OPEN
UNIVERSITY OF
CYPRUS

OPEN UNIVERSITY OF CYPRUS

SCHOOL OF PURE AND APPLIED SCIENCE

OPTIMIZING IMAGE SEGMENTATION AND CLASSIFICATION METHODS IN THE PRESENCE OF INTENSITY HETEROGENEITY AND FEATURE COMPLEXITY

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

BY GREGORIS LIASIS

SUPERVISOR: STAVROS STAVROU

NICOSIA, MAY 2016

ISBN 978-9963-695-45-4

VALIDITY PAGE

DOCTORAL CANDIDATE: Gregoris Liasis

TITLE OF DOCTORAL DISSERTATION: Optimizing image segmentation and classification methods in the presence of intensity heterogeneity and feature complexity.

The present doctoral dissertation was completed in the context of the Doctoral Programme Information and Communication Systems at the School of Pure and Applied Sciences of the Open University of Cyprus and was successfully defended by the candidate on the 31st of May 2016.

EXAMINING COMMITTEE

Stavros Stavrou	Assoc. Professor at the School of Pure and Applied Sciences of the Open University of Cyprus (supervisor)
Constantinos Pattichis	Professor at the School of Pure and Applied Sciences of the University of Cyprus (chair)
Nicolaos Tsapatsoulis	Assoc. Professor at the School of Communication and Media Studies of the Cyprus University of Technology (member)
Efthyvoulos Kyriacou	Assoc. Professor at the School of Engineering and Applied Sciences Frederick University (member)
Michalis Talias	Assist. Professor at the School of Economics and Management of the Open University of Cyprus (member)

Signature: Stavros Stavrou

Assoc. Professor of Information and Communication Systems

Open University of Cyprus

……

Abstract

Research in computer vision and image processing is expanding in various fields. This expansion creates the need to develop robust methods for acquiring qualitative and quantitative information from natural or real world images that can be used without or little customization in a diversity of applications. In natural scenes, intensity heterogeneity and feature complexity often occurs, thus fully automated computerized image processing and analysis methods are considered challenging processes. Image segmentation and classification are common processes for extracting valuable information from the analysed images. In image segmentation methods, the presence of intensity heterogeneity may lead to overlaps between the regions of interest and background objects or present artefacts. In image classification methods, intensity heterogeneity may lead to misperceptions of the various classes.

The algorithms presented in this thesis, address image segmentation and classification in the presence of intensity heterogeneity and feature complexity, with applicability in mammographic and remote sensing image analysis. The presented algorithms extract quantitative and qualitative information from the analysed scenes acquired from the selected domains where the images often include intensity heterogeneity and feature complexity.

The proposed segmentation algorithms are based on active contour and level set methodologies. Two different approaches are followed for designing optimal models for automatic segmentation. In the first approach, pre-processing methods are designed to optimize the overall performance of active contour segmentation models. Image transformations are produced and indices from these representations are utilized to form optimized active contour segmentation models. Intensity heterogeneity is eliminated or reduced by designing spectral and spatial filters. The sensitivity of filter responses to the parameter selection is addressed by applying an automated scheme for tuning and optimizing the filter settings. Furthermore, the method addresses the variation of colour intensity values that may be present in the analysed image and often leads to inaccurate results, by utilizing information from various colour systems. This information is analysed and used to selectively detect regions of interest that very often are not delineated by the traditional active contour models. Moreover, active contour models often tend to be sensitive on initialization. Thus, a clustering method is developed and incorporated in the segmentation algorithm to design optimal initial level set contours and drive the

propagation process faster to achieve better segmentations. Additionally, statistical measures are used for designing optimum post-processing morphological filters to eliminate any misleading information that still exists in the final segmentation mask. In the second approach, the presence of intensity heterogeneity in complex scenes is addressed by designing an optimized localized active contour model. The proposed segmentation model is based on the decomposition of the image into intrinsic components and the use of a kernel convolution function for eliminating or reducing the intensity heterogeneity artefact. A new energy minimization active contour model for image segmentation where the presence of intensity heterogeneity is corrected during the propagation process of the active contours is suggested.

In the classification domain, statistical distributions of various texture descriptors and their combination are investigated with support vector machines for the development of an objective image and/or regions of interest classification framework. Integrating various sets of features is fundamental for achieving better classification performance thus a model for the optimal integration of multiple feature sets for image classification in the presence of intensity heterogeneity is proposed.

This thesis also suggests how the proposed methods may be combined in a single framework for the quantitative and qualitative image analysis in real world applications. The optimized active contour segmentation methods are used in mammographic images for automatic breast region delineation and abnormalities extraction. In satellite images, the proposed methods are applied for building boundaries and building shadows extraction. Subsequently, this information is utilized for estimating the height of the corresponding building structures. The classification framework is applied for breast density characterization and breast abnormalities detection. In addition, a no-reference image quality assessment model is developed based on the proposed classification model for extracting useful information about the quality of the images under evaluation. In all the proposed models, special attention is given on the automation of the implemented methods and their ability to deal with big data of various classes. Experimental results are shown at each developed stage of the thesis.

Περίληψη

Η έρευνα στην υπολογιστική όραση και επεξεργασία εικόνας επεκτείνεται σε διάφορους τομείς. Η επέκταση αυτή δημιουργεί την ανάγκη ανάπτυξης αποτελεσματικών μεθόδων για την εξαγωγή χρήσιμης ποσοτικά και ποιοτικά πληροφορίας από πραγματικές εικόνες που να έχουν παράλληλα την δυνατότητα να εφαρμοστούν με περιορισμένες η καθόλου προσαρμογές σε διάφορα πεδία. Οι σκηνές που αναπαριστώνται σε πραγματικές εικόνες συχνά παρουσιάζουν ετερογενής ένταση και πολύπλοκα χαρακτηριστικά. Αυτό καθιστά την ανάπτυξη αυτοματοποιημένων υπολογιστικών μεθόδων επεξεργασίας και ανάλυσης εικόνων ιδιαίτερα δύσκολη. Η κατάτμηση και ταξινόμηση εικόνων αποτελούν βασικές διαδικασίες εξαγωγής χρήσιμης πληροφορίας. Στις μεθόδους κατάτμησης η παρουσία ετερογενής έντασης μπορεί να οδηγήσει σε αλληλοεπικαλύψεις μεταξύ των περιοχών ενδιαφέροντος και του φόντου όπως επίσης και σε λανθασμένες κατατμήσεις. Στις μεθόδους ταξινόμησης η παρουσία ετερογενής έντασης μπορεί να οδηγήσει σε εσφαλμένη εκτίμηση της κατηγορίας που ανήκουν οι περιοχές ενδιαφέροντος.

Οι αλγόριθμοι που παρουσιάζονται στην διατριβή αυτή πραγματεύονται την κατάτμηση και ταξινόμηση εικόνων στην παρουσία ετερογενής έντασης και πολυπλοκότητας με δυνατότητα εφαρμογής στην ανάλυση μαστογραφιών και δορυφορικών εικόνων. Εξάγουν χρήσιμη πληροφορία τόσο ποσοτικά όσο και ποιοτικά από τις υπό ανάλυση σκηνές που έχουν ληφθεί από τους συγκεκριμένους τομείς όπου οι εικόνες συνήθως παρουσιάζουν ετερογενής ένταση και πολυπλοκότητα.

Οι προτεινόμενοι αλγόριθμοι κατάτμησης βασίζονται στις μεθοδολογίες ενεργών καμπυλών και στις συναρτήσεις συνόλου επιπέδου. Δύο διαφορετικές προσεγγίσεις ακολουθούνται για τον σχεδιασμό βέλτιστων αυτοματοποιημένων μοντέλων κατάτμησης. Στην πρώτη προσέγγιση αναπτύσσονται διαδικασίες προ-επεξεργασίας με στόχο να βελτιώσουν την συνολική απόδοση κατάτμησης των μοντέλων ενεργών καμπυλών. Οι εικόνες μετασχηματίζονται και δείκτες από τις αναπαραστάσεις αυτές κωδικοποιούνται και χρησιμοποιούνται για την ανάπτυξη βέλτιστων μοντέλων ενεργών καμπυλών κατάτμησης περιοχών ενδιαφέροντος. Η ετερογενής ένταση εξαλείφεται η μειώνεται με τον σχεδιασμό και την εφαρμογή χωρικών και φασματικών φίλτρων. Στην ανάπτυξη φίλτρων εικόνας είναι σημαντική η επιλογή των πιο κατάλληλων παραμέτρων και στην διατριβή αυτή προτείνεται μια αυτοματοποιημένη διεργασία προσέγγισης βέλτιστων ρυθμίσεων. Επιπλέον, η μέθοδος αντιμετωπίζει την διακύμανση της έντασης των χρωμάτων που μπορεί να υπάρχει στις εικόνες αξιολόγησης και συχνά οδηγεί σε ανακριβή

