**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**
POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

# A VECTOR-VALUED MULTIPHASE ACTIVE CONTOUR MODEL FOR IMAGE SEGMENTATION

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS DE LA COMPUTACIÓN

PRESENTA:
ERIK YIDELL CARBAJAL DEGANTE

TUTOR:
DR. BORIS ESCALANTE RAMÍREZ
FACULTAD DE INGENIERÍA

COMITÉ TUTORAL:
DR. EDGAR GARDUÑO ÁNGELES
DR. FERNANDO ARÁMBULA COSÍO
INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y SISTEMAS

CIUDAD UNIVERSITARIA, CD.MX. JULIO 2021

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**
POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

# A VECTOR-VALUED MULTIPHASE ACTIVE CONTOUR MODEL FOR IMAGE SEGMENTATION

A DISSERTATION SUBMITTED
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF COMPUTER SCIENCE

PRESENTS:
ERIK YIDELL CARBAJAL DEGANTE

ADVISOR:
DR. BORIS ESCALANTE RAMÍREZ
FACULTAD DE INGENIERÍA

SUPERVISORY COMMITTEE:
DR. EDGAR GARDUÑO ÁNGELES
DR. FERNANDO ARÁMBULA COSÍO
INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS
APLICADAS Y SISTEMAS

CIUDAD UNIVERSITARIA, CD.MX. JULY 2021

# *Agradecimientos y dedicatoria*

*Por y para Dios.*

*A mi madre, mi tía, mi hermana, Lunita y desde luego al resto de mi familia.*

*A los que me han acompañado y considero también como parte de mi familia.*

*A mis grandes amistades y compañeros.*

*A mi tutor, los miembros del comité y del jurado.*

*A la prestigiosa UNAM y al Conacyt.*

*- V1DE -*

# Resumen

La segmentación de imágenes es básicamente el punto de partida para el desarrollo de nuevas aplicaciones basadas en visión por computadora, tales como procesamiento de video, manejo de vehículos autónomos, detección de objetos, diagnóstico médico, animación, entre otros. Además, la inteligencia artificial es crucial a medida que contribuye con las bases de algoritmos más sofisticados mientras que datos masivos son aprovechados. Por otro lado, los métodos variacionales y sus raíces son considerados aún como una sólida alternativa que han sido eficientemente probados desde que surgieron.

Los contornos activos (AC) se han convertido en una herramienta popular dentro de los métodos variacionales. Básicamente, un conjunto de contornos activos móviles es utilizado para comparar diferentes regiones de una imagen con el fin de alcanzar un balance de regularidad, lo que nos permite particionar regiones de manera semiautomática y generar distintas escenas. Un funcional dado, relacionado al balance de intensidades, debe ser resuelto con el fin de obtener la varianza mínima entre todas las regiones resultantes. Esta solución permite a los AC ser modelados como un problema de optimización.

El trabajo de esta tesis se encarga de estudiar la base de los métodos variacionales para abordar el problema de segmentación multiclase. Nuestro objetivo es proporcionar diferentes versiones de un nuevo modelo de AC y abordar problemas prácticos en tareas de segmentación como degradaciones, inhomogeneidades de la imagen e incluso proponer una alternativa para los problemas relacionados con bases de datos limitadas, las cuales son útiles en la etapa de entrenamiento para algoritmos de aprendizaje. El estudio propuesto también destaca la importancia de las características de textura como buenos descriptores en tareas de clasificación o segmentación con aplicaciones potenciales en tecnologías de visión por computadora.

El trabajo actual también incluye un estudio de cómo superar las limitaciones que surgen de la naturaleza de los AC al proponer nuevos modelos híbridos. Por ejemplo, el conocimiento previo sobre un objeto de interés como la forma y la ubicación, son a menudo requisitos para

un proceso posterior de segmentación. Esto hace que la interacción del usuario juegue un papel clave en la influencia para discernir entre los métodos semiautomáticos de los totalmente automatizados. Estos últimos permiten a los ordenadores gestionar técnicas de segmentación de extremo a extremo (del inglés end-to-end), muy solicitadas en la actualidad.

Finalmente, el modelo propuesto es evaluado para la segmentación de múltiples regiones el cual revela cómo tratar con imágenes multicanal. Probamos nuestros métodos con datos públicos, privados y sintéticos, mientras que las imágenes médicas contribuyen en gran medida a los experimentos. Así mismo, proporcionamos la teoría para que el llamado modelo de AC multifase-vectorial se incluya tanto en las mejoras recientes de los métodos variacionales como en el área de inteligencia artificial. Mostramos la capacidad del método propuesto para combinarse con una alta confiabilidad a medida que los modelos híbridos ganan fuerza con el tiempo.

# Abstract

Image segmentation is basically the starting point in the development of further vision-based applications, i.e. video processing, autonomous driving, object detection, medical diagnosis, photo animation, etc. Moreover, artificial intelligence is crucial as long as it contributes with the basis of more sophisticated algorithms while big data is exploited. On the other hand, variational methods and their roots are still considered as a solid alternative that have been efficiently tested since they emerged.

Active contours (AC) became widely popular in segmentation among the variational methods. Basically, a set of active evolving curves is used to compare different regions trying to reach a balance of regularity which leads to a semiautomatic region partitioning into different scenes. Precisely, a given functional related to this balance must be solved in order to obtain the minimum variance quantity among all the resulting regions. This solution allows AC to be studied as an optimization problem.

The work in this thesis is concerned with variational methods to address the multiclass segmentation problem. Our aim is both to provide different versions of a novel active contour model, and to tackle practical segmentation issues such as degradations, image inhomogeneities or even short datasets available for training purposes. The proposed study also highlights the importance of texture features as good descriptors in segmentation tasks with potential applications in computer vision technologies.

The current work also includes a study of how to overcome the limitations emerged from the nature of AC by proposing new hybrid models. For instance, prior-knowledge information of the object of interest such as shape and location is often required for a subsequent process of image segmentation. This makes the user interaction to play a key role in the influence to distinguish semiautomatic from fully automated methods. The latter allows computers to manage end-to-end segmentation techniques, highly requested nowadays.

Finally, the assessment of the proposed model for multiple region segmentation reveals

how to deal with multichannel images. We tested our methods with public, private and synthetic data, whereas medical images highly contribute in the experiments. We hereby provide the theory for the so-called vector-valued multiphase AC model to be included in both recent improvements of variational methods and artificial intelligence. Furthermore, we show the ability of the proposed models to be combined with high reliability, whereas hybrid models gain strength over time.

# Contents

# List of Figures

# Chapter 1

# Introduction

*We may regard the present state of the universe as the effect of its past and the cause of its future. An intelligence which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intelligence were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom;*
*for such an intelligence nothing would be uncertain and,*
*the future just like the past would be present before its eyes.*
*- Pierre-Simon Laplace, 1814*
*'Laplace's Demon' (Essai philosophique sur les probabilités)*

In the past decade, computer vision technologies have evolved dramatically. Nowadays, several applications involve image processing such as face recognition, object identification, video analysis, medical diagnosis, vision in robots and autonomous vehicles, etc. Many computer vision tasks require an intelligent agent to accurately perform image segmentation and then enable the corresponding analysis of interpretation. Nowadays, such understanding uses powerful complex models based on artificial intelligence (AI) to figure out at unimaginable levels what a single pixel represents exactly in our reality. This fact drives algorithms to reach perfection in the assessment of certain tasks. Which lead us to wonder: How big is the current gap between how AI interprets and how we humans do? Figure 1.1 illustrates how AI and computer vision technologies overlap somehow.

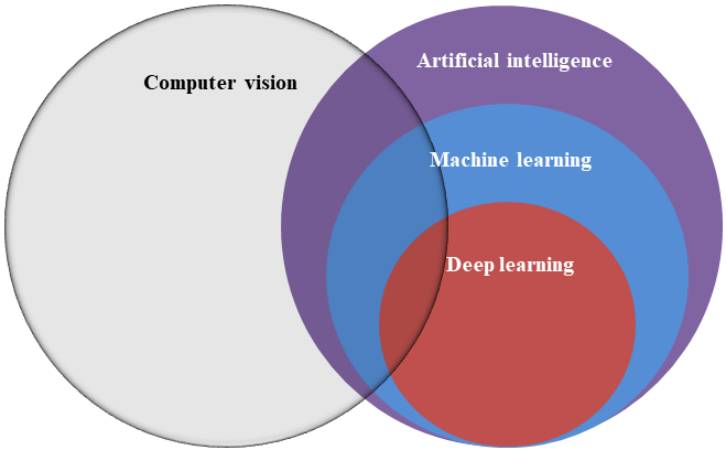



Figure 1.1: Relation between computer vision and artificial intelligence

## 1.1 Image segmentation

The problem of image segmentation (IS) dates back to the beginning of computer vision. In brief, segmentation consists of partitioning a given image into a collection of regions, or finding edges that delimit those regions for a subsequent characterization in meaningful scenes [Petrou, 2010]. Since the image acquisition nature provides an infinite number of possible scenes, several algorithms have been proposed in the literature to provide with an accurate identification of an object of interest. Features such as pixel intensity, color, texture or edges have been the most used characteristics to solve the segmentation challenge.

In fact, segmentation can be seen as a particular type of a pixel-wise classification problem, that can be split into two separate yet equally important groups:

- Semantic segmentation — Classification of all the pixels of an image into meaningful classes. Such classes are semantically interpretable and correspond to real-world categories.
- Instance segmentation — It is able to identify each object instance in an image. The main difference from semantic segmentation is that it can not categorize every pixel.

### 1.1.1 Approaches and techniques

In the design of IS algorithms, there are various techniques based on the image type to be processed and analyzed. They can be classified mainly into three broader categories as below

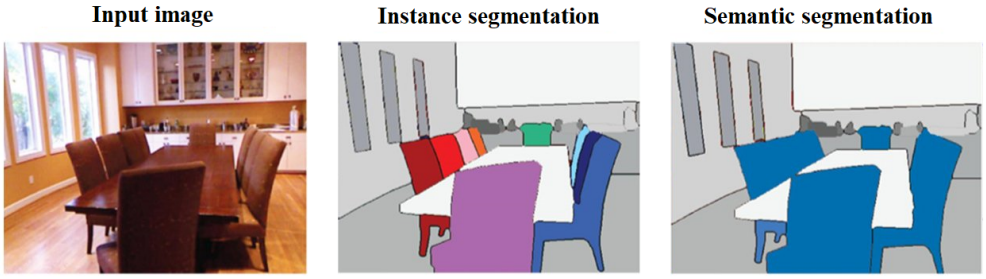



Figure 1.2: Segmentation task viewed as pixel-wise classification process to identify isolated objects or multiple classes.

[Gonzalez and Woods, 2018]:

1. Structural — The structural segmentation approach mainly focuses on a particular region to segment. This set of algorithms makes use of information such as density, distributions, histograms, color distribution, etc. Structural methods perform firstly the identification of the target area, which is linked to a specific segmentation problem.

2. Stochastic — In the stochastic segmentation approach, the primary information required is the pixel values rather than the structure of certain portion in the image. Those methods included in this group are highly effective in the case of analyzing large image datasets, where high uncertainty is faced in terms of the required region to be segmented.

3. Hybrid — As the name suggests, hybrid segmentation methods integrate both structural and stochastic approaches. Region-based and pixel-based information are required to boost the strength and augments accuracy of the segmentation process jointly.

According to the aforementioned IS approaches, the type of analysis needed produces several methods and the choice of the suitable algorithm depends on the nature of the problem. For instance, noise is an undesired element that is present in most of the applications and its study has became challenging since the origins of signal processing, hence some methods are forced to improve their robustness to deal well with common degradations.

- Thresholding Methods — These methods focus on finding a suitable value (threshold) commonly based on histograms. The resulting map provides with two or more disjoint regions, each region is associated to those pixels located within a threshold range.

- Edge-based Methods — Edge-based segmentation algorithms accomplish edge detection based on locating discontinuities in intensity, colour, texture, brightness, saturation, contrast, etc. To further improve the obtained results, a post-processing step is required to connect all the edges that correspond better with the borders that delimit the objects of interest.

- Region-based Methods — These methods consist of creating segments by dividing the image into various regions of similar characteristics. Hence, two techniques can be employed: Region growing or region splitting/merging.

- Clustering — Methods based on clusters can be considered as unsupervised algorithms unlike classification algorithms, where features, classes, or groups are not well defined. A clustering algorithm identifies groups in the data and assigns each data point to certain group based on similarity and location in the corresponding feature space.

- Watershed Method — In watershed segmentation, the input image is regarded as a topographic surface. The values in the surface are commonly represented by the intensity pixel levels or their gradient magnitude. Some watershed methods can be included as part of region-based methods due to their origins rely on mathematical morphology.

- PDE-based methods — Partial Differential Equation (PDE) models base their theory on variational calculus. The most relevant technique for image segmentation is known as active contours or snakes, that models the segmentation problem by using a PDE. Some famous methods of PDE used for image segmentation are Snakes, Level-Set, and the Mumford shah method.

- Neural Networks — Among the trainable techniques of machine learning, neural networks are supported and powered by modern deep learning technology. Basically, neural networks are a set of mutually connected networks that adjust their own parameters according to training and validation processes. The current increase in computer capabilities such as speed, reliability, adaptability and storage, sets neural networks as one of the most efficient and preferable methods for IS.

Moreover, IS methods may also follow the machine learning classification standpoint, this leads certain algorithms to be labeled as supervised or unsupervised methods. The supervised classification is based on the available data to describe a training set. The training set includes the number of known samples for each class whereas intraclass homogeneity helps to better discriminate among classes. On the other hand, unsupervised classification does not require

previous knowledge of the classes, but the algorithm computes automatically a batch of those similar features within the same class.

## 1.2 Active contours preliminaries

Variational methods for image processing address the problem of IS by determining the solution of an equation corresponding to the minimization of certain functional and finally obtain a PDE [Kass et al., 1988]. Within this scope, active contours possess several advantages like the easy extension to multiple dimensions. Moreover, the functional that controls convergence can be adapted to different region segmentation criteria instead of the traditional gradient-based benchmark.

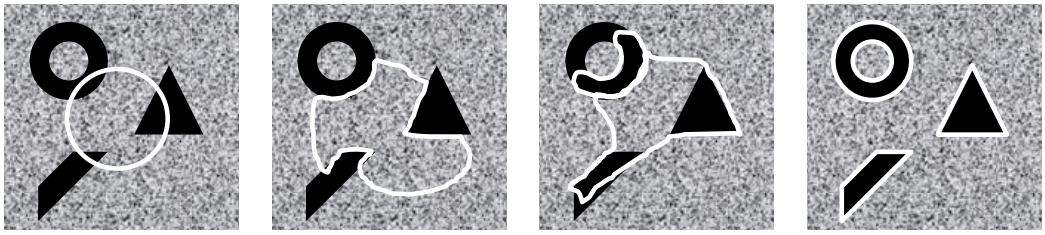


Figure 1.3: Active contour evolving process.

An important contribution to this area includes region-based methods of active contours, which were originally conceived to separate different regions by means of evolving contours related to a specific functional. Several advantages can be quoted since active contours use the theory of level sets by [Osher and Sethian, 1988], adapting the problem to a discrete model so they handle complex topological changes automatically and they can be implemented by efficient stable numerical schemes. Furthermore, active contours do not require a training stage making this tool an unsupervised technique with numerous applications.

### 1.2.1 A review of medical imaging applications

Active contours (AC) were initially conceived to separate two regions. The work of [Chan and Vese, 2001] represents the beginning of an era in the field of computer vision and pattern recognition. In the literature, numerous AC methods have been proposed and have drawn a lot of attention from many research areas being medical imaging one of the most studied applications. The non-invasiveness of image acquisition allows methods based on digital processing techniques to be widely used to diagnosis and ambulatory care purposes. For

instance, Chunming Li et al. introduced a region-based AC model in order to overcome the issues caused by inhomogeneities with applications to magnetic resonance imaging (MRI) segmentation in [Li et al., 2011]. A recent study of [Badshah et al., 2020] reveals the benefits of smoothing homogeneous regions by accentuating boundaries for an appropriate tumor detection in MRI and mammogram images. A similar application developed by [Moreno et al., 2014] proposed an efficient implementation of a binary partitioning function that accurately segments disjoint brain regions in MRI. This became useful for an accurate brain lesions detection as stated in [Nabizadeh and Kubat, 2017], where authors include an efficient texture description achieving satisfactory performance at reasonable computational cost compared to the state-of-the-art methods.

Computed tomography (CT) segmentation with AC models have also appeared on stage. Among the most relevant studies for Lung segmentation, the proposed by [Rebouças Filho et al., 2017] extends the traditional 2D AC technique to handle volumes and assist in the diagnosis of pulmonary disease. By including texture features to the segmentation models demonstrates a high improvement for measuring the cardiac movement in the works of [Barba-J et al., 2017, Olveres et al., 2017].

On the other hand, ultrasound (US) segmentation is still a challenging task for any IS model because of the low quality and contrast provided by US devices. Nonetheless, some AC-based techniques have shown good results when they include a pre-processing stage to deal with speckle noise. The use of local entropy suggested by [jing Zong et al., 2019] improves automatic US segmentation. The method presented by [Keatmanee et al., 2019] includes a refined initialization stage to aid in breast cancer diagnosis.

### 1.2.2 A review of industry applications

Alternative applications of CT image segmentation has been extended to the industry, such is the work presented by [Liu et al., 2013b] where authors include a 3D model of AC that can be widely applied in many areas of non-destructive testing and non-destructive evaluation such as aviation, railway, manufacturing, military industry, satellite, etc. In the case of satellite imagery, remote sensing is a relevant field of image processing, and IS has gained a lot of attention since it has become the main required task for remote analysis. Among the applications to this area, [Liu et al., 2013a] submitted a proposal of an AC algorithm to extract geospatial objects with irregular shape, while [Han and Wu, 2017] used the cross entropy measurement embedded in the classic AC model for river segmentation.

Tracking is based on matching the object appearance between successive frames of a video sequence, [Paragios and Deriche, 2000] suggested to use AC in their work because shape information can be preserved along the data frames, their work achieved good results for detecting and tracking multiple moving objects in data frames. The mathematical morphology of AC allows adapting models to multiple dimensions, in [Yagi et al., 2000] authors proposed road tracking with 3D shape reconstruction for smart vehicles. A similar technique applied to biomedicine was developed by [Zimmer et al., 2002] to videomicroscopy for tracking migrating cells with applications to drug testing, while [Ouabida et al., 2017] performed iris segmentation and tracking in eye motion video datasets.

### 1.2.3 Main drawbacks

Despite the huge range of applications existing in the literature, the certainty level of reaching a local/global minimum and the speed of convergence is one of the main unknowns of AC. The certainty varies according to the model and application which actually depends on how close the initial contour is placed to the object of interest. Thus, a good initialization plays a very important role in the final segmentation result and the number of iterations required. Another important aspect to consider is the fact that user interaction is often required for performing parameter tunning, this process helps guiding the correct evolution of curves and regulate the movement. In this thesis, we present a successful reformulation to address the main limitations of AC while providing an automated model for multiple class segmentation that can handle texture cues.

## 1.3 Thesis outline

Chapter 2 presents a general description of variational methods and techniques in computer vision, which has served as basis of the current work. In Chapter 3 we provide a review of the traditional AC model and its different versions, which have been widely studied in the era of image processing. The proposed model is addressed and extended in theory in Chapter 4 whereas the Chapter 5 provides the results reached in this work and future trends. Finally, Chapter 6 concludes this work while we briefly discuss the ethics implications of the current AI-based systems and scopes.

