NON-PARAMETRIC COUPLED SHAPE PRIORS FOR SEGMENTATION OF DEFORMABLE OBJECTS IN TIME-SERIES IMAGES USING PARTICLE FILTERS

by

Naeimeh Atabakilachini

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Master of Science

Sabanci University
August 2018

NON-PARAMETRIC COUPLED SHAPE PRIORS FOR SEGMENTATION OF DEFORMABLE OBJECTS IN TIME-SERIES IMAGES USING PARTICLE FILTERS

APPROVED BY

Assoc. Prof. Dr. Müjdat ÇETİN   ....

(Thesis Supervisor)
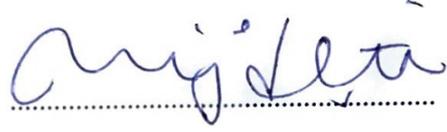
Assoc. Prof. Dr. Devrim ÜNAY   ....

Asst. Prof. Dr. Lavdie RADA   ....

DATE OF APPROVAL: ....31/07/2018................

...to my parents

Acknowledgments

As I conclude my Masters Thesis, I feel a sense of accomplishment and honor to have been associated with Sabanci Universitys Faculty of Engineering and Natural sciences and its excellent faculty members.

My graduate study at Sabanci has truly been a learning experience, which has enriched me with key technical and life skills, perseverance, self-confidence, and last but not the least, a strong network of friends and mentors.

Firstly, I would like to extend my sincerest gratitude to my advisor, Müjdat Çetin, whose continued support, guidance, encouragement, counseling and mentorship allowed me to accomplish my academic aspirations. I seek life-long inspiration from his work ethics and his academic values.

My journey would not have been possible without the relentless support of Ertunç Erdil, who shared his expertise and thorough knowledge with me. His valuable feedback and problem solving skills enabled me to handle all major impediments in reaching at the desired outcomes.

My Masters was funded by the TÜBİTAK[1]. Scholarship, and I would like to especially thank them for providing me this platform to show case my talents and realize my potential. I honestly believe that TÜBİTAK is doing a commendable job by providing this platform to aspiring academicians who aim to contribute towards cutting-edge research.

During the course of my journey, I made some special friends from the VPA Lab and SPIS Lab, who became part of my extended family at Sabanci. I am grateful to them for their unconditional help and for making my time so memorable.

Finally, I want to thank my family, who have always supported me in my endeavors and have stood beside me through thick and thin. This would not have

---

[1]Scientific and Technological Research Council of Turkey

been possible without their unconditional love.

NON-PARAMETRIC COUPLED SHAPE PRIORS FOR SEGMENTATION OF DEFORMABLE OBJECTS IN TIME-SERIES IMAGES USING PARTICLE FILTERS

Naeimeh Atabakilachini

EE, M.Sc. Thesis, 2018

Thesis Supervisor: Müjdat Çetin

Abstract

Segmentation is usually the first step in image processing and directly impacts the success or failure of the image analysis algorithm. It turns into a very challenging problem when the observed image suffers from insufficiencies, such as high level of noise, clutter, data loss and, occlusion. The effect of prior knowledge has been widely studied in the curve evolution-based models and it has been proved that utilization of some kind of prior information obtained by exploiting the known features of the object to be segmented, can aid the result of the segmentation process.

In this thesis, shape, which is a favorable attribute of the object to be segmented, is used to form the prior information. The proposed method has been developed based on a sampling approach using sequential Monte Carlo (Particle Filters) and In order to enrich the segmentation model, a new term is introduced which we refer to as coupled shape priors. By involving a curve evolution step into the sampling process, the coupled shape prior term, takes part in the proposed energy functional defined for the curve evolution step and incorporates the temporal shape dependencies. The proposed method has been evaluated on three different datasets, a deforming synthetic dataset, hand gesture dataset and 2-photon microscopy images

of dendritic spines and, according to both visual and quantitative results, it has been demonstrated that it has a successful performance in segmentation of deforming objects whose shapes come from multi-modal shape densities. Also it has been shown that the proposed method is able to handle low quality images, highly noisy images, images with data loss, and occluded images.

ZAMAN SERİSİ GÖRÜNTÜLERİNDEKİ DEFORME OLAN NESNELERİN PARÇACIK SÜZGECİ VE PARAMETRİK OLMAYAN BİRLEŞİK ŞEKİL ÖNSELİ KULLANILARAK BÖLÜTLENMESİ

Naeimeh Atabakilachini

EE, M.Sc. Thesis, 2018

Tez danışmanı: Müjdat Çetin

Özet

Bölütleme, genellikle bir görüntü işlem algoritmasının başarımını doğudan etkileyen bir adımdır. Yüksek seviye gürültü, veri kaybı ve nesnelerin üst üste gelmesi gibi durumlarda bölütleme problemi daha zorlu bir problem haline gelir. Şekil ön bilgisini kullanmanın etkisi eğri geliştirme tabanlı yöntemlerde çok çalışılmış olup, bu tür yöntemlerin bölütleme sonuçlarını önemli ölçüde iyileştirildiği kanıtlanmıştır.

Bu tezde, bölütlenecek olan nesneye ait önemli bir özellik olan şekil, bölütleme sürecinde bir ön bilgi olarak kullanılmaktadır. Önerilen yöntem, örnekleme tabanlı sıralı Monte Carlo (parçacık filtresi) yöntemlerine dayanmaktadır. Bölütleme sonuçlarını iyileştirmek için birlştirilmiş şekil ön bilgisini kullanan bir terim önerilmiştir. Örnekleme sürecine eğri geliştirme süreci eklenerek, birleştirilmiş şekil önseli terimi, zamansal şekil bağlantılarını içeren enerji fonksiyonuna eklenir. Önerilen yöntem şu üç farklı veri kümesinde test edilmiştir: zaman içinde deforme olan sentetik veri kümesi, el hareketleri veri kümesi ve 2-foton mikroskop veri kümesidir. Görsel ve sayısal sonuçlara göre önerilen yöntem, zaman icinde şekil değiştiren ve önsel dağılımı çok doruklu olan veri kümelerinde başarılı bölütleme sonuçları üretmiştir. Bunun yanında önerilen yöntemin düşük kaliteli, yüksek gürültülü ve veri kaybı olan

görüntülerde yüksek başarım sağladığı gösterilmiştir.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this thesis, we consider the problem of segmentation of objects with dynamically evolving shapes from observed image sequences. We propose a principled dynamic segmentation approach that involves ideas from shape priors and Sequential Monte Carlo sampling and demonstrate its effectiveness on several potential applications. In this chapter we define and motivate the problem addressed, list the technical contributions of this thesis, and provide its outline.

1.1 Motivation

Image segmentation is a classic problem in computer vision and among the most studied problems in image understanding and computer vision. In a broad sense, it is defined as the process of partitioning the image pixels into meaningful groups with homogeneous attributes. The typical goal of image segmentation is to distinguish between background and objects in the foreground or delimit object boundaries. Generally, segmentation is the first task in image analysis and in many applications, the success or failure of the segmentation algorithm has a direct impact on the success or failure of the image analysis algorithm.

Image segmentation has been widely applied in many disciplines and machine vision tasks such as object detection (e.g., pedestrian detection, localization of objects in satellite images), recognition (e.g., face recognition, fingerprint recognition, license plate recognition), medical imaging, occlusion boundary estimation within motion or stereo systems, image compression, image editing, image retrieval and is a building block for many applications such as self-driving cars, robotics, and human-computer interaction.

There are many scene attributes based on which, the segmentation is usually performed, such as intensity, color, or texture similarities, pixel continuity, and higher level knowledge about the object model. Many existing segmentation approaches exploit either the discontinuities (which are called edge-based methods) or homogeneity (which are called region-based methods) in the image.

Particularly, early approaches for image segmentation were based on detecting edges using filters, such as the Sobel filter [1] and the Canny filter [2]. Object boundaries were identified by classifying the pixels as edge/non-edge pixels based on a threshold. Region growing, split-and merge, and certain histogram-based techniques were also among the developed schemes for image segmentation [3]. A more comprehensive list can be found in [4] and [5].

There are factors that can affect the results of a segmentation method, such as object orientation, shape variability, the presence of extraneous details, degradations due to poor illumination, noise, occlusions, and sampling artifacts. Unfortunately, generic low-level segmentation algorithms often do not provide accurate and desired segmentation results, since purely low-level assumptions such as intensity or texture homogeneity and strong edge contrast are not sufficient to separate objects from the background in a scene. Thus, when we are dealing with noisy images, weak or missing edges, cluttered data, and occlusions, these approaches are not successful.

Curve evolution-based image segmentation models are one of the developed methods for image segmentation and have been used extensively for image segmentation since their introduction in [6]. In curve evolution-based approaches, image segmentation is posed as an optimization problem and an energy functional is defined based on features of the image. Minimizing the energy refines the curve and captures the boundary of the object of interest. These methods are effective over a broad class of images. In order to define the energy function, some methods have used boundary information for the object of interest [7],[8] and some have used regional information, such as intensity statistics [9], [10], [11].

In many recent active contour models, some type of prior knowledge is used to enrich their models and generate more accurate results. This makes the segmentation algorithm more robust to imperfections in the observed image data. Prior knowledge can be formed by exploiting known features of the object of interest.

2

One of these features that have been utilized in many works is the *shape*, to which we refer as the geometric outline of an object in 2D or 3D.

One can distinguish various kinds of shape knowledge [12].

- low-level shape priors which favor smaller boundary length.

- mid-level shape priors which favor thin and elongated structures.

- high-level shape priors which favor similarity to previously observed shapes.

There are numerous existing automated segmentation methods that enforce constraints on the underlying shapes.

In [13] a mathematical formulation to constrain an implicit surface to follow global shape consistency while preserving its ability to capture local deformations is proposed. In [14] and [15], an average shape and modes of variation through principal component analysis (PCA) is used in the model which captures the variability of shapes. However, this technique can handle only unimodal, Gaussian-like shape densities. As a solution to this limitation, [16] and [17] have introduced methods based on nonparametric shape densities learned from training shapes. In these works, the assumption is that the training shapes are drawn from an unknown shape distribution and this distribution is estimated by extending a Parzen density estimator to the space of shapes. They formulate the segmentation problem as a maximum *a posteriori* (MAP) estimation problem, where they use a nonparametric shape prior.

Another challenge that arises in some applications is to segment an object which is not static, in particular, objects that show variation in their shapes in different cases. For instance, a human liver has a different shape depending on the subject and their age. We also can define non-static object as one whose shape evolves over time. An example for this case can be left ventricle of the heart which goes through deformation across a cardiac cycle. Consequently, difficulties are faced when we want to model and infer its shape. In order to provide promising results in these challenging tasks, many segmentation approaches have been proposed in the literature which can be found in [18] in a categorized form.

However, typical active contour based algorithms are set in a static framework, therefore this can be a major drawback when segmenting or tracking a dynamically

3

deforming contour over time is desired, since they do not incorporate the deformation dynamics into their segmentation or tracking approach.

1.2   Problem Statement and Thesis Contribution

In this thesis, the problem that we focus on is the segmentation of deforming objects in time series images. In other words, the objects that we deal with undergo deformation over time and we propose a new method to segment the object of interest in the presence of high level of noise and data loss. Another challenge that we address in our approach is to effectively model multi-modal shape densities. In many applications, object shapes come from multiple classes. For example, the presence of different objects in a natural scene which includes cars, planes, trees, etc. poses a difficult segmentation task and the algorithm should ideally be able to recognize the class of the object in the scene in order to exploit information about the shape of that particular object type. Our approach considers segmentation problems that involve limited and challenging image data together with complex and multi-modal shape densities.

In order to overcome these limitations, we incorporate shape priors into our active contour-based segmentation method. By forming a training dataset which contains objects from different classes at different time points, the dynamics of the deformation is learned and shape prior knowledge is imposed to produce more accurate results.

In our preliminary work, we introduce a coupled shape prior term that is included in the energy functional we propose for image segmentation. This term incorporates the dependence between the shapes of an object in consecutive time points. Therefore, if information about the shape of the object to be segmented in the previous time point is provided, the coupled shape prior force can aid the algorithm to evolve the curve towards the accurate class of the object, exploiting the information from the previous time point.

As the extension of our work, we concentrate on the problem in a more general form, considering that the knowledge that we have about the shape of the object from previous time points is imperfect. In order to construct our method in this setting, we use particle filters, also known as sequential Monte Carlo (SMC) meth-

ods. Particle filters enable the estimation of the posterior distribution by generating samples. In this way, we propose a complete dynamic segmentation approach.

1.3 Thesis Outline

This thesis contains 5 chapters and is organized as follows. Chapter 2 provides the background and preliminaries that underlie the work presented in subsequent chapters. It begins with an introduction about deformable models and then a description of the segmentation problem via Bayesian inference followed by a brief review of recursive estimation and particle filters. Chapter 3 presents the mathematical derivation of our dynamic coupled shape prior model based on particle filters. Experimental results including visual and quantitative results can be found in Chapter 4. In Chapter 5 we conclude by summarizing the contributions of this thesis. We also suggest some possible extensions and future research directions.

Chapter 2

Background

In this chapter, we provide background material required to develop our algorithm. In Section 2.1, we present a review about deformable models in image segmentation. In Section 2.2, we provide information about image segmentation via Bayesian inference, static and dynamic shape priors, and density estimation. In Section 2.3 we present an overview of recursive estimation and particle filters.

2.1   Deformable Models

Deformable models have become a landmark in computer vision and have been widely used in image segmentation and tracking. Deformable models are capable of controlling the geometry and smoothness of the segmented boundaries, and allowing variabilities of the objects. In deformable model formulations, energy functions usually consist of two terms which are responsible to evolve the curve. The external energy term is obtained from the image information, such as intensity level and intensity gradient so that the curve or surface is driven towards the desired images features, such as strong edges within an image. The internal energy term is designed to enforce the smoothness of the curve or surface during deformation.

Deformable models can be generally divided into two classes, depending on the definition of the curve and surface [19]. The classes are as follows:

- parametric deformable models through which the contour is represented explicitly.

- non-parametric deformable models that represent the surface implicitly as the level set of a higher dimensional scalar function.

Active contours or snakes [6] is one of the most popular deformable models, which is usually used in 2D piecewise continuous and smooth contour segmentation. In the Snake algorithm, the contour is defined as a parameterized curve with a fixed topology and the contour or the surface is represented by a finite set of parameters explicitly, such as the spatial positions of points on the curve. There are some disadvantages regarding this approach. One is that the connectivity of the points on the surface is difficult to maintain and they can change during the evolution, also they may cross each other if they come too close. Another concern is that in order to reconstruct the surface, the discretization should be fine enough. A solution that can be proposed to deal with such problems is to redistribute, add or remove the points if required after few time steps. However, this turns into a complicated task, especially in three dimensions. Moreover, in general, the parametric approach is not capable of handling topology changes. To overcome this problem, some constraints are imposed for detecting possible merging and splitting of contours [20].

Level set [21] is another important deformable model and it was developed to overcome major drawbacks of the classical explicit curve evolution models. It implicitly represents shape as the zero level of a level set function. In particular, the boundaries are embedded in a higher dimensional space. Then the higher dimensional level set function is evolved rather than the boundary itself. In this way, level set achieves flexible topological changes.