αποτελέσματα, αξιοποιώντας πληροφορίες από διάφορα χρωματικά συστήματα. Η ανάλυση των πληροφοριών αυτών χρησιμοποιείται για τον εντοπισμό επιλεγμένων περιοχών ενδιαφέροντος που συχνά δεν μπορούν να οριοθετηθούν από τα παραδοσιακά μοντέλα ενεργών καμπυλών. Επιπλέον, τα μοντέλων ενεργών καμπυλών συχνά τείνουν να παρουσιάζουν ευαισθησία στην αρχικοποίηση τους. Έτσι ο αλγόριθμος κατάτμησης περιλαμβάνει μια μέθοδο τμηματοποίησης για τον σχεδιασμό βέλτιστων αρχικών ενεργών καμπυλών που οδηγεί την διάδοση των ενεργών καμπυλών στα όρια των περιοχών ενδιαφέροντος γρηγορότερα παρουσιάζοντας ταυτόχρονα μεγαλύτερη ακρίβεια. Επιπρόσθετα αναπτύσσεται μια μέθοδος τελικής επεξεργασίας για την αφαίρεση τυχόν παραπλανητικών τεχνουργημάτων, με την εφαρμογή μορφολογικών τελεστών όπου οι παράμετροι τους καθορίζονται αυτοματοποιημένα μετά από στατιστική ανάλυση. Στην δεύτερη προσέγγιση, η παρουσία ετερογενής έντασης αντιμετωπίζεται με τον σχεδιασμό ενός βέλτιστου τοπικού μοντέλου ενεργών καμπυλών. Το προτεινόμενο αυτό μοντέλο κατάτμησης βασίζεται στην ανάλυση των ενδογενή συστατικών της εικόνας όπου με την χρήση συναρτήσεων συνέλιξης η ετερογενής ένταση που θεωρείται ως τεχνούργημα χαμηλών συχνοτήτων εξαλείφεται η μειώνεται. Το προτεινόμενο τοπικό μοντέλο ενεργών καμπυλών περιλαμβάνει έναν νέο όρο ελαχιστοποίησης ενέργειας που παρέχει την δυνατότητα διόρθωσης της ετερογενής έντασης κατά την πορεία εξέλιξης των ενεργών καμπυλών.

Στο πεδίο ταξινόμησης εικόνων η στατιστική κατανομή διαφόρων χαρακτηριστικών υφής και ο συνδυασμός τους διερευνάται διαμέσου των μηχανών υποστηρικτικών διανυσμάτων για την ανάπτυξη ενός αντικειμενικού μοντέλου κατηγοριοποίησης και χαρακτηρισμού. Ο συνδυασμός διαφόρων χαρακτηριστικών είναι θεμελιώδης για την καλύτερη απόδοση μεθόδων ταξινόμησης και έτσι στην έρευνα αυτή προτείνονται βέλτιστα μοντέλα συνδυασμού πολλαπλών χαρακτηριστικών για την ταξινόμηση περιοχών ενδιαφέροντος όπως παρουσιάζονται σε εικόνες που περιλαμβάνουν ετερογενής ένταση και πολυπλοκότητα.

Η διατριβή αυτή πραγματεύεται επίσης πως οι προτεινόμενες μέθοδοι μπορούν να συνδυαστούν για την ποσοτική και ποιοτική ανάλυση πραγματικών εικόνων. Οι βελτιστοποιημένες μέθοδοι ενεργών καμπυλών αξιολογούνται στις μαστογραφίες για την οριοθέτηση της περιοχής του μαστού και την κατάτμηση ανωμαλιών. Σε δορυφορικές εικόνες οι προτεινόμενες μέθοδοι εφαρμόζονται για την κατάτμηση κτιρίων και των σκιών τους. Η πληροφορία αυτή στην συνέχεια χρησιμοποιείται για τον υπολογισμό του ύψους των υπό αναφορά κτιρίων. Το μοντέλο ταξινόμησης εφαρμόζεται μετά την κατάτμηση της περιοχής του μαστού για τον χαρακτηρισμό της πυκνότητας του μαστού και τον εντοπισμό

ανωμαλιών που μπορεί να παρουσιάζονται. Επιπρόσθετα, ένα αντικειμενικό μοντέλο αξιολόγησης της ποιότητας δορυφορικών εικόνων αναπτύσσεται με βάση το προτεινόμενο μοντέλο ταξινόμησης που παρέχει χρήσιμη πληροφορία ως προς την ποιότητα των υπό αξιολόγηση εικόνων.

Σε όλα τα προτεινόμενα μοντέλα ιδιαίτερη έμφαση δίνεται στην αυτοματοποιημένη εφαρμογή τους και στην ικανότητα τους να διαχειρίζονται αποτελεσματικά μεγάλο αριθμό δεδομένων διαφορετικών κατηγοριών. Πειραματικά αποτελέσματα παρουσιάζονται σε όλα τα στάδια ανάπτυξης της διατριβής αυτής.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisor Professor Stavro Stavrou for his continuous and most helpful support. His constant guidance made my Ph.D study and related research possible.

I would also like to thank the members of the examining committee, Prof. Constantino Patihi, Prof. Michael Talia, Prof. Efthyvoulo Kyriacou and Prof. Nicolao Tsapatsousli for their valuable comments and encouragement.

I would also like to express my sincere thanks to Dr. Styliani Petroudi for the contribution and support provided in various fields of this research.

Finally, I would like to thank my family, to whom this thesis is dedicated to, for their love, concern, understanding and precious support throughout all these years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGGD Asymmetric Generalized Gaussian Distribution

BIRADS Breast Imaging Reporting and Data System

GA Genetic Algorithms

GGD Generalized Gaussian Distribution

GLCM Grey level Co-occurrence Matrices

CV Chan Vese

FFDM Full Field Digital Mammography

FR-IQA Full Reference Image Quality Assessment

HSV Hue Saturation Value

IQA Image quality assessment

LBF Local Binary Fitting

LBP Linear Binary Pattern

LCV Local Chan Vese

LIC Local Intensity Cluster

LSACM Local Statistical Active Contour Model

MIAS Mammographic Image Analysis Society Digital Mammogram Database

MRF Markov Random Fields

NIR Near Infrared

NLC Normalize Luminance Coefficient

NR-IQA No Reference Image Quality Assessment

OAC Optimized Active Contour

OLAC Optimized Localized Active Contour

PACS Picture Archiving Communication Systems

PC Piecewise Constant

PDE Partial Differential Equation

PS Piecewise Smooth

RGB Red Green Blue

SIFT Scale Invariant Feature Transform

SVM Support Vector Machine

Chapter 1: Introduction

1.1 Computer vision and image processing

An image is an array, or a matrix, of square pixels or picture elements arranged in columns and rows. Each coloured picture element has an assigned intensity and colour. Humans are able to perceive the structure of the scenes presented in images and process them easily. Image processing and computer vision are closely related disciplines concerned with the use of computerized algorithms to process, analyse, and understand digital images. Three types of processes low-level, middle-level and high-level are defined in this context. Low-level processes can be thought as pre-processing operations, aiming at reducing the presence of noise in images, enhancing the contrast, or sharpening the image. Low-level operations accept as input the under processed images and output images suitable to be used in the next level of operations which can be defined as information acquisition operations. Middle-level processing of images involves tasks such as segmentation and classification in order to detect and describe the various objects that are present in the analysed image. The output of middle-level operations are image attributes capable in characterizing the individual objects or the regions of interest that may exist in the under processed images. Finally, high-level processing involves the interpretation of the information provided by the middle-level analysis on all the recognized objects, giving a conceptual description of the analysed scene [1].

Image processing and computer vision algorithms have received increasing interest the recent years because of their numerous real-world applications. Researchers in computer vision and image processing are developing methods and algorithms for acquiring useful information for the shape, appearance, and illumination of objects in imagery. Active areas among others include image segmentation, classification, quality assessment, object enhancement and three-dimension modelling to produce numerical and symbolic information for the analysed scene. An additional important objective of image processing and computer vision is to enhance computationally the abilities of human vision and image understanding.

Image segmentation can be defined as the process to cluster an image into several similar regions of interest and is an important component in many image understanding algorithms and practical computer vision systems. Segmentation or partitioning an image into homogeneous areas is an important process and can assist the further analysis and thus the better understanding of the presented scenes. Many different unsupervised image

segmentation algorithms are developed and applied in science and practical applications. Segmentation using properties such as colour, intensity, contrast and texture are used to detect regions or objects of interest. In computer vision, image segmentation is one of the oldest and widely studied themes. The existing segmentation algorithms can be approximately categorized into region-based, data-clustering and edge-based. Image regions are defined as groups of connected pixels with similar properties. Region-based segmentation algorithms operate iteratively by grouping together pixels that are neighbours and have similar values and at the same time splitting the groups of pixels that are dissimilar in value. Edges refer to sharp discontinuities between homogeneous regions presenting the boundaries of different objects. Edge detection refers to the process of identifying and locating these boundaries. Clustering is a process based on statistical measurements for classifying regions in such a way that samples of the same group are more similar to one another than samples belonging to different groups. Initially the different centroids representing the different regions are identified. Subsequently, these centroids are used as reference and segmentation of regions is performed based on the similarity or dissimilarity with regard to the centroid. Furthermore, in image segmentation literature, several researchers are integrating or merging two or more image segmentation models in accordance with certain criteria for achieving better segmentation results. These algorithms are often defined as hybrid [1-2].