# Chapter 2

# An overview of calculus of variations

*Everything proceeds mathematically... if someone could have a sufficient insight into the inner parts of things, and in addition had remembrance and intelligence enough to consider all the circumstances and take them into account, he would be a prophet and see the future in the present as in a mirror*
*- Gotfried Leibniz, 1680.*

Analogous to the usual methods of classic calculus, calculus of variations deals with functionals, see diagram in Figure 2.1. Briefly, a functional is defined as a function of another function. Most real world problems can be formulated as optimization problems, i.e., minimization or maximization of functionals. The assumption that all things in nature seek out a lower energy state makes calculus of variations extensively applied in physics. For instance, the problem of finding the minimum energy that a particle reaches under certain condition can be solved with calculus of variations. In image processing, the goal of many applications is to recover an ideal image from a certain corrupted observation, which is known as an inverse problem. Calculus of variations provides in this sense, many advantages for developing image restoration, denoising, inpainting, segmentation, as the most relevant tasks that can be addressed as the optimization problem of finding the best solution from all feasible solutions.

In this chapter, we will cover an overview of calculus of variations and optimization processes following the description provided by [Spencer, 2016], the latter becomes of particular interest to this thesis.

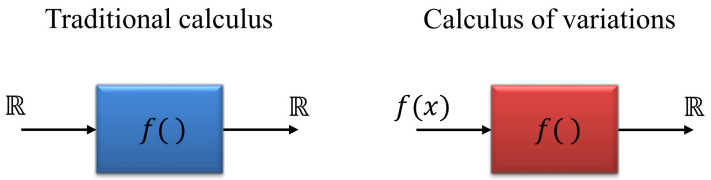



Figure 2.1: Two approaches of calculus.

## 2.1 Inverse problems modeling

As mentioned at the beginning of Chapter 2, an inverse problem begins with data that is the result of a process and the aim is to find the unknown input. Basically, inverse problems can be split in two groups: well-posed and ill-posed. Below are presented three main conditions to define well-posed inverse problems in the sense of Hadamara, [Petrov and Sizikov, 2005]:

1. A solution exists,
2. The solution is unique,
3. The solution depends continuously on the data (stability).

Problems that are not well-posed in the previous sense are considered ill-posed. In the universe of image processing, inverse problems are typically ill-posed due to issues with uniqueness and stability. Such is the case of denoising, consider a noisy image $z$ that contains additive Gaussian noise $\eta$ where the relation $z = I + \eta$ is satisfied. The inverse problem of finding $I$ is called restoration and can be addressed by using variational methods in the sense of minimizing certain functional. Hence, this problem becomes ill-posed due to the solution is not unique.

## 2.2 Variation of a functional

The first variation of a functional can be used to find an unknown function that maximizes or minimizes a functional. Classical solutions to minimization problems in calculus of variations are given by boundary value problems involving certain types of differential equations. Let $J(u) : \Omega \to \Re$ be a functional:

$$J(u) = \int_{\Omega} \mathscr{L}(x, u(x), \nabla u(x))dx\,, \tag{2.1}$$

where $\Omega$ denotes some normed linear space that provides a solution of the unknown function $u$, $\nabla u(x)$ denotes the gradient of $u$ and $dx$ is the $n-$differential element defined as $dx = dx_1, \cdots, dx_n$. We are concerned with the problem of minimizing the functional with respect to $u$:

$$\min_u J(u)\,. \tag{2.2}$$

The most important and necessary condition to be satisfied by any minimizer of a variational integral is the vanishing of its first variation $\delta J(u)$:

$$\delta J(u) = \frac{d}{d\epsilon} J(u + \epsilon\phi)\Big|_{\epsilon=0} = 0\,, \tag{2.3}$$

where $\phi \in \Omega$ is a test function and $\epsilon$ is a real parameter. That is, if $u$ is a minimizer of $J(u)$ with respect to $\delta u = \phi$, then Equation (2.3) must be satisfied for all $\phi$ with compact support in $\Omega$. Then we call $\delta J(u_0)$ the first variation of $J$ at $u_0$ in the direction of $\phi$, for some $u_0 \in \Omega$.

## 2.3 Minimization stage

Minimization problems can be analyzed by calculus of variations to characterize the equilibrium configurations of almost all continuous physical systems, ranging from elasticity, solid and fluid mechanics, electro-magnetism, gravitation, quantum mechanics, string theory, and many others [Olver, 2014]. Certain geometrical configurations, such as minimal surfaces, can be conveniently reformulated as optimization problems. Moreover, numerical approximations to the equilibrium solutions of such boundary value problems are based on a nonlinear finite element approach that reduces the infinite-dimensional minimization problem to a finite-dimensional problem.

### 2.3.1 The Gâteaux derivative

Let $J$ be a function on an open subset $U$ of a Banach space $V$, taking values in a second Banach space $Y$. Then we say that $J : U \to Y$ is Gâteaux differentiable at $u \in U$ in the direction of $\phi \in Y$, if the first directional derivative $J'(u;\phi)$ exists for each test function $\phi \in V$. That is:

$$\delta J(x) = \lim_{\epsilon \to 0} \frac{J(u + \epsilon\phi) - J(u)}{\epsilon} . \tag{2.4}$$

In other words, the Gâteaux derivative of $J(u)$ is just the derivative of $J(u+\epsilon\phi)$ with respect to $\epsilon$ as it tends toward zero.

#### 2.3.1.1 Remarkable definitions

- **Stationary points**. *Let $J : U \to \Re$ be a function with a solution space $U \subset V$. For some $\widetilde{u} \in U$, suppose $J$ is Gâteaux differentiable for all test functions $\phi \in V$. Then $\widetilde{u} \in U$ is said to be a stationary point of $J$ if $\delta J(\widetilde{u}) = 0$ for all $\phi \in V$.*

- **Local minimizer**. *A real-valued functional $J : U \to \Re$ defined in a normed space $V$, is said to have a local minimizer at the point $\widetilde{u}$, if there exists some $\epsilon > 0$ such that:*

  $$J(\widetilde{u}) \leq J(u), \forall u \in B_\epsilon(\widetilde{u}) \bigcap U ,$$

  *with $B_\epsilon(\widetilde{u}) = \{u \in V : ||u - \widetilde{u}|| < \epsilon\}$.*

- **Global minimizer**. The global minimizers of a functional $J(u)$ are obtained by adjusting the inequalities of local minimizer formulas.

  *A real-valued functional $J : U \to \Re$ is said to have a global minimizer at the point $\widetilde{u}$, if:*

  $$J(\widetilde{u}) \leq J(u), \forall u \in U .$$

  The equation $\delta J(u) = 0$ is called the Euler-Lagrange equation of the original minimisation problem in Equation (2.2). If $J(u)$ is a convex functional, and $U$ is a convex set, then every local minimizer of $J(u)$ is also a global minimizer.

### 2.3.2 The Divergence theorem

The divergence theorem is essential to obtain the Euler-Lagrange equations when minimizing a specific functional. Let $F$ be a continuous differentiable vector field in a domain $V \subset \Re^n$. Let $\Omega \subset V$ be a closed bounded region with boundary $\partial\Omega$. The integral of the divergence of $F$ over $\Omega$ and the surface integral of $F$ over the boundary $\partial\Omega$ are then related by:

$$\int_{\Omega}(\nabla \cdot F)dx = \int_{\partial\Omega}(F \cdot \vec{n})ds\,, \tag{2.5}$$

$$\nabla \cdot F = \frac{\partial F}{x_1} + \cdots + \frac{\partial F}{x_n}\,, \tag{2.6}$$

where $dx = dx_1, \cdots, dx_n$ and $ds$ indicate the integration operation with respect to surface area delimited by $\partial\Omega$, and $\vec{n}$ is the normal unit for each point $x \in \partial\Omega$.

#### 2.3.2.1 First variation minimization

Let us consider the problem of finding the first variation of a functional:

$$J(u) = \int_{\Omega}|\nabla u|dx\,, \tag{2.7}$$

defined in the domain $\Omega \subset \Re^2$. Recall that $\epsilon\phi$ consists of the parameter $\epsilon \to 0$ for the continuous differentiable test function $\phi$ in $\Omega$, then we compute:

$$\begin{aligned}\frac{d}{d\epsilon}J(u+\epsilon\phi)\Big|_{\epsilon=0} &= \frac{d}{d\epsilon}\int_{\Omega}|\nabla(u+\epsilon\phi)|dx\Big|_{\epsilon=0}\\ &= \int_{\Omega}\frac{\nabla(u+\epsilon\phi)}{|\nabla(u+\epsilon\phi)|}\cdot\phi dx\Big|_{\epsilon=0}\\ &= \int_{\Omega}\frac{\nabla u}{|\nabla u|}\cdot\nabla\phi dx\,.\end{aligned}$$

By using the relation based on the divergence theorem and integrating by parts:

$$\int_{\Omega}\frac{\nabla u}{|\nabla u|}\cdot\nabla\phi dx = \int_{\partial\Omega}\phi\frac{\nabla u}{|\nabla u|}\cdot\vec{n}ds - \int_{\Omega}\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right)\phi dx\,.$$

We require Equation (2.3) for all test functions $\phi$. This allows us to derive the following partial differential equation, known as Euler-Lagrange equation:

$$\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) = 0\,. \tag{2.8}$$

### 2.3.3 Euler-Lagrange equation for $L^2$

When the functional is a simple integral, the Euler-Lagrange equation gives us a powerful formula for fast calculation of the derivatives. The integral of Equation (2.1) is also known as the Lagrangian in honor of the mathematician Lagrange. We usually assume that the Lagrangian is a reasonable smooth function of all three of its arguments $\mathscr{L}(x, u, u')$.

As stated in the previous sections, the minimizers of an objective function defined on a finite dimensional space are initially characterized as critical points, where the gradient of the objective function vanishes. An analogous construction emerged in the infinite-dimensional context conducted by calculus of variations.

The derivative of the functional $J$ that includes the Lagrangian can be written as follows:

$$h(\epsilon) = J(u + \epsilon\phi) = \int_\Omega \mathscr{L}(x, u + \epsilon\phi, u' + \epsilon\phi')dx\,, \tag{2.9}$$

by assuming sufficient smoothness, it allows us to include the derivative inside the integral and solved by using the chain rule:

$$\begin{aligned} h'(\epsilon) = \frac{d}{d\epsilon}J(u + \epsilon\phi) = \int_\Omega \frac{d}{d\epsilon}\mathscr{L}(x, u + \epsilon\phi, u' + \epsilon\phi')dx\,, \\ = \int_\Omega \left[\phi\frac{\partial\mathscr{L}}{\partial u}(x, u + \epsilon\phi, u' + \epsilon\phi') + \phi'\frac{\partial\mathscr{L}}{\partial u'}(x, u + \epsilon\phi + u' + \epsilon\phi')\right]dx\,. \end{aligned}$$

Therefore, by setting $\epsilon = 0$:

$$\delta J = \int_\Omega \left[\phi\frac{\partial\mathscr{L}}{\partial u}(x, u, u') + \phi'\frac{\partial\mathscr{L}}{\partial u'}(x, u, u')\right]dx\,.$$

Integration by parts is applied to the second term in the previous equation. This leads to:

$$\delta J = \int_\Omega \phi\frac{\partial\mathscr{L}}{\partial u}dx + \left[\ \frac{\partial\mathscr{L}}{\partial u'}\phi'\right|_a^b - \int \frac{d}{dx}\left(\frac{\partial\mathscr{L}}{\partial u'}\right)\phi dx$$

$$\delta J = \int_\Omega \left(\frac{\partial\mathscr{L}}{\partial u} - \frac{d}{dx}\frac{\partial\mathscr{L}}{\partial u'}\right)\phi dx\,. \tag{2.10}$$

Equation (2.10) is the explicit formula for the variational derivative of the functional in Equation (2.1) with the Lagrangian $\mathscr{L}(x, u, u')$. Observe that the gradient $\delta J$ of a functional is a function. [1]

## 2.4 Gradient descent methods

Among the different methods for carrying out minimization of functionals, descent methods appear on stage. Let $F : \Omega \subset \Re^n \to \Re$ be a continuously differentiable function. Descent methods require an initialization $u^{(0)} \in \Re^n$ to be followed by an iteration scheme:

$$u^{(k)} = u^{(k-1)} - \alpha^{(k-1)} s^{(k-1)} \quad , \quad k = 1, 2, \dots , \tag{2.11}$$

where $s^{(k-1)}$ is a search direction and $\alpha^{(k-1)} > 0$ is the step length. This procedure leads to move closer to any apparent solution. The gradient descent is a particular case of the descent method, where the search direction is opposite to the gradient of $F$, that is $\nabla F(u^{(k-1)})$. In this sense, the function $F$ decreases faster in this direction. Basically, the gradient scheme is shown as follows:

$$u^{(k)} = u^{(k-1)} - \alpha^{(k-1)} \nabla F(u^{(k-1)}) \quad , \quad k = 1, 2, \dots . \tag{2.12}$$

The main characteristic of descent methods is that the iteration scheme reduces the value of the function for each $k$:

$$F(u^{(k)}) \leq F(u^{(k-1)}) . \tag{2.13}$$

This condition is satisfied by selecting the step length $\alpha^{(k-1)}$ appropriately. For instance, we refer to *time marching* when the step length is fixed for some time step $\tau$. Time marching is restricted in the sense that the stability is heavily dependent on choosing a small $\tau$, which increases the number of iterations required to converge to a steady state solution, hence $\nabla F(u) = 0$. The explicit time marching scheme is shown as follows:

$$u^{(k)} = u^{(k-1)} - \tau \nabla F(u^{(k-1)}) \quad , \quad k = 1, 2, \dots . \tag{2.14}$$

[1] We refer the reader to find more information about "Calculus of variations" in the cited work as well as in the course notes of [Svetitsky, 2005] and [Olver, 2014]

Despite its drawbacks in computational performance, its reliability and ease of implementation has made time marching very popular. It is possible to reduce the stability restrictions on $\tau$ by employing a semi-implicit scheme, i.e. the gradient is dependent on the current approximation of the solution: $\nabla F(u^{(k)}, u^{(k-1)})$. This means, an equation system has to be solved iteratively to obtain $u(k)$, which can possibly produce issues depending on the equation. See the graphic interpretation of the gradient descent in Figure 2.2.

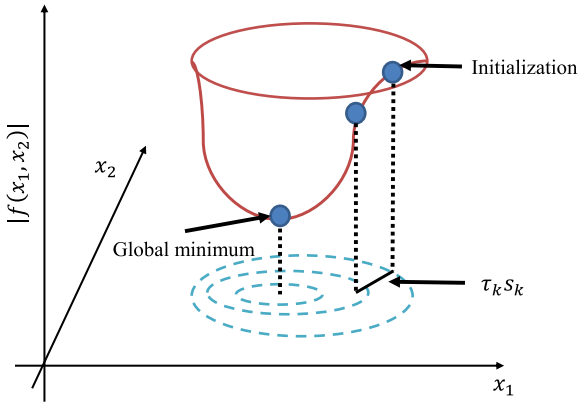



Figure 2.2: Gradient descent for global minimum searching.

Note that the steepest descent is a special case of gradient descent where certain step size is chosen to minimize the objective function value, see Algorithms 1 and 2. Roughly speaking, gradient descent refers to any class of algorithms that calculate the gradient of an objective function, then it moves down in the indicated direction. Moreover, algorithms are not forced to analyze in the direction targeted by the gradient. Other techniques like the Newton's method or the additive operator splitting scheme (AOS) are also available for such task.

**Algorithm 1** Gradient descent algorithm

1: Initialization: $u_0 \in \Re^n$
2: **for** $k = 0, 1, 2, \ldots$ **do**
3: $\quad u_{k+1} \leftarrow u_k + \tau_k s_k$
4: **end for**

**Algorithm 2** Steepest descent algorithm

1: Initialization: $u_0 \in \Re^n$
2: **for** $k = 0, 1, 2, \ldots$ **do**
3: $\quad u_{k+1} \leftarrow u_k - \tau_k \nabla F(u^{(k)})$
4: **end for**

### 2.4.1 Stepsize parameter

The stepsize parameter $\tau$ at iteration $k$ (also known as learning rate in the context of machine learning), specifies how aggressively an iterative algorithm proceeds in the search of a minimum. If the stepsize consists of small values, the convergence to a local/global minimum turns slow. On the other hand, if the stepsize is a high value, the algorithm becomes unstable, oscillating around the apparent minimum and possibly diverging. In the process of reaching convergence, $\tau$ can be considered as iteration-dependent rate $\tau_k$ by employing line search methods. The corresponding procedure often initializes $\tau_0$ at a high value to be aggressive in the beginning and subsequently decreases according to gradient observations to ensure stability.

#### 2.4.1.1 Fixed stepsize

The simplest choice of stepsize corresponds to a constant $\tau > 0$:

$$\tau_k = \tau \quad \forall k \geq 0\,. \tag{2.15}$$

Constant stepsize leads to stationary gradient schemes and not extra computation is required. Convergence is ensured while $\tau$ is sufficiently small for smooth functions with a single minimum.

#### 2.4.1.2 Optimal line search

The optimal line search algorithm computes the stepsize at each iteration as follows:

$$\tau_k = \underset{\tau \leq 0}{\operatorname{argmin}}\, F(u^{(k)} + \tau d_k)\,, \tag{2.16}$$

where $d_k$ has a positive projection in the direction of $-\nabla F(u^{(k)})$. This method seeks to find the values of $\tau_k$ over the half line in $[0, \infty)$. Optimal line search corresponds to a sequence of optimal local decisions that yields maximum decrease of $F$ at every iteration $k$, given the current state $u^{(k)}$ and the descent direction $d_k$.

#### 2.4.1.3 The Armijo rule

Also known as backtracking line search, the Armijo rule is presented as alternative solution to optimal line search that reduces complexity in the computation of $\tau$ at every iteration. The corresponding equation is defined as:

$$F(u^{(k)} + \tau d_k) = F(u^{(k)}) + \tau \nabla F(u^{(k)}) d_k + \gamma(\tau)\,. \tag{2.17}$$

While $\tau$ is small, the term $\tau \nabla F(u^{(k)}) d_k$ becomes dominant compared to $\gamma(\tau)$, hence the Armijo rule attempts to select the suitable value of $\tau$ such that $F(u^{(k)} + \tau d_k) \leq F(u^{(k)})$ holds.

# Chapter 3

# Active contours

Clearly, variational methods based their theory on calculus of variations. Basically, variational methods have several advantages since the theory allows reliable implementations through partial differential equations (PDE). They allow processing of remarkable visual features distinguishable by humans like gradients, curves, corners or edges and they are also effective to simulate various dynamic processes such as linear and non linear diffusion. In the computational sense, the literature of PDE is huge and it drives deeper analysis in both implementation and optimization processes. In general, the variational methods are closely related to stochastic processes that are formulated analogously by the Bayesian theory.

Variational methods applied to image processing, handle a digital image (naturally defined in a discrete domain) in the continuous domain in order to minimize a particular functional which finally leads to a particular equation. The existing solution can be formulated as follows:

$$\xi^* = \underset{\xi \in S}{\operatorname{argmin}} F(\xi)\,, \tag{3.1}$$

where $\xi^*$ is an optimizer of the functional $F$, defined in the corresponding space $S$. If $F$ is continuous and differentiable, the first variation leads to the Euler-Lagrange equation $\frac{\partial F}{\partial \xi} = 0$. The latter formula provides a necessary condition for $\xi^*$ to be an optimizer of $F$ such that $\frac{\partial F}{\partial \xi} = 0\Big|_{\xi^*=0}$. We might often find in the literature the following notation that describes a functional:

$$F(\xi) = \alpha \int_\Omega \Im(\xi) dx + \int_\Omega f(\xi) dx\,, \tag{3.2}$$

where $f$ is a fitting function to state a correspondence between the objective function and data. While $\Im$ is a term that suggests regularity. Note that $\xi$ is defined in a space of values $x$.

## 3.1 The Mumford-Shah functional

The formulation of Mumford-Shah functional proposed in early 90's by [Mumford and Shah, 1989] addresses the problem of energy minimization by computing an approximation of a piecewise function of an image. Therefore, this model has been studied deeply in the field of image processing for the tasks of restoration, denoising and segmentation.