In the following we discuss Active Contours and Surfaces and Level set methods in more detail.

2.1.1 Active Contours and Surfaces

Considering explicit representation, a parametric curve is represented using a function $C(s) = (x(s), y(s))$, where x, y denote the coordinates and $s \in [0, 1]$ is the arc-length parameter. Given an initialization, the parameterized curve deforms in order to minimize the following energy function [19]:

$$E(C) = E_{int}(C) + E_{ext}(C) \tag{2.1}$$

where:

$$E_{int}(C) = \alpha \int_0^1 \|C'(s)\|^2 ds$$

$$\beta \int_0^1 \|C''(s)\|^2 ds \tag{2.2}$$

$$E_{ext}(C) = \gamma \int_0^1 \|\nabla I(C(S))\| ds \tag{2.3}$$

where I is the image intensity and α, β, γ are corresponding weighting coefficients for each energy term. α and β control the tension and rigidity of the curve respectively.

As it can be seen in Equation (2.2) and (2.3), the internal energy $E_{int}(C)$ characterizes the tension or the smoothness of the contour and the external energy $E_{ext}(C)$ is responsible for attracting the curve towards the object appearing in the image [19], [22].

Parametric deformable models have the following advantages:

- They are capable of handling closed or open parametric curves or surfaces.

- They have low computational complexity.

- They can integrate prior knowledge.

On the other hand, requiring an initialization is a disadvantage of this approach. Balloons proposed in [23] use a pressure force to increase the attraction range which makes the model less sensitive to initialization. The approach was extended this approach by defining the external energy using a distance map in [24].

However, the most important limitation of parametric models is their weakness in coping with topological changes.

2.1.1.1 Geodesic Active Contours

As we discussed before, there is a possibility that contour points intersect or overlap while evolving the curve and to a way to avoid that is to generate a new curve with a similar geometry but with a new parameterization. Geodesic active contours (GAC), also called Geometric active contours, were introduced in [7] and [25] based on implicit representation, in order to overcome the need of re-parameterization. Being independent of parameterization is the reason which makes this model capable

of handling topological changes automatically. The energy functional of this method
is given by:

$$E(C) = \int_0^1 g(I(C(s)))ds \qquad (2.4)$$

and:

$$g = \frac{1}{1 + \|\nabla G_\sigma(s) * I(s)\|^2} \qquad (2.5)$$

where $G_\sigma(s) * I(s)$ is the convolution of the image I with a Gaussian filter and
the function g ,which is an edge indicator function, contains information about the
gradient of the image and I and serves as a stopping function.

The main idea of GAC is to couple the speed of deformation with the image data,
so that the evolution of the curve stops at object boundaries. A problem with this
model is that if the object boundary has gaps or does not provide high gradient, the
curve passes through the boundary and cannot be fitted to the accurate boundary.
Moreover, these method is still sensitive to local minima, since it is based on the
image gradient.

An alternative approach is to use image region characteristics. Earlier, energy-
based segmentation frameworks ,such as [26], define a functional which measures the
consistency of the segmented regions. Based on these framework, many level sets
approaches have been developed to combine image region statistics with boundary
measurements, such as [9], [15], which we discuss more in Section (2.1.1.2).

2.1.1.2 Active Contours Without Edges

We have mentioned that both the Snakes and GAC models are based on the
object's edges assuming that sharp gradients at the shape's boundaries define the
object. To address this problem an alternative approach is to use the information
of the image region. Many level set based approaches have been proposed to com-
bine region characteristics ,such as first and second order intensity statistics, with
boundary information.

Based on this approach, there are methods developed base on Mumford-Shah
functional (MS) [27] which is given as:

$$E(f, C) = \int_\Omega (f - I)^2 ds + \gamma \int_{\Omega - C} \|\nabla f\|^2 ds + \alpha \|C\| \qquad (2.6)$$

where Ω is the image domain and f is the piecewise smooth approximation of the image I. Putting constrains on the smoothness of the curve within the sub-regions and the length of the curve of the sub-regions, this functional minimizes the difference between I and f.

Based on a level-set formulation of the piecewise constant variant of MS, methods in [9] and [28] developed which proposes the following model, considering the image to be formed of two regions with distinct constant intensities:

$$\min_{S, c_1, c_2} \int_{\Omega \backslash S} |I(x) - c_1|^2 dx + \int_S |I(x) - c_2|^2 dx + \nu |\partial S| \qquad (2.7)$$

where S is the object to be segmented, $\Omega \backslash S$ denotes the background, $\nu |\partial S|$ is weighted boundary length of S, and c_1 and c_2 are mean intensity values inside and outside the regions.

Using the level set representation, the resulting model is [9],[28]:

$$\min_{\phi, c_1, c_2} \int_\Omega |I(x, y) - c_1|^2 H(\phi) dx + \int_\Omega |I(x, y) - c_2|^2 (1 - H(\phi)) dx + \nu \int_\Omega |\nabla H(\phi)| dx$$
$$(2.8)$$

where H is the Heaviside function. This model is known as Active Contours Without Edges (ACWE) or Chan-Vese (CV) and many approaches has been proposed based on it.

There are many drawbacks related to both edge-based and region-based models, especially in case of noisy and distorted images. Efforts to address these limitations are made in [29], [30], [31], [32], [33] which combine edge and region characteristics.

2.1.2   Curve Evolution Using Level Set Method

In this section more detailed information is provided regarding level set method and its usage in curve evolution.

As we mentioned before, classical parametrized active contours require extra processing to handle the issues such as overlapping or intersecting of the boundary

points. Level set theory was first introduced by Osher and Sethian in 1988 [34] and has been widely used in a variety of applications in computer vision and image processing and mainly in segmentation. The principal advantage of the level set method is the fact that it is based on an implicit curve representation and performs curve evolution regardless of topological changes such as splitting and merging [35].

Earlier, we showed the explicit representation of a curve in the plane as the function $C(p) = (x(p), y(p))$, where x, y denote the coordinates and $p \in [0, 1]$. Each value of p corresponds to a unique point on the curve (See Figure 2.1) and if C is a closed curve we have $C(0) = C(1)$. There are some basic properties of C listed below. basic properties of the parameterized curve C is listed below [36]:
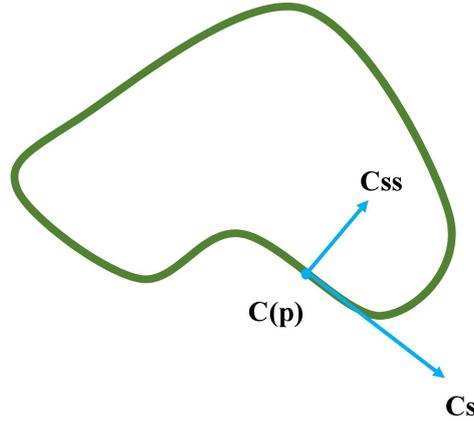
Figure 2.1: Explicit representation of a curve on the plane: $C(p)$ is a point on the curve C. Cs and Css represent the tangent vector and normal vector respectively

- The unit tangent vector T, which is defined as: $\overrightarrow{T} = \frac{C_p}{|C_p|}$, where $C_p = \frac{\partial C}{\partial p}$

- the curvature k, which is defined as: $C_{ss} = k\overrightarrow{N}$

Curvature is a measurement which shows the change in direction of C_s

Another representation of a curve can be by means of the level set of a function ϕ, which is

$$C = \{(x, y) | \phi(x, y) = 0\} \tag{2.9}$$

11

One of the alternative functions which can be defined as ϕ is the signed distance function, meaning that the value of ϕ is set as the positive distance of (x, y) to the curve C if the point is enclosed by the curve, and as the negative distance of the point to the curve C if the point is outside the curve (see Figure 2.2).
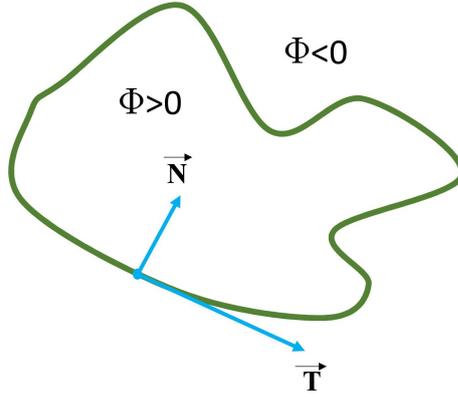
Figure 2.2: Implicit representation of a curve on the plane.

Tangent, normal and curvature are computed in the implicit framework using ϕ as follows [36]:

$$\overrightarrow{N} = -\frac{\nabla \phi}{|\nabla \phi|} \tag{2.10}$$

$$\overrightarrow{T} = \frac{\overline{\nabla} \phi}{|\nabla \phi|} \tag{2.11}$$

where the bar stands for:

$$\overline{\begin{bmatrix} a \\ b \end{bmatrix}} = \begin{bmatrix} b \\ a \end{bmatrix} \tag{2.12}$$

$$k = div(\frac{\nabla_\phi}{|\nabla_\phi|}) \tag{2.13}$$

Now the curve evolution in the implicit setting is formulated as follows defining the curve as $C := \{(x(s,t), y(s,t)) : \phi(x(s,t), y(s,t), t) = 0\}$, $\phi(x, y, t) : R^2 \longrightarrow R$:

The equation below demonstrates the relationship between the velocity V of points on the curve C at the time t:

$$\frac{dC}{dt} = V \overrightarrow{N} \tag{2.14}$$

And:

$$\frac{d\phi}{dt} = V |\nabla \phi| \tag{2.15}$$

Using all the obtained equations based on ϕ enables us to formulate the curve evolution according to ϕ. Thus, each point on the curve moves under the velocity V which is related to the level set by Equation (2.15).

2.2 Image Segmentation via Bayesian Inference and Shape Priors

Bayesian inference has been widely used in data analysis problems over the last decades. The Bayesian approach provides the means to incorporate prior knowledge in data analysis. The focus of Bayesian analysis is the posterior probability, which summarizes the degree of certainty about a given situation. Bayes' law states that the posterior probability is proportional to the product of the likelihood and the prior probability. The likelihood encompasses the information contained in the new data and the prior expresses the degree of certainty concerning the situation before the data are taken.

Although the posterior probability completely describes the state of certainty about any possible image, it is often necessary to select a single image as the *result* or *reconstruction*. A typical choice is the image that maximizes the posterior probability density, which is called the MAP estimate. Other choices for the estimator maybe more desirable, such as, the mean of the posterior density function. In situations where only very limited data are available, the data alone may not be sufficient to specify a unique solution to the problem. The prior introduced with the Bayesian method can guide the result toward a preferred solution. Choosing the prior is one of the most critical aspects of Bayesian analysis, since the MAP solution differs from the maximum likelihood (ML) solution solely because of the prior. In this section a variety of possible priors appropriate for image analysis are discussed.

Given an input image $I : \Omega \longrightarrow R$ on a domain $\Omega \subset R^2$, a segmentation C of the image can be found as a MAP estimation of the posterior probability as [12]:

$$P(C|I) = \frac{P(I|C)P(C)}{P(I)} \tag{2.16}$$

where $P(I|C)$ denotes the data likelihood for a given segmentation C and $P(C)$ denotes the prior probability. In order to maximize the posterior distribution in Equation(2.16), the following energy function is minimized:

$$E(C) = E_{data}(C) + E_{shape}(C) \tag{2.17}$$

where $E_{data}(C) = -\log P(I|C)$ and $E_{shape}(C) = -\log P(C)$ are known as the data fidelity term and the shape prior term. As mentioned earlier, rather than maximizing the posterior distribution, mean of this distribution or as in particle filtering [37], the whole posterior can be utilized.

Now, the data fidelity and the shape prior term need to be specified to proceed. There are various data terms proposed in literature. A data term proposed in [26] based on intensity statistics as:

$$E_{data}(C) = \sum_{i=1}^{k} \int_{\Omega_i} (I(x) - c_i)^2 dx \tag{2.18}$$

where I is the image intensity, $\Omega_1, \ldots, \Omega_k$ are pairwise disjoint regions separated by the boundary C and c_i is the average of intensities over ω_i.

There are more sophisticated data terms proposed in the literature which use the texture or the color information of the regions as likelihood (see [38], [39], [40], [35], [41]).

2.2.1 Static Shape Priors

Many prominent image segmentation methods are based on rather simple geometric shape priors, in which a penalty on the length of the curve to be segmented is used [26], [6]. In many applications, more specific knowledge about the shape of the object is provided. Various approaches of incorporating higher-level shape priors have been proposed in the literature.

In the parametric presentation framework, a training set of shapes can be constructed, represented by a spline curve of a fixed number of control points. Statistical

shape prior learned from this training dataset can be imposed in the segmentation problem, which ensures the similar shape family as the result.

An alternative to go about is to assume that the training shapes are from a Gaussian distribution. This assumption is quite popular considering the desirable properties of Gaussian distribution, which is defined as [12]:

$$P(z) = \frac{1}{|2\pi\Sigma_\perp|^{1/2}} exp(-\frac{1}{2}(z-\overline{z})^t \Sigma_\perp^{-1}(z-\overline{z})) \qquad (2.19)$$

where $\overline{z}$ denotes the mean control point vector and $\sum_\perp$ is a regularized covariance matrix introduced in [42]. Now the shape energy is given as:

$$E_{shape}(z) = -\log p(\hat{z}) \qquad (2.20)$$

where $\hat{z}$ is the shape vector upon similarity alignment with respect to the training shapes. As it can be seen, Equation (2.20) is invariant to similarity transformations. Examples of works using Gaussian shape priors can be seen in [43], [44], [42].

The assumption that the training shapes are Gaussian distributed is not valid in many real applications, which motivates using nonlinear shape density estimators. Here, nonlinear means that the permissible shapes are not simply given by a weighted sum of eigenmodes [12]. A classical approach for estimating nonlinear distributions is based on the Gaussian mixture model or the Parzen-Rosenblatt kernel density estimator (which we discuss more later). Authors in [45] propose a novel method for density estimation for computing the nonlinear statistics which is an extension of kernel PCA. The training shapes are approximated to a higher-dimensional feature space Y using a Gaussian distribution, on transformation $\psi : R^2 \longrightarrow Y$:

$$p_\psi(z) \propto exp(-\frac{1}{2}(\psi(z) - \psi_0^t)\Sigma_\psi^{-1}(\psi(z) - \psi_0)) \qquad (2.21)$$

which results in the following energy:

$$E(z) = \log p_\psi(\hat{z}) \qquad (2.22)$$

where ψ_0 and $\sum_\psi$ are the mean and covariance matrix of the transformed shapes as:

$$\psi_0 = \frac{1}{m} \sum_{i=1}^{m} \psi(z_i) \tag{2.23}$$

and

$$\sum_{\psi} = \frac{1}{m} \sum_{i=1}^{m} (\psi(z_i) - \psi_0)(\psi(z_i) - \psi_0)^{\top} \tag{2.24}$$

The energy $E(z)$ in (2.22) can be found by defining the corresponding Mercer kernel [46], [47] without explicitly specifying the nonlinear transformation ψ [45].

Learning the statistical priors can also be derived in the implicit representation framework benefiting from level sets' properties.