Classification is the process of labelling an image according to its visual content. Classification methods include artificial neural networks, decision-trees, support vector machines and fuzzy models. These methods can be divided into two major categories: supervised and unsupervised. If the image analyst predefines the classes, the method is called supervised classification. If the classes can be automatically clustered into sets of prototypes, where the analyst merely specifies the number of desired categories is called unsupervised classification. Classification algorithms typically employ two phases of processing, the training and testing. In the initial training phase, characteristic properties of typical image features are isolated and based on these, a unique description of each classification category is created. Subsequently, in the testing phase these feature space partitions are used to classify the unknown image features. The aim of image classification is to recognize objects or characterize regions based on collection of structural primitives known as features. The later has attracted significant attention in literature and many approaches have been developed in various application areas [1-6].

The available methods for image segmentation and classification are usually application specific. However, common issues needs to be resolved in most segmentation and

classification models applied in various fields, e.g. processing of images in mammographic and satellite applications. Mammographic and satellite image analysis is difficult since these images are complex in nature and rarely have simple features. Moreover, segmentation and classification algorithms applied in mammographic and satellite images are often affected by the presence of artefacts and intensity heterogeneities [3-4].

An X-ray mammogram is a projection of a compressed breast with an intensity range between the low intensities of the background to high intensities corresponding to very dense glandular or muscle tissue. An ideal mammogram would present the breast region with various intensity values and a constant background region with zero intensity value. However, this rarely occurs and in digitised mammograms low and high noise regions exist in the background. High intensity noise regions include labels, scanning artefacts, wedges and tape regions, while low intensity noise regions are a result of the X-ray imaging process [3]. The most common viewpoints of a mammogram is the Medio-Lateral Oblique view (MLO) and the Cranio-Caudal view (CC). The pectoral muscle that exists in MLO view is presented as a region with high intensity values in the left or right upper corner of the breast region and the edge of the pectoral muscle appears as a boundary between two relatively homogeneous regions. Furthermore, the breast edge is difficult to distinguish as a result of the differences in the optical properties of the different tissues. More details on the properties of the breast edge and the effects of the digitisation process are given in Chapter 5. Abnormalities, if they exist, are often presented as heterogeneous regions of different sizes, shapes and margins [5]. More details on the properties of breast abnormalities as presented in mammographic images are discussed in Chapter 10. Figure 1 presents a digitised film mammogram including all the noticed characteristics.

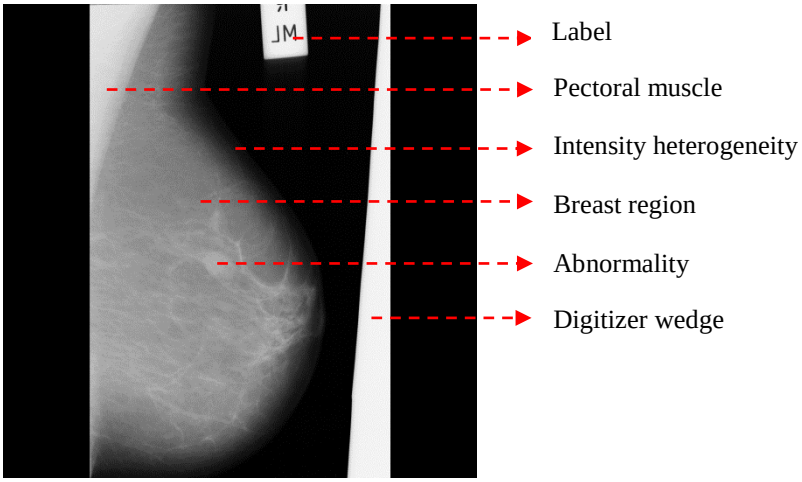


Figure 1: Complexity, intensity heterogeneity and artefacts in a digitised mammogram

In recent years, new aerial sensor devices and platforms are capable of providing high-resolution or very high-resolution remote sensing data sometimes at less than a 30-cm

ground sampling distance. This type of images is an important source for interpreting land use and urban objects. However, the large-scale difference of urban objects and spectral heterogeneity leads to a complex automatic image segmentation and classification process. Built-up areas on satellite images display diverse features, different colours and intensities. Buildings may be obscured by environmental objects and/or occlude each other producing intensity heterogeneities and complexity [6]. More details on the properties of satellite images are discussed in Chapters 6 and 7. Figure 2 presents satellite images of complex urban scenes acquired from Google Earth.

Figure 2: Complexity and intensity heterogeneity in satellite images

Classification and segmentation are closely related disciplines, as the former is another form of component labelling that can result in segmentation of various features in a scene. Segmentation and classification methods can be used for useful and meaningful information acquisition. Many advances were made in various computer vision fields the recent years. In remote sensing imagery three-dimensional surface/building models can be created using information about the presented objects derived from image segmentation and classification processes and used for urban planning, including telecommunication purposes and natural disaster assessment. In telecommunications, image segmentation and classification methods are also used for the development of global positioning systems. Finally, segmentation and classification methods are utilised in medical images such as screening X-rays, ultrasounds and MRIs for abnormalities detection.

1.2 Motivation, aim and objectives of the thesis

Despite all the advances in image processing, the attempt of a computer or any other electronic device interpreting an image effectively at a similar or very near level of a human interpretation, is a challenging task and it's still an open research field. Digital or digitised images are now available in various forms, wavelengths and dimensions, creating the need for novel solutions for automated image processing.

Computer vision and image processing of natural images is a complex task because valuable information must be extracted from the analysed scenes that usually include intensity heterogeneity and feature complexity. A variety of objects with irregular shapes, different colours, and intensities must be effectively processed. The features such as shape, illumination and colour distribution must be reconstructed while in several cases unknown

variables should be estimated. The huge volume of digital images generated by remote sensors, cameras or screening programs leads to an additional challenging aspect of the problem which is to develop automated image processing algorithms and methods that can be used with no or very little customization in different computer vision fields. The work presented in this thesis is concerned with the optimization of image segmentation and classification methods in the presence of intensity heterogeneity and feature complexity for automatically acquiring and processing useful information. This thesis is involved with the development of an optimized segmentation framework to address intensity heterogeneity and scene complexity that often occurs in natural images. Moreover, this work is concerned with the development of an objective classification framework that can be used for image quality assessment, for the characterization of the analysed image as a whole or for the classification of regions of interest that may exist in the processed scene. The design of pre-possessing filters for regions of interest enhancement is also investigated in this work. Finally, the extracted information from all the proposed methods is analysed for better image processing understanding. Image processing and computer vision fields, such as digitized X-ray mammography and satellite imagery are selected for evaluating the proposed methods.

X-ray mammography is currently the modality used for population screening in identifying breast cancer, which is one of the most common causes of death amongst women. Breast segmentation and classification plays a critical role in the extraction of useful information from mammograms. Different image processing techniques have been applied for the analysis of mammograms and segmentation and/or classification of the breast region or other regions of interest such as abnormalities are important processing steps for computer aided detection systems (CAD). Automated image analysis of mammograms is difficult since the acquisition process or the physical anatomy of the breast often leads to complexities and intensity heterogeneities within or near the regions of interest.

Segmentation and classification of monocular satellite images can serve land use analysis for the creation and update of maps, geographic information system databases and for the development of urban monitoring applications. Since satellite images present complex urban scenes, the manual processing of such images is tedious and time-consuming, thus very expensive, leading to the need for designing effective methods for automated segmentation and/or classification processing. Objects on satellite images display heterogeneities, diverse features, different colours and intensities. For successful regions of interest segmentation and/or classification, this complexity has to be handled properly.

Medical and satellite image analysis needs very good understanding of the subject and the specific application to achieve good results. Any developed algorithms and their ability to be used effectively in these different fields, where the acquired images vary, is challenging and serves the aim and the objectives of this work.

In this thesis, computer vision and imaging processing problems are addressed using the following approach. First a detailed problem definition is defined where the constraints and specifications of the problem are summarised. Subsequently, a variety of techniques is implemented and their performance is evaluated using realistic data, for selecting the most appropriate scheme to resolve the specific computer vision problem. Finally, the selected methods are optimized by addressing any limitations.

1.3 Thesis contributions

This work investigates image segmentation and classification in the presence of intensity heterogeneity and feature complexity. Intensity heterogeneity is defined as an inherent artefact caused by uneven illuminations. Feature complexity is closely related to the number and placement order of elements in the visual stimulus that makes their description difficult. Complexity is caused by the presence of a large scale of different intensities and colours representing the regions or objects of interest in the analysed scene. Curve evolution schemes like active contours, deformable models and level sets are thoroughly evaluated for the development of an image segmentation framework. These methods are selected because of their robustness to image noise and their ability to cope well with topological changes, thus can handle the complexity that may exists in natural images. Limitations of the examined methods like the sensitivity on curve initialization or the difficulty to extract selectively regions of interest in the presence of intensity heterogeneity and feature complexity in the analysed images are addressed by optimizing the traditional active contour models. Furthermore, a localized active contour model able to decrease the presence of intensity heterogeneity artefact and leads to better segmentations of natural images is also proposed. Following the work performed on image segmentation and the development of optimized solutions in the presence of intensity heterogeneity, this work investigates various classification methods and their ability to describe effectively the image as a whole and/or the regions of interest within the analysed scenes. After performing several experiments, it was found that textural features describe well images from diverse areas and different image modalities. It was also observed that integrating different set of texture features is fundamental for achieving better performance in various classification applications. Thus, a classification framework that suggests an optimal

integration of multiple feature sets using support vector machines (SVMs) is proposed. SVMs were chosen because of their ability to generalize well in high dimensional spaces. All the developed methods are evaluated against practical applications such as medical and remote sensing imagery, where often intensity heterogeneity and feature complexity occurs.

1.3.1 Mammogram breast region segmentation

Segmentation of the breast region is most important for further mammogram processing and analysis. Segmentation of the breast region from the background is made difficult because the compressed breast edge is mostly composed of adipose tissue that is radiolucent. The proposed algorithm uses level sets to establish the corresponding skin-air boundary. An initial pre-processing procedure is performed where the image contrast is improved and the breast edge is enhanced. Artefacts, labels and noise are also removed from the mammogram. Following, the breast boundary is identified using the level set formulation, by Chan and Vese. Post-processing has been also applied for spline smoothing and for line artefacts removal resulting in a smooth breast boundary. The results were quantitatively evaluated against the radiologists' ground truth.

1.3.2 Abnormalities classification using mammographic images

A classification framework is developed for investigating the use of different texture features from grey-level co-occurrence matrix in order to detect and classify abnormalities. Different texture features such as autocorrelation, contrast, cluster prominence, energy, entropy and homogeneity are evaluated on ground truth segmentation provided by expert radiologists. The ability of the corresponding features to characterize different abnormalities distinctly is defined, by comparing the range of values the features achieve for normal and abnormal regions of mammograms and using a support vector machine classifier. The different features are evaluated not only on the entire training set, but also on subsets that hold mammograms corresponding to different breast density classes. The values of these features range not only for normal and abnormal tissue, but also for a different tissue density. For the feature selection, the importance of each feature and its effect on the classification accuracy was investigated. The features that provided the best discrimination were selected for abnormality characterization.

1.3.3 Breast density classification using mammographic images

Breast density is an important risk for developing breast cancer. In this work, statistical distributions of different texture descriptors Scale Invariant Feature Transforms (SIFT), Local Binary Patterns (LBP), histograms of texton and their combination are investigated with Support Vector Machines (SVMs) for objective breast density classification. The breast density classification accuracy of the SVM classifiers modelled on the histograms of the three different set of texture features separately and their combination is evaluated on the Medical Image Analysis Society (MIAS) mammography database. The integration of different set of texture features proved to be important for achieving better performance in breast density classification. An effective method for the integration of multiple texture feature sets for breast density classification is developed. The evaluation results showed that this scheme achieved better accuracy on breast density classification, compared with the schemes that utilized the individual texture descriptor sets or the integration of descriptor sets at feature level.

1.3.4 Building structures segmentation using satellite images

An algorithm for building extraction in satellite images using level set and colour features is proposed. The previous work on image segmentation based on active contours without edges model is initially evaluated for its ability to detect building structures. It was found that for selectively extracting the building structures in complicate scenes of urban areas, the correct placement of the initial level set curves is important. Level set segmentation tends to be sensitive on initialization, thus proper initialization can yield better segmentation results. In this work, an effective procedure is performed using a k-means classifier for the design and development of optimal initial level set contours. Morphological features are incorporated for refining the obtained outlines. Finally, the outline of each building is extracted along with additional information for the processed scene, like the number of buildings, as well as the centre and area of each building. The optimization algorithm has been evaluated qualitative and quantitative against the original Chan-Vese model and proved to provide results that are more accurate. However, it was found that in several cases where spectral heterogeneity exists, over-detection or under-detection is still noticed. Thus, the RGB representation and the properties of the Hue, Saturation and Value (HSV) colour space have been analysed and used to optimize the extraction of buildings from satellite images in an active contour segmentation framework. Initially, the satellite image is pre-processed by applying a clustering technique using

colour features to eliminate vegetation areas and shadows that may adversely affect the performance of the algorithm. Subsequently, the image is decomposed in HSV representation and a new active contour model that has been developed is applied for building extraction, utilizing descriptors derived from the value and saturation images. A new energy term is encoded for biasing the contours to achieve better segmentation results. Additionally, statistical measures are used for designing optimum morphological filters to eliminate any misleading information that may still exist.

1.3.5 Shadow of building structures detection and building height estimation using satellite images

The presence of building shadows in satellite images can be processed and used for estimating the height of the buildings. The algorithm incorporates a Gabor filter designed to enhance the shadows presented in satellite or aerial images. Gabor filter response is very sensitive to the parameter selection thus an effective procedure has been developed using genetic algorithms for tuning and optimizing the Gabor filter parameters. The Gabor feature image is then utilized and an active contour model tailored to detect the shadow of buildings is proposed. Subsequently, the height of the buildings is estimated using the shadow length and the projection angle. Extensive experiments were done and the produced segmentation results of heterogeneous satellite images, verify that the suggested method works effectively. The proposed approach deals very well with complex-shaped buildings and avoids the possible disturbance of misleading information such as vegetation areas. Furthermore, by using the shadow mask for calculating the shadow length and by incorporating the projection angle in the suggested algorithm, the height of the corresponding building structure is estimated.

1.3.6 Optimized localized active contour segmentation model

An optimized localized active contour segmentation model with application to intensity heterogeneity and complex images is proposed. The algorithm utilizes the active contours and level set methodologies to extract boundary information from the acquired images. An optimized localized active contour segmentation model based on the decomposition of the image into intrinsic components and the use of a low pass convolution function for eliminating the intensity heterogeneity artefact is proposed. This work suggests a new energy minimization active contour model for image segmentation in the presence of intensity heterogeneity and feature complexity. In the proposed model, the low pass convolution function is used for constructing the local energy term of the active contour

model. The proposed model has been used in mammography images for automated breast region extraction and in satellite or aerial images for the automated detection of building boundaries. The algorithm has been evaluated using quantitative measures and promising results in terms of accuracy and robustness were achieved.

1.3.7 No-reference image quality assessment model

Satellite image archives represent a significant and rapidly expanding research resource. Provided images may undergo distortions during the preliminary acquisition process, compression, restoration, communication or final display that affect the level of detail of the captured scene. Image quality and information on whether an image is highly detailed or very smooth, plays an important role in several image processing applications. An objective or no-reference satellite image quality assessment model, using level of detail notion and feature appearance descriptors is proposed. The classification framework, as proposed in this work utilizes an integration feature descriptors scheme to characterize the quality of satellite images. The obtained results reveal the robustness of the method for automatically perceiving the quality of unknown images.

1.4 Overview of the thesis structure

The remainder of the thesis is organized as follows. Chapter 2 provides an introduction to image processing and computer vision. It summarizes the main applications of computer vision in different fields and the main characteristics of the most important image processing methods used today. In addition, the chapter is stressing the need to implement robust methods that can be used in different modalities and challenges in image analysis are discussed. A literature review and initial experiments are also presented in this chapter for supporting the selection of the most appropriate methods that can serve the objectives and rational of this work.

Chapter 3 introduces active contour image segmentation models. Emphasis is given on how these models can serve the detection of objects or regions of interest in the presence of intensity heterogeneity and feature complexity.

Chapter 4 presents image classification methods. Support vector machines using texture features for describing the image as a whole or the objects and the different region of interest that may exist in the analysed scenes, are also discussed.

In the remaining chapters, the implemented algorithms evaluated in real world applications are presented. Chapter 5 describes the proposed algorithm for mammogram breast-background segmentation. Part of this work was presented in [7] and [8]. Chapter 6

presents the proposed algorithm for building structure detection using active contours and colour features. Some of this work was presented in [9] and [10]. In chapter 7 satellite image analysis for shadow detection and building height estimation is performed. Some of this work was presented in [11]. An optimized localized active contour segmentation model with application to intensity heterogeneity and complex images is presented in chapter 8. Some of this work was presented in [12]. A classification model in the presence of intensity heterogeneity and feature complexity is presented in chapter 9. The proposed methods for breast density classification and abnormalities detection are discussed in chapter 10. Some of this work was presented in [13], [14] and [15]. A no-reference quality assessment method for satellite images is presented in chapter 11. Prior to the presentation of each proposed algorithm, there is an overview of the current literature on the subject. Experimental results are also presented for evaluating the performance of the proposed methods. Chapter 12, which is the closing chapter, summarizes the work presented in this thesis accompanied by discussion and conclusions and suggests how the proposed methods are combined for extracting useful information from the analysed scenes. Future work ideas based on the extension of the presented methods are also underlined.

Chapter 2: Computer Vision and Image Processing Analysis

2.1 Introduction

Computer vision and image processing is involved with the theory, design and implementation of computerized models that can automatically process visual data. This chapter is an introduction to computer vision and associated image processing functions applied in mammographic and satellite imagery. It introduces digital image acquisition and display, hardware and techniques with respect to the quality of the captured image. It discusses and illustrates experiments on image segmentation and classification applications. Finally, based on literature review and the initial experiments that are performed, this chapter concludes with the decisions that are made in selecting the image processing methods for further evaluation and optimization.

2.2 Computer vision applications

The aim of computer vision is to describe the real world as presented in one or more images and to reconstruct its properties, such as shape, illumination and colour distributions. Modelling the visual world in all of its rich complexity is considered difficult because is an inverse problem, in which the objective is to recover some unknowns based on insufficient information. Physics based and probabilistic models should be examined and evaluated thoroughly to disambiguate between the potential solutions [1-2, 16].

2.2.1 Computer aided detection systems

Significant research has been performed for the development of computerized applications that help healthcare professionals use medical images more effectively and efficiently. Medical imaging analysis offers valuable information on patients' health conditions. Various advancements have been made in imaging modalities, such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), diagnostic ultrasound and X-ray technology. Thus, a large number of medical images are produced, which physicians and/or radiologists must interpret. Computer-aided detection (CAD) systems aim to provide significant help to healthcare professionals by reducing the medical image reading or analysed time and improving the accuracy or consistency of their diagnosis. CAD systems are developed to be used by the healthcare professionals as a "second opinion" for their diagnosis for a diversity of health diseases [17-18].

Computer aided mammography refers to the application of Computer Aided Detection and Diagnosis algorithms to assist a radiologist in the diagnosis of breast cancer based on mammograms. The aim of CAD systems in screening mammograms is to increase the accuracy and efficiency in detecting breast abnormalities. Their main use is for the screening programmes, where a large number of mammograms have to be processed while the difficulty of their interpretation requires robust and reliable results. Furthermore, the rapid development of digital mammography increases the need for utilizing CAD systems in everyday image processing like zooming, contrast enhancement, edge detection, image subtraction and fully automated detection methods. Mammographic CAD systems such as ImageChecker of R2 Technologies Inc., CADx SecondLook and Kodak Mammography CAD Engine have been FDA approved and are commercially available. Several studies have been performed for the practical evaluation of CAD mammography systems [19-21]. The results in these studies highlight that the clinically effectiveness of CAD systems is a still an open research field. It has been agreed that CAD systems are useful for detecting cancers that are clinically missed. However, drawbacks still exist since a significant number of false positive detections have been noticed and true abnormalities were very often detected only in one mammographic view, thus there is still need to design more robust methods of computerized mammographic image analysis to be incorporated in CAD systems [22].

2.2.1.1 Mammographic images

X-ray mammography remains the most cost-effective imaging modality for national screening. It is considered as the key screening tool for the early detection of breast abnormalities [22].

An (X-ray) mammogram is a plain-film radiogram of the breast. It is a two-dimensional projection of a three-dimensional structure. A mammogram is a projective 2-D image of the compressed breast, formed by the attenuation characteristics of X-rays through tissue. X-ray has good Signal-to-Noise Ratio (SNR) and high spatial resolution. Thus, it can be used not only for detecting small abnormalities but also for visualising very fine details of abnormalities, such as speculations and areas of local tissue density. Moreover, X-ray mammography is currently the only imaging modality suitable for detecting micro-calcifications whose frequency is one of the earliest indications of breast cancer, due to the high absorption characteristics of calcium. A mammogram results from the attenuation and transmission of a known flux of X-rays through tissue. The different types of tissue inside the breast have different attenuation coefficients, which is why they attenuate and absorb

different proportions of the X-ray beams. These differences are finally translated into different intensities in the exposed film. In film-screen mammography, the single emulsion film is in contact with an intensifying screen that acts as a photon-amplifier.

In screening mammography, ideally two different views Cranio-Caudal (CC) and Medio-Lateral-Oblique (MLO) of a compressed breast are obtained to maximise breast cancer detection likelihood. In the CC view images, the breast is compressed between two plates in the head to toe direction. In the MLO view images, the breast is compressed between two plates placed in the shoulder to the opposite hip direction. The CC view image shows "spread" tissue and maximises the effectiveness of detecting abnormalities by improving contrast. The MLO view image shows the breast including the pectoral muscle and maximises the effectiveness of detecting abnormalities near the axillae.

A new type of mammography equipment called full field digital mammography offers potential advantages over conventional film-based mammography. With digital mammography, X-rays pass through the breast exactly the same way as in conventional mammography, but the exiting X-ray photons are converted directly to electrons and recorded by means of an electronic digital detector instead of the film. The image acquisition and the image visualisation are performed independently so that the optimisation process in between may improve the quality of generated images. Digital mammography achieves higher contrast resolution and better quality images. In addition, using digital mammography reduces the cost for developing X-ray films and provides a basis for reliable and fast retrieval of mammograms from digital archives. Moreover, CAD systems can be directly integrated.

2.2.2 Remote sensing applications

Extracting qualitative and quantitative information from remote sensing data is of great practical interest for various applications. In telecommunication and remote sensing applications such as natural disaster monitoring, urban change detection, urban scene reconstruction, cartography update, urban inventory and wireless planning applications, computer vision and image processing models applied in remote sensing images are utilised for land use analysis to recover the three-dimensional shape or appearance of the existing objects [6, 23-25]. Moreover, various applications have been using high-resolution satellite imagery for the creation and update of maps or geographical information system databases, digital elevation extraction, and most notably, consumer level mapping over large areas on platforms that are used by hundreds of millions of people every day such as Google Earth, Google maps and Bing maps.

2.2.2.1 Remote sensing images

The first sub-meter resolution earth observation satellite system commercially available is IKO-NOS that released in 1999. Since then, many additional high-resolution satellites have been launched and many countries now have their own high-resolution satellites.

Currently the three highest resolution commercial satellites are GeoEye-1, WorldView-1 and WorldView-2 owned by U.S. commercial operator, DigitalGlobe of Longmont Colorado. GeoEye-1 provides 0.41 m panchromatic and 1.65 m multispectral images in 15.2 km swaths. GeoEye-1 can image up to 60 degrees off nadir. WorldView-1 is a high-capacity, panchromatic imaging system featuring 0.46 m resolution imagery. With a nominal swath width of 17.6 km at nadir and revisit time of 1.7 days on average. WorldView-1 is capable of collecting up to 750,000 square kilometres per day of 0.46 m imagery. The most recent satellite is WorldView-2 that launched in 2009 and has a panchromatic band with 0.46 m resolution and eight multispectral bands with 1.84 m resolution. This satellite system introduced new opportunities for high-resolution thematic mapping. However, it should be noticed that all of the imagery provided by the above satellite systems are resampled to 0.5 m panchromatic and 2.0 m multispectral for commercial use because of US government maximum resolution restriction.

2.3 Image quality and intensity heterogeneity

Before the application of image processing functions or the development of computer vision solutions an image must be acquired by an optical device and converted into an entity that can be processed [26]. The objects in natural scenes have the ability to reflect or absorb energy and this attribute is utilized in the image acquisition process. The energy of interest is light or more generally electromagnetic waves also known as photons. The energy is focused by an optical device and subsequently a sensor is used to measure and record the amount of the reflected energy. The photon particle can be described, by its energy. To capture an image, an energy source to illuminate the scene is needed. The sun and visual light most often are acting as the energy source however, other frequencies such as radar, infrared, X-ray and ultrasound are also used. Illumination is an important aspect of the quality of the captured image. Depending on the nature of the source, illumination energy is reflected from objects or transmitted through objects. The lit of the scene must be carefully considered in the image acquisition process for avoiding too dark or too bright regions in the acquired image. If for example the illumination is pointing directly to the image acquisition system, then the captured image will be very bright ending up losing

important information. Even when the illumination is not directed toward the camera, bright regions in the image might still occur. These highlights are often a result of a shiny region surface, which reflects most of the illumination. Furthermore, the light from different points or from the surroundings of the region of interest might create a useless image. Lens and sensors are the basic components of the image acquisition system. The lenses are made with glass and used to focus the incoming light to the sensor. A high number of light rays with slightly different incident angles collide with each point on the objects surface and some of these are reflected toward the optics. The lenses are responsible in focusing the parallel-reflected rays in order to form the image of the object that will be captured. Sensors are the components responsible to transform illumination energy into digital images. A sensor is placed exactly at the position of the object that the image will be acquired. The parallel rays intersect at a point denoted as the focal point. The distance from the centre of the lens which is denoted as the optical centre to the plane where all parallel rays intersect, is denoted as the focal length. The line on which optical centre and focal point lie is the optical axis. The distance from the object to the lens and the actual size of the object needs to be considered during image acquisition process to avoid unfocused images. The sensor consists of pixels, and each pixel has a certain size. As long as the rays from one point stay inside one particular pixel, this pixel is focused. If rays from other points also intersect the pixel in question, then the pixel will receive light from more points, the resulting pixel value will be a mixture of light from different points, and the acquired image will be unfocused. An object can be moved a distance further or closer to the lens and remain in focus. This range is denoted as the depth-of-field. A smaller depth-of-field can be achieved by increasing the focal length. However, this has the consequence that the captured area will be reduced. The observable area is denoted as the field-of-view of the image acquisition system. The field-of-view depends, besides the focal length, also on the physical size of the image sensor. A final parameter influencing the depth-of-field is the aperture. The aperture is a flat circular object with a hole in the centre with adjustable radius. The aperture is located in front of the lens and used to control the amount of incoming light. In the worst case where for example the aperture only allows rays through the optical centre then an infinite depth-of-field will be created. This can be avoided with the lower shutter speed to ensure enough light is passed through the system for creating the image, otherwise the acquired image will be blurry.

The last process in an image acquisition system is the use of a sensor to record the image. An image sensor consists of a two-dimension array of cells. A pixel is assigned for each cell for measuring the amount of incident light, transformed it into a voltage and

subsequently converts it into a digital number using an analogue to digital converter. A device known as shutter controls the exposure time to the incident light. An underexposed or overexposed image will be captured if the exposure time is too low or too high respectively. A cell is sensitive to incident light hitting the cell, however it is not sensitive to where exactly the light hits the cell. In order to acquire images where the structures of the shapes are preserved, the size of the cells should be infinitely small for avoiding the blocky shape effect. The number of pixels used to represent an image is called the spatial resolution of the image. A high resolution means that a large number of pixels are used, resulting in fine details in the image. A high-resolution image acquired system needs a large amount of memory capacity for storage. The resolution of an image can be changed by a process called image resampling. However, by applying such process to the acquired image, loss of information may occur. Finally, the numbers of photons hitting a cell are represented using digital numbers. Since the human eye is able to distinguish the exact number of photons, this number is quantified. The most common quantization is the representation of using one byte or 8 bits where a charge of 0 volt will be quantized to 0 and the highest charge will be quantized to 255. Other grey-level quantization is also used. The effect of changing the grey-level quantization also called the grey-level resolution. Down to 16 grey levels the image will frequently still look realistic, but with a clearly visible quantization effect. The grey-level resolution is usually specified in number of bits. Typical grey-level resolutions are 8, 10, and 12 bit leading to 256, 1024, and 4096 grey levels. To transform the information from the sensor into an image, cell content is now converted into a pixel value. Such a value as was mentioned above is interpreted as the amount of light hitting a cell during the exposure time and is denoted as the intensity of a pixel. It is visualized as a shade of grey between 0 and 255 values where 0 represents black and 255 represents white. Thus, a grey-scale image is 2D array of pixels and can be represented as $f(x, y)$, where x is the horizontal position of the pixel and y the vertical position. In the case of an overexposed image, a number of cells might have charges above the maximum measurable charge. These cells are all quantized to 255 leading to the effect that is known as the saturated effect. When a cell is saturated, it can affect also the neighbour pixels by increasing their charges. This is known as blooming and therefore must be avoided.

A colour image is represented using a mixture of the primary colours red, green, and blue. In order to produce a colour image in this domain every pixel is associated with three values for the three primary colour components. This colour model is called the RGB model. Besides colour models whose components describe the value of a specific primary

colour, there are models that have three entirely different variables: the most common amongst these is the HSV, or hue, saturation, value model. The hue identifies the colour and an angle in the so-called colour circle represents it. In the colour circle, 0^o corresponds to red, 120^o to green, 240^o to blue, and all other colours are embedded additively between them. The saturation identifies how pure the colour component is and with the maximum number 1 a pure colour is represented, while with the lowest number 0 the absent of a colour is represented. Shades of colours are within the range 0 to 1. Finally, the value component refers to the luminance value of the colour.

Despite all of the advances in image acquisition systems, all the above issues distance to object, motion of object, zoom, focus, depth-of-field, focal length, shutter, aperture, sensor and digital conversions are still important attributes that may affect the quality of the acquired image. These attributes may produce intensity or spectral heterogeneities and may cause the loss of valuable information [1-2, 16].

2.3.1 Imaging quality factors

Effective image processing and analysis algorithms require images of acceptable quality. In those cases where the quality of the image to be analysed is low, then any information about the quality or the factors responsible for reducing the quality of the image may help the optimization of image processing algorithms to overcome any introduced limitations. The quality of the provided images may be reduced during the preliminary acquisition process as previously discussed but also during compression, restoration, communication or final display. The most important quality factors that should be considered are the contrast, noise, spatial resolution, sharpness, vigneting, distortion and chromatic aberration.

Image contrast refers to the fractional difference in optical density of brightness between two regions of an image. The combination of the capture object along with the lighting condition is responsible for the image contrast and it has a significant impact on the texture of the image.

Image noise is defined as an uncertainty or imprecision of the recording of the image signal. It is denoted as a random variation of image density and in digital images is visible as pixel level variations. The noise can also be appear randomly on smooth surfaces and may significantly degrade image quality. The sensitivity of the image acquisition system, the time length of the exposure and the temperature are factors that may affect the amount of noise in the image. It is expected that some degree of noise is always present in any natural image.

Sharpness determines the amount of detail an imaging system can reproduce and it is responsible for defining the boundaries between zones of different tones or colours. Resolution and acutance are the most important factors responsible to perceive the sharpness of an image. Acutance describes how quickly image information transitions at an edge. High acutance results in sharp transitions and clearly defined borders. Resolution describes the camera's ability to distinguish between closely spaced elements of detail.

The sensitivity and structure of the lens in the image acquisition system are responsible for vignetting, distortion and chromatic aberration. Vigneting appears as a progressive darkening toward the edges of the image. Distortion appears as otherwise straight lines bending inwards or outwards. Chromatic aberration appears as colour deterioration along high contrast edges.

In computer vision or image processing applications, a challenging task is the development of methods to assess and enhance the image quality. It is vital that the designed information acquisition algorithms and techniques be able to overcome limitations related to image acquisition process and low quality image factors.

2.4 Image processing

The aim of low-level image processing operations is to enhance the quality of the image, remove any misleading information that exists in the analysed scene and highlight the region or object of interest. In middle-level operations, the aim is to extract useful information about the analysed scene. This information may include various attributes or features about the objects or regions of interest. Finally, high-level operations aim to understand the acquired information as close as humans do, or in some cases even better, using artificial intelligence or machine learning methods. The following paragraphs present important developments in these three imaging or computer vision processing categories, with respect to the objectives and specific findings of this work.

2.4.1 Image quality enhancement

In low-level operations, a considerable amount of work has been devoted in enhancing the quality of the image. As was previously discussed the brightness and contrast of the image are important quality factors. Using the histogram to visualize the set of lightness values in an image, statistics such as the minimum, maximum and average intensity values may help to correct the brightness and contrast of a given image. Histogram equalization or variations of this like locally adaptive histogram equalization (histogram equalization

applied in small moving windows of the image) are very popular methods and used extensively to correct these issues due to their simplicity and effectiveness [27]. An intensity mapping function able to create a flat histogram and perceive or enhance the contrast should be defined. The histogram of grey levels in a window around each pixel is generated first. The cumulative distribution of the grey-levels values, that is the cumulative sum over the histogram, is then used to map the input pixel grey-level value to an output pixel grey-level value. If a pixel has a value lower than all other pixels in the surrounding window, the output is maximally black and if it has median value in its window the output is 50% grey. Methods based on histogram equalization have been widely used for contrast and brightness enhancement in a variety of applications, such as medical image processing, radar signal processing and satellite image processing [27-30].

To enhance further the quality of the image several methods have been designed and implemented for sharpening details, accentuating edges, or removing noise. Various disciplines such as probability theory, statistics, partial differential equations, linear and nonlinear filtering, spectral and multi-resolution analysis have been extensively evaluated towards this direction [31-33].

An image can be filtered either in the frequency or in the spatial domain. In the spatial domain, the corresponding process can be described with the following equation:

$$g(i,j) = h(i,j) \odot f(i,j) \tag{2.1}$$

where $h(i,j)$ is the filter function and is approximated using a kernel or mask, i.e. a limited size matrix of numbers, which consist the filter coefficients. The input image $f(i,j)$ is convolved (multiplying together two arrays of numbers, generally of different sizes, but of the same dimensionality) with the kernel.

Filters such as mean, median, Gaussian and conservative are considered as smoothing filters and used for noise reduction. Laplacian or Laplacian of Gaussian filter is used for edge detection and un-sharp filters can be used for edge enhancement.

In the frequency domain, the image is first transformed using the Fourier function, and then is multiplied with the filter function as can be seen in the following equation:

$$G(i,j) = H(i,j)F(i,j) \tag{2.2}$$

where $F(i,j)$ is the image in the Fourier domain and $H(i,j)$ the filter function. Finally, the result is transformed back to the spatial domain using the inverse Fourier function.

The form of the filter function determines the effects of the operator. There are three different kinds of filters in the frequency domain and these are the low-pass, high-pass and band-pass filters. A low-pass filter attenuates high frequencies and retains low frequencies unchanged. This is equivalent to a smoothing filter in the spatial domain. A high-pass filter is equivalent to an edge enhancement or edge detection filter in the spatial domain, because edges contain many high frequencies. Areas of rather constant grey-levels are consisted of mainly low frequencies and are therefore suppressed. A band-pass filter attenuates very low and very high frequencies, but retains a middle range band of frequencies. Band-pass filtering can be used to enhance edges (suppressing low frequencies) and in parallel reducing the noise (attenuating high frequencies).

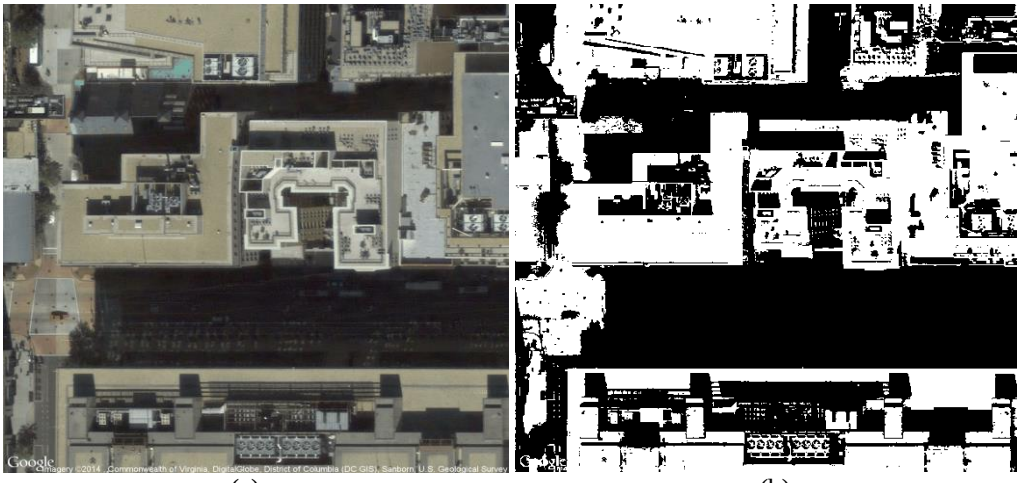
Edges and singularities are very important information for image analysis and understanding. The multi-resolution transforms, where the image is represented on different resolution and thus different quality levels, achieve good scattering for edges and singularities. In natural images, a significant portion of the information is embedded in edges and singularities, thus multi resolution transforms have been extensively used for image de-noising. Examples of such image de-noising methods can be found in [34-38].

2.4.2 Image segmentation

After the image quality enhancement, the groups of pixels in the digital image presenting the objects to be analysed must be defined. This process is known as image segmentation and is considered as a middle-level image processing operation since is responsible for detecting and extracting the regions of interest in the analysed scene. Many definitions have been published in the literature for image segmentation. H. Barrow and J. Tennenbaum, in 1978 defined the segmentation as "the process of partitioning an image into semantically interpretable regions." R. Haralick and L. Shapiro in 1992 described the segmentation as "the partitioning of an image into a set of non-overlapping regions whose union is the entire image. The purpose of segmentation is to decompose the image into parts that are meaningful with respect to a particular application." Usually image segmentation is an initial and vital process to the overall image understanding process. Following the above definition, it is expected that the selection of the most appropriate segmentation method be closely related with the application to be used. A thorough investigation must be performed in order to find which method or combination of methods from the three different segmentation categories edge-based, region-based and data

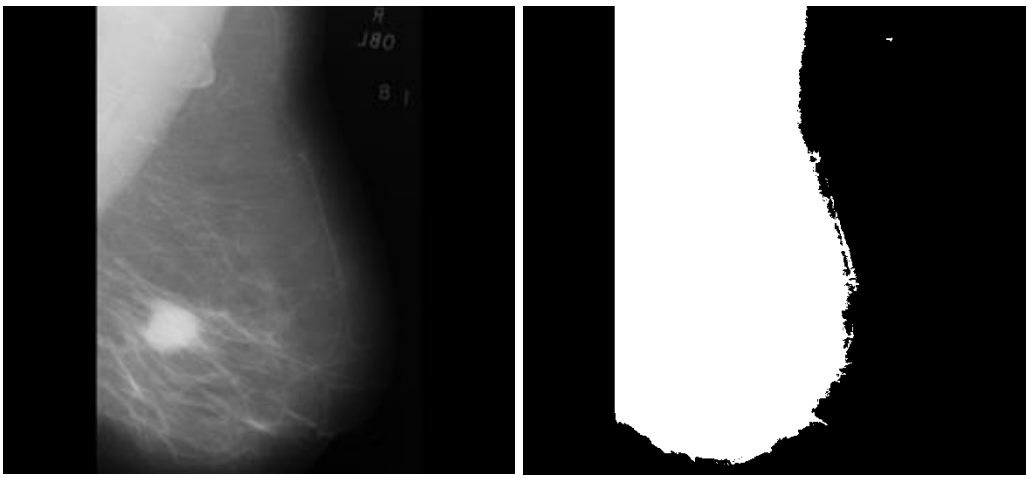
clustering, as was previously discussed, is suitable for a specific application. After a literature review for identifying the most recently advances in this field, a number of algorithms have been implemented and evaluated towards the above direction.

Region algorithms based on threshold are very popular because of their simplicity. These algorithms are using the principal that the regions or objects of interest have distinctive feature values within a specific range. The image intensity or the gradient magnitude is often used and the range may be selected manually or automatically. Otsu's threshold method [39], along with several variations for achieving better results has been extensively used in multidiscipline applications [40-41]. In the following figure segmentation results of the Otsu's threshold method applied in remote sensing and medical images are presented.



(a) (b)

(c) (d)



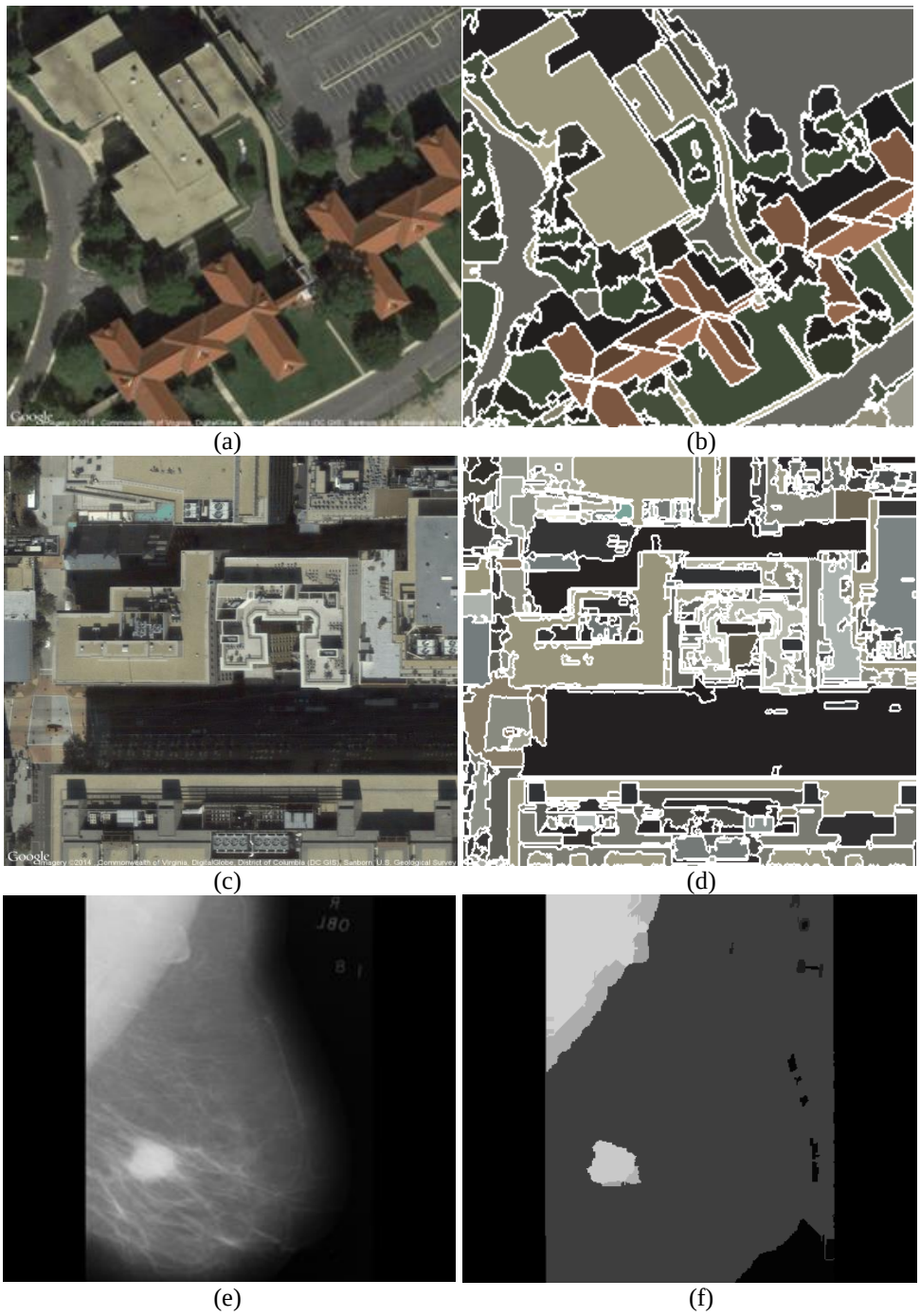
(e) (f)

Figure 3: Otsu's threshold segmentation method

(a), (c), (e) Original images, (b), (d), (f) Segmentation results

An edge can be defined as a local discontinuity in the pixel values and can be thought as a boundary between regions. An edge segmentation method can be designed by finding the magnitude of the gradient and then apply a threshold. Edge based methods such as Canny, Sobel, Prewitt, Robert and Laplacian of Gaussian were evaluated thoroughly in a variety of applications. Algorithms that are using threshold values are strongly depended on the correct calculation of the threshold and have difficulties to create smooth boundaries of the region of interest and control the leakage or over-segmentation of the image. Edge segmentation operations are based on discrete pixel features, thus the detected boundaries of the regions of interest are often incomplete or non-continues. Furthermore, edge based and region based threshold methods also suffer if the regions of interest in the analysed image are not clearly defined with strong boundaries. Therefore, algorithms based on threshold or edge detection are not usually applied individually but instead are used as a pre-processing step to help a more sufficient segmentation algorithm [42].

Approaches based on clustering techniques using features such as intensity, colour or texture are widely used for image segmentation. The analysed image is divided into small regions or patches and the features are integrated based on fitting mixture models, mode finding, graph cuts, variational formulations and level sets [43-45]. In this category Felzenszwalb and Huttenlocher's graph based region merging [43], Shi and Malik's Normalized Cuts [44] and Comaniciu and Meer's Mean Shift [45] seems to be very popular and have been used in several different applications in the recent years. In the following figure, segmentation results using a mean shift clustering method are presented.



(a) (b)

(c) (d)

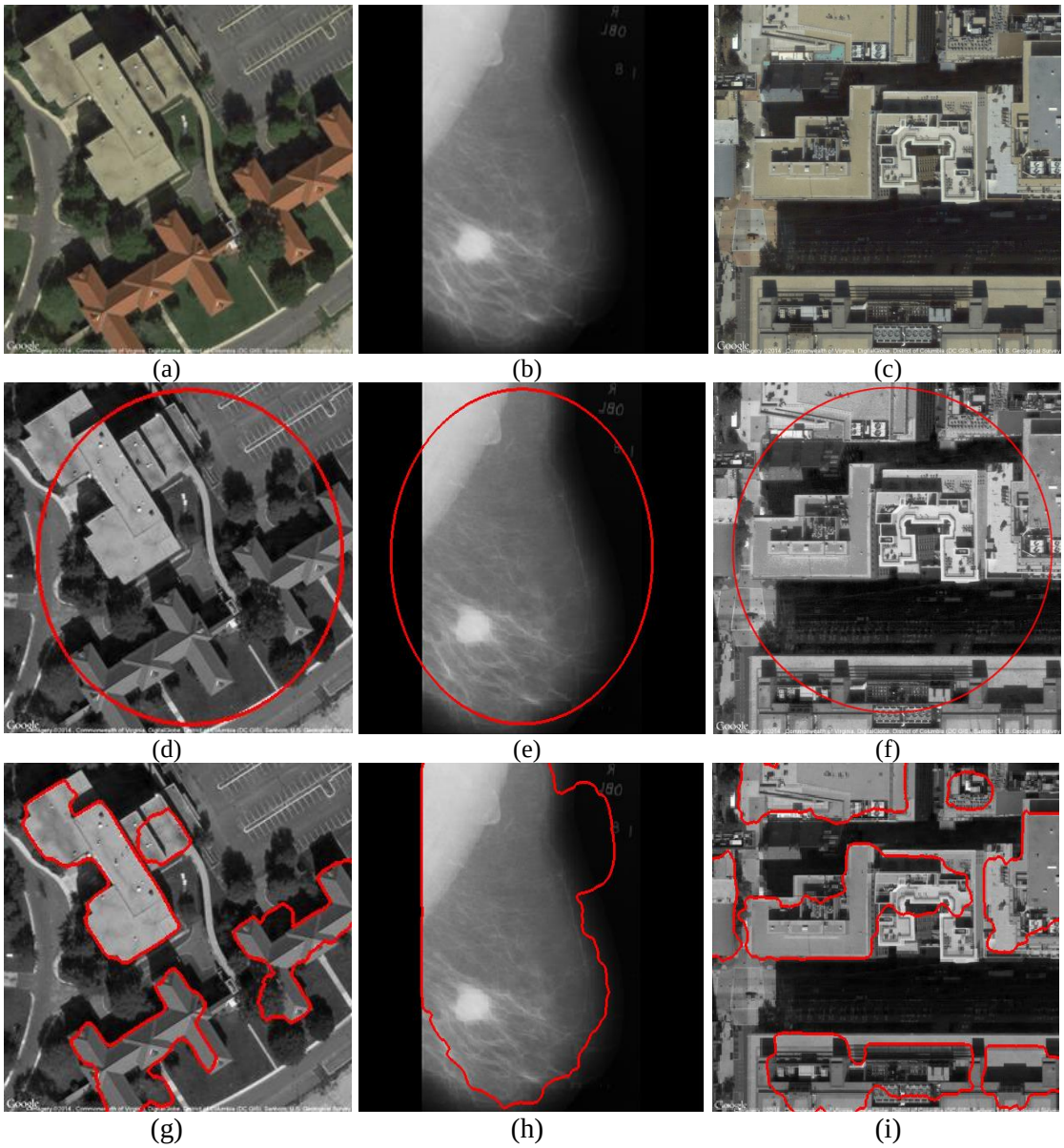
(e) (f)

Figure 4: Mean shift segmentation method

(a), (c), (e) Original image, (b), (d), (f) Mean shift segmentation

Over-segmentation is a common issue across feature clustering image segmentation approaches. Smooth changes in the selected features to be clustered such as texture or intensity due to shading can cause dissimilar patches to appear, despite the fact that they may belong to the same image region. This problem is even more difficult to be resolved in noisy images at the presence of intensity or spectral heterogeneity.

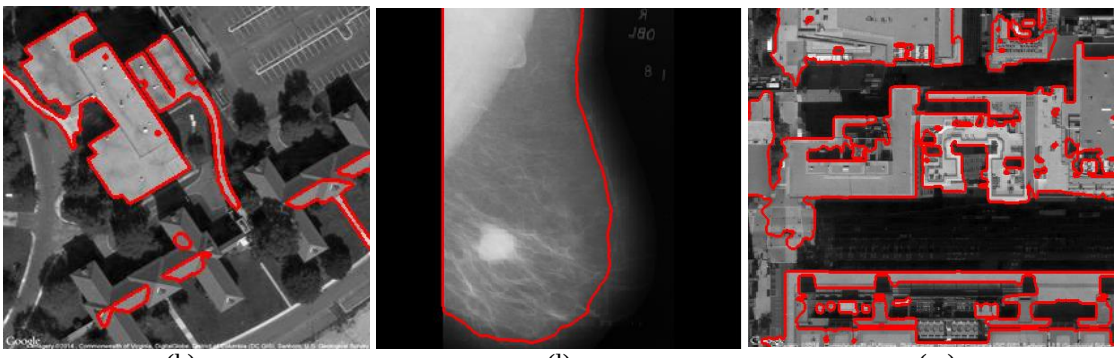
Active contours are extensively used in the image segmentation field to delineate regions of interest that may exist in the analysed image [1-2, 46-54]. Active contour models were firstly introduced by Kass et al. [47] and can be defined as piecewise polynomial curves that match a deformable model to an image by means of energy minimization. Due to several limitations like dealing with topological changes and progress into boundary concavities, Osher and Sethian [48] introduced the level set approach as a method for computing and analysing the evolution of a contour using partial differential equations. A variety of features such as edges, intensity values and texture are utilized to compute the force function responsible to evolve the curve and segment the image in active contour models. The Chan and Vese (C-V) active contours without edges algorithm [49] and the Caselles geodesic active contours [50] are two popular algorithms for image segmentation based on active contours and level sets. In the following figure, segmentation results derived by applying the above active contour models are presented.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)



(k) (l) (m)

Figure 5: Active contour level set segmentation methods

(a), (b), (c) Original images, (d), (e), (f) Contour initialization, (g), (h), (i) Caselles algorithm segmentation results, (k), (l), (m) Chan and Vese algorithm segmentation results

Caselles algorithm is an edge-based method where the gradient of the image is used for computing the force function and thus the curve is driven to regions with high gradient. C-V algorithm is a region-based method and it tends to separate the image into two homogeneous regions according to the mean intensity values. The method does not use an edge function based on gradient information to drive the curve and delineate objects, thus is able to detect objects with or without weak boundaries. The performance of traditional active contour models is highly dependent on initial curve placement. This is a difficult task for images with intensity or spectral heterogeneity of complicated scenes and can often result in misleading segmentation. However, the initial experiments that were performed in this work revealed that active contour models implemented using level set method seems to have better potential than other methods with images with intensity or spectral heterogeneity. Thus, active contour and level set models were selected for further investigation in this thesis. These include algorithms such as Bernard [51], Lankton [52], Li [53] and Shi [54]. These algorithms are variations or extensions of the Caselles [50] and Chan-Vese [49] models that are considered as “traditional” implementations for addressing intensity heterogeneity drawback. These algorithms will be further discussed in chapters 3 and 8.

2.4.3 Image classification

Image classification is the process of discriminating and characterizing a pixel, a group of homogeneous pixels or the analysed image as a whole into different classes. The aim of classification is to predict class labels for the analysed samples based on features that describe the samples. Classification processes can be divided initially into supervised and unsupervised methods based on the information of the different sample classes and their characteristics that may exist and used as a priori [55].

In a supervised classification method, a prior knowledge of the imagery homogeneous representative samples of the different information classes of interest is used. These samples are referred to as training data. Spectral features able to discriminate the classes are extracted for the pixels comprising these data and used to train a model to recognize similar data for each class. A numerical signature is determined for each training class and each pixel in the image is compared to these signatures and labelled to the most closely class. Thus, in a supervised classification, the information classes are first identified and then spectral classes that represent them are determined. In an unsupervised classification method, spectral classes are grouped first, based solely on features able to discriminate the different classes that may exist. Unsupervised classification relies on clustering algorithms to segment the training data into prototype classes automatically. Statistical groupings or structures in the data are used to determine the existing classes. