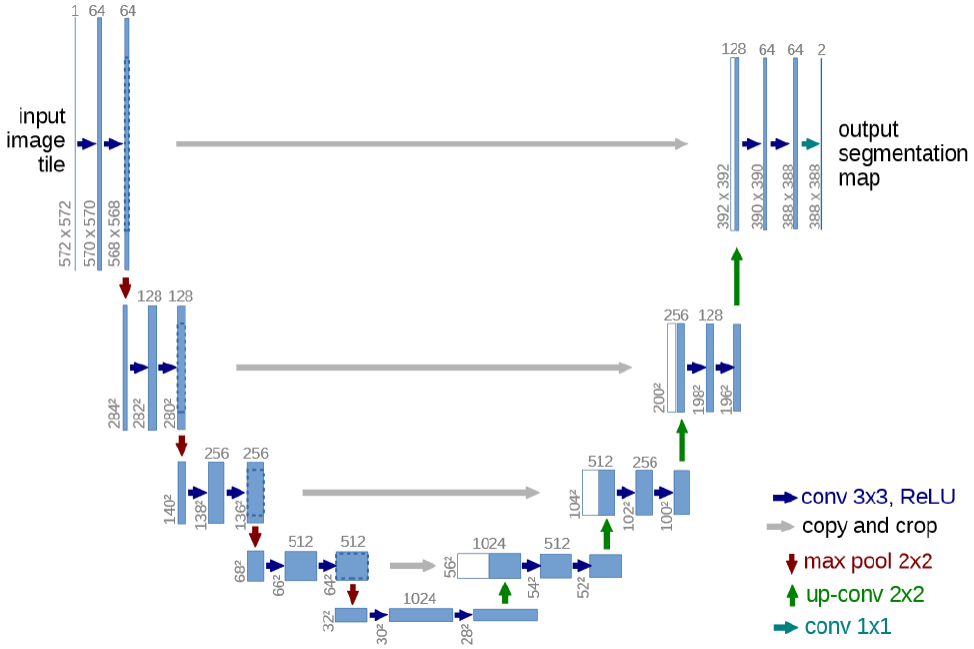


Figure 7.1: U-Net architecture (taken from Figure 1 from [1])

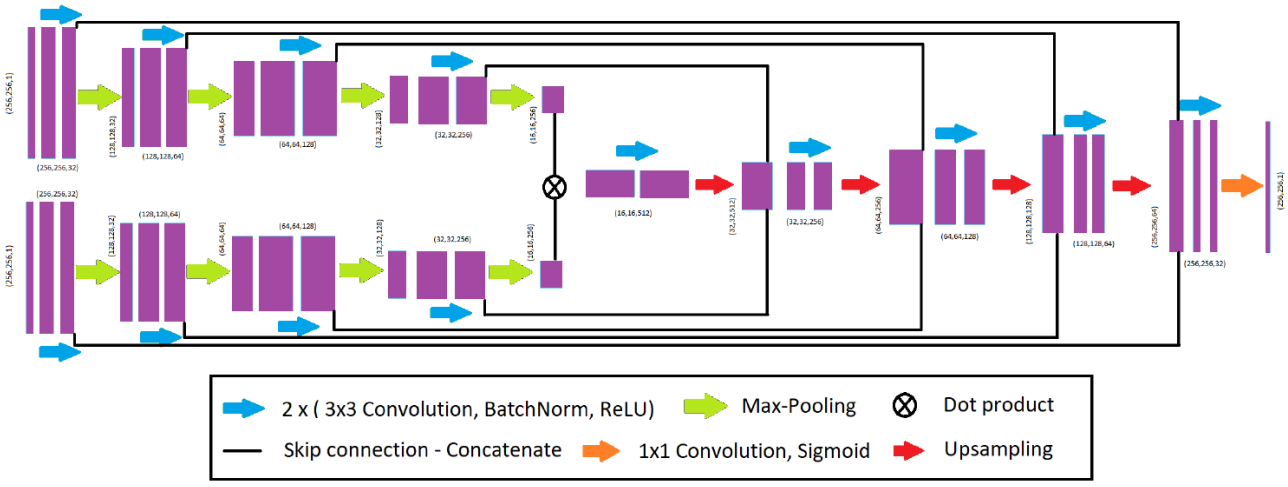


Figure 7.2: Structure of CNN used for tumour segmentation. Network has two downward paths one for input image z and other for D_g. This network is based on U-Net [1].

Unsupervised Algorithm

Marker set in form of a user input is usually obtained by a user clicking points on the target object. This network in Figure 7.2 requires a user input M as it needs to pass through geodesic distance D_g (6.3) through network and also to evaluate loss function.

Now we donot use ground truth labels for segmentation results. Instead, we implement an unsupervised approach by using the purposed variational functional 6.18 as a loss function. In the following loss function, N represents total training images. $u_\Theta(z, M)$ is the output of our network with weights Θ and input image as z and marker set as M. We denote the jth output of the network with jth input as $u(j) = u_\Theta(z(j), D(j))$. The input image variance is σ with μ, θ as positive parameters.

$$\mathrm{L}(\Theta) = \sum_{j=1}^{N} \int_\Omega |\nabla u^j| dx + \int_\Omega \mu [\frac{e^{\frac{(-|z^j - c_2^j|)}{2\sigma^2}}}{e^{\frac{(-|z^j - c_1^j|)}{2\sigma^2}} + e^{\frac{(-|z^j - c_2^j|)}{2\sigma^2}}} - \frac{e^{\frac{(-|z^j - c_1^j|)}{2\sigma^2}}}{e^{\frac{(-|z^j - c_1^j|)}{2\sigma^2}} + e^{\frac{(-|z^j - c_2^j|)}{2\sigma^2}}}] u^j dx + \theta \int_\Omega D_g^j u^j dx \tag{7.1}$$

7.1.3 Predicted Results from Deep learning network

We use BraTS dataset [131, 132] of brain tumour MR images to show the performance of our trained deep learning model. We used 100 images for training. The predicted results on testing images are given below in Figure 7.3 and Figure 7.4. In the top row are the input testing images and the bottom row represents the pixels in the image predicted as tumour.