As the first step, a distance or dissimilarity measure for two shapes represented implicitly using $L_2 - distance$ over Ω can be defined as [14]:

$$\int_{\Omega} (\phi_1 - \phi_2)^2 dx \tag{2.25}$$

where ϕ_i, $i = 1, 2$ is the signed distance function. As it can be seen in Equation (2.25), the measurement is dependent on the Ω, which is a drawback.

An alternative can be the methods proposed and used in [48] and [49] as:

$$d^2(\phi_1, \phi_2) = \int_{\Omega} (H\phi_1(x) - H\phi_2(x)^2 dx \tag{2.26}$$

which uses the area of the symmetric difference.

Since the measure in Equation (2.26) is not dependent on the image size, it is more favorable in general case. Dealing with shapes of same size in our work, we have used the distance measure in Equation (2.25) for the sake of simplicity. These measures can be used as shape prior as well.

Specifying the dissimilarity measurement, we concentrate on the shape representation and kernel density estimation in the level set domain.

Regarding explicitly represented shapes, principal component analysis (PCA) is used in [43] to reconstruct a the space of familiar shapes using a set of training shapes. Selecting corresponding landmarks between shapes is challenging in this ap-

proach. Initially landmarks were selected manually. Automated landmark selection was proposed in [50] afterwards which was not efficient though.

Desirable properties of implicit-based shape representations led to the development of methods using level sets pioneered by [14], which constructs shape priors out of a set of training shapes using signed distance functions. In order to address the problem regarding the signed distance function, which is not being closed under linear operations, [32] proposed a method which reduces the shape distance from sample shape along with its distance from the space of signed distance functions. Therefore, preserving a reasonable signed distance function, a variational framework is generated using maximum likelihood estimation.

Assuming Gaussian distribution as the shape distribution makes these methods perform poorly while dealing with more complex shape variation. Alternatively, non-parametric density estimation approaches are developed to address this issue which learn an unknown density function using a set of samples without enforcing any structure on the density to be estimated.

Considering a finite-dimensional density estimation and given a set of sample shapes $X_{ii=1...N}$, Parzen-Rosenblatt kernel density estimator is defined as [51], [52]:

$$P(\widetilde{x}) = \frac{1}{N} \sum_{i=1}^{N} k(x - x_i, \Sigma) \tag{2.27}$$

where $k(x, \Sigma)$ is a m-dimensional Gaussian kernel with covariance matrix Σ.

In applications which multiple random variables density estimation is desired, multivariate version of the Parzen-Rosenblatt kernel density estimator is applied. Let us assume a N-dimensional random vector as $\mathbf{X} = (X_{i1}, X_{i2}, \ldots, X_{iN})$ where $X_{ij\,(i=1,\ldots,N),(j=1,\ldots,M)}$ is an one dimensional random variables which denotes ith observation of jth random variable. The joint PDf of $\mathbf{X}$ is given as:

$$f(\mathbf{X}) = \frac{1}{N} \sum_{i=1}^{N} k(\mathbf{x} - \mathbf{X}_i, \sigma_M) \tag{2.28}$$

which means:

$$f(\mathbf{X}) = \frac{1}{N} \sum_{i=1}^{N} k((x_1 - X_{i1}, ..., x_M - X_{iM}), \sigma_M) \tag{2.29}$$

where k is a multivariate kernel. The resulting kernel forms can be a multiplication of separate kernels and written as:

17

$$f(\mathbf{X}) = \frac{1}{N} \sum_{i=1}^{N} \prod_{j=1}^{M} k(d(x^j - x_i^j), \sigma_j) \tag{2.30}$$

Applying the Parzen-Rosenblatt kernel density estimator on the space of signed distance functions, the following is obtained:

$$p(\phi) \propto \frac{1}{N} \sum_{1}^{N} exp(-\frac{1}{2\sigma^2} d^2(H\phi, H\phi_i)) \tag{2.31}$$

where different shape distances can be integrated in the above equation, such as the measurement in Equation (2.26).

2.2.2 Gradient Descent Evolution

Learning statistical distribution using Equation (2.31) can be used to enhance the level set based segmentation process. Modeling the level set-based segmentation problem using Bayes' rule as [12]:

$$p(\phi|I) = \frac{p(I|\phi)p(\phi)}{p(I)} \tag{2.32}$$

where I denotes the image and ϕ denotes the level set function. MAP estimation of the posterior distribution in Equation(2.32) amounts to minimizing the following:

$$E(\phi) = E_{data}(\phi) + E_{shape}(\phi) \tag{2.33}$$

with

$$E_{shape}(\phi) = -\log p(\phi) \tag{2.34}$$

This segmentation model allows the statistical density estimator to incorporate the similarity between the evolving curve and the training shapes.

In order to minimize the Equation (2.33) results in the evolution below using gradient descent [12]:

$$\frac{\partial \phi}{\partial t} = -\frac{1}{\alpha} \frac{\partial E_{data}}{\partial \phi} - \frac{\partial E_{shape}}{\partial \phi} \tag{2.35}$$

where:

$$\frac{\partial E_{data}}{\partial \phi} = -\frac{\sum \alpha_i \frac{\partial}{\partial \phi} d^2(H\phi, H\phi_i)}{2\sigma^2 \sum \alpha_i} \tag{2.36}$$

18

and:

$$\alpha_i = exp(-\frac{1}{2\sigma^2}d^2(H\phi, H\phi_i)) \tag{2.37}$$

The shape force in Equation(2.36), is based on weighted training shapes by α_i, which itself is related to the dissimilarity between te evolving curve and each training shapes.

2.2.3   Dynamic Shape Priors

Shape priors are developed and utilized in order to improve the segmentation results of a known class of objects, especially in presence of noise, clutter, and occlusion. Static shape priors are not desired when dealing with deformable objects though, since they do not consider temporal coherence between object's shape. Here, the concept of dynamic shape priors emerges. The idea behind dynamic shape priors is exploiting the temporally deforming dynamics and integrating the information of the shape of the object in preceding frames, rather than considering each frame independently. In order to incorporate the temporal information as prior, the segmentation problem can be modeled within the Bayesian framework for implicitly represented shapes and curve evolution can be obtained by gradient descent, which results in a data force based on the image intensity and a shape force based on the previous segmentation results.

A temporal statistical shape model is introduced in [53], which performs the Bayesian inference in a low-dimensional formulation within the subspace spanned by the largest principal eigenmodes of a set of sample shapes [12]. Using PCA, approximation of N implicitly represented training shapes are provided by a shape vector, $\boldsymbol{\alpha}$. Additionally, a transformation parameter has been introduced as θ. Based on these parameters, for consecutive images I_{t-1} and I_t, a Bayesian model is proposed to compute the most likely deformation and transformation parameters at time t, given the deformation and the transformation parameters at time $t - 1$, which is:

$$p(\boldsymbol{\alpha}_t, \theta_t | I_t, \hat{\boldsymbol{\alpha}}_{1:t-1}, \hat{\theta}_{1:t-1}) = \frac{p(I_t | \boldsymbol{\alpha}_t, \theta_t)p(\boldsymbol{\alpha}_t, \theta_t | \hat{\boldsymbol{\alpha}}_{1:t-1}, \hat{\theta}_{1:t-1})}{p(I_t | \hat{\boldsymbol{\alpha}}_{1:t-1}, \hat{\theta}_{1:t-1})} \tag{2.38}$$

In this work, a single Gaussian parametric model is used for estimating the shape distribution, which enforces a smooth, unimodal distribution for the joint likelihood. Therefore, it performs poorly in case of complicated shape distributions. On the other hand, PCA is able to just handle small shape deformations and is quite inadequate when shape variations if the object being tracked undergoes large deformations.

Senegas et al. introduced dynamical shape model to the segmentation of cardiac images in order to model cardiac dynamics using Sequential Monte Carlo sampling [54]. The proposal distribution used in this work is only based on the observation model. Therefore, it does not incorporate the transition model. Sun et al. proposed learning the cardiac dynamics [55] which utilizes a statistical shape model trained by a set of representative left ventricle shapes observed in cardiac MRI images. The limitation regarding this approach is that the introduced proposal distribution is based on the transition model and it does not take into account the observation model.

There are algorithms for tracking which form the shape prior using kernel PCA rather than linear PCA, such as [56] and [57] and improvement in their performances have been demonstrated.

The segmentation model proposed in [58] uses multiple shape priors and minimizes a joint energy which is defined based on image and a labeling function. In this way, the shape information based on the object's class is chosen. In case of tracking deforming objects, this method can not provide shape priors dynamically, though.

In Chapter 3, we will discuss how to learn the relationships of shapes across temporal frames and use it as a shape prior to segment the deformable objects in time series images based on Sequential Monte Carlo sampling.

2.3 Recursive Estimation and Particle Filtering

Particle filtering methods that incorporate shape information of the object to be segmented are widely used in the field of image analysis. They are used in different tasks such as, tracking in clutter, tracking multiple targets and segmentation. Isard and Blake in [59] and [60] used particle filtering for contour tracking using the *CONDENSATION* filter. In these works, *CONDENSATION* means conditional

density propagation. The conditional pdf or posterior pdf is represented by the weights of N samples and is propagated over time. Particle filters also fit well to multiple target tracking problems because they are capable of handling multi-modal densities. This chapter provides a brief overview of recursive Bayesian filtering followed by information about particle filtering and its application in computer vision.

In general, state estimation is the process of estimating quantities which are not directly observable, but can be inferred using data from other observable quantities called measurements [61].

Let us assume x_t the state of a system at time point t. In some cases there is no direct way to acquire information about x_t and a related observed measurement, y_t, is provided. Also let us assume that x_t is stochastically generated from the state x_{t-1}.

A dynamic system can be modeled using two approaches. First approach models the dynamic system using two probability distributions: transition probability and measurement probability which are respectively given as [61]:

$$p(x_t|x_{0:t-1}, y_{0:t-1}) \tag{2.39}$$

and

$$p(y_t|x_{0:t}, y_{0:t-1}) = p(y_t|x_t) \tag{2.40}$$

Assuming that the system is a Markov Model (or a Hidden Markov Model, since the states are estimated from the measurements) we can rewrite Equation(2.39):

$$p(x_t|x_{0:t-1}, y_{0:t-1}) = p(x_t|x_{t-1}) \tag{2.41}$$

The second approach is to define the dynamic system as a set of two equations: the state transition equation and the measurement equation.

$$x_t = f(x_{t-1}, \mu_{t-1}) \tag{2.42}$$

and

$$y_t = g(x_t, \epsilon_t) \tag{2.43}$$

where μ_{t-1} is the state noise and ϵ_t is measurement noise.

In order to estimate the states recursively, the following two steps are followed alternately at each time point:

Predict: The next state is predicted as:

$$p(x_{t-1}|y_{0:t-1}) \longrightarrow p(x_t|y_{0:t-1}) \tag{2.44}$$

Update: The current measurements are imposed:

$$p(x_t|y_{0:t-1}) \longrightarrow p(x_t|y_{0:t}) \tag{2.45}$$

This approach is the main idea behind the recursive filters. Bayes Filter is the most general recursive filter, which estimates a state using the previous state and the measurements. The prediction step in Bayes Filter is finding the prior distribution of the state at time t without knowing the new measurement and can be modeled by the Chapman-Kolmogorov equation as:

$$p(x_t|y_{0:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{0:t-1})dx_{t-1} \tag{2.46}$$

In the update step, the posterior distribution is computed from the predicted density and the new measurement is imposed as well.

$$
\begin{aligned}
p(x_t|y_{0:t}) &= \frac{p(y_{0:t}|x_t)p(x_t)}{p(y_{0:t})} \\
&= \frac{p(y_t, y_{0:t-1}|x_t)p(x_t)}{p(y_t, y_{0:t-1})} \\
&= \frac{p(y_t|y_{0:t-1}, x_t)p(y_{0:t-1}|x_t)p(x_t)}{p(y_t|y_{0:t-1})p(y_{0:t-1})} \\
&= \frac{p(y_t|y_{0:t-1}, x_t)p(x_t|y_{0:t-1})p(y_{0:t-1})p(x_t)}{p(y_t|y_{0:t-1})p(y_{0:t-1})p(x_t)} \\
&= \frac{p(y_t|x_t)p(x_t|y_{0:t-1})}{p(y_t|y_{0:t-1})}
\end{aligned}
\tag{2.47}
$$

and:

$$p(y_t|y_{0:t-1}) = \int p(y_t|x_t)p(x_t|y_{0:t-1})dx_t \tag{2.48}$$

The drawback of the Bayes Filter is that in many cases the integrals involved are intractable, therefore it is far from being practical. In order to address this issue sampling-based techniques have been developed for estimation purposes. One

popular example is the Sequential Importance Sampling or Particle Filters. This technique is capable of estimating a hidden state of a dynamical system, X_t, given observations by generating a set of samples or particles, $\{x_t^i\}_{i=1,...,N}$, approximating the posterior density by the set of particles. A particle can be consider as a hypothesis regarding the state to be estimated. In order to delineate the amount of the samples' contribution in the estimation process, a weight, $\{w_t^i\}_{i=1,...,N}$, is assigned to each. Particle filters consist of three steps which are prediction, update, and re-sampling.

In the prediction step, a new predicted particle, $\hat{x}_{t+1}$, is generated from the current state estimate x_t. To this end, a transition function f is defined based on x_t. This step can be written as:

$$\hat{x}_{t+1} = f(x_t) + \mu_t \tag{2.49}$$

where μ_t is the state noise.

For the update step, an update function, g, is defined to compute the distance between the measurement and the generated particle and based on that the corresponding weights are defined as:

$$w_{t+1} = g(\hat{x}_{t+1}, y_t) + \epsilon_t \tag{2.50}$$

where ϵ_t is the measurement noise.

In addition to classical prediction and update steps, the re-sampling step is introduced to address the weight degeneracy problem, which means concentration of the weights on a small number of particles compared to the rest after a relatively small number of iterations. Therefore, re-sampling is a step by which the samples are replaced with a new set of equally weighted samples generated from the highly weighted samples. Therefore, by generating samples more similar to samples with large weights, the particle set is more concentrated on the sub-region of the state space with high probability of being the true state [62].

Chapter 3

Methodology

In this chapter we describe our dynamic time series image segmentation approach in detail. First, the motivation of our proposed approach is introduced in Section 3.1. In Section 3.2, we present our general framework for dynamic segmentation model by introducing a new term, *coupled shape priors*. The coupled shape prior term enables the model to learn the dynamics of the object's deformation from a training dataset.

In the following sections, we start with a preliminary version of the work assuming that in our time series segmentation model, the information about the exact shape of the object's contour in the previous time point is provided. Based on this assumption, a cost function is constructed which incorporates the coupled shape prior as well. Afterwards, we upgrade our approach to a more generalized version considering that the information we have about the shape of the object in the previous time point is imperfect. We formulate a sampling algorithm by means of Sequential Monte Carlo (particle filters) and impose the coupled shape prior while estimating the shape of the object to be segmented dynamically.

3.1 Motivation for the proposed approach

Image segmentation can turn into a very complicated task in some applications. One of the challenges that is faced, is to segment regions with boundary insufficiencies, i.e., missing edges or lack of texture contrast between regions of interest and background. Presence of high noise or missing data are among the difficulties that must be dealt with in the segmentation process. Biomedical images are an example of challenging images for the segmentation task. The reason is that, gener-

ally biomedical images are low quality images and have low intensity contrast. For instance, we can mention images of dendritic spines or brain tissues. In order to deal with these challenges, prior information can be incorporated to obtain more promising results and improve the segmentation process.