Let $\Omega \in \Re^n$ be open and bounded region and $C$ a closed subset in $\Omega$ which consists of a finite set of smooth curves. The connected components of $\Omega \backslash C$ are denoted by $\Omega_i$ such that $\Omega = \Omega_i \bigcup C$. The function $u_0 : \Omega \to \Re$ is the given bounded image-function. Precisely, the Mumford-Shah formulation states that given a certain image $u_0$, find a decomposition $\Omega_i$ of $\Omega$ and an optimal piecewise smooth approximation $u$ of $u_0$ such that $u$ varies smoothly within each $\Omega_i$. The minimizer of the energy functional would identify the approximate edges of $u_0$. The solution $u$ is formed by regions $\Omega_i$ with boundary $C$.

$$F^{MS}(u, C) = \int_{\Omega/C} (\alpha |\nabla u|^2 + \beta |u - u_0|^2) dxdy + length(C)\,. \tag{3.3}$$

The problem can be further simplified by restricting the segmented image to piecewise-constant functions inside each component $\Omega_i$. The reduced form leads to the so-called minimal partition problem. A new functional is minimized to determine the boundary of the object region:

$$F^{MS}(u, C) = \sum_i \int_{\Omega_i} |u_0 - c_i|^2 dxdy + \nu |C|\,. \tag{3.4}$$

The previous energy functional in Equation (3.4) is minimized by setting $c_i = mean(u_0)$ in $\Omega_i$. Nevertheless, to improve the performance of minimization methods in image processing, both local and global information of the image must be included in the functional equation. This is a trendy technique managed by classic active contour methods.

## 3.2 Snakes by Kass

Active contours also called Snakes, are dynamic curves that evolve according to the image content in order to delineate and identify distinct objects. The evolving procedure is accomplished by constraints from a given image. In [Kass et al., 1988], authors developed initially a method based on PDE to delineate and extract different objects in an image. Their model allows user guidance by adding energy terms to the formula. This level of interaction eases the searching of usable energy functions to reach local or global minima.

Let $\Omega$ be a bounded open subset of $\Re^2$ with $\delta\Omega$ its boundary. Let $u_0$ be a given image with certain active parametric contour represented by $C(s) : [0,1] \rightarrow \Re^2$ in terms of $x-$ and $y-$ axes, $C(s) = (x(s), y(s))$. The Kass formulation of active contours is written as:

$$F^{Kass}(C) = \int_0^1 (E_{internal}(C) + E_{image}(C) + E_{constraint}(C))ds\,, \tag{3.5}$$

where the total energy is split into three parts. The internal energy $E_{internal}$ controls the contour deformation. The external force includes $E_{image}$ to attract the contour towards the object, and $E_{constraint}$ guides evolution to behave in certain manner.

### 3.2.1 Geometric active contours

Generally, edge guided schemes use the assumption that region boundaries correspond to high image transitions (visually highlighted edges) in addition to geometric constraints. The study developed by [Kass et al., 1988] has been a strong precursor of edge detection by curve evolution. A special case of snakes arises when setting the energy constraint to zero while using derivatives of the contour as description of the internal forces.

$$F^{GMAC}(C) = \alpha \int_0^1 |C'|^2 ds + \beta \int_0^1 |C''|^2 ds - \lambda \int_0^1 |\nabla u_0(C)|^2 ds\,, \tag{3.6}$$

where $C'$ denotes the first derivative with respect to $C(s)$ and $C''$ denotes the second derivative. The terms $\alpha$ and $\beta$ are both user defined weights to control sensitivity. For large values of $\alpha$ the change in distance is penalized among the points in the contour, whereas large values of $\beta$ penalize oscillations.

### 3.2.2 Geodesic active contours

Geodesic AC were introduced by [Caselles et al., 1997]. A geodesic curve measures the minimum distance path between given points. The energy functional of the model, which has to be minimized, is expressed as follows:

$$F^{GDAC}(C) = \int_0^1 g(|\nabla u_0(C(s))|)ds\,, \tag{3.7}$$

where $g(\nabla u_0) = 1/(1+\gamma|\nabla z|^2)$. The value $g$ should be small near the object boundaries defined by abrupt changes of intensity and controlled by $\gamma$.

## 3.3 Active contours without edges

Models that base their theory on snakes rely on some edge-function to stop the contour evolution. This fact might cause downsides since objects to be detected have to show clear boundaries defined by gradients. Moreover, the boundary gets vanished when is discretized, it leads that the stopping function will never be zero on the boundary.

The active contour model without edges was first introduced by [Chan and Vese, 2001] and does no rely on an edge stopping function, but instead relies on the Mumford-Shah technique, where the initial image is approximately piecewise constant. Basically, a functional that includes information about the energy of certain curve $C$ is proposed, the optimal partition is achieved when modeling the functional as a minimization problem. The solution provides the convergence of $C$ and thus encloses the object of interest. In other words, segmentation is performed by balancing the content between the inner and outer regions produced by the curve partition.

Let us consider two existing regions in an image $u_0(x,y)$ with distinct approximate constant intensities $u^i$ and $u^o$. Let $C$ be the evolving curve in $\Omega$, as the boundary of an open subset $\omega$ such that $\omega \subset \Omega$ and $C = \delta\omega$. Furthermore, the region inside $C$ is defined as $\omega$ and the region outside $C$ is defined as $\bar{\omega}$. The fitting energy formula comprises two terms:

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

where $c_1$ and $c_2$ are the average pixel values of inside and outside the contour respectively.

The fitting energy term is minimized when the contour $C$ is exactly situated on the boundary of the object. For instance, the difference between $u_0(x, y)$ inside the contour and $c_1$ tends to zero, and the difference between $u_0(x, y)$ outside the contour and $c_2$ tends to zero as well. This can be easily seen at Figure 3.1.

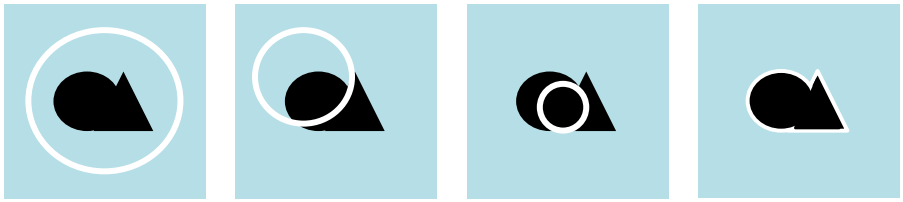



Figure 3.1: Sample image to demonstrate the balance of inner and outer regions according to $C$.

We can say then, the boundary of the object enclosed by $C$ is the minimizer of the fitting energy term:

$$\inf_c \{F_1(C) + F_2(C)\} \approx 0\,. \tag{3.9}$$

Moreover, the classic model suggests to include some regularization terms: the length of the curve and the area of the resulting region.

$$\begin{aligned} F^{CV}(C, c_1, c_2) &= \mu \cdot length(C) + \nu \cdot area(inside(C)) \\ +\lambda_1 \int_{inside(C)} |u_0(x, y) - c_1|^2 dxdy &+ \lambda_2 \int_{outside(C)} |u_0(x, y) - c_2|^2 dxdy\,, \end{aligned} \tag{3.10}$$

where $\mu \geq 0$ and $\nu \geq 0$ are constants to penalize the length and the area respectively. The fixed parameters $\lambda_1$ and $\lambda_2$ penalize the average terms of the inner and outer regions, respectively. Finally, the minimization of the given energy functional leads to the iterative procedure of contour evolution.

## 3.4 Level set formulation

The main idea behind level set stated by [Osher and Sethian, 1988] is to reformulate AC theory according to an evolutionary non-linear PDE. Such method simplifies computation

in order to avoid complex curves, including shapes of time-varying topology. Unlike most numerical approaches, the level set method is implemented in a fixed coordinate system. Hence, the computer capabilities are fixed and does not change throughout the iterations. In this sense, the level set formulation enables numerical computations involving complex surfaces on a fixed grid without the need of parametrization.

Let $C$ be the contour to be represented by a zero level set of a Lipschitz continuous function $\phi : \Omega \to \Re$, so:

$$C = \delta\omega = \{(x, y) \in \Omega : \phi(x, y) = 0\}. \tag{3.11}$$

The level set method performs partitions of regions in the image domain. The boundary of those regions are defined as the zero level set of the underlying function $\phi$. Specifically, the $xy$ plane is placed at the zero-level of $\phi$. All points on the $xy$ plane located inside the boundary correspond to the points where $\phi > 0$ whereas the points outside the boundary are those where $\phi < 0$. The latter statement is explicitly formulated as shown in Equation (3.12). In addition, Figure 3.2 illustrates how this method can easily deal well with topological transitions. In traditional numerical methods, region splitting and merging would require the algorithm to identify those transitions. Moreover, contour parametrization will introduce additional points as long as complexity increases.

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

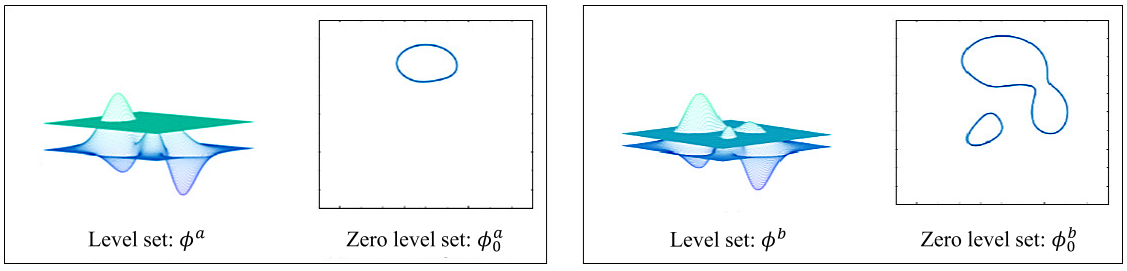



Figure 3.2: A plane in $xy$ cuts the 3D surface to generate the curve $\phi$

Consider the level set formulation with the Heaviside function $H$ and $\delta_0$ as the one-dimensional Dirac measure:

$$H(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{if } z < 0 \end{cases} \quad , \quad \delta_0 = \frac{d}{dz} H(z)\,, \tag{3.13}$$

since $H$ is not differentiable in zero, it is a common practice to use regularized versions of $H$ and $\delta$ functions denoted by $H_\epsilon$ and $\delta_\epsilon$:

$$H_\epsilon(\phi) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\left(\frac{\phi}{\epsilon}\right)\right) \quad , \quad \delta_\epsilon(\phi) = \frac{\epsilon}{\pi(\epsilon^2 + \phi^2)}\,, \tag{3.14}$$

where $H_\epsilon \to 0$ as fast as $\epsilon \to 0$. The Chan-Vese energy functional $F^{CV}$ given in Equation (3.10) can be rewritten as follows:

$$\begin{aligned} F^{CV}(\phi, c_1, c_2) =& \mu \int_\Omega \delta_0(\phi(x,y))|\nabla\phi(x,y)|dxdy \\ &+ \nu \int_\Omega H(\phi(x,y))dxdy \\ &+ \lambda_1 \int_\Omega |u_0(x,y) - c_1|^2 H(\phi(x,y))dxdy \\ &+ \lambda_2 \int_\Omega |u_0(x,y) - c_2|^2 (1 - H(\phi(x,y)))dxdy\,. \end{aligned} \tag{3.15}$$

The minimal partition problem allows us to denote $c_1$ and $c_2$ as functions of $\phi$:

$$c_1(\phi) = \frac{\int_\Omega u_0(x,y)H(\phi(x,y))dxdy}{\int_\Omega H(\phi(x,y))dxdy} \qquad c_2(\phi) = \frac{\int_\Omega u_0(x,y)(1 - H(\phi(x,y)))dxdy}{\int_\Omega (1 - H(\phi(x,y)))dxdy} \tag{3.16}$$

The formulas of Equation (3.16) are translated into the inner and outer averages respectively. Finally, the minimization can be solved by alternating between updating $\phi$ and the region average $c_1$ and $c_2$. In order to update $\phi$, the values $c_1$ and $c_2$ are kept fixed, then $F$ is minimized with respect to $\phi$. This allows us to deduct the associated Euler-Lagrange equation for $\phi(t, x, y)$ for an artificial time $t \geq 0$. Furthermore, since the initial contour is $\phi(0, x, y) = \phi_0(x, y)$, we express the solution as follows:

$$\begin{aligned} \frac{\partial \phi}{\partial t} = \delta_\epsilon \left\{ \mu \cdot div\left(\frac{\nabla \phi}{|\nabla \phi|}\right) - \nu - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2 \right\} = 0 \qquad & in \qquad \Omega\,, \\ \phi(0, x, y) = \phi_0(x, y) \qquad & in \qquad \Omega\,, \\ \frac{\delta_\epsilon(\phi)}{|\nabla \phi|} \frac{\delta \phi}{\delta \vec{n}} = 0 \qquad & on \qquad \delta\Omega\,. \end{aligned} \tag{3.17}$$

Here, $\vec{n}$ denotes the exterior normal to the boundary $\delta\Omega$, and $\delta\phi/\delta\vec{n}$ denotes the normal derivative of $\phi$ at the boundary. The term $div(\frac{\nabla\phi}{|\nabla\phi|})$ represents the divergence of the unit normal to the level curve $\phi$ at every point. This contour is then evolved according to an artificial time $t$ with a second-order nonlinear PDE. It is worth noting that, from Equation (3.17), the evolution of $\phi$ is performed due to two types of forces: the length is minimized by the curvature motion, and two intensity-based forces minimize the deviation of intensity from the mean in a segmented region.

The energy functional in Equation (3.15) is nonconvex and may have more than one local minimums. However, since $H_\epsilon$ and $\delta_\epsilon$ are non-zero everywhere, the Equation (3.17) acts on all level sets of the function $\phi$. Hereby, the segmentation proceeds regardless the choice of the initial contour $\phi_0$.

### 3.4.1 Convex relation modeling

Avoiding the presence of local minima is still challenging for AC models. Particularly, this problem emerges in the framework of piecewise-constant segmentation methods since it involves minimization of functionals over a characteristic set of functions. An algorithm to assist in the search of global minimum is introduced by [Chan et al., 2006] for the case of the two-phase piecewise-constant segmentation problem with fixed intensity constants.

Let us rewrite the solution of the functional $F^{CV}$ in Equation (3.17) as in the next form:

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi) \left\{ \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} - \lambda r(x, y) \right\}, \tag{3.18}$$

where $r(x, y) = (u_0 - c_1)^2 - (u_0 - c_2)^2$. The stationary solution of Equation (3.18) corresponds to a particular energy equation after applying a gradient descent scheme:

$$\int_\Omega |\nabla \phi| dxdy + \lambda \int_\Omega r(x, y)\phi dxdy\,. \tag{3.19}$$

In general, this energy function does not have a minimizer as it is homogeneous of first degree in $\phi$. Minimization is additionally restricted to $\{0 \leq \phi \leq 1, \forall (x,y) \in \Omega\}$, so the following theorem is introduced:

**Theorem**. *For any given fixed $c_1, c_2 \in \Re$, a global minimizer for $MS(\cdot, c_1, c_2)$ can be found by carrying out the following convex minimization:*

$$\min_{0 \leq \phi \leq 1} \left\{ \int_\Omega |\nabla \phi| dxdy + \lambda \int_\Omega r(x,y)\phi dxdy \right\}, \tag{3.20}$$

*and then, setting $\Sigma = \{x : \phi \geq \mu\}$ for a.e. $\mu \in [0,1]$.*

The global minimizer can only be found when $c_1$ and $c_2$ are fixed. Typically, when the fitting function incorporates parameters that have to be optimized, the joint problem becomes non-convex, which may cause issues in the search of a global minimum.

## 3.5 Effects of parameter selection

The resulting PDE from Equation (3.17) is often implemented by finite differences or successive over-relation (SOR) as the most used techniques. However, parameter tunning plays a crucial role in order to quickly converge in a particular segmentation task. Parameters required for the $F^{CV}$ functional are listed below:

| **Parameter** | **Effect** |
|---|---|
| $\phi_0$ | Initialization of the contour |
| $\mu$ | Length of the contour |
| $\nu$ | Area inside the contour |
| $\lambda_1$ | Priority to evolve inside |
| $\lambda_2$ | Priority to evolve outside |

- Initialization. One of the most important criteria required in the theory of AC is the selection of the initial contour (location and shape), that even comprises its own field of study. The zero level set starts evolving from the initial shape provided. The speed of convergence depends on how closer is the initialization situated to the object of interest. Figure 3.3 shows three different approaches of initialization. The dark gray object represents the target to segment. This topic will be addressed again and we will provide another alternatives in Chapter 4.

- Length and area. The term $\mu$ penalizes the contour length. Larger values of $\mu$ might

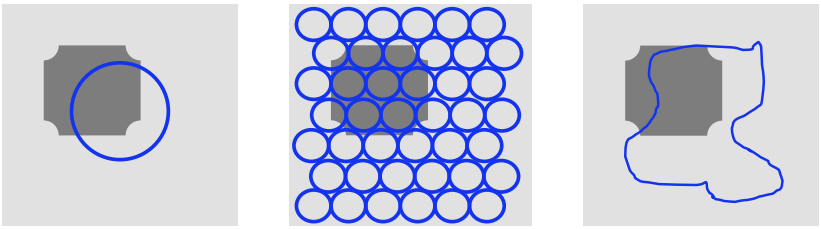


Figure 3.3: Different initialization functions. Left: A concentric circle of certain radius. Middle: A mosaic of circles. Right: A random shape.

cause the contours to have smooth boundaries, whereas smaller values allow more precise delineation such as borders or corners. See example in Figure 3.4, where two sets of circles are enclosed by the AC depending on $\mu$. The circles can be segmented individually or assemble clusters.

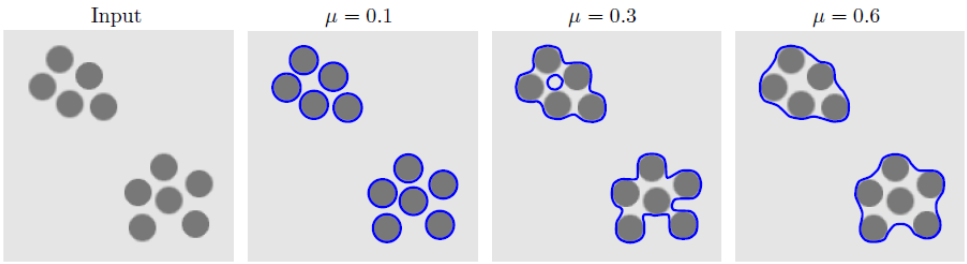



Figure 3.4: Effect of varying $\mu$. Figure extracted from [Getreuer, 2012].

The term $\nu$ penalizes the area inside the contour. It is common to set $\nu = 0$ in the literature, although by setting $\nu \neq 0$ the area of the contour can be forced to either shrink for $\nu > 0$, or tend to expand for $\nu < 0$. See Figure 3.5

- Weights $\lambda$. Also called average terms, often represented by $\lambda_1$ and $\lambda_2$ for the inner and outer regions respectively in the single contour case studied previously. Most of times, $\lambda_1 = \lambda_2 = 1$ which forces the algorithm to evolve without following a special rule. The relation $\lambda_1 > \lambda_2$, assigns priority to the inner region (foreground) to deal better with the variance of pixels than the outer region (background), which forces the contour to shrink. On the other hand, adjusting $\lambda_1 < \lambda_2$ forces the contour to expand towards the outer direction.

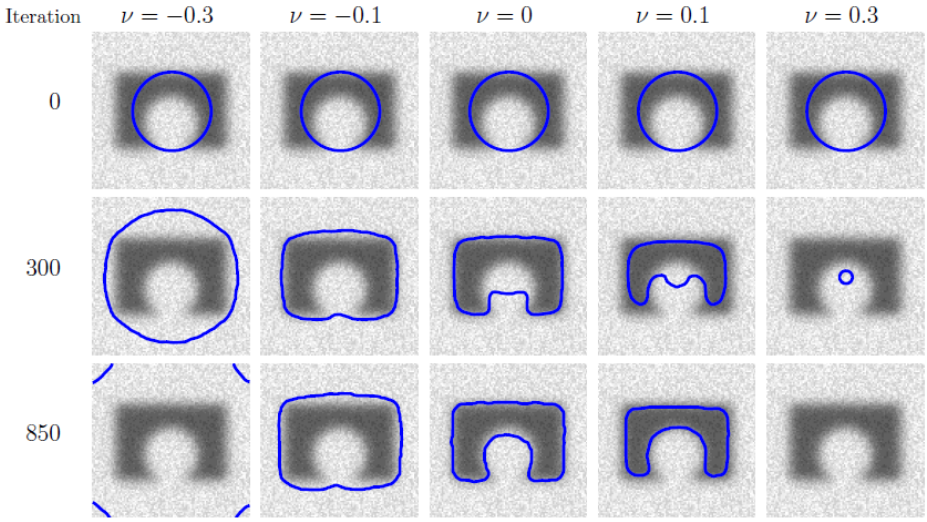



Figure 3.5: Effect of varying $\nu$. Figure extracted from [Getreuer, 2012].

## 3.6 Vector-valued active contour model

The notion of vector edges brought by [Sapiro and Ringach, 1996] is used here to define an stopping edge-function. Consequently, the vector-valued active contour model can be applied to multivalued images. Speaking generally, edges are localized in those regions where the gradient of an image luminance is high. However, extra image attributes such as color, texture or motion, might provide additional information to help in localizing image discontinuities. The process of measuring various image attributes at a single image location leads to multivalued representation methods.

Several approaches for detecting edges in multivalued images attempt to combine the response of a single-valued edge detector applied separately to each of the image components or layers and the way that each component responds is heuristic. Even tough the theory of multivalued active contours suggests to analyze each component separately when dealing with multi-channel image representation such as RGB, multispectral, textured images, etc.

In this section, the Chan-Vese classic method is extended to vector-valued images as shown in [Chan et al., 2000]. The algorithm uses the level set method to determine the boundary of the object of interest. Let $u_0^i$ be the $i-$th channel of an image on $\Omega$, with $i = 1, \ldots, N$ channels and $C$ the evolving curve. Consider that each channel contains alternative versions of the same image. Let $\hat{c^+} = (c_1^+, c_2^+, \ldots, c_N^+)$ and $\hat{c^-} = (c_1^-, c_2^-, \ldots, c_N^-)$ be two constant vectors whose components correspond to each channel $u_0^i$ for the inner $\hat{c^+}$ and outer $\hat{c^-}$ region. The extension case follows the classic Equation (3.10) and can be expressed as seen

bellow:

$$
\begin{aligned}
F^{vvCV}(C, \hat{c^+}, \hat{c^-}) = \mu \cdot length(C) + \nu \cdot area(inside(C)) \\
+ \frac{1}{N}\sum_{i=1}^{N} \lambda_i^+ \int_{inside(C)} |u_0^i(x,y) - c_i^+|^2 dxdy \\
+ \frac{1}{N}\sum_{i=1}^{N} \lambda_i^- \int_{outside(C)} |u_0^i(x,y) - c_i^-|^2 dxdy \,,
\end{aligned}
\tag{3.21}
$$

where $\lambda_i^+$ and $\lambda_i^-$ are arrays of parameters that define priority of inner and outer regions of the $i-$th layer respectively. As in the single-valued case, the vector-valued model looks for the best energy balance averaged over all channels. In this way, the model can detect discontinuities present in at least one channel. Rewriting the Equation (3.21) by following the level set theory we have:

$$
\begin{aligned}
F^{vvCV}(\phi, \hat{c^+}, \hat{c^-}) = \mu \int_\Omega \delta(\phi(x,y))|\nabla\phi(x,y)|dxdy + \nu \int_\Omega H(\phi(x,y))dxdy \\
+ \frac{1}{N}\sum_{i=1}^{N} \lambda_i^+ \int_\Omega |u_0^i(x,y) - c_i^+|^2 H(\phi(x,y))dxdy \\
+ \frac{1}{N}\sum_{i=1}^{N} \lambda_i^- \int_\Omega |u_0^i(x,y) - c_i^-|^2 (1 - H(\phi(x,y)))dxdy \,.
\end{aligned}
\tag{3.22}
$$

Minimizing the energy with respect to the constants $\hat{c_i^+}, \hat{c_i^-}$, it leads to:

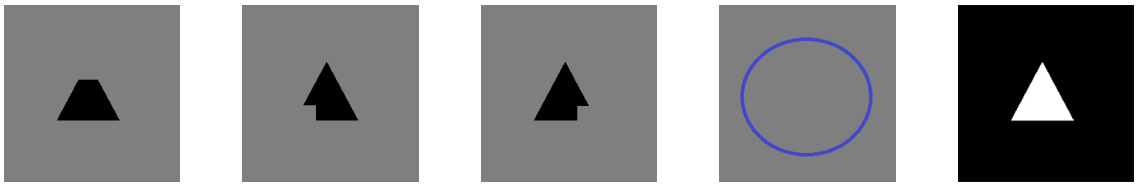
$$
c_i^+(\phi) = \frac{\int_\Omega u_0^i(x,y) H(\phi(x,y))dxdy}{\int_\Omega H(\phi(x,y))dxdy} \qquad c_i^-(\phi) = \frac{\int_\Omega u_0^i(x,y)(1 - H(\phi(x,y)))dxdy}{\int_\Omega (1 - H(\phi(x,y)))dxdy} \,. \tag{3.23}
$$

Note that the contour $\phi$ is the same for all layers which leads to combined changes along all images. It is worth pointing out that the Equations (3.22) and (3.23) are extended versions of the classic model based on level set in Equations (3.15) and (3.16), respectively. The minimization of $F^{vvCV}(\phi, \hat{c^+}, \hat{c^-})$ with respect to $\phi$ yields the following Euler-Lagrange equation:

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon \left\{ \mu \cdot div\left(\frac{\nabla \phi}{|\nabla \phi|}\right) - \nu - \frac{1}{N}\sum_{i=1}^{N} \lambda_i^+ (u_0^i - c_i^+)^2 + \frac{1}{N}\sum_{i=1}^{N} \lambda_i^- (u_0^i - c_i^-)^2 \right\} = 0 \quad in \quad \Omega\,,$$
$$\phi(0,x,y) = \phi_0(x,y) \quad in \quad \Omega\,,$$
$$\frac{\delta_\epsilon(\phi)}{|\nabla \phi|}\frac{\delta \phi}{\delta \vec{n}} = 0 \quad on \quad \delta\Omega\,. \tag{3.24}$$

### 3.6.1 The occlusion problem case

We call disocclusion the process of recovery the occluded areas in an image [Masnou, 2002]. In digital image processing, it is common to face missing information produced by deformation like noise, compression or simply by the nature of the image acquisition process. The latter issue highly depends on the system equipment technology since some objects can be only captured in different modalities. Such is the case of multispectral analysis, where a multichannel image is generated by capturing the same target at different wavelength across the electromagnetic spectrum. Color representation in the RGB space is another basic example but widely used in real world. For instance, the Figure 3.6 shows the segmentation result of an object with occluded areas along different channels. Figures 3.6a, 3.6b and 3.6c show a triangle with a missing peak each. The vector-valued method performs segmentation over the three channels with a single initial contour in Figure 3.6d and the final reconstructed triangle in Figure 3.6e.



(a) (b) (c) (d) (e)

Figure 3.6: Segmentation example of occluded peaks in triangle

## 3.7 Multiphase active contour model

The need of detecting as many objects as possible in a scene, implies to address a multiphase approach in order to extend the traditional active contour model based on level set. Such proposal is helpful to identify more than two distinct regions with multiple boundaries that would be difficult to represent using a single level set. This extension is performed by incorporating additional level set functions $\phi_i$ that divide the domain $\Omega$ into a finite number of partitions [Vese and Chan, 2002].

The multiphase formulation requires $m = log_2(n)$ level set functions to represent $n$ phases or classes, and segment up to $n$ regions. For the $2-$nd dimensional case, a set of pixels of $\Omega$ strictly belongs to only one phase. A number of $n = 2^m$ phases are obtained and determined by including the vector level set function $\Phi = \{\phi_1, \ldots, \phi_m\}$ and the vector Heaviside function $H(\Phi) = \{H(\phi_1), \ldots, H(\phi_m)\}$ in the classic energy functional. The generalized multiphase functional is expressed as follows:

$$F_n^{MP}(\bar{c}, \Phi) = \sum_{1 \leq k \leq n} \int_\Omega (u_0 - c_k)^2 \xi_k dxdy + \sum_{1 \leq i \leq m} \mu_i \int_\Omega |\nabla H(\phi_i)| \,. \tag{3.25}$$

In the case of multiphase version, $k$ denotes each region and $i$ is a label linked to the level set used. An array of constants $\bar{c} = (c_1, ..., c_n)$ expresses the region average. The characteristic function $\xi_k$ delimits the regions $k$. Parameter $\mu_i$ regularizes the surface shape of $\phi_i$, while the value $\nu$ is set to zero and is not included.

A comparison of the active contour-based methods described in this section is displayed in Figure 3.7. The classic version of a single level set is shown in Figure 3.7a. The vector-valued case which handles multiple images by using a single level set is shown in Figure 3.7b. The multiphase model that produces various disjoint regions is shown in Figure 3.7c and Figure 3.7d for two and three level sets respectively.

### 3.7.1 The four color theorem meets the 4-phase level set case

The Four-Color Theorem is helpful to understand that the correct separation of adjacent regions has to be characterized by visual intensity differences, similar to displaying contiguous regions with distinct colours. Based on this theorem, it becomes easier to determine that four colours (or phases) are required to separate the distinct cells from the foreground, background and other surrounding regions.

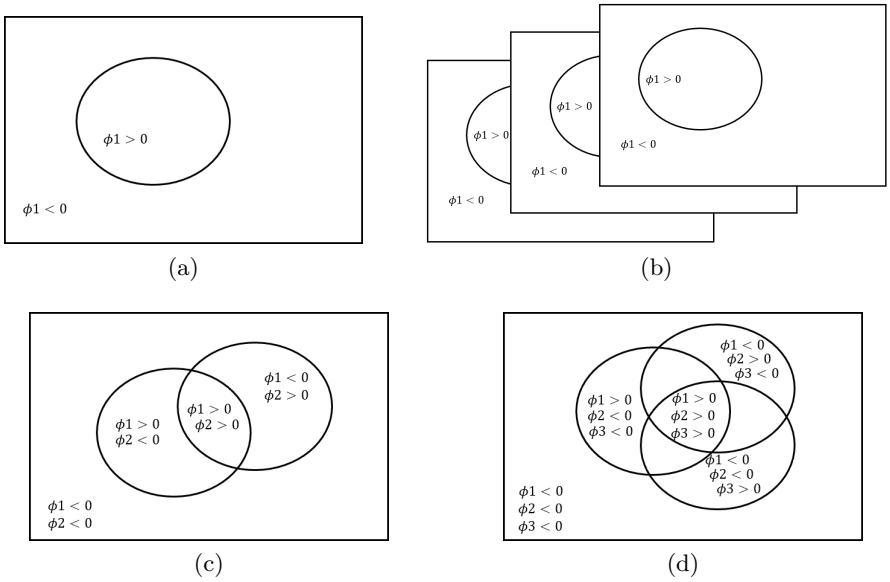



Figure 3.7: Comparison of different versions of active contours. A single level set in (a) and the extension to multichannel images in (b). The 4-PLS case in (c) and the extended multiphase version of three level sets in (d).

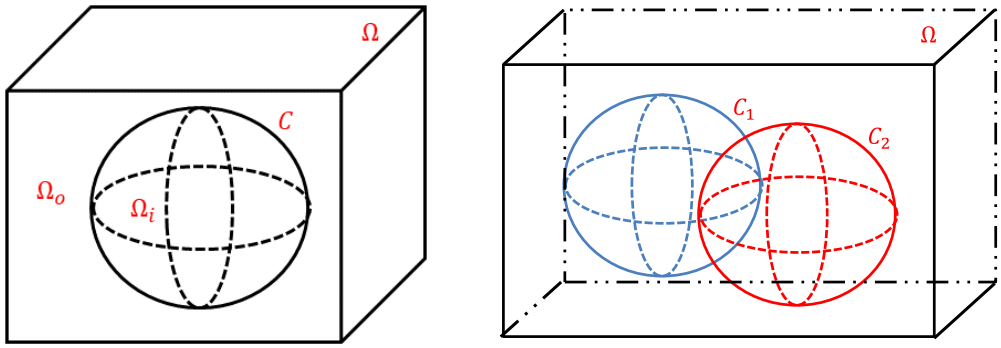
The assumption of $m = 2$ contours used, generates $n = 4$ regions. For instance, let us consider a curve $C$ as the union of two level sets $\{\phi_1 = 0\}$ and $\{\phi_2 = 0\}$. Four regions divide the domain by disjoint sets as: $\{\phi_1 > 0, \phi_2 > 0\}$, $\{\phi_1 < 0, \phi_2 > 0\}$, $\{\phi_1 > 0, \phi_2 < 0\}$ and $\{\phi_1 < 0, \phi_2 < 0\}$ as shown in Figure 3.7c. We can rewrite the energy of Equation (3.25) in terms of the number of classes as follows:

$$
\begin{aligned}
F_4^{MP}(c,\Phi) = &\int_\Omega (u_0 - c_{11})^2 H(\phi_1)H(\phi_2)dxdy + \int_\Omega (u_0 - c_{10})^2 H(\phi_1)(1 - H(\phi_2))dxdy \\
&+ \int_\Omega (u_0 - c_{01})^2 (1 - H(\phi_1))H(\phi_2)dxdy + \int_\Omega (u_0 - c_{00})^2 (1 - H(\phi_1))(1 - H(\phi_2))dxdy \\
&+ \mu_1 \int_\Omega |\nabla H(\phi_1)| + \mu_2 \int_\Omega |\nabla H(\phi_2)| \,.
\end{aligned} \tag{3.26}
$$

The energy functional in Equation (3.26) is a particular case of the multiphase model, the so-called 4-phase level set model (4-PLS) where $\Phi = \{\phi_1, \phi_2\}$ and $c = (c_{11}, c_{10}, c_{01}, c_{00})$. This energy functional can be minimized by the gradient descent method for $\Phi$ which leads to two independent PDEs for $\phi_1$ and $\phi_2$, respectively.

In general, the level set theory eases variational methods to extend the 2D implementation

into a 3D since the topological changes in the boundary of higher dimension structures, such as splitting or merging, are dealt with automatically. Among the simplest approaches to face 3D segmentation challenge, is by using a 2D segmentation model applied to each slice of a volume. Nonetheless, there are many downsides to this methodology, for example, different slices require different parameters to reach the suitable segmentation. On the other hand, to redefine the model with a bounded open subset of $\Re^3$ and $v_0$ as an input volume seems to be more adequate. For the purpose of illustration, Figure 3.8 displays two possible initial contours for 3D binary segmentation in 3.8a and 3D multiclass segmentation 3.8b.




(a) (b)

Figure 3.8: 3D initial spherical contours.

# Chapter 4

# Proposed active contour model

*When man reasoneth, he does nothing else but conceive a sum total, from addition of parcels; or conceive a remainder, from subtraction of one sum from another: which, if it be done by words, is conceiving of the consequence of the names of all the parts, to the name of the whole....*

*- Thomas Hobbes, 1651*

*Leviathan, V.*

In previous chapters, we have mainly discussed the theory of variational methods where active contours have emerged with potential applications to image processing. The approach addressed in this section is an attempt to overcome the downsides produced by using a single AC such as the limitation of only two classes (binary) segmentation or the occlusion problem case. Moreover, the fact to restrict methods to only use pixel intensities might ignore the rest of potentially helpful information. For instance, texture represents one of the most important properties linked to images. Relevant image features such as homogeneity, contrast, depth, periodicity, etc. are visually revealed and highlighted by texture-based approaches [Marr, 1982, Jain and Farrokhnia, 1991].

Among the first ideas about the importance of texture, [Gibson, 1950] established that texture produces a psychological stimulus in the human perception, in the way that texture provides enough information to assume a focal length value on the surface as well as local orientation is specified. These intuitive measurements are the result of the vision process since our perception is a compilation of the environment and the interaction with it. Perception collects data to create an optical arrangement in relation to changes in texture density which simulates the so-called optical flow.

The current proposal suggests to extend the multiphase segmentation approach to adapt any texture representation. Basically, a set of AC yields different image partitions and evolves following the average texture content. This idea allows us to enclose a region with similar texture characteristics within each curve, instead of using intensities that leads to contour overflow, an usual effect caused when dealing with regions that are no clearly delimited by edges.

## 4.1 Vector-valued multiphase active contour model

Let $L$ be the multichannel representation of an image $I$, which consists of $g = 1, \ldots, N$ layers. The proposed formulation states to guide and control multiple AC in the level set representation $\Phi = \{\phi_1, ..., \phi_m\}$ , for each layer according to a weighted sum of $L$:

$$\hat{F_n}(\bar{c_k}, \Phi) = \sum_{1 \leq k \leq n} \frac{1}{N} \sum_{g=1}^{N} \lambda_{k,g} \int_{\Omega} (L_g - c_{k,g})^2 \xi_k dxdy + \sum_{1 \leq i \leq m} \mu_i \int_{\Omega} |\nabla H(\phi_i)|\,, \tag{4.1}$$

where the term $\mu_i$ is introduced to provide flexibility for the $i-$th level set used. The value $\lambda_{k,g}$ is a 2D array of constants that assign priority to regulate the evolution of the AC in the current layer $g$ for the resulting region $k$. Similarly, an array of constants $c_{k,g}$ that expresses the average content within the corresponding region and layer. The two-dimensional array for $\lambda_{k,g}$ and $c_{k,g}$ is shown below:

$$\lambda_{k,g} = \begin{bmatrix} \lambda_{1,1} & \lambda_{2,1} & \ldots & \lambda_{n,1} \\ \lambda_{1,2} & \lambda_{2,2} & \ldots & \lambda_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{1,N} & \lambda_{2,N} & \ldots & \lambda_{n,N} \end{bmatrix} \quad , \quad c_{k,g} = \begin{bmatrix} c_{1,1} & c_{2,1} & \ldots & c_{n,1} \\ c_{1,2} & c_{2,2} & \ldots & c_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1,N} & c_{2,N} & \ldots & c_{n,N} \end{bmatrix}, \tag{4.2}$$

where $k \in [1, n]$, $g \in [1, N]$ and $i \in [1, m]$. The example in Figure 4.1 shows the partition obtained by using 2 contours and 3 layers with the $\lambda_{k,g}$ values corresponding to each region.

Finally, we introduce the characteristic function $\xi_k$ which yields disjointed regions in terms of the Heaviside function as shown in Equation (4.3). This formula leads to different terms according to all possible combinations by computing the product of the Heaviside function of the corresponding level set. For instance, with two level sets, the resulting values enclosed by $\phi_1 > 0$ and excluded by $\phi_2 < 0$ are described as $H(\phi_1)(1 - H(\phi_2))$, which represents the region exclusively delimited by $\phi_1$.

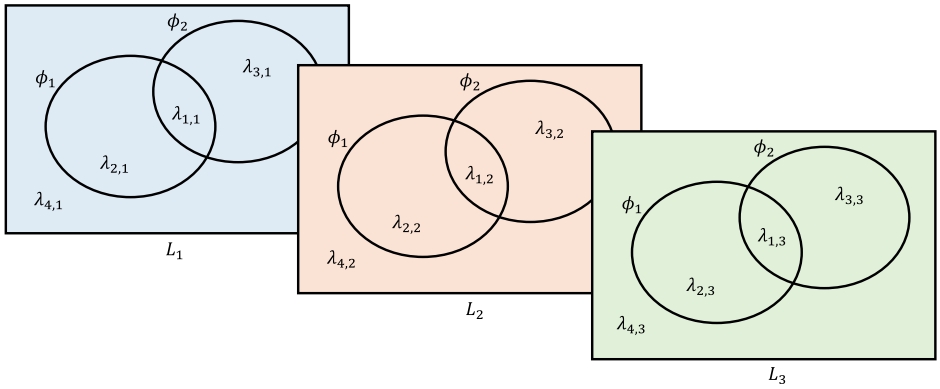



Figure 4.1: Displaying the case of $N = 3$ layers and $m = 2$ contours which generates $n = 4$ regions. The size of $\lambda_{k,g}$ in the illustration is $4 \times 3$.

$$\xi_k = \prod_{i:\phi_i>0 \text{ in } k} H(\phi_i) \prod_{i:\phi_i<0 \text{ in } k} \Big(1 - H(\phi_i)\Big). \tag{4.3}$$

### 4.1.1 The 4-PLS vector case

A special case of the vector-valued multiphase AC arises when two contours are implemented: the so-called **4-PLS vector**. Let $C$ be a resulting curve of the union $\{\phi_1 = 0\}$ with $\{\phi_2 = 0\}$, a total of $m = 2$ level set produces $n = 4$ regions. By means of Equation (4.3), the characteristic function generates four terms $\xi = \{H(\phi_1)H(\phi_2), H(\phi_1)(1-H(\phi_2)), (1-H(\phi_1))H(\phi_2), (1-H(\phi_1))(1-H(\phi_2))\}$. Each term corresponds to a single region with a respective weight that assigns priority in the curve evolution. The main functional contemplates four region terms plus two curvature terms as the Equation (4.4) suggests.

$$\begin{aligned}
\hat{F}_4(\bar{c_k}, \Phi) = &\frac{1}{N}\sum_{g=1}^{N} \lambda_{1,g} \int_\Omega (L_g - c_{1,g})^2 H(\phi_1)H(\phi_2)dxdy \\
&+ \frac{1}{N}\sum_{g=1}^{N} \lambda_{2,g} \int_\Omega (L_g - c_{2,g})^2 H(\phi_1)(1 - H(\phi_2))dxdy \\
&+ \frac{1}{N}\sum_{g=1}^{N} \lambda_{3,g} \int_\Omega (L_g - c_{3,g})^2 (1 - H(\phi_1))H(\phi_2)dxdy \\
&+ \frac{1}{N}\sum_{g=1}^{N} \lambda_{4,g} \int_\Omega (L_g - c_{4,g})^2 (1 - H(\phi_1))(1 - H(\phi_2))dxdy \\
&+ \mu_1 \int_\Omega |\nabla H(\phi_1)| + \mu_2 \int_\Omega |\nabla H(\phi_2)|.
\end{aligned} \tag{4.4}$$

As stated throughout the Chapter 3, the procedure for solving the resulting functional $F_4(\bar{c_k}, \Phi)$ in Equation (4.4) involves minimizing with respect to $\{\phi_1, \phi_2\}$, separately. Two PDEs are obtained for an artificial time $t \geq 0$:

$$\frac{\partial \phi_1}{\partial t} = \delta(\phi_1)\Bigg\{\mu_1 \text{div}\left(\frac{\nabla \phi_1}{|\nabla \phi_1|}\right) + \frac{1}{N}\sum_{g=1}^{N}\Bigg[\left[\lambda_{2,g}(L_g - c_{2,g})^2 - \lambda_{4,g}(L_g - c_{4,g})^2\right]\left(1 - H(\phi_2)\right) - \left[\lambda_{1,g}(L_g - c_{1,g})^2 - \lambda_{3,g}(L_g - c_{3,g})^2\right]H(\phi_2)\Bigg]\Bigg\}, \quad (4.5)$$

$$\frac{\partial \phi_2}{\partial t} = \delta(\phi_2)\Bigg\{\mu_2 \text{div}\left(\frac{\nabla \phi_2}{|\nabla \phi_2|}\right) + \frac{1}{N}\sum_{g=1}^{N}\Bigg[\left[\lambda_{3,g}(L_g - c_{3,g})^2 - \lambda_{4,g}(L_g - c_{4,g})^2\right]\left(1 - H(\phi_1)\right) - \left[\lambda_{1,g}(L_g - c_{1,g})^2 - \lambda_{2,g}(L_g - c_{2,g})^2\right]H(\phi_1)\Bigg]\Bigg\}, \quad (4.6)$$

where $\bar{c_k} = (c_{1,g}, c_{2,g}, c_{3,g}, c_{4,g})$ corresponds to the average content in the regions delimited by $\xi$. Note from Equations (4.5) and (4.6) that even when working with multiple layers, additional contours are not required throughout the process.

## 4.2 Shape constraint as optional parameter

In practice, information about the target is often known a priori. By incorporating this information into a model involves clear advantages in many aspects, e.g. robustness, accuracy, speed, reliability, etc. This is translated into better results for challenging data that are beyond the scope of conventional approaches. Shape-based segmentation methods use prior knowledge for its own benefit. In the literature of shape-based methods, active shape models (ASM) and active appearance model (AAM) [Cootes et al., 1995, Cootes et al., 2001] have emerged as potentially trainable segmentation techniques.

The aim to include a shape constraint parameter in the main functional is to control each element of $\phi_i$ that approximates the shape given by $\phi_i^s$. This can be done by performing the difference between shapes as suggested in [Rousson and Paragios, 2002, Barba-J et al., 2017]

$$F_f(\Phi) = \int_\Omega (\Phi - \Phi^s)^2 H(\Phi) dxdy\,, \tag{4.7}$$

where $\Phi^s = \{\phi_1^s, ..., \phi_m^s\}$ points out to the target shape and $\Phi = \{\phi_1, ..., \phi_m\}$ is the estimated shape. The solution of the given functional in the PDE form produces:

$$\frac{\partial \Phi}{\partial t} = -\big[2(\Phi - \Phi^s)H(\Phi) + (\Phi - \Phi^s)^2\delta(\Phi)\big]\,. \tag{4.8}$$

Let us interpret now the result of minimization in Equation (4.8). The first term $2(\Phi - \Phi^s)H(\Phi)$ measures the difference between the projection of the considered shape and the shape prior interface. The second force $(\Phi - \Phi^s)^2\delta(\Phi)$ aims at decreasing the area, therefore we can ignore this component in order to obtain a scale-invariant measurement.

## 4.3 Considerations of implementation

The iterative process requires a time step related to the speed of convergence. Such process might be stopped by the user by entering manually the number of possible iterations, assuming that at the end of the process, the desired segmentation is reached. An alternative stopping criterion is to accomplish a stationary process in spite of the method to select a stepsize. See the alternatives at the end of Chapter 2. The $j-$th iterative form is depicted in the same sense of Equation (2.14) as follows:

$$\phi^{(j+1)} = \phi^{(j)} + \Delta_t(-\nabla F_n)\,. \tag{4.9}$$

The explicit discretization is obtained for $\phi_1$ and $\phi_2$ from Equations (4.5) and (4.6) at each $j-$th iteration:

$$\phi_1^{(j+1)} = \phi_1^{(j)} + \Delta_t\delta_\epsilon(\phi_1^{(j)})\Big\{\mu_1\cdot\kappa_1^{(j)} + \frac{1}{N}\sum_{g=1}^{N}\Big[\Big[\lambda_{2,g}(L_g - c_{2,g}^{(j)})^2 - \lambda_{4,g}(L_g - c_{4,g}^{(j)})^2\Big] - \Big[\lambda_{1,g}(L_g - c_{1,g}^{(j)})^2 - \lambda_{3,g}(L_g - c_{3,g}^{(j)})^2\Big]\Big]\Big\}, \tag{4.10}$$

$$\phi_2^{(j+1)} = \phi_2^{(j)} + \Delta_t \delta_\epsilon(\phi_2^{(j)}) \Big\{ \mu_2 \cdot \kappa_2^{(j)} + \frac{1}{N} \sum_{g=1}^{N} \Big[ \Big[ \lambda_{3,g}(L_g - c_{3,g}^{(j)})^2 - \lambda_{4,g}(L_g - c_{4,g}^{(j)})^2 \Big] - \Big[ \lambda_{1,g}(L_g - c_{1,g}^{(j)})^2 - \lambda_{2,g}(L_g - c_{2,g}^{(j)})^2 \Big] \Big] \Big\}, \quad (4.11)$$

where $\delta_\epsilon$ is an approximation of the discrete Delta function. Values $\kappa_1$ and $\kappa_2$ suggest the curvature level of $\phi_1$ and $\phi_2$, respectively. These values can be computed by a finite difference scheme. The entire computation of the level set functions is terminated after an arbitrary number of iterations or when an stopping criterion is satisfied. The way to check whether the solution is stationary or not, is by means of a simple tolerance threshold:

$$\sum_{1 \le i \le m} |\phi_i^{(j)} - \phi_i^{(j-1)}| < TOL\,. \quad (4.12)$$

The latest formula states that the difference between two consecutive iterations of each active contour must be lower than certain value for considering the process terminated. The computation is taken over a narrow band around the segmenting contour.

### 4.3.1 General computation of curvature $\kappa$ and component $c_{k,g}$

In numerical implementations, the use of meshes naturally defined by square pixels or voxels is a very common approach. Finite differences are operations to discretize and approximate derivatives by using a lattice $\{(x_p, y_q)|1 \le p \le P, 1 \le q \le Q\}$. The curvature consists precisely of measuring the level of variation for each AC. We define the finite difference of a function $\phi$ by:

$$\begin{aligned} D_x^- \phi_{p,q} &= \frac{\phi_{p,q} - \phi_{p-1,q}}{\Delta h}\ ,\ D_x^+ \phi_{p,q} = \frac{\phi_{p+1,q} - \phi_{p,q}}{\Delta h}\ ,\ D_x^o \phi_{p,q} = \frac{D_x^- \phi_{p,q} + D_x^+ \phi_{p,q}}{2} \\ D_y^- \phi_{p,q} &= \frac{\phi_{p,q} - \phi_{p,q-1}}{\Delta h}\ ,\ D_y^+ \phi_{p,q} = \frac{\phi_{p,q+1} - \phi_{p,q}}{\Delta h}\ ,\ D_y^o \phi_{p,q} = \frac{D_y^- \phi_{p,q} + D_y^+ \phi_{p,q}}{2} \end{aligned} \quad (4.13)$$

here, we call $D^-$, $D^+$ and $D^o$ as the backward, forward and central differences respectively, $\Delta h$ is often set to 1 representing the separation between two successive pixels. Hence, the curvature $\kappa$, Equation (4.14), can be expressed through the scheme of differences from Equations (4.13). Note that multiple AC lead to multiple curvatures, which are calculated independently

to the rest of the contours and are given by the formula:

$$\kappa = \frac{(D_x^+ D_x^- \phi_{p,q})(D_y^o \phi_{p,q})^2 - \frac{1}{2}(D_x^+ D_y^+ \phi_{p,q})(D_x^o \phi_{p,q})(D_y^o \phi_{p,q}) + (D_y^+ D_y^- \phi_{p,q})(D_x^o \phi_{p,q})^2}{(D_x^o \phi_{p,q})^2 + (D_y^o \phi_{p,q})^2}. \tag{4.14}$$

On the other hand, components $c_{k,g}$ play a key role since the balance of regularity between the partitions on the image content are saved in them. Basically, $c_{k,g}$ measures the average content within the specific region delimited by $\xi_k$ from Equation (4.3). The term is often regularized according to the size of the region given by:

$$c_{k,g} = \frac{\int_\Omega L_g \xi_k dxdy}{\int_\Omega \xi_k dxdy}, \tag{4.15}$$

where $k$ ranges between 1 up to the total number of partitions $n$, and $g$ varies from 1 up to the total number of channels $N$ in the image $L$. For the purpose of illustration, we give the details of $c_{k,g}$ for the **4-PLS vector** model by solving the equations:

$$\begin{aligned} c_{1,g} &= \frac{\int_\Omega L_g H(\phi_1) H(\phi_2) dxdy}{\int_\Omega H(\phi_1) H(\phi_2) dxdy} \quad , \quad c_{2,g} = \frac{\int_\Omega L_g H(\phi_1)(1 - H(\phi_2)) dxdy}{\int_\Omega (H(\phi_1)(1 - H(\phi_2)) dxdy} \\ c_{3,g} &= \frac{\int_\Omega L_g (1 - H(\phi_1)) H(\phi_2) dxdy}{\int_\Omega (1 - H(\phi_1)) H(\phi_2) dxdy} \quad , \quad c_{4,g} = \frac{\int_\Omega L_g (1 - H(\phi_1))(1 - H(\phi_2)) dxdy}{\int_\Omega (1 - H(\phi_1))(1 - H(\phi_2)) dxdy} \end{aligned} \tag{4.16}$$

### 4.3.2 Main algorithm

Algorithm 3 reveals the iterative procedure of the **4-PLS vector** model.

**Algorithm 3** 4-PLS vector algorithm

1: Initialization: $\phi_1^{(0)}$, $\phi_2^{(0)}$
2: Parameter selection: $\lambda_{k,g}$, $\mu_1$, $\mu_2$, $\Delta_t$
3: **for** $j = 0, 1, 2, \ldots$ **do**
4: Compute: $\kappa_1^{(j)}$, $\kappa_2^{(j)}$
5: Compute: $c_{k,g}^{(j)}$
6: Solve Equations: (4.10) and (4.11)
7: **if** Equation (4.12) == true, **then**
8: **Break**
9: **end if**
10: **end for**

## 4.4 Steered Hermite coefficients as texture description

The main idea behind oriented filters is to analyze images by applying steered versions of 2D cartesian filters at different orientations. Steerable filters are implemented by a linear combination of a set of basis filters [Freeman and Adelson, 1991]. This principle is handled by the 2D Steered Hermite transform which allows adapting the analysis direction to the local image content according to a maximum oriented energy criterion [van Dijk and Martens, 1997]. Following this statement, the steered Hermite coefficients (SHCs) $L^{\theta}_{p-q,q}$ are generated as a consequence of linearly combining the cartesian Hermite coefficients (HCs) $L_{p-k,k}$.

$$L^{\theta}_{p-q,q}(x_0, y_0) = \sum_{k=0}^{p} L_{p-k,k}(x_0, y_0) R_{p-k,k}(\theta)\,, \tag{4.17}$$

$$R_{p-q,q}(\theta) = \sqrt{\binom{q}{p}} \cos^{p-q}(\theta) \sin^{q}(\theta)\,, \tag{4.18}$$

where $R_{p-q,q}(\theta)$ are known as the Cartesian angular functions that reveal the directional selectivity of the filter. Values $(p-q)$ and $q$ denote the analysis order in $x-$ and $y-$ axes for $p = 0, ..., \infty$ and $q = 0, ..., p$.

Finally, the HCs can be obtained by convolving an image $I(x, y)$ with the analysis filters $D_{p-q,q}(x, y)$. The analysis filters are the result of combining a Gaussian window and the Hermite polynomials [Martens, 1990]. This combination depicts a localized description of an image with filter properties of rotational symmetry and spatial separability $D_{p-q,q} = D_{p-q}(x) D_q(y)$ .

$$L_{p-q,q}(x_0, y_0) = \int_x \int_y I\,(x, y)\, D_{p-q,q}\,(x_0 - x, y_0 - y)\, dxdy\,. \tag{4.19}$$

In addition, the SHCs preserve compacted energy whereas the local coefficients are oriented into the direction of maximum energy, see the example of Figure 4.2. This produces a more compact representation along the first orders and allows multi-scale decomposition for analyzing images at different resolutions [Estudillo-Romero and Escalante-Ramirez, 2011, Silvan-Cardenas and Escalante-Ramirez, 2006].

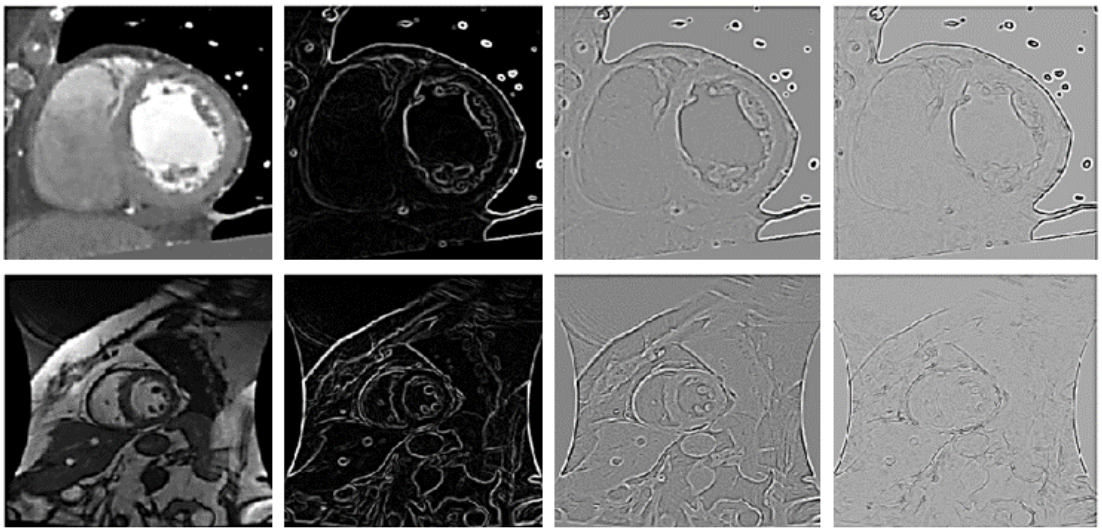


Figure 4.2: Illustration of the steered Hermite coefficients of a CT image in the first row and MRI in the second row. Each column represents the corresponding coefficient $L^{\theta}_{0,0}$, $L^{\theta}_{1,0}$, $L^{\theta}_{2,0}$ and $L^{\theta}_{3,0}$.

# Chapter 5

# Applications to image segmentation

*Be ashamed to die until you have won some victory for humanity.*
*- Horace Mann, 1859*
*Address at Antioch College*

## 5.1 First approach on texture segmentation

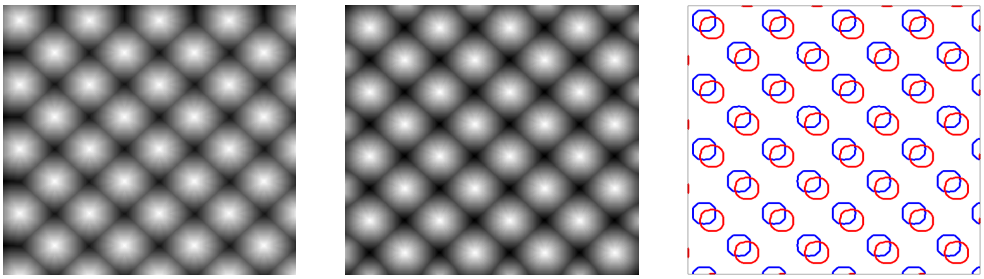
The current section begins by presenting the early contributions in the development of the segmentation model examined in the preceding chapter. The results shown here have been partially submitted and published in [Carbajal-Degante et al., 2018] and extended in theory.

### 5.1.1 Initialization alternatives

As stated in previous sections, initialization has a significant role in contributing to the final segmentation result. Generally speaking, the speed of convergence depends on how close the initial curve is placed to the object of interest. Moreover, automated initialization is of special concern when identification methods require fast and precise segmentation. In this sense, an alternative approach to tackle the initialization issue is by using convolutional neural networks (CNNs). Such operation takes advantage of the available data to build a trainable model. Nevertheless, we will get into more details on this alternative when we will discuss about CNNs in next section.

Our first approach to carry out the initialization is by creating a mosaic tile of circular patterns according to a periodicity criterion. In the current implementation, the level set

function takes positive values inside the curve and negative values outside, simulating a distance map. We define our initialization to follow the formula $\phi^{(0)} = \sin(\pi x f) \cdot \sin(\pi y f) + \psi$ where $\psi$ represents a shift which causes overlapping between contours and $f$ is the oscillating frequency in terms of pixels, see Figure 5.1 The $x-$ and $y-$ axes yield column and row vectors, respectively.

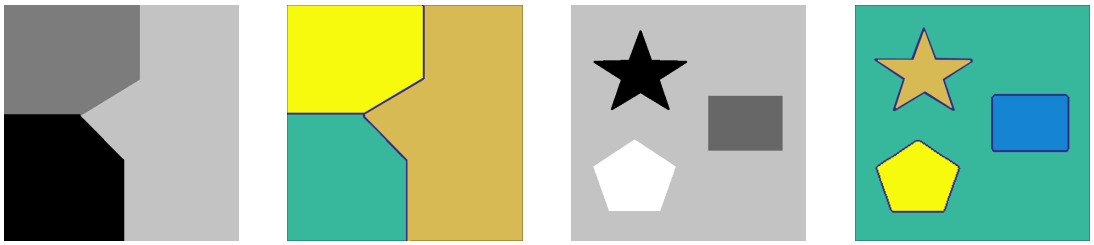


(a) (b) (c)

Figure 5.1: Initial distance map for the level set $\phi_1^{(0)}$ in (a) and $\phi_2^{(0)}$ in (b). Image (c) shows overlapped contours in the $xy-$plane.

### 5.1.2 Segmentation of synthetic images

We carried out a segmentation experiment of two synthetic images of three and four distinct regions respectively, see Figure 5.2. Both images were gradually corrupted with Gaussian noise of mean $\mu_n = 0$ and variance $\sigma_n^2$ in the range of $[0, 0.5)$. In order to compare the classic multiphase model $F_4^{MP}$ of Equation (3.26) and the texture-based model $\hat{F}_4$ of Equation (4.4), we computed the correct segmentation rate (CSR) which is a relation between the number of pixels that have been correctly segmented and the total number of pixels in the image for each class. Plots in Figure 5.3 display the CSR while noise increases.



(a) (b) (c) (d)

Figure 5.2: Synthetic images without noise and segmentation results for three regions (a)-(b), and four regions in (c)-(d).

Both classic and texture-based methods were initially conceived to segment up to four regions. Note that the algorithms performed a perfect segmentation of three regions in low level or absent noise, see graphics in Figure 5.3(a). Results are also displayed for four regions in Figure 5.3(b). Curves exhibit an evident CSR decrease as the level of noise increases. Nevertheless, the texture-based method shows higher levels of robustness to noise than the classic method in most cases.

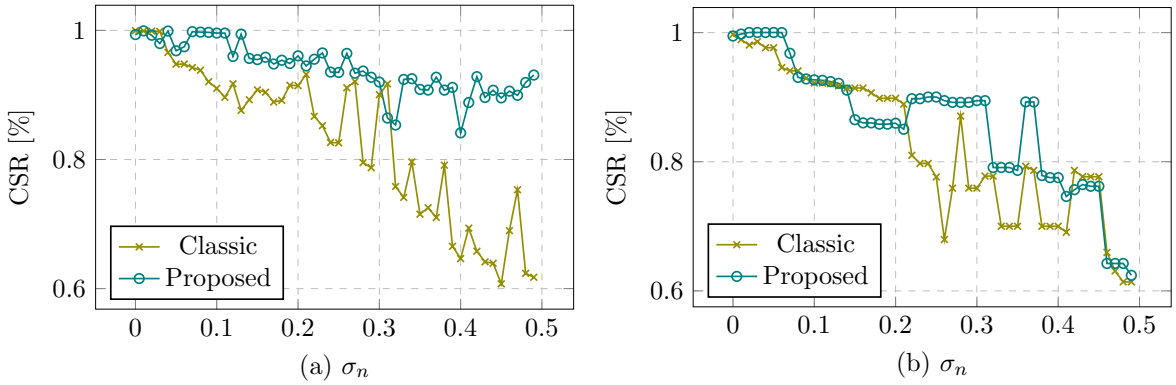


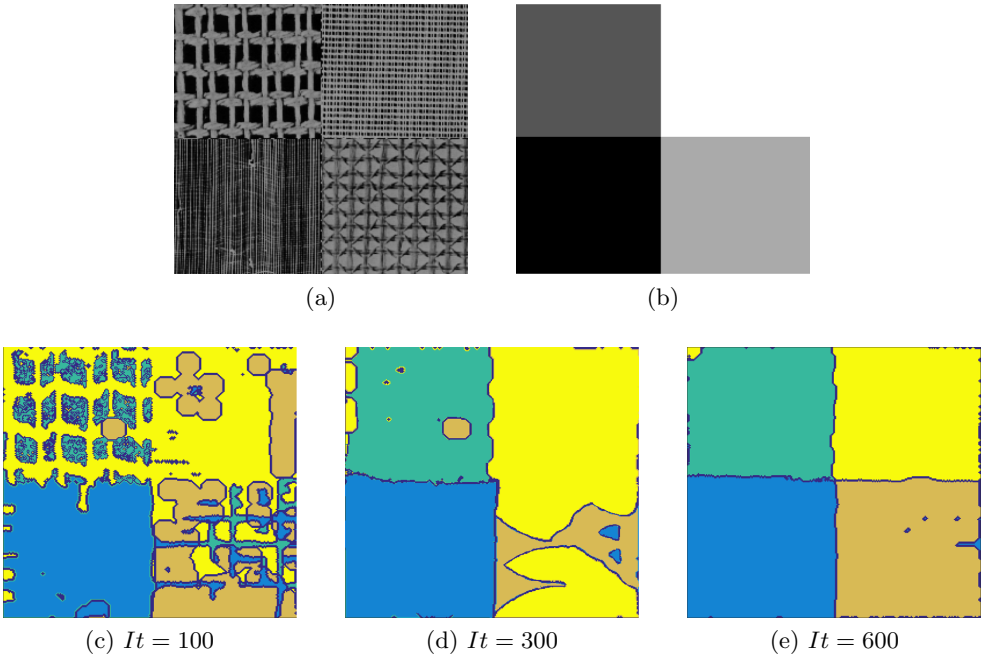

Figure 5.3: Segmentation results of synthetic images corrupted by Gaussian noise. $\sigma_n$ ranges in $[0, 0.5)$. Graphic for three regions in (a) and four regions in (b).

### 5.1.3 Segmentation of mosaic of textures

Several texture-segmentation algorithms are assessed in a synthetic tile of textures to evaluate their performance in ideal conditions. The model described as $\hat{F}_4$ in Equation (4.4) was tested in a mosaic of four different textures chosen from the Brodatz database [Brodatz, 1966]. Figure 5.4 presents a visual comparison among the original mosaic, ground truth and various segmentation results during the process. The final segmentation was achieved after 600 iterations and reached a CSR = 0.97.

### 5.1.4 Brain structure segmentation

The correct delineation of brain structures in magnetic resonance imaging (MRI) has been extensively studied in recent years and still remains as a crucial problem in medicine [Dolz et al., 2015, González-Villà et al., 2016]. The current experiment is merely for illustration purposes to highlight the effect of varying $\mu$ from Equation (4.4). We performed segmentation to detect three main tissue types in brain: gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) [Kapur et al., 1996, Despotović et al., 2015, Li et al., 2011]. In

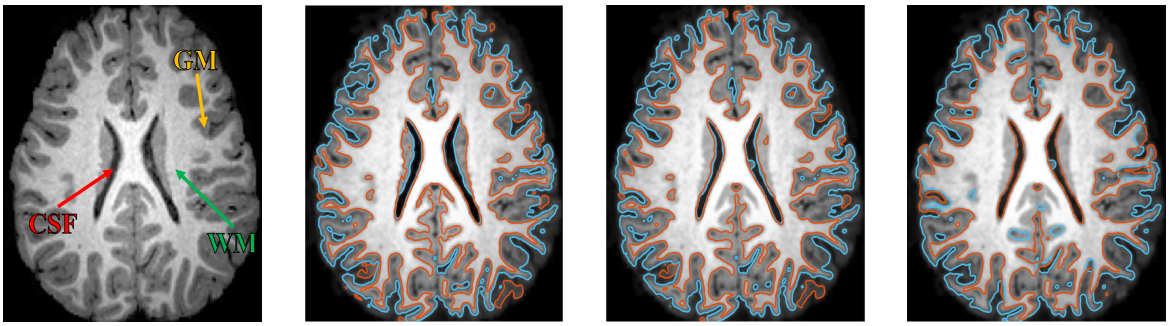


(a) (b)

(c) $It = 100$ (d) $It = 300$ (e) $It = 600$

Figure 5.4: Segmentation of mosaic of textures (a), ground truth (b) and segmentation results at different iterations from (c)-(e).

Figure 5.5, segmentation results reveal the effect of varying $\mu$, the bigger is $\mu$ the smoother the boundary.




(a) (b) $\mu = 0.05$ (c) $\mu = 0.1$ (d) $\mu = 0.5$

Figure 5.5: Effect of varying $\mu$. An MRI brain image in (a) and segmentation results with different values of $\mu$ (b)-(d).

## 5.2 Hybrid segmentation approach

Deep learning is a specific subfield of machine learning which places great emphasis on learning from massive data sets. The term *deep* is a reference to several successive operations that lead to compact representations. For instance, convolutional neural networks (CNNs) have been considered as one of the most powerful tools in the recent development of artificial intelligence-based systems, and it has become relevant to the deep learning field [Chollet, 2017]. Nowadays, there is a wide range of CNNs architectures available in the literature. We provide an overview of the use of CNNs to built new hybrid models in the following sections, as originally discussed in [Carbajal-Degante et al., 2020].

### 5.2.1 The U-Net autoencoder

The U-Net is a CNN which consists of a contracting path and an expansive path. It is an auto-encoder architecture originally designed to perform semantic image segmentation in [Olaf Ronneberger and Brox, 2015]. Due its good performance, U-Net has been commonly used as a starting point for the development of a wide variety of models. Seeking to further improve its performance, several modifications have been proposed [Zhou et al., 2018, Garcia-Uceda et al., 2019, Özgün Çiçek et al., 2016, Oktay et al., 2018]. This model is based on three main processes, see Figure 5.6:

- The encoder: Throughout this process, feature maps are generated by means of applying different convolutional filters. In each stage, the spatial dimension of the feature map is halved while the number of channels is duplicated. This drives the model to analyze context at different scales.
- The decoder: During this process, the feature map generated by the encoder goes through convolution and upsampling operations. The latter is controlled by transposed convolutions. Finally, the model can learn to recover the spatial information and gradually generates the segmentation map that corresponds to the original image.
- Skip connections: These are used to connect the feature map from the encoder directly to the corresponding decoder stage so that the information at every scale produces the segmentation map, preventing the loss of spatial details.

Maximum likelihood provides a strategy to estimate the parameters of the model. This is

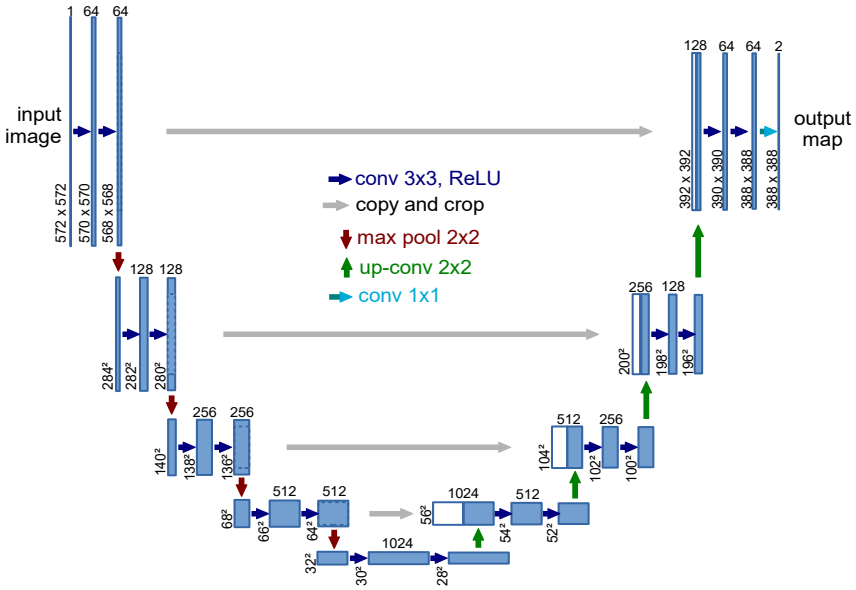



Figure 5.6: Scheme of the U-Net architecture and the operations performed in the input image to generate a segmentation map as output. Illustration extracted from [Olaf Ronneberger and Brox, 2015].

equivalent to minimize the cross entropy between the model and the empirical distributions. In the case of categorical distributions (multiclass segmentation), this corresponds to:

$$CCE = -\sum_{i}^{C}\sum^{N} \widehat{p}_{data}(y_i|x_i)[\log p_{model}(y_i|x_i)]\,, \tag{5.1}$$

where $p_{model}(y_i|x_i)$ is the probability from the model distribution for a given pixel $x_i$ to be assigned to the class $y_i$. The value $\widehat{p}_{data}(y_i|x_i)$ is the probability from the empirical distribution of $x_i$ corresponding to $y_i$. The optimization process is done by means of the iterative method of gradient descent.

A method to efficiently compute the modification in each parameter is the back-propagation algorithm which is used by most artificial neural networks. The error in the last layers helps to adjust the parameters in the previous layers. In this sense, information about adjustment flows in the opposite direction which yields inference.

#### 5.2.1.1 General limitations of CNNs

Mainly, the fundamental aspect that machine learning analyzes is the relation between optimization and generalization. Optimization refers to the process of adjusting a model to obtain

the best performance possible on training data, whereas generalization refers to how well the trained model performs on new data. The goal is to get good generalization, unfortunately we can only adjust the model based on its training data.

Optimization and generalization are correlated at first iterations. The lower the loss on training data, the lower the loss on test data. While this is happening, the model is said to be **underfit** since the network has not yet learned all relevant patterns in the training process. After a certain number of iterations, generalization stop improving while the model begins to learn patterns that are misleading when it comes to new data. In this sense, we say the model is starting to **overfit**.

To prevent a model from learning irrelevant patterns during training, the best solution is to get more training data, which is a common problem found in real applications. Labeled data is not always available and then to manually delineate all images needed represents a tedious task. The process of increasing the data size still remains challenging and widely discussed in the literature.

### 5.2.2 Intensity-based hybrid model

The main contributions of the hybrid-IB model are listed below:

- Perform a high precision automated segmentation that combines the efficiency of a CNN and AC. Hybrid models provide a solution to mutually overcome the limitations of each tool separately.
- Demonstrate that the proposed model helps the network to redefine the segmentation result with a low amount of training data.
- Reduce the variability of convergence in the 4-PLS which is highly dependent on initialization.

First of all, the input image is analyzed via the U-Net after being trained with a dataset that contains all possible classes to segment (up to 4). Afterwards, the output of the U-Net yields a pre-segmentation result which provides the estimated shape and position of each object of interest. These attributes are subsequently assigned to the initialization stage of the 4-PLS model ($F_4^{MP}$ of Equation (3.26)) as the zero level set. Parameters employed in the second stage should be chosen by the user in order to control the evolution of each curve

leading to a refined segmentation at the end of the process. The whole process is depicted in Figure 5.7.

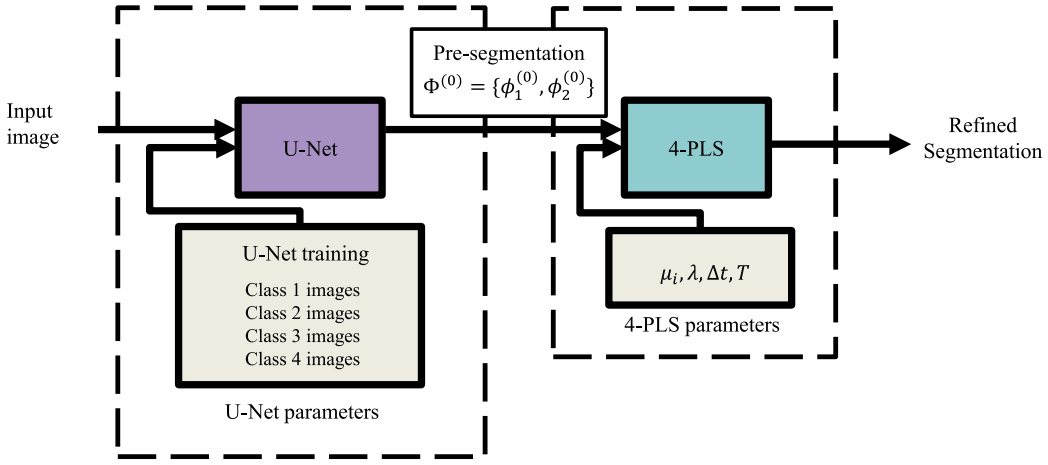



Figure 5.7: Intensity-based hybrid model

### 5.2.3 Texture-based hybrid model

The texture-based (TB) hybrid proposal is inspired in the Hybrid-IB model, which contains the same two main modules directly connected with an extra texture module embedded, see Figure 5.8. Let us assume that a certain image is given as input to the system. The first module comprises the U-Net architecture with a ResNet (Residual Networks) as a backbone to assist the texture-based multiphase AC model.

The U-Net yields a pre-segmentation result after being trained with the available data with up to four classes. The corresponding pre-segmentation provides detailed information of the apparent position and estimated shape for each class, this information is subsequently forwarded to the input of the second module. It is worth pointing out that the pre-segmentation accuracy fully depends on the training-data size, however only a coarse location estimation is needed at this stage.

The *4-PLS Texture* module handles two inputs: a pre-segmentation result to initialize the the zero level set and the SHCs of the input image to exploit the wide texture description retrieved. Parameters of the 4-PLS Texture allow to control the evolution of each curve leading to a refined segmentation at the output of the entire system. In this sense, the proposed hybrid model provides an automated segmentation mechanism end-to-end.

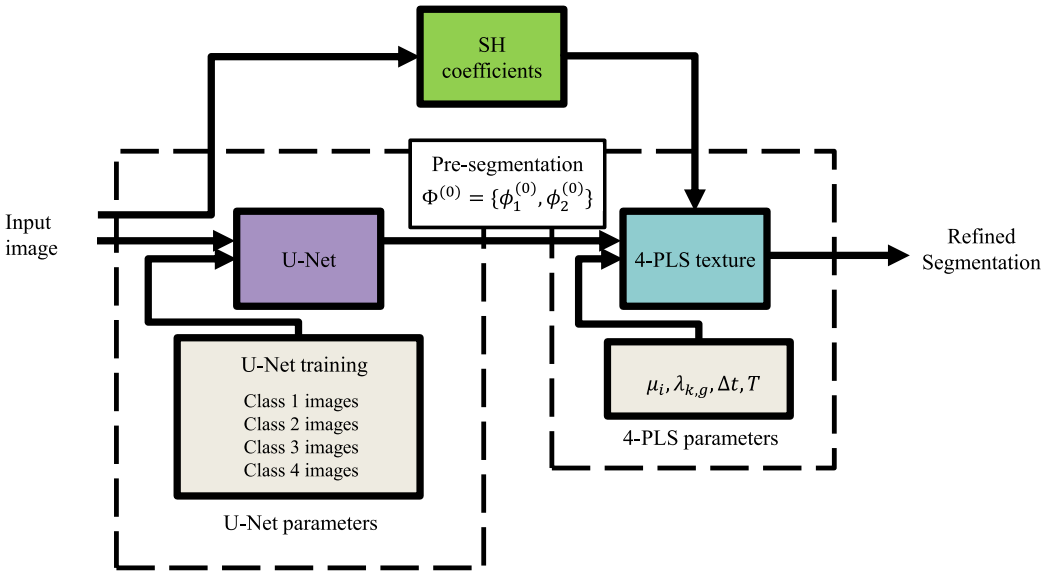



Figure 5.8: Block diagram of the Hybrid-TB model.

## 5.3 Heart ventricles segmentation

Cardiovascular diseases (CVDs) are of great concern for everyone. According to the World Heart Federation, CVDs are currently ranked first in morbidity and mortality within almost two thirds of the world population. Statistics from the World Health Organization report that in 2016 there were more than 15 million deaths due to these diseases and it is projected to remain the single leading cause of death by 2030 [WHO, 2019].

The study of the heart and its diseases usually makes use of different imaging techniques. Nuclear cardiology (NC), cardiac computed tomography (CT), cardiac magnetic resonance imaging (MRI) and echocardiography (ECHO) have proven of enormous advantage for clinicians [Kumar et al., 2012, Petitjean and Dacher, 2011]. Qualitative and quantitative examination is routinely executed by experts, but also a precise quantification step is necessary by the computer-based measurement systems.

This section addresses the problem of left (LV) and right ventricle (RV) segmentation, since automatic segmentation remains one of the most important tasks in computed aided cardiac diagnosis. Active contours have shown to be efficient for this task, however they often require a good initialization, which drives the tool substantially dependent on a prior knowledge or manual process. Here, we show the application of the aforementioned hybrid models.

### 5.3.1 Materials and experiments

Five different techniques were assessed throughout this experiment. The first method consists of the **U-Net** module which was trained with all available data. Since overfitting is caused by having a very limited dataset to learn from, data augmentation (DA) increases the number of training data from the existing samples via random transformations that yield believable-looking images. By including DA to classic U-Net leads to **U-Net-DA** with a similar configuration as recommended in [Baumgartner et al., 2018]. Third, the version called **4-PLS Texture-UI** is a non trainable method based on texture that requires user interaction (UI) to manually initialize contours while using parameters by default (set all to 1), this option arises from Equation (4.4). Finally, the current proposal combines U-Net-DA and 4-PLS for intensity-based analysis which leads to **(Hybrid-IB)** and the extended texture-based case **(Hybrid-TB)**.

Two different datasets were used to compare the techniques mentioned previously:

1. *CT images.* The dataset used in this experiment consists of 228 images extracted from different tomographic cardiac studies provided by a local hospital with a CT Siemens dual source scanner (128 channels). These images have a resolution of 300×300 pixels with manual annotations of left and right ventricles done by a clinical specialist. Throughout this experiment, the dataset was split into 80% and 20% for training and testing purposes respectively. This leads to 45 randomly selected images for assessment. Parameters used for this experiment are shown in Table 5.1.

2. *MRI images.* The dataset used for the experiment with magnetic resonance images was taken from Sunnybrook Cardiac Data (SCD) [Radau et al., 2009] and publicly available [SUN, 2019]. SCD was originally created for myocardium segmentation and extended later to a cardiac LV segmentation challenge. It consists of 420 LV images labeled by their experienced cardiologists. For the rigth ventricle case, a manual segmentation was drawn by a local cardiologist expert. Finally, endocardium ground truth for both LV and RV was used. A total of 119 images of 256×256 pixels were randomly selected and divided into 80% for training and 20% for testing, leading to 24 randomly selected images for assessment. Parameters used in this experiment are shown in Table 5.2.

The loss of categorical cross entropy was computed and the minimum validation error along 100 epochs was selected to carry out the proposed experiments. We report the interval before 50 epochs where the minimum was found for all cases. Figure 5.9 shows the loss curves

Table 5.1: Parameters of CT experiments for each tool. Hybrid methods gather U-Net and U-Net-DA parameters.

| **U-Net** | | **U-Net-DA** $[-v, +v]$ | | **Hybrid IB** | | **Hybrid TB** | |
|---|---|---|---|---|---|---|---|
| Input size | 256×256 | Rotation | 10 | $\lambda_2$ | -3 | $\lambda_{2,g}$ | -3 |
| Batch size | 16 | Width shift | 0.1 | $\lambda_3$ | 0.5 | $\lambda_{3,g}$ | 0.5 |
| Dropout rate | 0.5 | Height shift | 0.1 | $\mu_1$ | 8 | $\mu_1$ | 1 |
| Optimizer | Adam | Shear range | 0.05 | $\mu_2$ | 3 | $\mu_2$ | 1 |
| Total weights | 31,046,339 | Zoom range | 0.15 | dt | 5 | dt | 2 |

Table 5.2: Parameters of MRI experiments for each tool. Hybrid methods gather U-Net and U-Net-DA parameters.

| **U-Net** | | **U-Net-DA** $[-v, +v]$ | | **Hybrid IB** | | **Hybrid TB** | |
|---|---|---|---|---|---|---|---|
| Input size | 256×256 | Rotation | 10 | $\lambda_2$ | 1 | $\lambda_{2,g}$ | 1 |
| Batch size | 16 | Width shift | 0.1 | $\lambda_3$ | -0.5 | $\lambda_{3,g}$ | -1 |
| Dropout rate | 0.5 | Height shift | 0.1 | $\mu_1$ | 1 | $\mu_1$ | 1 |
| Optimizer | Adam | Shear range | 0.05 | $\mu_2$ | 1 | $\mu_2$ | 2 |
| Total weights | 31,046,339 | Zoom range | 0.15 | dt | 5 | dt | 2 |

of U-Net and U-Net(DA) for training and validation sets for both CT and MRI experiments. Note from Figure 5.9 that convergence is faster when DA is not used, although U-Net-DA yields a better generalization due to the increase of training set that contributes to better performance on new data.

### 5.3.2 Results and discussion

We carried out a validation process to compare results against manual annotations, see Table 5.3. We used the Dice coefficient which suggests more similar contour areas when values are closer to 1. Likewise, Hausdorff distance calculates separation between boundaries. In this sense, Hausdorff indicates more alike boundaries when values tend to 0.

Both Dice coefficient and Hausdorff distance showed that U-Net segmentation improves when data augmentation is used for the case of CT, as expected. Alternatively, 4-PLS Texture-UI demonstrates a good performance for RV due to the user interaction who provides very alike shape initialization. Finally, hybrid proposals present the best segmentation for LV and overcome the results of CNN-based methods for RV.

On the other hand, for the case of MRI segmentation, the same argument holds for both ventricle assessments. Our proposals exhibit an improved result in comparison to CNN-based methods leading to the best LV segmentation, but failed to overcome substantially the non-trainable 4-PLS Texture-UI method for RV evaluation.

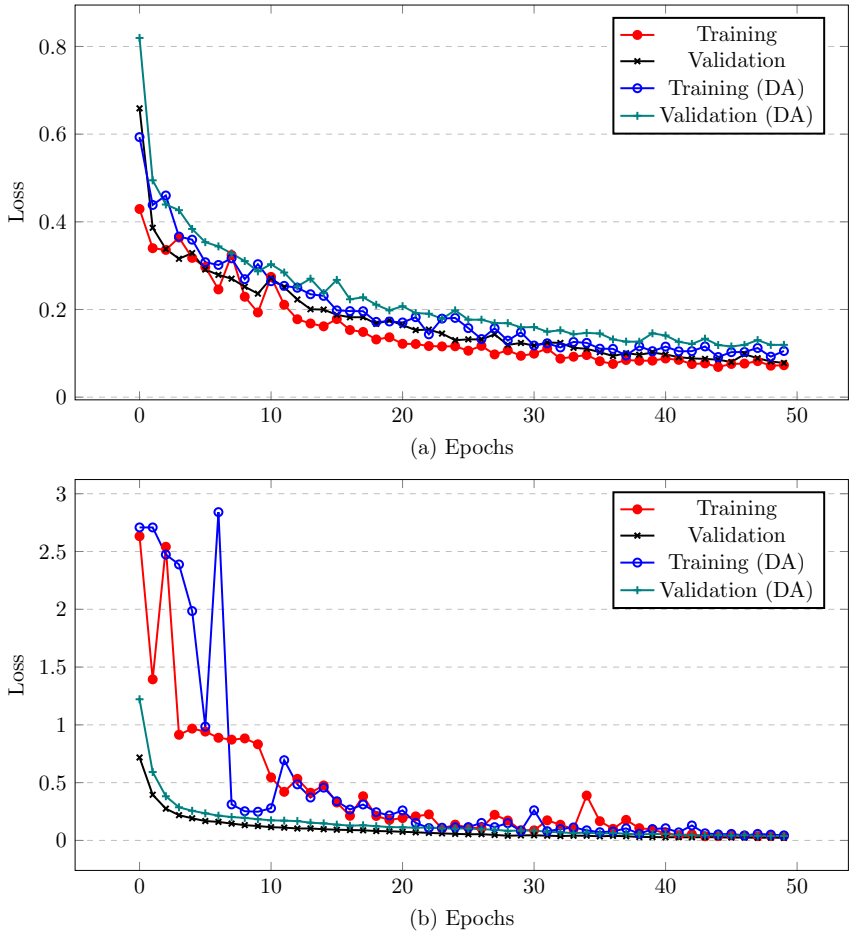



Figure 5.9: Training and validation loss for CT experiment in (a). Training and validation loss for MRI experiment in (b).

A visual comparison among methods for a single CT slice is shown in Figure 5.10. We observe that CNN-based methods fail to attain a full delineation, while methods that include AC succeed in filling missing information like small holes as if a morphological operation of dilation was applied. Moreover, hybrid proposals find it easier to avoid abrupt contour deformations due to the ability to adjust $\mu_1$ and $\mu_2$ parameters. Finally, similar performance can be seen in Figure 5.11 for a single MRI image. Methods facing the low contrast of this modality often struggle to differentiate ventricles. In this sense, texture-based models are better to differentiate tissue and do better segmentation jobs than intensity-valued techniques.

Table 5.3: Average performance for CT and MRI experiments

| Methods | 45 CT images | | | | 24 MRI images | | | |
|---|---|---|---|---|---|---|---|---|
| | Left Ventricle | | Right Ventricle | | Left Ventricle | | Right Ventricle | |
| | Dice | Hausdorff | Dice | Hausdorff | Dice | Hausdorff | Dice | Hausdorff |
| U-Net | 0.890 | 4.100 | 0.844 | 5.375 | 0.865 | 2.980 | 0.821 | 3.063 |
| U-Net-DA | 0.924 | 3.761 | 0.854 | 5.480 | 0.880 | 2.985 | 0.819 | 3.360 |
| 4-PLS Texture-UI | 0.905 | 3.875 | **0.890** | **4.306** | 0.905 | 2.677 | **0.873** | **2.891** |
| Hybrid IB | 0.927 | 3.698 | 0.860 | 5.037 | 0.907 | 2.825 | 0.848 | 3.046 |
| Hybrid TB | **0.933** | **3.621** | 0.861 | 5.356 | **0.915** | **2.524** | 0.854 | 3.368 |

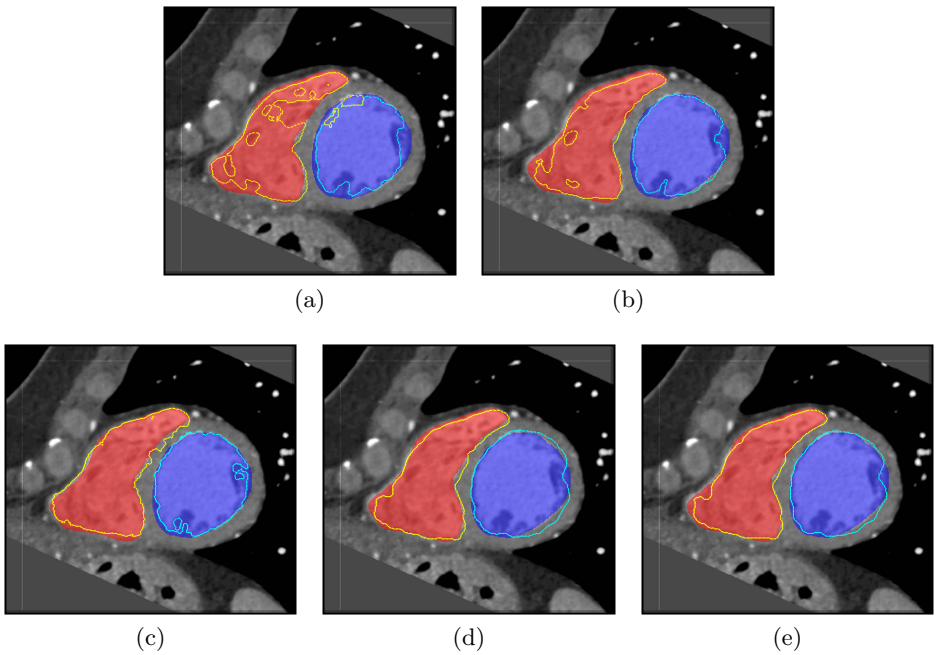
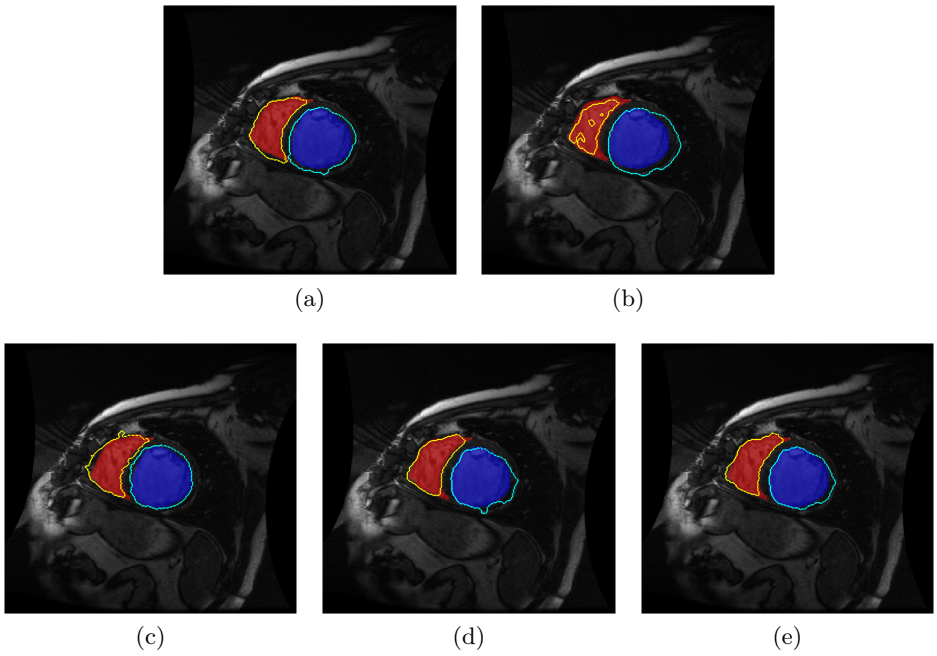


Figure 5.10: Graphic segmentation details of a CT slice. Expert annotations in blue for LV and red for RV. Results in cyan contour for LV and yellow contour for RV. U-Net in (a), U-Net-DA in (b), 4-PLS Texture-UI in (c), Hybrid IB in (d) and Hybrid TB in (e).



(a) (b) (c) (d) (e)

Figure 5.11: Graphic segmentation details of a MRI slice. Expert annotations in blue for LV and red for RV. Segmentation results delineated in cyan for LV and yellow for RV. U-Net in (a), U-Net-DA in (b), 4-PLS Texture-UI in (c), Hybrid IB in (d) and Hybrid TB in (e).

## 5.4 Future prospect

Both AC and CNN methods share a lot in common despite the fact that contrasting approaches are addressed at first sight. First of all, we discussed that the theory of calculus of variation involves analyzing functions to accomplish the proper minimization process, which is finally translated into convex solutions for reaching local or global minimum values. In this sense, the first variation of a functional is entirely computed by using derivatives. On the other hand, when working with any specified CNN architecture available in the literature, the loss function and optimizers are the keys to configure the learning process. These two elements (hyperparameters) are listed below:

- **Loss function** is the quantity that must be minimized in the training stage. Loss function represents a measure of success for certain required task.
- **Optimizer** determines how the network will be updated based on the loss function. Optimizers are actually implemented based on a variant of the gradient descent method.

A neural network with multiple outputs may have multiple loss functions, but the gradient-descent process must be based on a single scalar loss value. For the case of multiloss networks, the results of losses are averaged into a single scalar, see diagram of a classic CNN in Figure 5.12.

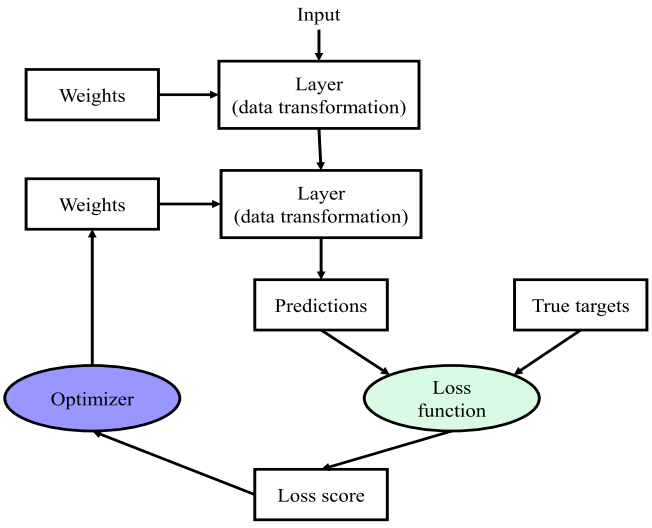



Figure 5.12: Basic diagram of a CNN meets loss function.

One of the most fundamental operations in deep learning is to use the score provided by the loss function as a feedback signal to update the weights, in a direction that the loss

score is decreasing for the current task. This adjustment is the job of the optimizer, which is closely related to the backpropagation algorithm. In practice, a CNN function consists of many tensor operations chained together, each of which has a simple derivative. Based on calculus, the chain of functions can be derived using the chain rule. Backward propagation is opposite to the forward propagation, see Figure 5.13, and it starts with the final loss value and works backward from the top to the bottom layers by applying the chain rule. This operation computes the contribution that each parameter has to the loss value.

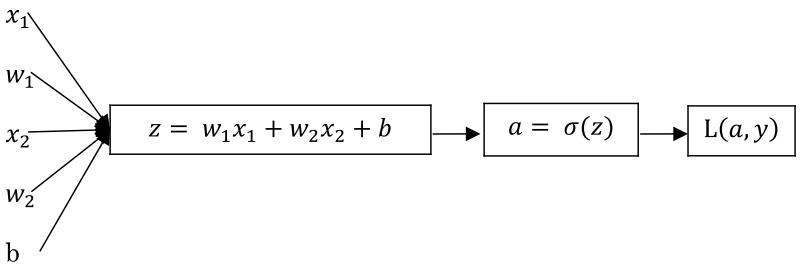



Figure 5.13: Basic diagram of the forward propagation process of a single-layer neural network.

In Figure 5.13, the term $L(a, y)$ represents the loss function in terms of the prediction $a$ and the true class $y$, $\sigma(z)$ is the activation function and $z$ is basically the linear combination of weights $w$, inputs $x$ and biases $b$. The process of updating weights is performed by $\frac{dL}{dw} = (\frac{dL}{da})(\frac{da}{dz})(\frac{dz}{dw})$.

The entire process of choosing the right objective function is extremely important since the network will take any shortcut to minimize the loss. In the case that the objective function is unable to accomplish a task with success, the CNN will produce undesired operations. All CNN-based methods will be as ruthless in lowering their loss function, so unintended side effects are the results of a bad choice.

When we face common problems such as classification, regression and sequence prediction, there is a simple guideline we can follow to choose the correct loss. For instance, binary cross entropy is often used for a binary classification problem, while categorical cross entropy for a multi-class classification problem, as we used in this work. Regarding the choice of a loss function, note that it is not always possible to directly optimize for the metric that measures success on a certain task. Loss functions need to be computable given only a mini-batch of data or a single data point and must be differentiable, otherwise it would be impossible to use backpropagation to train the network. The proposed multiphase formula based on level set and all its variants seems to meet these requirements with promising results.

# Chapter 6

# Conclusions

The proposal presented in this work conceives one of the possible solutions to overcome the limitation of both trainable and non-trainable methods. A model based on active contours was initially suggested which is able to segment distinct regions and be extended to multiple image representation like broad texture descriptions. This improvement allows us to admit texture features, whereas patterns are highlighted and well recognizable by the vector-valued multiphase formula proposed here. We also observed possible advantages with this framework in terms of robustness. For instance, texture-based approaches show signs of confidence when images are corrupted with Gaussian noise, as we showed in the first experiments.

We believe that the inclusion of hybrid models in the recent improvements of AI leads to fast and end-to-end automated tools, which are often in high demand today. Our experimental results demonstrate the effectiveness of our proposed methods, particularly with respect to reliability and implementation. In our main assessments, the accuracy was compared with the default state-of-the-art methods for segmenting left and right cardiac ventricles in both CT and MRI datasets, showing hybrid texture-based model an improvement over the rest of the methods in most cases.

Nonetheless, we face possible issues related to the complexity as more contours are requested. The number of regions to segment is linked to the number of active contours used. In this sense, it is evident that the computational cost increases as active contours are added to the system. Moreover, the number of parameters also augments. Another important issue to mention is the parameter setup, which is specially designed for the required dataset and application. Selecting default values is an option, but it does not guaranty the best performance. Nevertheless, this is a generalization problem which is commonly faced by sev-

eral machine learning-based algorithms. Unless stated otherwise, more attention should be directed towards overcoming issues of parameter tuning automatically.

Further refinements to the current formulation could yield boosted performance with many potential applications. Future work could involve how to address 3D data more efficiently due to the advantages brought by level sets, as well as to demonstrate the fulfilled requirements in order to be included within the optimization stage of neural network architectures.

## 6.1 Discussion: A brief of ethics in artificial intelligence

This supplementary section concerns a brief overview of implications in artificial intelligence-based systems. Among the whole range of significant implications in the field, ethics play a key role since human beings face several challenging situations, that must be translated into machine language to provide the desired solution in the decision-making process to avoid unintended side effects. Is there any dispute or disagreement regarding our interpretation?

The Turing test is an experiment for measuring the machine's ability to exhibit intelligence. Let us suppose that certain machine which is provided with AI tries to perform the role of human, while a human helps in the process of deciphering if it is a real human. In the case where the machine can cheat the human interrogator, then we can say that the machine has a certain degree of intelligence. Even tough, the latter statement does not necessarily imply that the machine can reason exactly like a human, but it would be almost impossible to distinguish if it actually does, since the human was previously deceived. Rather than cheating the interrogator successfully, the main idea behind this test is to evaluate how well a machine mimics the human reasoning, good enough to engage a natural conversation. In such a case, we could wonder that doing something like a human, makes you human?

Throughout the decade of 2020's, AI is projected to create 2.9 trillion of business value and 6.2 billion hours-working productivity globally. It means that under a partnership scheme between humans and AI, they will work to improve cognitive performance, which includes learning, decision-making and reaching new experiences, etc. This partnership is expected to reduce errors while providing comfort in performing a certain customized task. "The goal is to be more efficient with automation, while being complemented by a human touch and common sense to manage the risks of decision automation," says Svetlana Sicular, research vice president at Gartner [Gar, 2020].

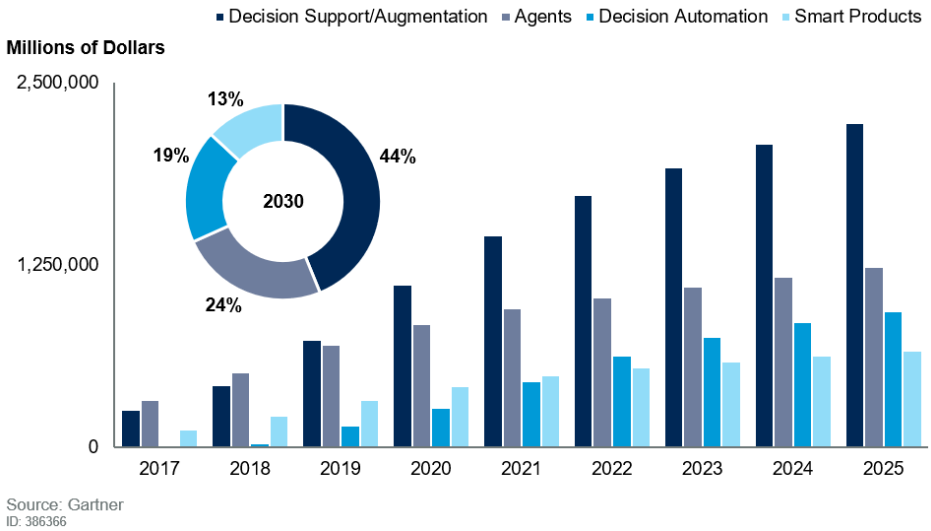



Figure 6.1: Worldwide Business Value by AI Type (Millions of Dollars). [Gar, 2020]

At this moment and so on, we humans will have to deal with the idea that some decisions will be taken by an AI. As in any other case, disagreements will arise due to the fact that such decisions might affect a group of people or the simple reason that we cannot understand the logic behind. Such is the hypothetical case of an autonomous car that faces the situation of saving the driver from an accident through two possible exits, one path that leads him to kill a child or another that leads him to kill an old man. Is AI ready to face such a brave situation? If an AI can provide an explanation of its actions, are human beings able to understand it? Undoubtedly, it is a dilemma that even the humans find it difficult to analyze, although we force machines to decide in a couple of seconds by following mathematical rules instead of human rules which are forgotten temporarily.

The concept of *black boxes* dates back to the battle of computationalists to provide eloquently a neural network definition by highlighting their lack of logic, which goes further than programming issues. In this sense, the lack of logic of AI-based systems relies on questioning their decision and compare them to those human-based decisions. Humans are influenced by various factors such as feelings of empathy and fear, reasoning, impulses and instincts, which in many cases help to perceive the risks of exercising certain actions. The dilemma arises when considering whether humans could trust in a system made up of black boxes whose responses could be found outside the human intellect. "We are heading towards a black future full of black boxes," says Selmer Bringsjord, scientist in the department of cognitive science at Rensselaer Polytechnic Institute.

Nevertheless, current AI systems are designed to optimize and this is a fundamental operation for solving problems. Indeed, different versions of neural networks found in the literature are completely based on optimizing certain loss function to find the minimum error between a true class and its prediction. In addition, the efficiency is measured in time or epochs (iteration cycles), so real-time and high-accurate systems are mostly required nowadays. Based on the aforementioned statement, the fact of proposing a stage that attempts to explain its actions, makes these systems have to delve into themselves and waste computing power trying to discern the correct way in which its actions could be explainable for us in understandable terms. In this way, an optimization-teaching paradox emerges for such systems: a machine which is designed to optimize cannot be undone (goes against optimizing) just because we do not understand a decision among all its processes.

At this time, even if someone could provide a reasonable explanation of the operation and decision-making of AI-based systems, it is highly probable that information remains incomplete, or perhaps we will reach a certain point where asking machines for an explanation is something that we will never be able to conceive. Undoubtedly, AI has reached us, enough to doubt of our own reasoning.

# Bibliography

[SUN, 2019] (2019). Sunnybrook cardiac data 2009.cardiac mr left ventricle segmentation challenge. Accessed: 2019-11-31.

[WHO, 2019] (2019). World health organization. cardiovascular diseases (cvd). 2019. Accessed: 2020-03-11.

[Gar, 2020] (2020). Gartner says ai augmentation will create 2.9 trillion of business value in 2021. Accessed: 2019-11-31.

[Badshah et al., 2020] Badshah, N., Rabbani, H., and Atta, H. (2020). On local active contour model for automatic detection of tumor in mri and mammogram images. *Biomedical Signal Processing and Control*, 60:101993.

[Barba-J et al., 2017] Barba-J, L., Escalante-Ramírez, B., Vallejo Venegas, E., and Arámbula Cosío, F. (2017). A 3d hermite-based multiscale local active contour method with elliptical shape constraints for segmentation of cardiac mr and ct volumes. *Medical & Biological Engineering & Computing*, 2017.

[Baumgartner et al., 2018] Baumgartner, C. F., Koch, L. M., Pollefeys, M., and Konukoglu, E. (2018). An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation. In Pop, M., Sermesant, M., Jodoin, P.-M., Lalande, A., Zhuang, X., Yang, G., Young, A., and Bernard, O., editors, *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*, pages 111–119, Cham. Springer International Publishing.

[Brodatz, 1966] Brodatz, P. (1966). Texture: a photographic album for artist and designers.

[Carbajal-Degante et al., 2020] Carbajal-Degante, E., Avendaño, S., Ledesma, L., Olveres, J., and Escalante-Ramírez, B. (2020). Active contours for multi-region segmentation with a convolutional neural network initialization. In Schelkens, P. and Kozacki, T., editors,

*Optics, Photonics and Digital Technologies for Imaging Applications VI*, volume 11353, pages 36 – 44. International Society for Optics and Photonics, SPIE.

[Carbajal-Degante et al., 2018] Carbajal-Degante, E., Olveres, J., and Escalante-Ramírez, B. (2018). A multiphase active contour model based on the Hermite transform for texture segmentation. In Schelkens, P., Ebrahimi, T., and Cristóbal, G., editors, *Optics, Photonics, and Digital Technologies for Imaging Applications V*, volume 10679, pages 364 – 372. International Society for Optics and Photonics, SPIE.

[Caselles et al., 1997] Caselles, V., Kimmel, R., and Sapiro, G. (1997). Geodesic active contours. *International Journal of Computer Vision*, 22:61–79.

[Chan et al., 2006] Chan, T., Esedoglu, S., and Nikolova, M. (2006). Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal of Applied Mathematics*, 66:1632–1648.

[Chan et al., 2000] Chan, T. F., Sandberg, B., and Vese, L. A. (2000). Active contours without edges for vector-valued images. *Journal of Visual Communication and Image Representation*, 11(2):130 – 141.

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

[Chollet, 2017] Chollet, F. (2017). *Deep Learning with Python*. Manning Publications Co., USA, 1st edition.

[Cootes et al., 2001] Cootes, T., Edwards, G., and Taylor, C. (2001). Active appearance models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23:681 – 685.

[Cootes et al., 1995] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1995). Active shape models: their training and application. *Comput. Vis. Image Underst.*, 61(1):38–59.

[Despotović et al., 2015] Despotović, I., Goossens, B., and Philips, W. (2015). Mri segmentation of the human brain: Challenges, methods, and applications. *Computational and Mathematical Methods in Medicine*, 2015:23.

[Dolz et al., 2015] Dolz, J., Massoptier, L., and Vermandel, M. (2015). Segmentation algorithms of subcortical brain structures on mri for radiotherapy and radiosurgery: A survey. *IRBM*, 36(4):200 – 212.

[Estudillo-Romero and Escalante-Ramirez, 2011] Estudillo-Romero, A. and Escalante-Ramirez, B. (2011). Rotation-invariant texture features from the steered Hermite transform. *Pattern Recognition Letters*, 32(16):2150 – 2162. Advances in Theory and Applications of Pattern Recognition, Image Processing and Computer Vision.

[Freeman and Adelson, 1991] Freeman, W. T. and Adelson, E. H. (1991). The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9):891–906.

[Garcia-Uceda et al., 2019] Garcia-Uceda, A., Selvan, R., Saghir, Z., and de Bruijne, M. (2019). A joint 3d unet-graph neural network-based method for airway segmentation from chest cts.

[Getreuer, 2012] Getreuer, P. (2012). Chan-Vese Segmentation. *Image Processing On Line*, 2:214–224.

[Gibson, 1950] Gibson, J. J. (1950). *The Perception Of The Visual World*. Boston: Houghton Mifflin.

[Gonzalez and Woods, 2018] Gonzalez, R. and Woods, R. (2018). *Digital Image Processing*. Pearson.

[González-Villà et al., 2016] González-Villà, S., Oliver, A., Valverde, S., Wang, L., Zwiggelaar, R., and Llado, X. (2016). A review on brain structures segmentation in magnetic resonance imaging. *Artificial Intelligence in Medicine*, 73:45–69.

[Han and Wu, 2017] Han, B. and Wu, Y. (2017). A novel active contour model based on modified symmetric cross entropy for remote sensing river image segmentation. *Pattern Recognition*, 67:396 – 409.

[Jain and Farrokhnia, 1991] Jain, A. K. and Farrokhnia, F. (1991). Unsupervised texture segmentation using gabor filters. *Pattern Recogn.*, 24(12):1167–1186.

[jing Zong et al., 2019] jing Zong, J., shuang Qiu, T., dong Li, W., and mei Guo, D. (2019). Automatic ultrasound image segmentation based on local entropy and active contour model. *Computers and Mathematics with Applications*, 78(3):929 – 943.

[Kapur et al., 1996] Kapur, T., Grimson, W. L., Wells, W. M., and Kikinis, R. (1996). Segmentation of brain tissue from magnetic resonance images. *Medical Image Analysis*, 1(2):109 – 127.

[Kass et al., 1988] Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *INTERNATIONAL JOURNAL OF COMPUTER VISION*, 1(4):321–331.

[Keatmanee et al., 2019] Keatmanee, C., Chaumrattanakul, U., Kotani, K., and Makhanov, S. S. (2019). Initialization of active contours for segmentation of breast cancer via fusion of ultrasound, doppler, and elasticity images. *Ultrasonics*, 94:438 – 453.

[Kumar et al., 2012] Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S., Schabath, M., Forster, K., Aerts, H., Dekker, A., Fenstermacher, D., Goldgof, D., Hall, L., Lambin, P., Balagurunathan, Y., Gatenby, R., and Gillies, R. (2012). Radiomics: the process and the challenges. *Magnetic resonance imaging*, 30(9):1234–1248.

[Li et al., 2011] Li, C., Huang, R., Ding, Z., Gatenby, J. C., Metaxas, D. N., and Gore, J. C. (2011). A level set method for image segmentation in the presence of intensity inhomogeneities with application to mri. In *IEEE Transactions on Image Processing 20 (7) (2011) 2007–2016. Sen, S.K. Pal / Information Sciences 191 (2012) 169–191 191.*

[Liu et al., 2013a] Liu, G., Sun, X., Fu, K., and Wang, H. (2013a). Interactive geospatial object extraction in high resolution remote sensing images using shape-based global minimization active contour model. *Pattern Recognition Letters*, 34(10):1186 – 1195.

[Liu et al., 2013b] Liu, L., Zeng, L., and Luan, X. (2013b). 3d robust chan–vese model for industrial computed tomography volume data segmentation. *Optics and Lasers in Engineering*, 51(11):1235 – 1244.

[Marr, 1982] Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Henry Holt and Co., Inc., USA.

[Martens, 1990] Martens, J. B. (1990). The Hermite transform-theory. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(9):1595–1606.

[Masnou, 2002] Masnou, S. (2002). Disocclusion: a variational approach using level lines. *IEEE Transactions on Image Processing*, 11(2):68–76.

[Moreno et al., 2014] Moreno, J. C., Surya Prasath, V., Proença, H., and Palaniappan, K. (2014). Fast and globally convex multiphase active contours for brain mri segmentation. *Computer Vision and Image Understanding*, 125:237 – 250.

[Mumford and Shah, 1989] Mumford, D. and Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5):577–685.

[Nabizadeh and Kubat, 2017] Nabizadeh, N. and Kubat, M. (2017). Automatic tumor segmentation in single-spectral mri using a texture-based and contour-based algorithm. *Expert Systems with Applications*, 77:1 – 10.

[Oktay et al., 2018] Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., and Rueckert, D. (2018). Attention u-net: Learning where to look for the pancreas.

[Olaf Ronneberger and Brox, 2015] Olaf Ronneberger, P. F. and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597.

[Olver, 2014] Olver, P. (2014). Introduction to the calculus of variations.

[Olveres et al., 2017] Olveres, J., Carbajal-Degante, E., Escalante-Ramírez, B., Vallejo, E., and García-Moreno, C. M. (2017). Deformable models for segmentation based on local analysis. *Accepted at Mathematical Problems in Engineering*, 2017.

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

[Ouabida et al., 2017] Ouabida, E., Essadique, A., and Bouzid, A. (2017). Vander lugt correlator based active contours for iris segmentation and tracking. *Expert Systems with Applications*, 71:383 – 395.

[Paragios and Deriche, 2000] Paragios, N. and Deriche, R. (2000). Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(3):266–280.

[Petitjean and Dacher, 2011] Petitjean, C. and Dacher, J.-N. (2011). A review of segmentation methods in short axis cardiac MR images. *Medical Image Analysis*, 15(2):169–184.

[Petrou, 2010] Petrou, M. (2010). *Image Processing : the Fundamentals*. Wiley.

[Petrov and Sizikov, 2005] Petrov, Y. and Sizikov, V. (2005). *Well-Posed, Ill-Posed, and Intermediate Problems with Applications*.

[Radau et al., 2009] Radau, P., Lu, Y., Connelly, K., Paul, G., Dick, A., and Wright, G. (2009). Evaluation framework for algorithms segmenting short axis cardiac mri.

[Rebouças Filho et al., 2017] Rebouças Filho, P. P., Cortez, P. C., da Silva Barros, A. C., C. Albuquerque, V. H., and R. S. Tavares, J. M. (2017). Novel and powerful 3d adaptive crisp active contour method applied in the segmentation of ct lung images. *Medical Image Analysis*, 35:503 – 516.

[Rousson and Paragios, 2002] Rousson, M. and Paragios, N. (2002). Shape priors for level set representations. In *Computer Vision — ECCV 2002*, pages 78–9, Berlin, Heidelberg. Springer Berlin Heidelberg.

[Sapiro and Ringach, 1996] Sapiro, G. and Ringach, D. (1996). Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Transactions on Image Processing*, 5:1582–1586.

[Silvan-Cardenas and Escalante-Ramirez, 2006] Silvan-Cardenas, J. L. and Escalante-Ramirez, B. (2006). The multiscale Hermite transform for local orientation analysis. *IEEE Transactions on Image Processing*, 15(5):1236–1253.

[Spencer, 2016] Spencer, J. (2016). *Variational methods for image segmentation*. PhD thesis, University of Liverpool.

[Svetitsky, 2005] Svetitsky, B. (2005). Notes on functionals.

[van Dijk and Martens, 1997] van Dijk, A. M. and Martens, J.-B. (1997). Image representation and compression with steered Hermite transforms. *Signal Processing*, 56(1):1 – 16.

[Vese and Chan, 2002] Vese, L. A. and Chan, T. F. (2002). A multiphase level set framework for image segmentation using the mumford and shah model. *International Journal of Computer Vision*, 50(3):271–293.

[Yagi et al., 2000] Yagi, Y., Brady, M., Kawasaki, Y., and Yachida, M. (2000). Active contour road model for smart vehicle. In *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, volume 3, pages 811–814 vol.3.

[Özgün Çiçek et al., 2016] Özgün Çiçek, Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). 3d u-net: Learning dense volumetric segmentation from sparse annotation.

[Zhou et al., 2018] Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., and Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. *CoRR*, abs/1807.10165.

[Zimmer et al., 2002] Zimmer, C., Labruyere, E., Meas-Yedid, V., Guillen, N., and Olivo-Marin, J. . (2002). Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: a tool for cell-based drug testing. *IEEE Transactions on Medical Imaging*, 21(10):1212–1221.