It is evident that the unsupervised algorithm produces successful lesion segmentation. Since it can be time consuming and expensive to acquire ground truth labels, the intended unsupervised algorithm is capable of producing satisfactory results.

7.1.4 Conclusion

We proposed to use the variational segmentation model in chapter 6 as a loss function in a UNet network. We trained the network in an unsupervised way without marked labels on tumour images. Visual results represent that a adequate level of tumor segmentation can be achieved in a automatic manner, once the model is trained.

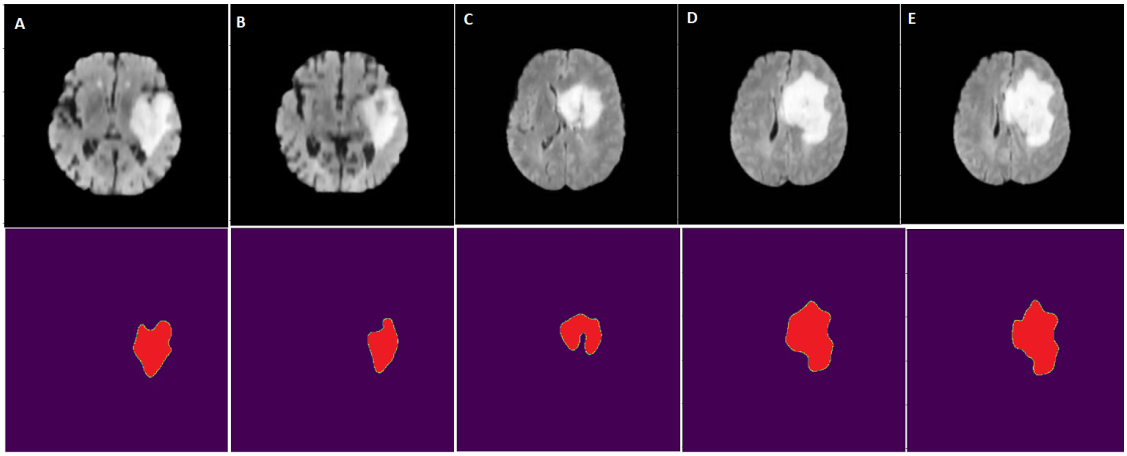


Figure 7.3: A collection of results using our test data from the BraTS dataset. The top row has the input testing images and the bottom row represents the pixels in the image predicted as tumour.

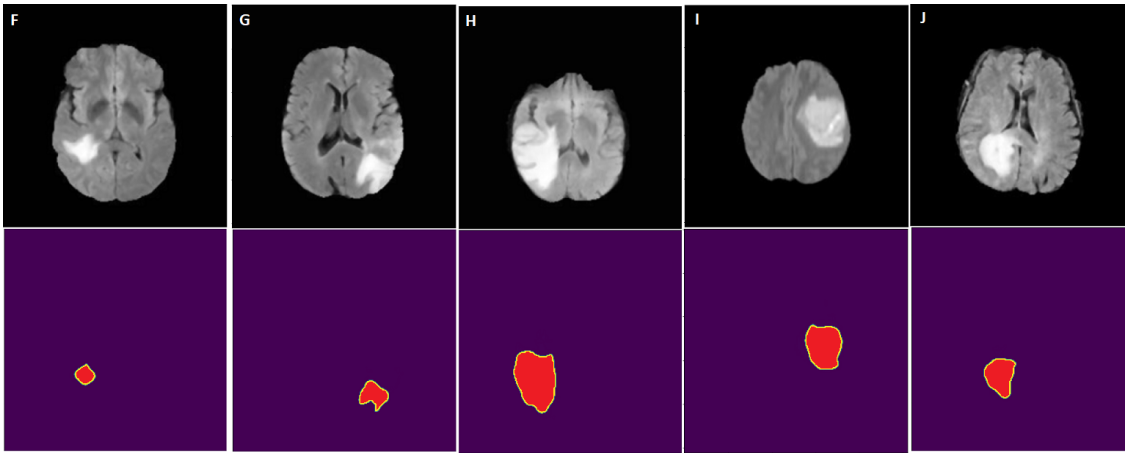


Figure 7.4: A collection of results using our test data from the BraTS dataset. The top row has the input testing images and the bottom row represents the pixels in the image predicted as tumour.

7.2 Future Work

There are several directions to work for further research, here we discuss some topics for extension:

- **Arterial spin labeling (ASL) for penumbra estimation in stroke patients:** In chapter 4,5 we discuss problems associated with the imaging pipeline in MRP / CTP for core and penumbra estimation. Another type of technique to measure tissue perfusion in terms of penumbra and core is Arterial spin labeling (ASL)[133]. ASL has a number of benefits over other perfusion techniques and is now regularly used in clinical settings. ASL is perfect for research and clinical studies due to its non-invasiveness and capacity to sensitively measure tissue perfusion in terms of penumbra and core [133]. However, certain ASL artefacts (e.g. arterial transit artefact, blurring, increased venous signal, etc.) lead to diagnostic pitfalls in interpreting ASL CBF, Tmax maps [134]. Reworking on the ASL image pipeline by correcting ASL artefacts and comparing ASL results with the perfusion MRI or CTP would be an ideal extension of the work presented in the thesis.