The statistically learned shape prior seems to be vital for handling deficiencies such as noise, clutter, and occlusion during the segmentation process. However, it may not be sufficient in case of evolving objects in a sequence of images, since it does not take the temporal shape dependencies into account. Therefore, an effective segmentation algorithm should be capable of engaging the dynamics of the system if the object goes through a deformation over time.

In this thesis, our aim is to address these issues through our presented approach. We concentrate on the problem of segmenting objects with dynamically evolving shapes over time. Thus, prior knowledge which has been obtained from the past time points seems to be a reasonable choice to benefit from in the segmentation algorithm. The information that we want to include in our approach is the shape of the object. If a dependence is noticed between the shapes of the object over time, shape information will be of great assistance based on which we build an estimate about the shape of the object in the future.

In the following sections, we formulate our approach integrating the shape prior knowledge into the dynamic segmentation framework for deformable objects in time series images. We introduce a new term - to which we refer as the coupled shape prior term - that appears as part of the cost function and imposes the system dynamics. In order to construct the cost function, we start with the assumption that we are provided with the segmentation of the object in the previous time point and this assumption leads us to two different scenarios.

- We are provided with *perfect* segmentation of the object in the previous time point (preliminary version).

- We are provided with *imperfect* segmentation of the object in the previous time point (generalized version).

Now we proceed to the development of two different versions of our approach for these scenarios separately.

3.2   Mathematical formulation

3.2.1   Preliminary version

Let us assume that we have m training shapes $\mathbf{C^{(t-1)}} = \{C_1^{(t-1)}, C_2^{(t-1)}, \dots$ $, C_m^{(t-1)}\}$ at time $(t-1)$. Let us also assume that we have m training shapes $\mathbf{C^{(t)}} = \{C_1^{(t)}, C_2^{(t)}, \dots, C_m^{(t)}\}$ at time (t), where $C_i^{(t-1)}$ and $C_i^{(t)}$ are boundaries of the same object in two consecutive time points for each $i \in 1, \dots, m$. In this preliminary approach, we assume a sequence with two time points only, for simplicity. It is also assumed that the shapes in both $C^{(t-1)}$ and $C^{(t)}$ have already been aligned so that the shape variations due to pose differences are removed. Then, the posterior probability of $C^{(t)}$ given intensity images $y^{(t-1)}$ at time point $(t-1)$ and $y^{(t)}$ at time point (t) is written as

$$p(C^{(t)}|y^{(t-1)}, y^{(t)}) = \int p(C^{(t)}, C^{(t-1)}|y^{(t-1)}, y^{(t)})\mathrm{d}C^{(t-1)} \tag{3.1}$$

Using Bayes' rule and the assumption that $y^{(t-1)}$ and $y^{(t)}$ are independent conditioned on the corresponding boundaries, we get

$$
\begin{aligned}
&p(C^{(t)}|y^{(t-1)}, y^{(t)}) \\
&\propto \int p(y^{(t-1)}, y^{(t)}|C^{(t)}, C^{(t-1)})p(C^{(t)}, C^{(t-1)})\mathrm{d}C^{(t-1)} \\
&= \int p(y^{(t-1)}|C^{(t-1)})p(y^{(t)}|C^{(t)})p(C^{(t)}|C^{(t-1)})p(C^{(t-1)})\mathrm{d}C^{(t-1)} \\
&\propto p(y^{(t)}|C^{(t)}) \int p(C^{(t)}|C^{(t-1)})p(C^{(t-1)}|y^{(t-1)})\mathrm{d}C^{(t-1)}.
\end{aligned}
\tag{3.2}
$$

As it can be seen in Equation (3.2), there are three probability densities to be calculated and one can proceed with various assumptions on these densities. As a preliminary approach, we continue with the following set of assumptions:

- $p(C^{(t)}|C^{(t-1)})$ is learned non-parametrically from the training dataset.

- The data term proposed in [9] is used to obtain $p(y^{(t)}|C^{(t)})$.

- Considering that we are provided with the perfect segmentation of the object in the previous time point, $p(C^{(t-1)}|y^{(t-1)})$ is substituted with a delta function. (Of course this is not a realistic assumption in practice. We use it to derive a

preliminary version of our approach here and we relax this assumption later in the thesis.)

Let us explain the items above in detail. As we mentioned earlier, we assume that the posterior density $p(C^{(t-1)}|y^{(t-1)})$, is a delta function, which can be written as follows:

$$p(C^{(t-1)}|y^{(t-1)}) = \delta(C^{(t-1)} - C') \tag{3.3}$$

where $\delta(.)$ is the Dirac delta function and C' is the perfect segmentation of the object at time point $(t-1)$ which we assume that we already have. Thus, we can rewrite the Equation (3.2) as:

$$p(C^{(t)}|y^{(t-1)}, y^{(t)}) \propto p(y^{(t)}|C^{(t)})p(C^{(t)}|C'). \tag{3.4}$$

Eventually, we define the energy function to be minimized by taking the negative logarithm of Equation (3.4) as

$$E(C^{(t)}) = -\log p(y^{(t)}|C^{(t)}) - \log p(C^{(t)}|C'). \tag{3.5}$$

For the first term $-\log p(y^{(t)}|C^{(t)})$ in Equation (3.5), we use the piecewise-constant version of the Mumford functional [26] proposed in [9] which is given by:

$$\begin{aligned} -\log p(y^{(t)}|C^{(t)}) = &\int_{C_{in}^{(t)}} (I(x) - m_{in})^2 \mathrm{d}x \\ &+ \int_{C_{out}^{(t)}} (I(x) - m_{out})^2 \mathrm{d}x \end{aligned} \tag{3.6}$$

where m_{in} (m_{out}) are the mean intensities inside (outside) of curve $C^{(t)}$. We learn the second term $p(C^{(t)}|C')$ in Equation (3.5) nonparametrically from the training dataset using Parzen density estimation as follows:

$$p(C^{(t)}|C') \propto \frac{1}{m} \sum_{i=1}^{m} k(d_{L_2}(C^{(t)}, C_i^{(t)}), d_{L_2}(C', C_i^{(t-1)}), \sigma) \tag{3.7}$$

where $k(\cdot, \cdot, \sigma)$ is a 2D Gaussian with standard deviation σ and $d_{L_2}(\cdot, \cdot)$ is the L_2 distance metric.

We minimize the energy function in Equation (3.5) using gradient descent. In order to propagate the curve, we use the level set representation of binary shapes

which we denote by ϕ. We use the sign convention $\phi < 0$ for inside the curve and $\phi > 0$ for outside the curve. The overall gradient flow of the energy function is the sum of the two terms, one based on the data term and the other based on the coupled shape prior term. The partial derivative of the data term is computed as follows:

$$\frac{-\partial \log p(y^{(t)}|\phi^{(t)})}{\partial \phi^{(t)}} = \beta\Big[- (I(x) - m_{in})^2 + (I(x) - m_{out})^2\Big]\vec{N} \qquad (3.8)$$

where $\vec{N}$ is the outward curve normal. Also, the partial derivative of the coupled shape prior term is given by

$$
\frac{\partial \log p(\phi^{(t)}|\phi')}{\partial \phi^{(t)}} = \frac{1}{p(\phi^{(t)}|\phi')}\frac{1}{\sigma^2}\frac{1}{n}
$$
$$
\sum_{i=1}^{n} k(d_{L_2}(\phi^{(t)}, \phi_i^{(t)}), d_{L_2}(\phi', \phi_i^{(t-1)}), \sigma)(\phi_i^{(t)} - \phi^{(t)}). \qquad (3.9)
$$

We start the segmentation with the level set representation of an initial curve and update the curve with the gradient of the energy function in Equation (3.5). Finally, when the curve converges at some point, $C^{(t)}$ can be found from $\phi^{(t)}$ by thresholding at zero.

3.2.2 Generalized version

In the previous section, we developed a preliminary approach which was motivated by incorporating shape information from previous time points to segment an object in the current time point. In particular, a coupled shape prior term has been constructed (learned from manually segmented training data) for providing the shape force in the curve evolution process and the resulting energy function has been minimized using gradient descent and level sets.

Up to this point, we had the assumption that we are provided with the perfect segmentation of the shape of the object in the previous time point. This assumption led to the elimination of $p(C^{(t-1)}|y^{(t-1)})$ term from our formulation, where $C^{(t-1)}$ is the boundary of the object at time $t-1$ and $y^{(t-1)}$ is the intensity image of the object at time $t-1$. However, this assumption is clearly not applicable in practice. We now describe how we extend our preliminary study by using imperfect segmentations

28

from previous time points and characterizing the posterior distribution of the curve. This extends our framework to a complete dynamic estimation setting.

To have a set of more comprehensive assumptions, let us redefine our training dataset as $\mathbf{C^{(t)}} = \{C_1^{(t)}, C_2^{(t)}, \ldots, C_m^{(t)}\}$, where $C_i^{(t)}$ is the boundary of the ith training shape at time t and m is the number of the shapes. Note that, as mentioned before, the training shapes in all time points have already been aligned to remove the shape variations due to pose differences.

Similar to our derivation in the preliminary work, we can construct the posterior probability density function of $C^{(t)}$ given $y^{1:(t-1)}$ and $y^{(t)}$, where $y^{1:(t-1)}$ denotes all the intensity images from time point 1 to $(t-1)$ and $y^{(t)}$ denotes the intensity image at time point t, using Bayes' rule and assuming that $y^1, \ldots, y^{(t)}$ are independent conditioned on the corresponding shape boundaries. Now we can write:

$$
\begin{aligned}
&p(C^{(t)}|y^{1:(t)}) \\
&= p(C^{(t)}|y^{1:(t-1)}, y^{(t)}) \\
&= \frac{p(y^{(t)}|C^{(t)})p(C^{(t)}|y^{1:(t-1)})}{p(y^{(t)}|y^{1:(t-1)})}
\end{aligned}
\tag{3.10}
$$

We also can write:

$$
\begin{aligned}
p(C^{(t)}|y^{1:(t-1)}) &= \int p(C^{(t)}, C^{(t-1)}|y^{1:(t-1)})\mathrm{d}C^{(t-1)} \\
&= \int p(C^{(t)}|C^{(t-1)})p(C^{(t-1)}|y^{1:(t-1)})\mathrm{d}C^{(t-1)}
\end{aligned}
\tag{3.11}
$$

$p(y^{(t)}|y^{1:(t-1)})$ can also be computed by:

$$
\begin{aligned}
p(y^{(t)}|y^{1:(t-1)}) &= \int p(y^{(t)}, C^{(t)}|y^{1:(t-1)})\mathrm{d}C^{(t)} \\
&= \int p(y^{(t)}|C^{(t)})p(C^{(t)}|y^{1:(t-1)})\mathrm{d}C^{(t)}
\end{aligned}
\tag{3.12}
$$

By plugging in the Equation (3.11) and (3.12) in Equation (3.10), we obtain:

$$
\begin{aligned}
&p(C^{(t)}|y^{1:(t)}) \\
&= \frac{p(y^{(t)}|C^{(t)}) \int p(C^{(t)}|C^{(t-1)})p(C^{(t-1)}|y^{1:(t-1)})\mathrm{d}C^{(t-1)}}{\int p(y^{(t)}|C^{(t)})p(C^{(t)}|y^{1:(t-1)})\mathrm{d}C^{(t)}}
\end{aligned}
\tag{3.13}
$$

which is the more general version of the formulation we had in Equation (3.2).

In the new setting, our aim is to estimate the probability density obtained in Equation (3.13). The framework of particle filtering enables us to approximate a state distribution, $p(x)$, using a set of weighted samples and update to a new state distribution at each time step using a recursive model. However, it is usually difficult or even impossible to directly and efficiently sample from the posterior distribution. Therefore, an importance density, $q(x)$, is proposed which is easier to sample from.

Let us assume that the sample-based representation of the posterior distribution in Equation (3.13) at time point $t-1$ is provided. In other words, a set of samples (particles) as $\mathbf{C^{(t-1)}} = \{C_1^{(t-1)}, C_2^{(t-1)}, \ldots, C_n^{(t-1)}\}$ are drawn from the posterior distribution in Equation (3.13) at time point $t-1$, where $C_i^{(t-1)}$ denotes the ith sample at time $t-1$, $i \in 1, \ldots, n$. In order to approximate $p(C^{(t)}|y^{1:(t)})$ we can propose an importance density such that it factorizes as:

$$q(C^{(t)}|y^{1:(t)}) = \\ q(C^{(t)}|C^{(t-1)}, y^{1:(t)})q(C^{(t-1)}|y^{1:(t-1)}) \tag{3.14}$$

Therefore, the particle set at time t can obtained by augmenting particles set at $t-1$ with the particles drawn from the importance density $q(C^{(t)}|C^{(t-1)}, y^{1:(t)})$ [63]. By weighting the newly obtained particles, we can provide an estimate for the posterior density $p(C^{(t)}|y^{1:(t)})$ which is expressed as follows:

$$p(C^{(t)}|y^{1:(t)}) = \sum_{i=1}^{n} W_i^{(t)}\delta(C^{(t)} - C_i^{(t)}) \tag{3.15}$$

where $W_i^{(t)}$ is the normalized importance weights of the ith particle is defines as:

$$W_i^{(t)} = W_i^{(t-1)}\frac{p(y^{(t)}|C_i^{(t)})p(C_i^{(t)}|C_i^{(t-1)})}{q(C_i^{(t)}|C_i^{(t-1)}, y^{(t)})} \tag{3.16}$$

This equation provides the weights of each particle in the particle set $\mathbf{C^{(t-1)}}$ in a recursive fashion.

Notice that the recursive weight in Equation (3.16) is derived assuming the boundaries to be a hidden Markov process, the observations to be independent conditioned on the corresponding shape boundaries, and the importance density to be dependent only on the previous boundary and current observation.

Now we can update from $p(C^{(t-1)}|y^{1:(t-1)})$ to $p(C^{(t)}|y^{1:(t)})$ by sampling $C_i^{(t)}$ from $q(.|C_i^{(t-1)}, y^{(t)})$ and calculating the weights using Equation (3.16).

We now have a recursive model to estimate the desired posterior distribution. An issue that needs to be considered at this point, is the choice of the importance density which is critical to design a successful particle filter. One of the approaches that is useful to do this is to propose a Gaussian density with a mean dependent on the current observation [64]. We proposed this density as a Gaussian distribution with a standard deviation σ below:

$$q(C^{(t)}|C^{(t-1)}, y^t) = N(f(C^{(t-1)}, y^{(t)})), \sigma) \tag{3.17}$$

where $f(C^{(t-1)}, y^{(t)})$ is defined as a function of $C^{(t-1)}$ and $y^{(t)}$ and represents a curve evolution operation on each sample at time point $t - 1$.

As mentioned before, in order to estimate the posterior density at t generating samples from $q(.|C_i^{(t-1)}, y^{(t)})$ is required. Our proposed importance density suggests that each sample at time point $t - 1$ (which is a binary image) will be refined to a curve using the observation (the intensity image) at time point t and yields to samples at time point t.