- **Segmentation extensions** : The model we have suggested in chapter 6 is for segmenting grayscale medical images, but this could be generalised to segment colour images as well.

- **Video Segmentation**: The models discussed in chapter 6 and chapter 4 can also be applied to video segmentation or object tracking. Using the segmentation from one frame/image slice can provide initialisation for the next frame/image slice, allowing the tracking of objects over different time points.

- **Deep Learning Segmentation for AIF**: The model in chapter 4 for AIF segmentation may have a nice deep learning implementation, similar to that discussed in chapter 7. To identify AIF, CNN network can be trained in a unsupervised way by implementing the variational segmentation as loss function in deep learning architecture.

Bibliography

[1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. volume 9351, pages 234–241, 10 2015.

[2] Fernando Calamante. Arterial input function in perfusion mri: A comprehensive review. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 74:1–32, 2013.

[3] Sukhdeep Singh Bal, Ke Chen, Fan-Pei Gloria Yang, and Giia-Sheun Peng. Arterial input function segmentation based on a contour geodesic model for tissue at risk identification in ischemic stroke. *Medical Physics*, 49(4):2475–2485, 2022.

[4] Shengyu Fan, Yueyan Bian, Erling Wang, Yan Kang, Danny J. J. Wang, Qi Yang, and Xunming Ji. An automatic estimation of arterial input function based on multi-stream 3d cnn. *Frontiers in Neuroinformatics*, 13, 2019.

[5] Leif Ostergaard, Robert M. Weisskoff, David A. Chesler, Carsten Gyldensted, and Bruce R. Rosen. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. part i: Mathematical approach and statistical analysis. *Magnetic Resonance in Medicine*, 36(5):715–725, 1996.

[6] Sukhdeep Singh Bal, Fan Pei Gloria Yang, Yueh-Feng Sung, Ke Chen, Jiu-Haw Yin, and Giia-Sheun Peng. Optimal scaling approaches for perfusion mri with distorted arterial input function (aif) in patients with ischemic stroke. *Brain Sciences*, 12(1), 2022.

[7] Philip A. Barber, Andrew M. Demchuk, Jinjin Zhang, and Alastair M. Buchan. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. *The Lancet*, 355:1670–1674, 2000.

[8] Fanpei Yang, Sukhdeep Singh Bal Bal, Yueh-Feng Sung, and Giia-Sheun Peng. Mathematical framework of deconvolution algorithms for quantification of perfusion parameters. *Acta neurologica Taiwanica*, 29(3):79—85, September 2020.

[9] Fanpei Yang, Sukhdeep Singh Bal Bal, Yueh-Feng Sung, and Giia-Sheun Peng. Mathematical framework of deconvolution algorithms for quantification of perfusion parameters. *Acta neurologica Taiwanica*, 29(3):79—85, September 2020.

[10] R. Wirestam, L. Knutsson, J. Risberg, S. Börjesson, E.-M. Larsson, L. Gustafson, U. Passant, and F. Ståhlberg. Attempts to improve absolute quantification of cerebral blood flow in dynamic susceptibility contrast magnetic resonance imaging: a simplified t1-weighted steady-state cerebral blood volume approach. *Acta Radiologica*, 48(5):550–556, 2007. PMID: 17520432.

[11] Denis Peruzzo, Alessandra Bertoldo, Francesca Zanderigo, and Claudio Cobelli. Automatic selection of arterial input function on dynamic contrast-enhanced mr images. *Computer Methods and Programs in Biomedicine*, 104(3):e148–e157, 2011.

[12] Suk Jae Kim, Jeong Pyo Son, Sookyung Ryoo, Mi-Ji Lee, Jihoon Cha, Keon Ha Kim, Gyeong-Moon Kim, Chin-Sang Chung, Kwang Ho Lee, Pyoung Jeon, and Oh Young Bang. A novel magnetic resonance imaging approach to collateral flow imaging in ischemic stroke. *Annals of Neurology*, 76(3):356–369, 2014.

[13] Johannes Kaesmacher, Panagiotis Chaloulos-Iakovidis, Leonidas Panos, Pasquale Mordasini, Patrik Michel, Steven D. Hajdu, Marc Ribo, Manuel Requena, Christian Maegerlein, Benjamin Friedrich, Vincent Costalat, Amel Benali, Laurent Pierot, Matthias Gawlitza, Joanna Schaafsma, Vitor Mendes Pereira, Jan Gralla, and Urs Fischer. Mechanical thrombectomy in ischemic stroke patients with alberta stroke program early computed tomography score 0–5. *Stroke*, 50(4):880–888, 2019.

[14] Nida M. Zaitoun and Musbah J. Aqel. Survey on image segmentation techniques. *Procedia Computer Science*, 65:797–806, 2015. International Conference on Communications, management, and Information technology (ICCMIT'2015).

[15] Michael Roberts, Ke Chen, and Klaus L. Irion. A convex geodesic selective model for image segmentation. *Journal of Mathematical Imaging and Vision*, 61(4):482–503, May 2019.

[16] Mohamed T. Bennai, Zahia Guessoum, Smaine Mazouzi, Stéphane Cormier, and Mohamed Mezghiche. Multi-agent medical image segmentation: A survey. *Computer Methods and Programs in Biomedicine*, 232:107444, 2023.

[17] E. S. Brown, T. Chan, and X. Bresson. Completely convex formulation of the chan-vese image segmentation model. *International Journal of Computer Vision*, 98:103–121, 2011.

[18] Liam Burrows, Ke Chen, and Francesco Torella. Selective segmentation of a feature that has two distinct intensities. *Journal of Algorithms & Computational Technology*, 15:17483026211007776, 2021.

[19] Laurent Condat. Discrete total variation: New definition and minimization. *SIAM Journal on Imaging Sciences*, 10(3):1258–1290, 2017.

[20] N. Badshah and Ke Chen. Image selective segmentation under geometrical constraints using an active contour approach. *Communications in Computational Physics*, 7:759–778, 2009.

[21] Noor Badshah, Ke Chen, Haider Ali, and Ghulam Murtaza. Coefficient of variation based image selective segmentation model using active contours. *East Asian Journal on Applied Mathematics*, 2(2):150–169, 2012.

[22] Stanley Osher and James A Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.

[23] Chunxiao Liu, Michael Kwok-Po Ng, and Tieyong Zeng. Weighted variational model for selective image segmentation with application to medical images. *Pattern Recognition*, 76:367–379, 2018.

[24] Jack A. Spencer and Ke Chen. A convex and selective variational model for image segmentation. *Communications in Mathematical Sciences*, 13:1453–1472, 2015.

[25] Jacques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, page 49–52, 1902.

[26] A. N. Tikhonov and V. IA. Arsenin. *Solutions of ill-posed problems / Andrey N. Tikhonov and Vasiliy Y. Arsenin ; translation editor, Fritz John.* Winston ; distributed solely by Halsted Press Washington : New York, 1977.

[27] V. Klema and A. Laub. The singular value decomposition: Its computation and some applications. *IEEE Transactions on Automatic Control*, 25(2):164–176, 1980.

[28] Leonid Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60:259–268, 11 1992.

[29] Yousef Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, second edition, 2003.

[30] Sandip Mazumder. Chapter 3 - solution to a system of linear algebraic equations. In Sandip Mazumder, editor, *Numerical Methods for Partial Differential Equations*, pages 103–167. Academic Press, 2016.

[31] Ranjana Dwivedi and Vinay Kumar Srivastava. 12 - fundamental optimization methods for machine learning. In Tilottama Goswami and G.R. Sinha, editors, *Statistical Modeling in Machine Learning*, pages 227–247. Academic Press, 2023.

[32] J. Weickert, B.M.T.H. Romeny, and M.A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410, 1998.

[33] Tom Goldstein, Ernie Esser, and Richard Baraniuk. Adaptive primal-dual hybrid gradient methods for saddle-point problems. 05 2013.

[34] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. 2011.

[35] Wei Deng and Wotao Yin. On the global and linear convergence of the generalized alternating direction method of multipliers. *Journal of Science and Computation*, 66(3):889–916, mar 2016.

[36] Daniel O'Connor and Lieven Vandenberghe. On the equivalence of the primal-dual hybrid gradient method and douglas–rachford splitting. *Mathematical Programming*, 179:1–24, 08 2018.

[37] David Mumford and Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.

[38] T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[39] Christian Gout, Carole Guyader, and Luminita Vese. Segmentation under geometrical conditions using geodesic active contours and interpolation using level set methods. *Numerical Algorithms*, 39:155–173, 01 2005.

[40] Xiangrong Wang and Jieyu Zhao. Image segmentation using improved potts model. In *2008 Fourth International Conference on Natural Computation*, volume 7, pages 352–356, 2008.

[41] Ke Wei, Ke Yin, Xue-Cheng Tai, and Tony Chan. New region force for variational models in image segmentation and high dimensional data clustering. *Annals of Mathematical Sciences and Applications*, 3, 04 2017.

[42] Andreas Weinmann and Martin Storath. Iterative Potts and Blake–Zisserman minimization for the recovery of functions with discontinuities from indirect measurements. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2176):20140638, 2015.

[43] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proceedings of IEEE International Conference on Computer Vision*, pages 694–699, 1995.

[44] Muthukrishnan R. Edge detection techniques for image segmentation. *International journal of computer science and information technology*, 3:259–267, 12 2011.

[45] Lavdie Rada. A new variational model with dual level set functions for selective segmentation. *Communications in Computational Physics*, 12, 07 2012.

[46] Jack A. Spencer and Ke Chen. A convex and selective variational model for image segmentation. *Communications in Mathematical Sciences*, 13:1453–1472, 2015.

[47] Luigi Ambrosio, Nicola Fusco, and John E. Hutchinson. Higher integrability of the gradient and dimension of the singular set for minimisers of the Mumford–Shah functional. *Calculus of Variations and Partial Differential Equations*, 16:187–215, 2003.

[48] Laurent Condat and Peter Richtárik. Randprox: Primal-dual optimization algorithms with randomized proximal updates, 2022.

[49] Xiaohao Cai, Raymond Chan, and Tieyong Zeng. Image segmentation by convex approximation of the Mumford-Shah model. *UCLA CAM Report*, 01 2012.

[50] Mila Nikolova, Selim Esedoglu, and Tony Chan. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal of Applied Mathematics*, 66:1632–1648, 01 2006.

[51] Abdul K. Jumaat Ke Chen. A reformulated convex and selective variational image segmentation model and its fast multilevel algorithm. *Numerical Mathematics: Theory, Methods and Applications*, 12(2):403–437, 2018.

[52] Lavdie Rada and Ke Chen. Improved selective segmentation model using one level-set. *Journal of Algorithms & Computational Technology*, 7(4):509–540, 2013.

[53] J. Weickert, B.M.T.H. Romeny, and M.A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410, 1998.

[54] Fernando Calamante. Arterial input function in perfusion MRI: A comprehensive review. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 74:1–32, 2013.

[55] L Ostergaard, AG Sorensen, KK Kwong, RM Weisskoff, C Gyldensted, and BR Rosen. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. part ii: Experimental comparison and preliminary results. *Magnetic resonance in medicine*, 36(5):726—736, November 1996.

[56] Fernando Calamante, David G. Gadian, and Alan Connelly. Delay and dispersion effects in dynamic susceptibility contrast MRI: Simulations using singular value decomposition. *Magnetic Resonance in Medicine*, 44(3):466–473, 2000.

[57] Matus Straka, Gregory W. Albers, and Roland Bammer. Real-time diffusion-perfusion mismatch analysis in acute stroke. *Journal of Magnetic Resonance Imaging*, 32(5):1024–1037, 2010.

[58] Shengyu Fan, Yueyan Bian, Erling Wang, Yan Kang, Danny J. J. Wang, Qi Yang, and Xunming Ji. An automatic estimation of arterial input function based on multi-stream 3d cnn. *Frontiers in Neuroinformatics*, 13:49, 2019.

[59] Nils Daniel Forkert, Jens Fiehler, Thorsten Ries, Till Illies, Dietmar Möller, Heinz Handels, and Dennis Säring. Reference-based linear curve fitting for bolus arrival time estimation in 4d mra and mr perfusion-weighted image sequences. *Magnetic Resonance in Medicine*, 65(1):289–294, 2011.

[60] Denis Peruzzo, Alessandra Bertoldo, Francesca Zanderigo, and Claudio Cobelli. Automatic selection of arterial input function on dynamic contrast-enhanced mr images. *Computer Methods and Programs in Biomedicine*, 104(3):e148–e157, 2011.

[61] Matus Straka, Gregory Albers, and Roland Bammer. Real-time diffusion-perfusion mismatch analysis in acute stroke. *Journal of magnetic resonance imaging : JMRI*, 32:1024–37, 11 2010.

[62] Kim Mouridsen, Søren Christensen, Louise Gyldensted, and Leif Østergaard. Automatic selection of arterial input function using cluster analysis. *Magnetic Resonance in Medicine*, 55(3):524–531, 2006.

[63] A. Gregory Sorensen, William A. Copen, Leif Østergaard, Ferdinando S. Buonanno, R. Gilberto Gonzalez, Guy Rordorf, Bruce R. Rosen, Lee H. Schwamm, Robert M. Weisskoff, and Walter J. Koroshetz. Hyperacute stroke: Simultaneous measurement of relative cerebral blood volume, relative cerebral blood flow, and mean tissue transit time. *Radiology*, 210(2):519–527, 1999. PMID: 10207439.

[64] Jessy J. Mouannes-Srour, Wanyong Shin, Sameer A. Ansari, Michael C. Hurley, Parmede Vakil, Bernard R. Bendok, John L. Lee, Colin P. Derdeyn, and Timothy J. Carroll. Correction for arterial-tissue delay and dispersion in absolute quantitative cerebral perfusion dsc mr imaging. *Magnetic Resonance in Medicine*, 68(2):495–506, August 2012.

[65] Cory Lorenz, Thomas Benner, Poe Jou Chen, Chloe Joan Lopez, Hakan Ay, Ming Wang Zhu, Nina M. Menezes, Hannu Aronen, Jari Karonen, Yawu Liu, Juho Nuutinen, and A. Gregory Sorensen. Automated perfusion-weighted mri using localized arterial input functions. *Journal of Magnetic Resonance Imaging*, 24(5):1133–1139, 2006.

[66] Laurent Condat. Discrete total variation: New definition and minimization. *SIAM Journal on Imaging Sciences*, 10(3):1258–1290, 2017.

[67] Michael Roberts, Ke Chen, and Klaus L. Irion. A convex geodesic selective model for image segmentation. *Journal of Mathematical Imaging and Vision*, 61(4):482–503, May 2019.

[68] T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[69] N. Badshah and Ke Chen. Image selective segmentation under geometrical constraints using an active contour approach. *Communications in Computational Physics*, 7:759–778, 2009.

[70] E. S. Brown, T. Chan, and X. Bresson. Completely convex formulation of the chan-vese image segmentation model. *International Journal of Computer Vision*, 98:103–121, 2011.

[71] Meng Tang, Tian Chen, Xiao Zhang, and Xiao Huang. Gre t2* -weighted mri: Principles and clinical applications. *BioMed research international*, 2014:312142, 04 2014.

[72] Fernando Calamante, Morten Mørup, and Lars Kai Hansen. Defining a local arterial input function for perfusion mri using independent component analysis. *Magnetic Resonance in Medicine*, 52(4):789–797, 2004.

[73] Sukhdeep Singh Bal, Fan Pei Gloria Yang, Yueh-Feng Sung, Ke Chen, Jiu-Haw Yin, and Giia-Sheun Peng. Optimal scaling approaches for perfusion mri with distorted arterial input function (aif) in patients with ischemic stroke. *Brain Sciences*, 12(1), 2022.

[74] Ona Wu, Leif Østergaard, Walter J. Koroshetz, Lee H. Schwamm, Joanie O'Donnell, Pamela W. Schaefer, Bruce R. Rosen, Robert M. Weisskoff, and A. Gregory Sorensen. Effects of tracer arrival time on flow estimates in mr perfusion-weighted imaging. *Magnetic Resonance in Medicine*, 50(4):856–864, 2003.

[75] M.R. Smith, H. Lu, S. Trochet, and R. Frayne. Removing the effect of svd algorithmic artifacts present in quantitative mr perfusion studies. *Magnetic Resonance in Medicine*, 51(3):631–634, 2004.

[76] A reformulated convex and selective variational image segmentation model and its fast multilevel algorithm. *Numerical Mathematics: Theory, Methods and Applications*, 12(2):403–437, 2018.

[77] Anthony Winder, Christopher D. d'Esterre, Bijoy K. Menon, Jens Fiehler, and Nils D. Forkert. Automatic arterial input function selection in ct and mr perfusion datasets using deep convolutional neural networks. *Medical Physics*, 47(9):4199–4211, 2020.

[78] U. Arioz, K.K. Oguz, U. Baysal, and A. Cila. Multislice brain mapping and quantification of perfusion mri data. In *Conference Proceedings. 2nd International IEEE EMBS Conference on Neural Engineering, 2005.*, pages 9–12, 2005.

[79] Pedro Ramos-Cabrer, Francisco Campos, Tomás Sobrino, and José Castillo. Targeting the ischemic penumbra. *Stroke*, 42(1_suppl_1):S7–S11, 2011.

[80] M Kasam. Su-e-e-05: The role of conventional mr physics, resting state functional mr physics and mr spectroscopy for the in-situ monitoring of infarct and peri-infarct tissue and for the precise diagnosis of acute ischemic stroke. *Medical Physics*, 40(6Part4):118–118, 2013.

[81] Jean J. Chen, Michael R. Smith, and Richard Frayne. The impact of partial-volume effects in dynamic susceptibility contrast magnetic resonance perfusion imaging. *Journal of Magnetic Resonance Imaging*, 22(3):390–399, 2005.

[82] Bruce C.V. Campbell, Søren Christensen, Christopher R. Levi, Patricia M. Desmond, Geoffrey A. Donnan, Stephen M. Davis, and Mark W. Parsons. Comparison of computed tomography perfusion and magnetic resonance imaging perfusion-diffusion mismatch in ischemic stroke. *Stroke*, 43(10):2648–2653, 2012.

[83] Fernando Calamante, David G. Gadian, and Alan Connelly. Delay and dispersion effects in dynamic susceptibility contrast mri: Simulations using singular value decomposition. *Magnetic Resonance in Medicine*, 44(3):466–473, 2000.

[84] Jean J. Chen, Michael R. Smith, and Richard Frayne. The impact of partial-volume effects in dynamic susceptibility contrast magnetic resonance perfusion imaging. *Journal of Magnetic Resonance Imaging*, 22(3):390–399.

[85] Adam E. Hansen, Henrik Pedersen, Egill Rostrup, and Henrik B.W. Larsson. Partial volume effect (pve) on the arterial input function (aif) in t1-weighted perfusion imaging and limitations of the multiplicative rescaling approach. *Magnetic Resonance in Medicine*, 62(4):1055–1059, 2009.

[86] André Ahlgren, Ronnie Wirestam, Emelie Lind, Freddy Ståhlberg, and Linda Knutsson. A linear mixed perfusion model for tissue partial volume correction of perfusion estimates in dynamic susceptibility contrast mri: Impact on absolute quantification, repeatability, and agreement with pseudo-continuous arterial spin labeling. *Magnetic Resonance in Medicine*, 77(6):2203–2214, 2017.

[87] Linda Knutsson, Emelie Lindgren, André Ahlgren, Matthias J.P. van Osch, Karin Markenroth Bloch, Yulia Surova, Freddy Ståhlberg, Danielle van Westen, and Ronnie Wirestam. Reduction of arterial partial volume effects for improved absolute quantification of dsc-mri perfusion estimates: Comparison between tail scaling and prebolus administration. *Journal of Magnetic Resonance Imaging*, 41(4):903–908, 2015.

[88] Hiroshi Matsuda, Takashi Ohnishi, Takashi Asada, Zhi-jie Li, Hidekazu Kanetaka, Etsuko Imabayashi, Fumiko Tanaka, and Seigo Nakano. Correction for partial-volume effects on brain perfusion spect in healthy men. *Journal of Nuclear Medicine*, 44(8):1243–1252, 2003.

[89] Matus Straka, Gregory W. Albers, and Roland Bammer. Real-time diffusion-perfusion mismatch analysis in acute stroke. *Journal of Magnetic Resonance Imaging*, 32(5):1024–1037, 2010.

[90] Nils Daniel Forkert, Jens Fiehler, Thorsten Ries, Till Illies, Dietmar Möller, Heinz Handels, and Dennis Säring. Reference-based linear curve fitting for bolus arrival time estimation in 4d mra and mr perfusion-weighted image sequences. *Magnetic Resonance in Medicine*, 65(1):289–294, 2011.

[91] V.G. Kiselev. On the theoretical basis of perfusion measurements by dynamic susceptibility contrast mri. *Magnetic Resonance in Medicine*, 46(6):1113–1122, 2001.

[92] K A Rempp, G Brix, F Wenz, C R Becker, F Gückel, and W J Lorenz. Quantification of regional cerebral blood flow and volume with dynamic susceptibility contrast-enhanced mr imaging. *Radiology*, 193(3):637–641, 1994. PMID: 7972800.

[93] Ona Wu, Leif Østergaard, Robert M. Weisskoff, Thomas Benner, Bruce R. Rosen, and A. Gregory Sorensen. Tracer arrival timing-insensitive technique for estimating flow in mr perfusion-weighted imaging using singular value decomposition with a block-circulant deconvolution matrix. *Magnetic Resonance in Medicine*, 50(1):164–174, 2003.

[94] Linda Knutsson, Freddy Ståhlberg, and Ronnie Wirestam. Absolute quantification of perfusion using dynamic susceptibility contrast mri: pitfalls and possibilities. *Magma (New York, N.Y.)*, 23(1):1—21, February 2010.

[95] Birgitte F. Kjølby, Irene K. Mikkelsen, Michael Pedersen, Leif Østergaard, and Valerij G. Kiselev. Analysis of partial volume effects on arterial input functions using gradient echo: A simulation study. *Magnetic Resonance in Medicine*, 61(6):1300–1309, 2009.

[96] Guillaume Duhamel, Gottfried Schlaug, and David C. Alsop. Measurement of arterial input functions for dynamic susceptibility contrast magnetic resonance imaging using echoplanar images: Comparison of physical simulations with in vivo results. *Magnetic Resonance in Medicine*, 55(3):514–523, 2006.

[97] Kenya Murase, Keiichi Kikuchi, Hitoshi Miki, Teruhiko Shimizu, and Junpei Ikezoe. Determination of arterial input function using fuzzy clustering for quantification of cerebral blood flow with dynamic susceptibility contrast-enhanced mr imaging. *Journal of Magnetic Resonance Imaging*, 13(5):797–806, 2001.

[98] Matthias J.P. van Osch, Evert-jan P.A. Vonken, Chris J.G. Bakker, and Max A. Viergever. Correcting partial volume artifacts of the arterial input function in quantitative cerebral perfusion mri. *Magnetic Resonance in Medicine*, 45(3):477–485, 2001.

[99] Anthony Winder, Christopher D. d'Esterre, Bijoy K. Menon, Jens Fiehler, and Nils D. Forkert. Automatic arterial input function selection in ct and mr perfusion datasets using deep convolutional neural networks. *Medical Physics*, 47(9):4199–4211, 2020.

[100] Matthias J.P. van Osch, Jeroen van der Grond, and Chris J.G. Bakker. Partial volume effects on arterial input functions: Shape and amplitude distortions and their correction. *Journal of Magnetic Resonance Imaging*, 22(6):704–709, 2005.

[101] I.C. van der Schaaf, Evertjan Vonken, Annet Waaijer, Birgitta K. Velthuis, Marcel J. Quist, and Thijs L J van Osch. Influence of partial volume on venous output and arterial input function. *AJNR. American journal of neuroradiology*, 27 1:46–50, 2006.

[102] Angelos Konstas, Gregory Goldmakher, and Micheal Lev. Theoretic basis and technical implementations of ct perfusion in acute ischemic stroke, part 2: Technical implementations. *American Journal of Neuroradiology*, 30(5):885–892, 2009.

[103] Jose Bernal, Maria d.C. Valdés-Hernández, Javier Escudero, Anna K. Heye, Eleni Sakka, Paul A. Armitage, Stephen Makin, Rhian M. Touyz, Joanna M. Wardlaw, and Michael J. Thrippleton. A four-dimensional computational model of dynamic contrast-enhanced magnetic resonance imaging measurement of subtle blood-brain barrier leakage. *NeuroImage*, 230:117786, 2021.

[104] N.D. Forkert, P. Kaesemann, A. Treszl, S. Siemonsen, B. Cheng, H. Handels, J. Fiehler, and G. Thomalla. Comparison of 10 ttp and tmax estimation techniques for mr perfusion-diffusion mismatch quantification in acute stroke. *American Journal of Neuroradiology*, 34(9):1697–1703, 2013.

[105] H Tei, S Uchiyama, and T Usui. Clinical-diffusion mismatch defined by nihss and aspects in non-lacunar anterior circulation infarction. *Journal of neurology*, 254(3):340—346, March 2007.

[106] Zhe Cheng, Xiaokun Geng, Gary B Rajah, Jie Gao, Linlin Ma, Fenghai Li, Huishan Du, and Yuchuan Ding. Nihss consciousness score combined with aspects is a favorable predictor of functional outcome post endovascular recanalization in stroke patients. *Aging and disease*, 12(2):415—424, April 2021.

[107] Harold Adams, Patricia Davis, Enrique Leira, Kevin Chang, and Birgitte Bendixen. Baseline nih stroke scale score strongly predicts outcome after stroke: A report of the trial of org 10172 in acute stroke treatment (toast). *Neurology*, 53(1):126—131, July 1999.

[108] Patrick D. Lyden. Using the national institutes of health stroke scale: A cautionary tale. *Stroke*, 48:513–519, 2017.

[109] Maxim Mokin, Christopher T. Primiani, Adnan H. Siddiqui, and A. Turk. Aspects (alberta stroke program early ct score) measurement using hounsfield unit values when selecting patients for stroke thrombectomy. *Stroke*, 48:1574–1579, 2017.

[110] William J. Powers, Colin P. Derdeyn, José Biller, Christopher S. Coffey, Brian L. Hoh, Edward C. Jauch, Karen C. Johnston, S. Claiborne Johnston, Alexander A. Khalessi, Chelsea S. Kidwell, James F. Meschia, Bruce Ovbiagele, and Dileep R. Yavagal. 2015 american heart association/american stroke association focused update of the 2013 guidelines for the early management of patients with acute ischemic stroke regarding endovascular treatment. *Stroke*, 46(10):3020–3035, 2015.

[111] Jiaming Liu, Yu Sun, Xiaojian Xu, and Ulugbek S. Kamilov. Image restoration using total variation regularized deep image prior. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7715–7719, 2019.

[112] Huiyi Hu, Justin Sunu, and Andrea Bertozzi. Multi-class graph mumford-shah model for plume detection using the mbo scheme. volume 8932, pages 209–222, 01 2015.

[113] Antonio Criminisi, Toby Sharp, and Andrew Blake. Geos: Geodesic image segmentation. volume 5302, pages 99–112, 10 2008.

[114] Liu XiangYang, Yong YuJie, and Zhang XiaoFeng. Image segmentation based on geodesic distance combined with region and edge gradient. In *2020 4th International Conference on Computer Science and Artificial Intelligence*, CSAI 2020, page 91–96, New York, NY, USA, 2021. Association for Computing Machinery.

[115] Xue Bai and Guillermo Sapiro. Geodesic matting: A framework for fast interactive image and video segmentation and matting (preprint). *International Journal of Computer Vision*, 82:113–132, 04 2009.

[116] Alexis Protiere and Guillermo Sapiro. Interactive image segmentation via adaptive weighted distances. *IEEE Transactions on Image Processing*, 16(4):1046–1057, 2007.

[117] Jing Yuan, Egil Bae, and Xuecheng Tai. A study on continuous max-flow and min-cut approaches. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2217–2224, 2010.

[118] Ivar Ekeland and Roger Témam. *Convex Analysis and Variational Problems*. Society for Industrial and Applied Mathematics, 1999.

[119] Ernie Esser, Xiaoqun Zhang, and Tony F. Chan. A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. *SIAM Journal on Imaging Sciences*, 3(4):1015–1046, 2010.

[120] Xavier Bresson, Thomas Laurent, David Uminsky, and James H. von Brecht. Multiclass total variation clustering, 2013.

[121] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40:120–145, 2011.

[122] Mingqiang Zhu and Tony Chan. An efficient primal-dual hybrid gradient algorithm for total variation image restoration. *UCLA CAM Report*, 05 2008.

[123] Silvia Bonettini and Valeria Ruggiero. On the convergence of primal–dual hybrid gradient algorithms for total variation image restoration. *Journal of Mathematical Imaging and Vision*, 44:236–253, 11 2012.

[124] Xiaohao Cai, Raymond Chan, and Tieyong Zeng. A two-stage image segmentation method using a convex variant of the mumford–shah model and thresholding. *SIAM Journal on Imaging Sciences*, 6(1):368–390, 2013.

[125] Jonathan M. Borwein and Jon D. Vanderwerff. *Convex Functions: Constructions, Characterizations and Counterexamples*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010.

[126] Dan Luo, Wei Zeng, Jinlong Chen, and Wei Tang. Deep learning for automatic image segmentation in stomatology and its clinical application. *Frontiers in Medical Technology*, 3, 2021.

[127] Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. Image segmentation using deep learning: A survey, 2020.

[128] Liam Burrows, Ke Chen, and Francesco Torella. Using deep image prior to assist variational selective segmentation deep learning algorithms. *Seventeenth International Symposium on Medical Information Processing and Analysis*, page 34, 2021.

[129] Liam Burrows, Ke Chen, and Francesco Torella. A deep image prior learning algorithm for joint selective segmentation and registration. *Variational Methods in Computer Vision*, pages 411–422, 04 2021.

[130] Liam Burrows, Theljani Anis, and Ke Chen. On a variational and convex model of the blake-zisserman type for segmentation of low-contrast and piecewise smooth images. *Journal of Imaging*, 10 2021.

[131] Bjoern H. Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, Levente Lanczi, Elizabeth Gerstner, Marc-André Weber, Tal Arbel, Brian B. Avants, Nicholas Ayache, Patricia Buendia, D. Louis Collins, Nicolas Cordier, Jason J. Corso, Antonio Criminisi, Tilak Das, Hervé Delingette, Çağatay Demiralp, Christopher R.

Durst, Michel Dojat, Senan Doyle, Joana Festa, Florence Forbes, Ezequiel Geremia, Ben Glocker, Polina Golland, Xiaotao Guo, Andac Hamamci, Khan M. Iftekharuddin, Raj Jena, Nigel M. John, Ender Konukoglu, Danial Lashkari, José António Mariz, Raphael Meier, Sérgio Pereira, Doina Precup, Stephen J. Price, Tammy Riklin Raviv, Syed M. S. Reza, Michael Ryan, Duygu Sarikaya, Lawrence Schwartz, Hoo-Chang Shin, Jamie Shotton, Carlos A. Silva, Nuno Sousa, Nagesh K. Subbanna, Gabor Szekely, Thomas J. Taylor, Owen M. Thomas, Nicholas J. Tustison, Gozde Unal, Flor Vasseur, Max Wintermark, Dong Hye Ye, Liang Zhao, Binsheng Zhao, Darko Zikic, Marcel Prastawa, Mauricio Reyes, and Koen Van Leemput. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024, 2015.

[132] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin Kirby, John Freymann, Keyvan Farahani, and Christos Davatzikos. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific Data*, 4, 09 2017.

[133] Patricia Clement, Jan Petr, Mathijs B. J. Dijsselhof, Beatriz Padrela, Maurice Pasternak, Sudipto Dolui, Lina Jarutyte, Nandor Pinter, Luis Hernandez-Garcia, Andrew Jahn, Joost P. A. Kuijer, Frederik Barkhof, Henk J. M. M. Mutsaerts, and Vera C. Keil. A beginner's guide to arterial spin labeling (asl) image processing. *Frontiers in Radiology*, 2, 2022.

[134] Deepasree Jaganmohan, Somnath Pan, Chandrasekharan Kesavadas, and Bejoy Thomas. A pictorial review of brain arterial spin labelling artefacts and their potential remedies in clinical studies. *The Neuroradiology Journal*, 34(3):154–168, 2021. PMID: 33283653.