In order to implement it we perform L iterations of gradient descent on the energy function we introduced in Equation (3.5), where the sample at time point $t - 1$ and observation at time point t are served as the segmented boundary in the previous time point and the test image in the current time point, respectively. The energy function in Equation (3.5) employs the coupled shape priors. Thus, the dynamics of the model is exploited through our proposed importance density. Adding the step of curve evolution, enriches our method in generating more likely particles.

Another issue that arises in designing our particle filter based approach is to generate samples in the first time point, in other words, estimating $p^{(1)}$. To address this issue, we use the approach proposed by [65] which introduces a MCMC based

sampling approach.

Now we describe how our overall algorithm works.

- Generating sufficiently many samples for the posterior density in the first time point (sample from $p^{(1)}$), using the method in [65].

- Computing weights for generated samples as $\frac{p^{(1)}}{q^{(1)}}$, which indicates the amount of contribution that each sample has in the estimation process.

- Providing an estimation for $p^{(1)}$ using Equation (3.15).

- Applying curve evolution to each samples by minimizing the energy function proposed in [66] and described in Section (3.2.1) using gradient descent.The gradient flow of this energy function is the sum of the two terms, one based on the data term and the other based on the coupled shape prior term. Therefore, temporal dynamic evolution of the shapes is considered in generating samples for approximating the posterior distribution in the next time point.

- Calculating the weights for the newly generated samples using the recursive relationship in Equation (3.16).

- Estimating the posterior distribution based on the weighted samples.

The sampling, weight computation, and posterior density estimating steps can be repeated up to the desired time points. It should be noted that, in practice, the estimation step mentioned in the above algorithm is performed by obtaining the weighted average of the samples. The reason of choosing weighted averaging rather than MAP estimation is that our proposed algorithm is able to distinguish the shape class of the object to be segmented and samples are generated within the detected shape class. Therefore, weighted averaging is more preferable.

Chapter 4

Experimental Results

In this chapter we evaluate the performance of our proposed segmentation algorithm which we described in Chapter 3. We perform two sets of experiments based on preliminary version and generalized version of our work described in the previous chapter and present the results on three different datasets, a synthetic dataset, a hand gesture dataset, and a dendritic spine dataset. In the experiments, we demonstrate the effects of coupled shape priors which we introduced as dynamic shape priors in comparison with the Chan and Vese force in [9] and the static shape priors proposed in [17]. In order to demonstrate the performance of our method, we provide visual results along with quantitative analysis.

In the following sections, first we introduce our datasets and then we present the experimental results.

4.1   Datasets

As we already stated, there are three datasets on which the performance of our approach has been evaluated. The following is the description of each dataset.

- **Synthetic dataset**: This dataset includes 35 intensity images and their corresponding binary images in 4 consecutive time points. There are 4 classes of shapes, 8 shapes in class 1, 8 shapes in class 2, 11 shapes in class 3 and 8 shapes in class 4. The objects go through deformation in the future time points and turn into a different shape by the end of the 4th time point. In order not to face any pose difference, the dataset has been created such that images are aligned. some examples of this dataset are shown in Figures (4.1), (4.2), (4.3), and (4.4).
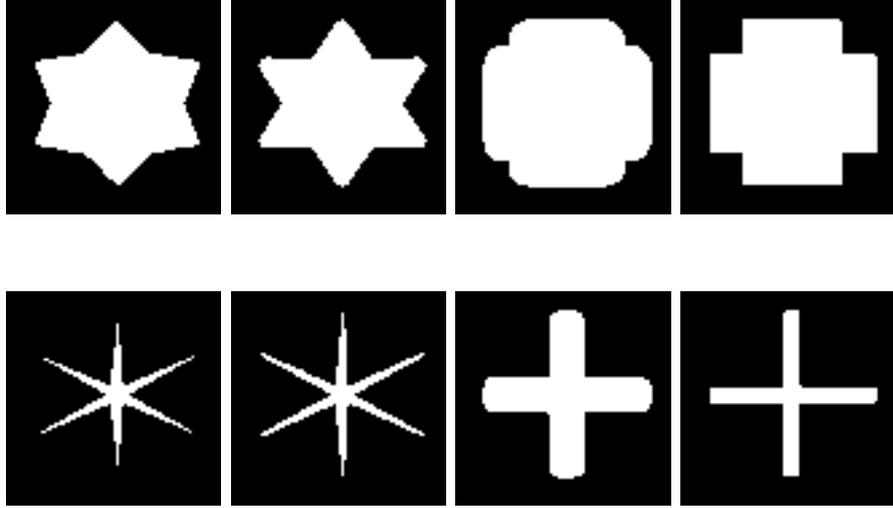
Figure 4.1: Samples of binary images in the synthetic dataset (class 1): each row depicts a shape of a single object in 4 different time points.
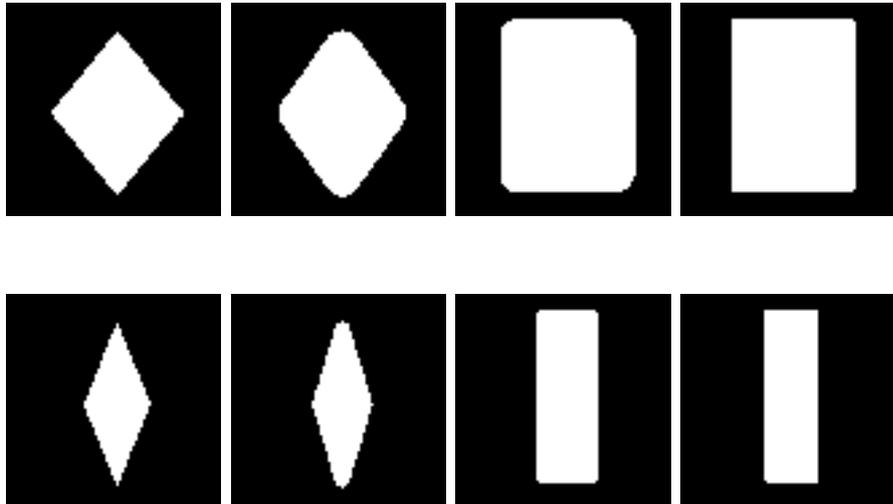
Figure 4.2: Samples of binary images in the synthetic dataset (class 2): each row depicts a shape of a single object in 4 different time points.
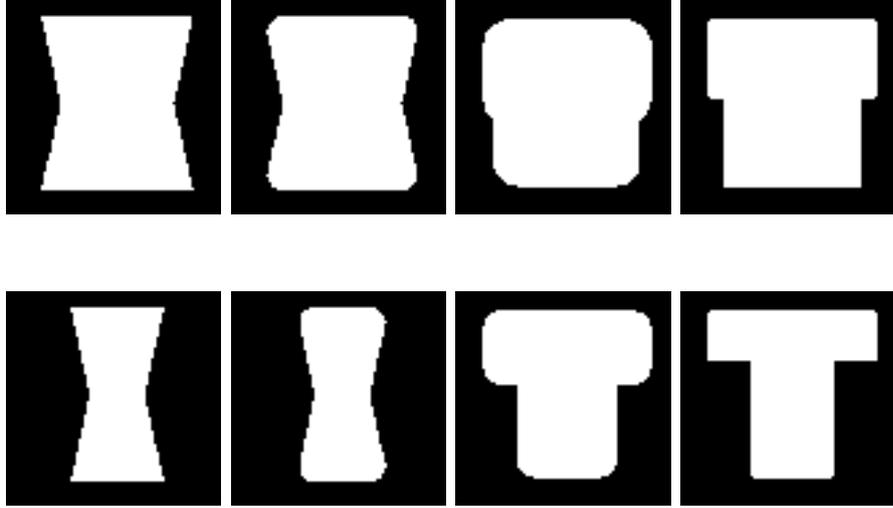
Figure 4.3: Samples of binary images in the synthetic dataset (class 3): each row depicts a shape of a single object in 4 different time points.
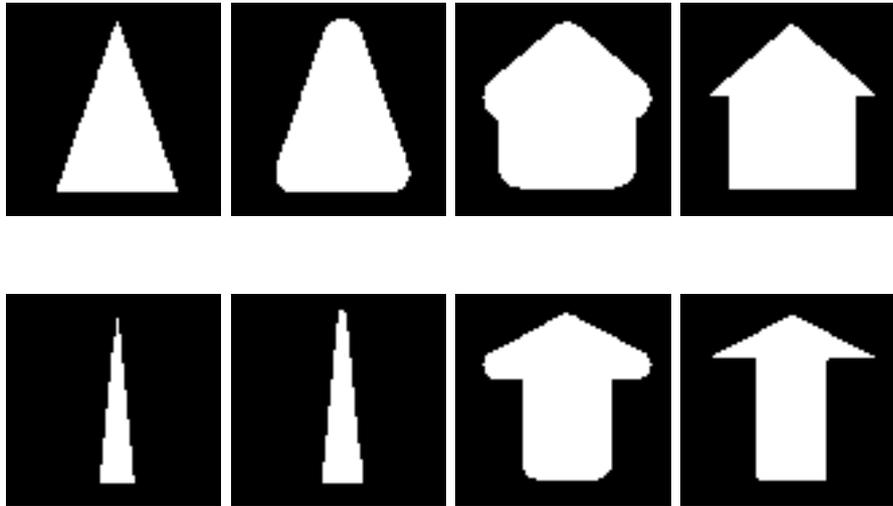
Figure 4.4: Samples of binary images in the synthetic dataset (class 4): each row depicts a shape of a single object in the 4 different time points.

As it can be in Figures (4.1), (4.2), (4.3), and (4.4), the difference between shapes within each class is their thickness.

Corresponding to each binary image, we generate its intensity image by adding noise. We have manipulated some of the instances further in order to represent occlusion and data loss. The challenge that we need to overcome in this dataset is the high level of noise, highly deforming shapes, multi-modal shape distribution, occlusion, and data loss. Later on we show the performance of the proposed algorithm on this generated test dataset. Figures (4.5), (4.6), (4.7), and (4.5) show some samples of the test dataset.
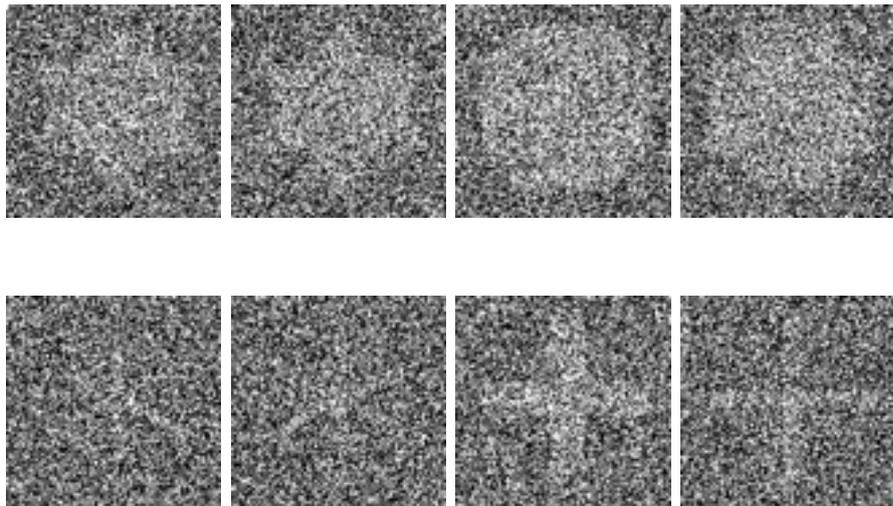
Figure 4.5: Samples of intensity images in the synthetic dataset (class 1): each row depicts a shape of a single object in 4 different time points.
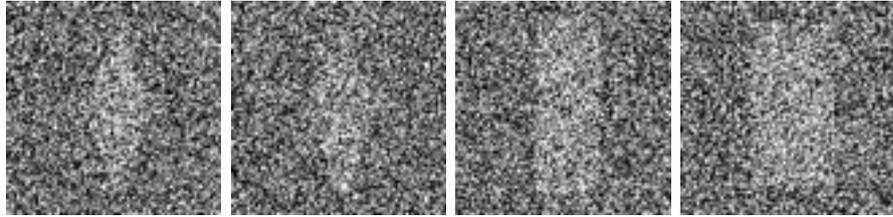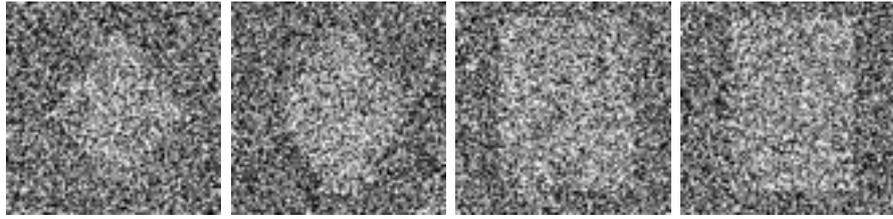
Figure 4.6: Samples of intensity images in the synthetic dataset (class 2): each row depicts a shape of a single object in 4 different time points.
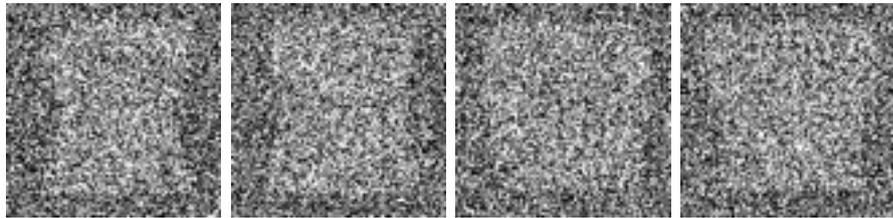
Figure 4.7: Samples of intensity images in the synthetic dataset (class 3): each row depicts a shape of a single object in 4 different time points.
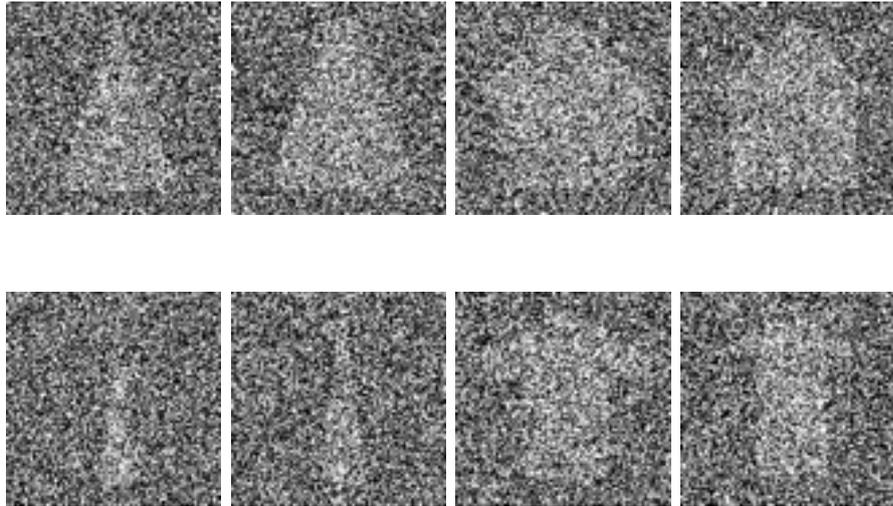
Figure 4.8: Samples of intensity images in the synthetic dataset (class 4): each row depicts a shape of a single object in the 4 different time points.

- **Hand gesture dataset**: In order to construct this dataset, we have used Thomas Moeslund's Gesture Recognition Database which is used in [67]. In this dataset we have chosen 77 images in 2 time points and 2 image classes. 40 images belong to class 1 and 37 belong to class 2. The deformation dynamics are defined by us in this dataset. In each image a hand gesture is depicted and it is assumed that if a hand shows a specific gesture in the first time point, it will show a corresponding gesture in the next time point. We also generated a training dataset including the binary images of the images in the database. All the images have been aligned manually to prevent pose differences. Examples of this dataset are depicted in Figure (4.9).

(a) Shape class 1   (b) Shape class 2
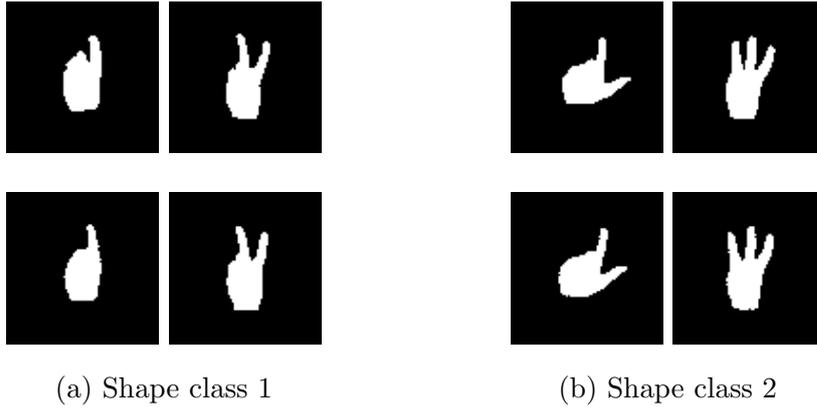
Figure 4.9: Samples of binary images (training dataset) in the hand dataset: each row depicts a shape of a single object in 2 different time points.

Figure (4.10) shows samples from the test dataset.

(a) Shape class 1   (b) Shape class 2
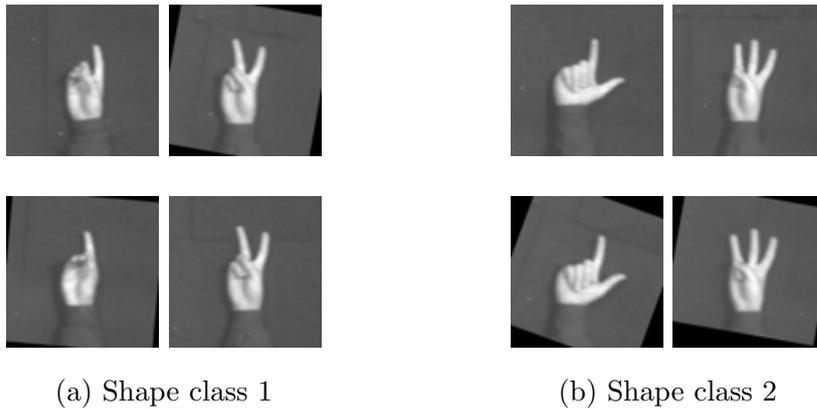
Figure 4.10: Samples of intensity images (test dataset) in the hand dataset: each row depicts a shape of a single object in 2 different time points.

- **Dendritic spine dataset**: Dendritic spines are small protrusions covering the surface of a dendrite and have two main parts, spine head and spine neck [68]. To construct this dataset we have used 2-photon microscopy images of a dendrite in consecutive time points an example of which is shown in Figure (4.11). These images have been provided by the Neuronal Structure and Function Laboratory of the Champalimaud Neuroscience Foundation, Lisbon. The spines are detected by a domain expert in temporal images, and we center-crop and align them manually to generate our test dataset. The training dataset is also constructed using the manually segmented spines by the expert.

Some examples are shown in Figure (4.12). Figure (4.12a) and (4.12b) present samples from the training and the test dataset, respectively. In this dataset we deal with challenging real data which can provide a more realistic evaluation of our proposed method. 2-photon microscopy images are low quality images and very challenging to deal with in a segmentation task. The reason is that, the dendrite and the spine appear to have similar intensity range in these images. Moreover, in the case of spines with neck, the spine neck does not exhibit strong contrast in comparison to the background. Therefore, a segmentation method can fail to capture the actual boundary of the spine.
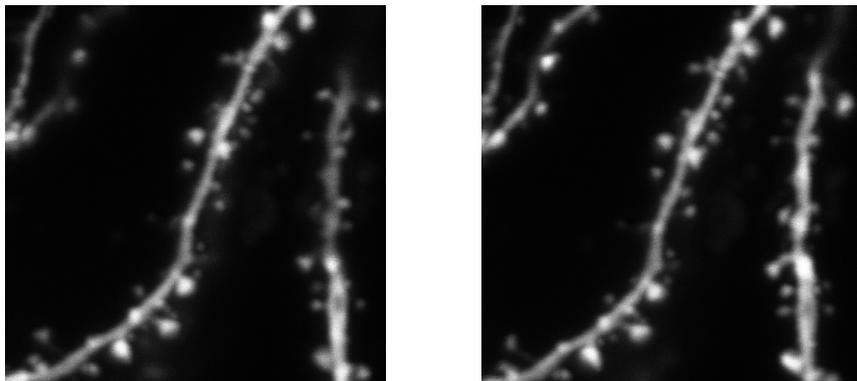
Figure 4.11: 2-photon microscopy images of a dendrite in consecutive time points.
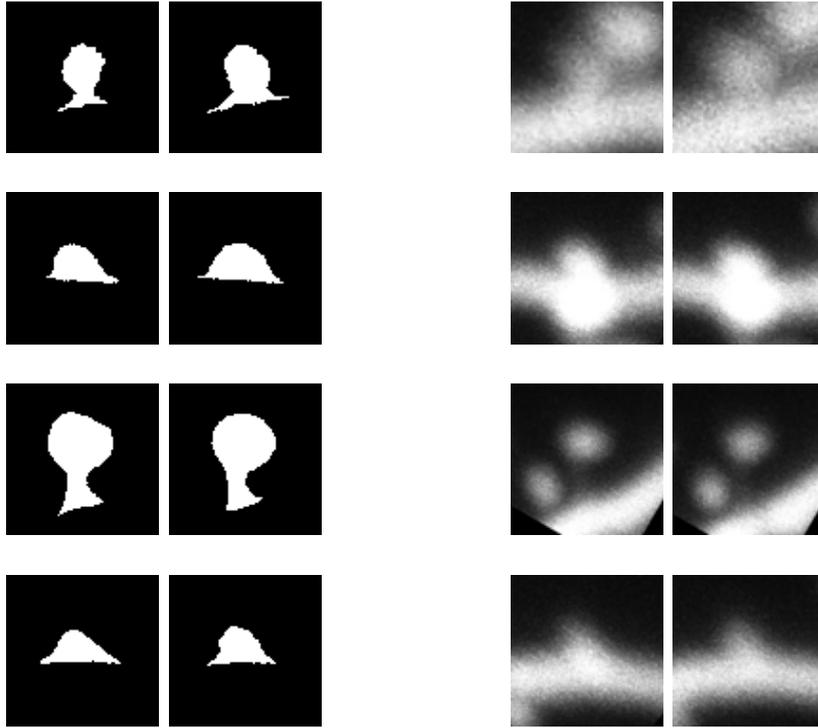
(a) Binary images (training dataset)   (b) Intensity images (test dataset)

Figure 4.12: Samples of corresponding binary and intensity images in the dendritic spine dataset: each row depicts a shape of a single object in the 2 different time points.

4.2   Experimental results

In the following sections we provide experimental results, both visual and quantitative, in order to demonstrate the performance of our algorithm. We perform two set of experiments, which are based on preliminary and generalized version of the work described in the 3 and presented in Section (4.2.1) and (4.2.2), respectively.

We obtain quantitative results based on the Dice Score measurement [69]. Dice score measures the similarity between two sets X and Y. These sets are binary images representing the ground truth and segmentation results of the segmentation algorithm. Dice score outputs values between 0 and 1 where 1 indicates perfect match between X and Y. Dice score is computed as follows:

$$D(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|}. \tag{4.1}$$

4.2.1 Results based on preliminary work

Our preliminary approach has been evaluated on the spine dataset. In Table (4.1), we present the quantitative (Dice Score) results obtained by applying our proposed approach on the spine dataset and compare these results with Kim's approach [17].

Table 4.1: Dice Score results on the spine dataset

	Preliminary Method	**Kim et al. [17]**
Spine 1	0.6386	0.6379
Spine 2	0.7520	0.7470
Spine 3	0.6343	0.6366
Spine 4	0.7847	0.6727
Spine 5	0.6627	0.6655
Spine 6	0.7937	0.6673
Spine 7	0.6689	0.6906
spine 8	0.7905	0.6789
Spine 9	0.7488	0.6414
Spine 10	0.7414	0.6643
Spine 11	0.7536	0.6693
Spine 12	0.7299	0.6435
Spine 13	0.7703	0.6753
Spine 14	0.7523	0.5977
Spine 15	0.6872	0.7057
Spine 16	0.6908	0.7145
Spine 17	0.6965	0.5955
Spine 18	0.8041	0.7352
Spine 19	0.8079	0.6626
Spine 20	0.7230	0.6473
Spine 21	0.7963	0.6566
Spine 22	0.7903	0.7696
Spine 23	0.7162	0.6700
Average	**0.7362**	**0.6715**

We also provide some visual results obtained by the proposed approach and the approach of Kim et al. [17] (see Figure 4.13). Also the ground truth can be found in the same figure to have a more precise comparison between the performances.
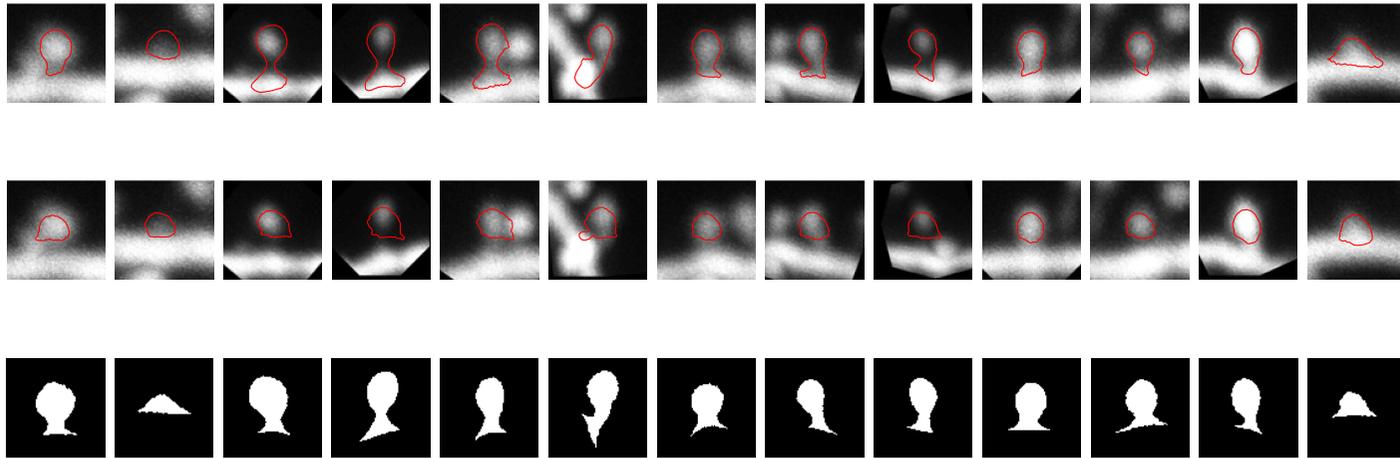
Figure 4.13: Examples of visual segmentation results: The first row is obtained by the preliminary version. The second row is obtained by Kim's method. The third row is the ground truth.

The visual results also demonstrate that the proposed approach is more successful in firstly, identifying the shape class, secondly providing a better representation of the spine's shape and obtaining the boundaries.

4.2.2 Results based on generalized work

We present the results obtained by the generalized version of our work on the three datasets.

4.2.2.1 Synthetic dataset

As we mentioned earlier, the challenges we face in this dataset is the high level of noise which causes the boundaries of the object not to appear clearly in the image. In this section we present the results obtained by applying our proposed approach on the synthetic dataset. We compare these results with Chan and Vese algorithm [9] and Kim et al. [17]. The quantitative results are provided in the Tables (4.2) and (4.3) :

Table 4.2: Dice Score results on synthetic dataset

	Chan and vese [9]	Kim et al. [17]	Generalized Method
image 1	0.1045	0.5747	0.6453
image 2	0.0721	0.6378	0.6166
image 3	0.1323	0.5926	0.6365
image 4	0.0994	0.7205	0.7051
image 5	0.1681	0.7553	0.6950
image 6	0.2311	0.6205	0.6150
image 7	0.2574	0.7569	0.7038
image 8	0.6633	0.1942	0.6907
image 9	0.2412	0.5021	0.7268
image 10	0.1077	0.7112	0.7294
image 11	0.1044	0.7421	0.7588
image 12	0.1859	0.8486	0.8076
image 13	0.2481	0.7411	0.8807
image 14	0.2386	0.9102	0.8200
image 15	0.2517	0.8912	0.6582
image 16	0.3141	0.8079	0.8370
image 17	0.1888	0.5815	0.6243
image 18	0.6269	0.0821	0.6169
image 19	0.6344	0.0869	0.6223
image 20	0.0805	0.6856	0.6316
image 21	0.0987	0.7074	0.6484
image 22	0.6374	0.1191	0.6758
image 23	0.6907	0.1357	0.7097

Table 4.3: Dice Score results on synthetic dataset

	Chan and vese [9]	**Kim et al. [17]**	**Generalized Method**
image 24	0.4327	0.5291	0.8213
image 25	0.1498	0.6034	0.8661
image 26	0.2846	0.6737	0.8668
image 27	0.4327	0.4322	0.7983
image 28	0.1367	0.5542	0.6930
image 29	0.1178	0.4021	0.6938
image 30	0.2326	0.7083	0.7273
image 31	0.1627	0.7162	0.8212
image 32	0.1995	0.7462	0.8859
image 33	0.2246	0.7248	0.8599
image 34	0.1814	0.7629	0.7517
image 35	0.7517	0.2964	0.8532
Average	**0.282**	**0.5341**	**0.7407**

Here, visual results are provided in Figure (4.14). In these figures, we provide the results obtained by the Chan and Vese followed by the result obtained by the Kim et al.. Later on, the results of our proposed approach are presented. Lastly, the ground truth is provided.
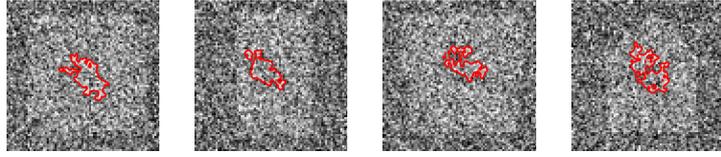
(a) The results obtained by Chan and Vese method
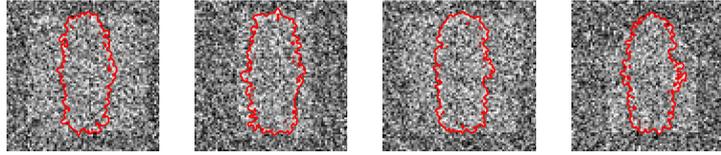
(b) The results obtained by Kim et al.'s method
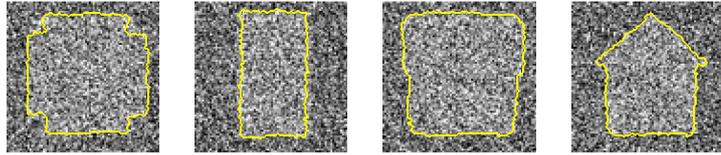
(c) The results obtained by the proposed method

(d) The ground truth

Figure 4.14: Examples of visual segmentation results on 4 different test images in the synthetic dataset

By comparing the provided results, we can see that the Chan and Vese method can not handle the high level of noise in the images and terminates the curve evolving process after a few iterations (approximately 50 iterations in our case). The Kim's approach can continue evolving the curve using the shape priors, however, it fails to evolve the curve towards the true class of the shape. The proposed approach is successful in detecting the shape class and handling noise.

In order to demonstrate more advantages of our proposed algorithm, we have performed more experiments based on different scenarios which are provided in the following.

• **Scenario 1**

In this problem, we consider that we are provided with a test image with missing data. An example of such test image and its corresponding ground truth can be seen in the Figure (4.15).
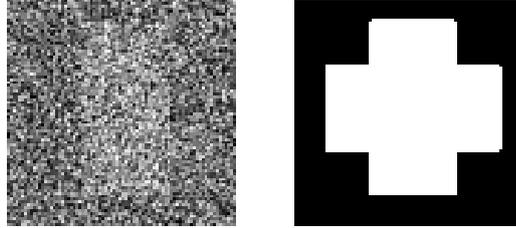
Figure 4.15: Sample test image with data loss and its corresponding ground truth

The result of the proposed algorithm and is shown in the Figure (4.16):
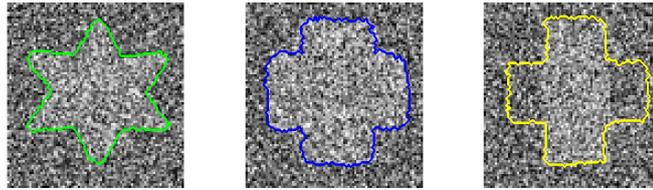
Figure 4.16: The result of the proposed algorithm on a test image with data loss : from left images are estimated curves for the shape in 2nd, 3rd, and 4th time points, respectively.

As it can be seen, the proposed approach is successful in capturing the missing part of the shape.

• **Scenario 2**

Let us assume that we have two test images with missing parts and they are similar to each other in terms of shape and size. An example of this case in shown in Figure (4.17). The result of our algorithm on the similar looking images are provided in Figure (4.18).
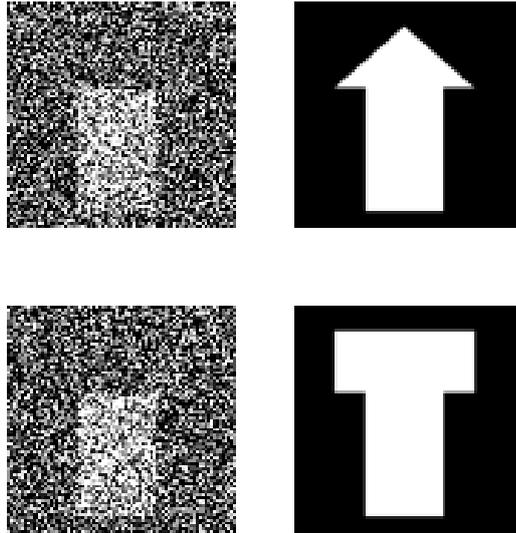
Figure 4.17: Similar looking test images after data loss, which belong to different shape classes and their corresponding ground truth. First and second row are shape1 and shape 2, respectively
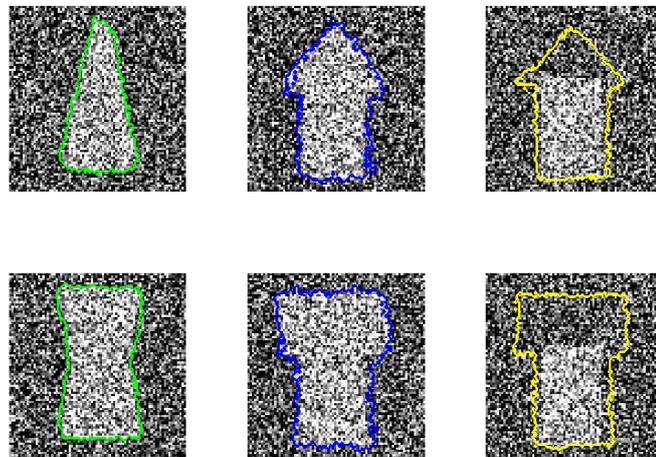
Figure 4.18: The result of the proposed algorithm on Similar looking test images after data loss, which belong to different shape classes. First and second row are shape1 and shape 2, respectively. Their estimated curves in the previous time points are also provided.

As it can be seen in Figure 4.18, despite the similarity in the data, our proposed algorithm can identify the class of each object in the test images and evolve the curve towards the corresponding shape class.

- **Scenario 3**

In this section, our aim is to evaluate the performance of the algorithm when we have data loss in two or more consecutive time points in the test images.

Let us assume that segmenting the shape in $4th$ time point is desired and the data is partially missed. Also, the data that we are provided with in the $3d$ time point is partially missed. An example of this scenario is provided in Figure (4.19).
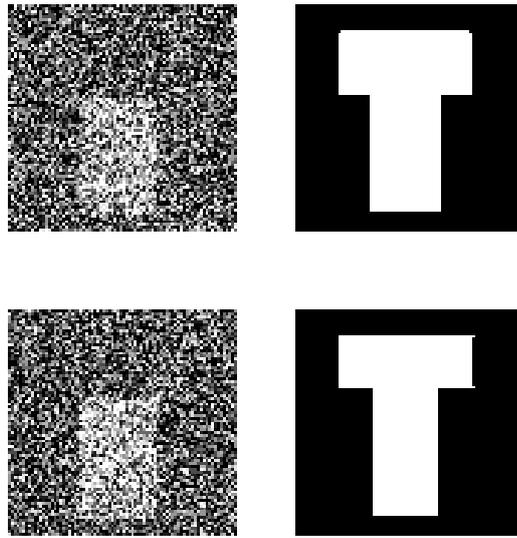
Figure 4.19: A test image with data loss in two consecutive time points: The first and second rows represent the data of the same object in 3d and 4th time point beside their corresponding ground truth

According to the results shown in Figure (4.20) obtained by our algorithm, the proposed method is successful in segmenting the test data in $4th$ time point, although the data loss occurred in consecutive time points.

Figure 4.20: The segmentation result on a test image with data loss in two consecutive time points. Its estimated curves the in previous time points are also provided.

- *Scenario 4*

In this scenario, we deal with occluded test images. Examples are shown in the Figure 4.21
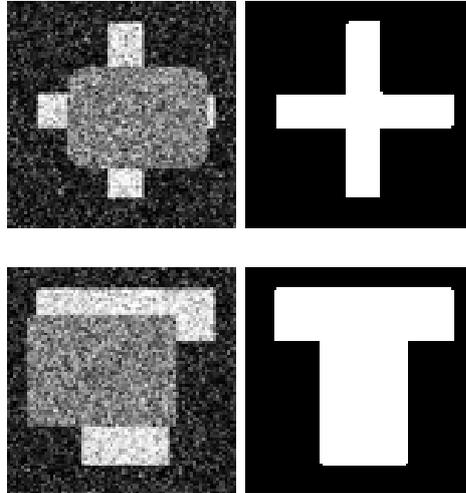
Figure 4.21: Examples of partially occluded the test images beside their corresponding ground truth

Figure 4.22, presents the results produced by our proposed method, which demonstrates the capability of the proposed method in coping with occluded images.
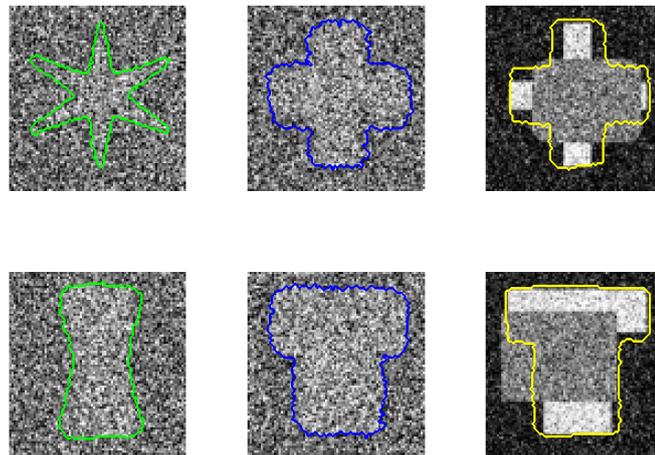
Figure 4.22: The segmentation result on partially occluded test images. Their estimated curves in the previous time points are also provided.

4.2.2.2 Hand dataset

The hand dataset is a real dataset, the deformation dynamics are defined synthetically though, and evaluating the proposed algorithm's performance on real data is beneficial. As already shown in Figure 4.10a and 4.10b, we do not deal with very noisy images in this dataset. Therefore, we expect other algorithms such as Chan and Vese or Kim et al. to be capable of delivering promising results, yet they fail to capture the accurate boundary of the hand in case of data loss. The proof of our assumption is provided in Figure (4.23) and Figure (4.24) . In Figure (4.23), all three methods (Chan and Vese, Kim et al. , and the proposed method) segment the hand shape successfully. The Figure (4.24) depicts a test image of which some parts are lost and its corresponding ground truth. Despite the similarity of the shape to the hand gestures in class 1, the test dataset is coming from class 2. The results of all the three methods are shown in Figure (4.25) and as expected the shape class of the object to be segmented is indistinguishable for Chan and Vese and Kim et al. However, the proposed method evolved the curve towards the right shape class using the dynamic shape priors.
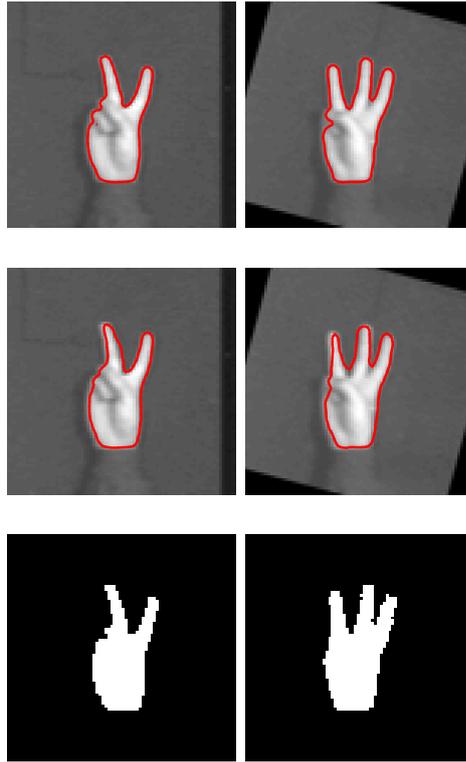
Figure 4.23: Examples of visual segmentation results on the hand dataset. First and second rows are obtained by Chan and Vese and Kim et al. method respectively. Each row represents a single hand in two different time points. The last row is the corresponding ground truth.

Figure 4.24: A test image with a missing data in the hand dataset and the corresponding ground truth
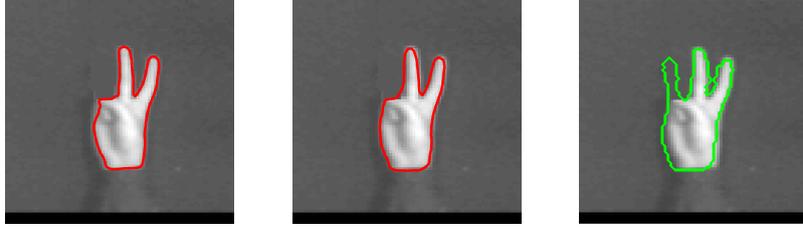
Figure 4.25: Visual segmentation results on a test image with missing data in the hand dataset: First, second and third images from left to right are obtained by Chan and Vese, Kim et al, and the proposed method respectively.

4.2.2.3   Dendritic Spine dataset

We have performed experiments based on our generalized formulation on the spine dataset. The goal is to we can demonstrate a better evaluation of our proposed dynamic method on real data with real dynamics. Moreover, it can provide a comparison between the preliminary and the generalized approach. However, a significant difference between their performance is not expected, since the general idea behind them is similar and they meant to be developed for distinct scenarios and applications (as mentioned earlier, the preliminary approach has access to the perfect segmentation in the previous time point and the generalized approach does not). The main difference that is expected to result in better performance of generalized version is that it employs the particle filtering which empowers it to exploit the shape variations and provide a better estimation of the shape of the object to be segmented. Table (4.4) and Figure (4.26) present the results which verify of our discussion as well.

Table 4.4: Dice Score results on the spine dataset

	Preliminary Method	**Generalized Method**
Spine 1	0.6386	0.6379
Spine 2	0.7520	0.7870
Spine 3	0.6343	0.5367
Spine 4	0.7847	0.7321
Spine 5	0.6627	0.6621
Spine 6	0.7937	0.8213
Spine 7	0.6689	0.7541
spine 8	0.7905	0.8123
Spine 9	0.7488	0.7611
Spine 10	0.7414	0.7533
Spine 11	0.7536	0.7892
Spine 12	0.7299	0.7961
Spine 13	0.7703	0.8369
Spine 14	0.7523	0.8191
Spine 15	0.6872	0.7057
Spine 16	0.6908	0.7694
Spine 17	0.6965	0.7855
Spine 18	0.8041	0.8566
Spine 19	0.8079	0.8712
Spine 20	0.7230	0.7325
Spine 21	0.7963	0.7721
Spine 22	0.7903	0.8035
Spine 23	0.7162	0.7584
Average	**0.7362**	**0.7632**
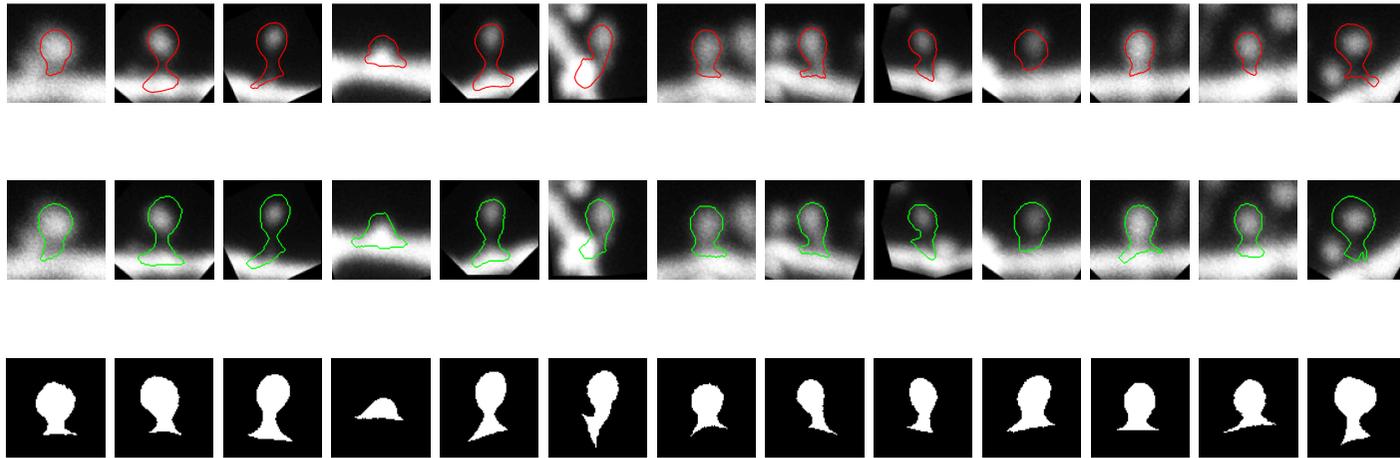
Figure 4.26: Examples of visual segmentation results: The first row is obtained by the preliminary version of the work. The second row is obtained by the generalized version of the work. The third row is the ground truth.

Chapter 5

Conclusion

5.1   Summary

The problem of segmenting deformable objects in time series images is encountered in many automated image analysis algorithms. It can turn into a very challenging task in cases of the high level of noise, weak exhibition of contrast between the foreground region of interest and the background, missing data and occlusion. Such challenges motivate the incorporation of prior knowledge into the segmentation algorithm.

In case of deformable objects, using a prior which just favors the length of the curve or a static shape prior appears to be insufficient. Imposing a prior based on the deformation dynamics should be provided in order to obtain promising results. Also, there are many applications in which the object to be segmented belongs to different shape categories. Thus, an effective segmentation algorithm should be capable of handling multi-modal prior shape densities as well.

In this thesis, we developed a methodology to address these issues. We have proposed a segmentation algorithm for deformable objects that employs the shape of the object in the previous time point, which we refer to as couple shape priors. Proposed coupled shape prior can learn the dynamics of the deformation non-parametrically and enhance the segmentation results by incorporating it in an active contour-based segmentation process. Also, it assists in evolving the curve towards the right class of shapes. In order to generalize our approach, we have proposed a sampling based algorithm using Sequential Monte Carlo (Particle Filters). By involving curve evolution, we generate samples based on the current observation in order to obtain the most accurate samples representing the posterior distribution. Afterwards, an

estimation of the posterior distribution is provided using weighted samples.

5.2 Future Work

In the following, we introduce some potential topics on which one can focus on taking this work further and suggest ways to increase the accuracy of the segmentation outputs.

As the main contribution of our work, we have proposed a dynamic shape prior which is capable of utilizing the dynamics of temporal shape deformation. Alternative shape priors can be introduced in order to learn and incorporate these dynamics into the segmentation process. This would lead to more sophisticated importance densities in the framework of particle filters as well.

As discussed in Chapter (4), training data imposes some constraints regarding learning the prior information. By introducing a training data augmentation approach, one can improve the ability of modeling the dynamics.

In our approach, L_2 distance between level set representation of the shapes has been used in order to find the distances between shapes. L_2 distance can be replaced by any other distance measure which is more sophisticated in computing the similarities between shapes.

Incorporating other kinds of prior knowledge than shapes is another way of improving the proposed algorithm. The intensity information in the previous time points can be favorable in some applications and can increase the robustness of the segmentation algorithm.

One can consider upgrading this work to a complete tracking algorithm by adding motion modeling as well. Since the proposed algorithm can handle the occlusion issue and benefits from the temporal deformation dynamics, it can easily fit in a tracking framework.

In the present work, the issues related to the effect of pose difference is not addressed and all the test and training images are assumed to be aligned to remove the shape variations for sake of simplicity. There have been approaches proposed to address this problem such as [70], which incorporates pose priors. This approach or similar ones can be integrated in our algorithm to update the pose of the evolving curve in each iteration.

Bibliography

[1] R. C. Gonzalez, R. E. Woods, S. L. Eddins *et al.*, *Digital image processing using MATLAB.* Pearson-Prentice-Hall Upper Saddle River, New Jersey, 2004, vol. 624.

[2] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.

[3] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," in *Applications of Artificial Intelligence II*, vol. 548. International Society for Optics and Photonics, 1985, pp. 2–10.

[4] N. M. Zaitoun and M. J. Aqel, "Survey on image segmentation techniques," *Procedia Computer Science*, vol. 65, pp. 797–806, 2015.

[5] Y. J. Zhang, "A survey on evaluation methods for image segmentation," *Pattern recognition*, vol. 29, no. 8, pp. 1335–1346, 1996.

[6] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International journal of computer vision*, vol. 1, no. 4, pp. 321–331, 1988.

[7] V. Caselles, "Geometric models for active contours," in *Image Processing, 1995. Proceedings., International Conference on*, vol. 3. IEEE, 1995, pp. 9–12.

[8] G. H. Ho and P. Shi, "Domain partitioning level set surface for topology constrained multiobject segmentation," in *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on.* IEEE, 2004, pp. 1299–1302.

[9] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on image processing*, vol. 10, no. 2, pp. 266–277, 2001.

[10] A. Tsai, A. Yezzi, and A. S. Willsky, "A curve evolution approach to medical image magnification via the mumford-shah functional," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2000, pp. 246–255.

[11] J. Kim, J. W. Fisher, A. Yezzi, M. Çetin, and A. S. Willsky, "A nonparametric statistical method for image segmentation using information theory and curve evolution," *IEEE Transactions on Image processing*, vol. 14, no. 10, pp. 1486–1502, 2005.

[12] D. Cremers, "Image segmentation with shape priors: Explicit versus implicit representations," *Handbook of Mathematical Methods in Imaging*, pp. 1909–1944, 2015.

[13] M. Rousson and N. Paragios, "Shape priors for level set representations," in *European Conference on Computer Vision.* Springer, 2002, pp. 78–92.

[14] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Biomedical Imaging, 2002. 5th IEEE EMBS International Summer School on.* IEEE, 2002, pp. 8–pp.

[15] A. Tsai, A. Yezzi Jr, W. Wells III, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "Model-based curve evolution technique for image segmentation," in *null.* IEEE, 2001, p. 463.

[16] D. Cremers, S. J. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation," *International journal of computer vision*, vol. 69, no. 3, pp. 335–351, 2006.

[17] J. Kim, M. Çetin, and A. S. Willsky, "Nonparametric shape priors for active contour-based image segmentation," *Signal Processing*, vol. 87, no. 12, pp. 3021–3044, 2007.

[18] M. Erdt, S. Steger, and G. Sakas, "Regmentation: A new view of image segmentation and registration," *Journal of Radiation Oncology Informatics*, vol. 4, no. 1, pp. 1–23, 2017.

[19] B. Xiang, "Knowledge-based image segmentation using sparse shape priors and high-order mrfs," Ph.D. dissertation, Châtenay-Malabry, Ecole centrale de Paris, 2013.

[20] S. O. R. Fedkiw and S. Osher, "Level set methods and dynamic implicit surfaces," *Surfaces*, vol. 44, p. 77, 2002.

[21] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *Computer Vision, 1995. Proceedings., Fifth International Conference on*. IEEE, 1995, pp. 694–699.

[22] ——, "Geodesic active contours," *International journal of computer vision*, vol. 22, no. 1, pp. 61–79, 1997.

[23] L. D. Cohen, "On active contour models and balloons," *CVGIP: Image understanding*, vol. 53, no. 2, pp. 211–218, 1991.

[24] L. D. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-d and 3-d images," *IEEE Transactions on Pattern Analysis and machine intelligence*, vol. 15, no. 11, pp. 1131–1147, 1993.

[25] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Gradient flows and geometric active contour models," in *Computer vision, 1995. proceedings., fifth international conference on*. IEEE, 1995, pp. 810–815.

[26] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on pure and applied mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[27] ——, "Boundary detection by minimizing functionals," in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 17. San Francisco, 1985, pp. 137–154.

[28] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *International journal of computer vision*, vol. 50, no. 3, pp. 271–293, 2002.

[29] L. H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 11, pp. 1061–1075, 1992.

[30] T. Riklin-Raviv, N. Kiryati, and N. Sochen, "Segmentation by level sets and symmetry," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1. IEEE, 2006, pp. 1015–1022.

[31] A. Yezzi, A. Tsai, and A. Willsky, "A statistical approach to image segmentation for bimodeal and trimodal imagery," in *Proceedings of International Conference on Computer Vision (ICCV)*, 1999.

[32] N. Paragios and R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 3, pp. 266–280, 2000.

[33] A. Chakraborty, L. H. Staib, and J. S. Duncan, "Deformable boundary finding in medical images by integrating gradient and region information," *IEEE Transactions on Medical Imaging*, vol. 15, no. 6, pp. 859–870, 1996.

[34] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations," *Journal of computational physics*, vol. 79, no. 1, pp. 12–49, 1988.

[35] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape," *International journal of computer vision*, vol. 72, no. 2, pp. 195–215, 2007.

[36] G. Sapiro, *Geometric partial differential equations and image analysis.* Cambridge university press, 2006.

[37] A. Blake and M. Isard, "Active contours: The application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion. secaucus," *Springer0Verlag, London*, 1998.

[38] J. Kim, J. W. Fisher, A. Yezzi, M. Cetin, and A. S. Willsky, "Nonparametric methods for image segmentation using information theory and curve evolu-

tion," in *Image Processing. 2002. Proceedings. 2002 International Conference on*, vol. 3.   IEEE, 2002, pp. 797–800.

[39] T. Brox, M. Rousson, R. Deriche, and J. Weickert, "Unsupervised segmentation incorporating colour, texture, and motion," in *International conference on computer analysis of images and patterns.*   Springer, 2003, pp. 353–360.

[40] M. Unger, T. Pock, W. Trobin, D. Cremers, and H. Bischof, "Tvseg-interactive total variation based image segmentation." in *BMVC*, vol. 31.   Citeseer, 2008, pp. 44–46.

[41] S. P. Awate, T. Tasdizen, and R. T. Whitaker, "Unsupervised texture segmentation with nonparametric neighborhood statistics," in *European Conference on Computer Vision.*   Springer, 2006, pp. 494–507.

[42] D. Cremers, F. Tischhäuser, J. Weickert, and C. Schnörr, "Diffusion snakes: Introducing statistical shape knowledge into the mumford-shah functional," *International journal of computer vision*, vol. 50, no. 3, pp. 295–313, 2002.

[43] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer vision and image understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[44] C. Kervrann and F. Heitz, "Statistical deformable model-based segmentation of image motion," *IEEE Transactions on Image Processing*, vol. 8, no. 4, pp. 583–588, 1999.

[45] D. Cremers, T. Kohlberger, and C. Schnörr, "Shape statistics in kernel space for variational image segmentation," *Pattern Recognition*, vol. 36, no. 9, pp. 1929–1943, 2003.

[46] R. Courant and D. Hilbert, "Methods of mathematical physics, vol. i," *Physics Today*, vol. 7, no. 5, pp. 17–17, 1954.

[47] B. J Mercer, "Xvi. functions of positive and negative type, and their connection the theory of integral equations," *Phil. Trans. R. Soc. Lond. A*, vol. 209, no. 441-458, pp. 415–446, 1909.

[48] T. Riklin-Raviv, N. Kiryati, and N. Sochen, "Unlevel-sets: geometry and prior-based segmentation," in *European Conference on Computer Vision*. Springer, 2004, pp. 50–61.

[49] G. Charpiat, O. Faugeras, and R. Keriven, "Approximations of shape metrics and application to shape warping and empirical shape statistics," *Foundations of Computational Mathematics*, vol. 5, no. 1, pp. 1–58, 2005.

[50] R. H. Davies, T. F. Cootes, and C. J. Taylor, "A minimum description length approach to statistical shape modelling," in *Biennial International Conference on Information Processing in Medical Imaging*. Springer, 2001, pp. 50–63.

[51] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics*, pp. 832–837, 1956.

[52] E. Parzen, "On estimation of a probability density function and mode," *The annals of mathematical statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.

[53] D. Cremers, "Dynamical statistical shape priors for level set-based tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1262–1273, 2006.

[54] J. Sénégas, T. Netsch, C. A. Cocosco, G. Lund, and A. Stork, "Segmentation of medical images with a shape and motion model: A bayesian perspective," in *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis*. Springer, 2004, pp. 157–168.

[55] W. Sun, M. Cetin, R. Chan, V. Reddy, G. Holmvang, V. Chandar, and A. Willsky, "Segmenting and tracking the left ventricle by learning the dynamics in cardiac images," in *Biennial International Conference on Information Processing in Medical Imaging*. Springer, 2005, pp. 553–565.

[56] D. Cremers, T. Kohlberger, and C. Schnörr, "Nonlinear shape statistics in mumfordshah based segmentation," in *European Conference on Computer Vision*. Springer, 2002, pp. 93–108.

[57] Y. Rathi, S. Dambreville, and A. Tannenbaum, "Statistical shape analysis using kernel pca," in *Image processing: algorithms and systems, neural networks, and*

machine learning, vol. 6064. International Society for Optics and Photonics, 2006, p. 60641B.

[58] D. Cremers, N. Sochen, and C. Schnörr, "Multiphase dynamic labeling for variational recognition-driven image segmentation," in *European Conference on Computer Vision*. Springer, 2004, pp. 74–86.

[59] A. Blake and M. Isard, "The condensation algorithm-conditional density propagation and applications to visual tracking," in *Advances in Neural Information Processing Systems*, 1997, pp. 361–367.

[60] M. Isard and A. Blake, "Contour tracking by stochastic propagation of conditional density," in *European conference on computer vision*. Springer, 1996, pp. 343–356.

[61] Y. Movshovitz-Attias, *Persistent Particle Filters For Background Subtraction*. Citeseer, 2010.

[62] A. Doucet, S. Godsill, and C. Andrieu, "On sequential monte carlo sampling methods for bayesian filtering," *Statistics and computing*, vol. 10, no. 3, pp. 197–208, 2000.

[63] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on signal processing*, vol. 50, no. 2, pp. 174–188, 2002.

[64] R. Van Der Merwe, A. Doucet, N. De Freitas, and E. A. Wan, "The unscented particle filter," in *Advances in neural information processing systems*, 2001, pp. 584–590.

[65] E. Erdil, S. Yildirim, M. Cetin, and T. Tasdizen, "Mcmc shape sampling for image segmentation with nonparametric shape priors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 411–419.

[66] N. Atabakilachini, E. Erdil, A. O. Argunsah, L. Rada, D. Unay, and M. Cetin, "Coupled shape priors for dynamic segmentation of dendritic spines," in *Sig-*

nal Processing and Communications Applications Conference (SIU), 2017 25th. IEEE, 2017, pp. 1–4.

[67] H. Birk and T. B. Moeslund, *Recognizing gestures from the hand alphabet using principal component analysis.* Aalborg Universitet, 1996.

[68] K. M. Harris, "Structure, development, and plasticity of dendritic spines," *Current opinion in neurobiology*, vol. 9, no. 3, pp. 343–348, 1999.

[69] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[70] M. G. Uzunbas, O. Soldea, M. Cetin, G. Ünal, A. Erçil, D. Unay, A. Ekin, and Z. Firat, "Multi-object segmentation using coupled nonparametric shape and relative pose priors," in *Computational Imaging VII*, vol. 7246. International Society for Optics and Photonics, 2009, p. 72460H.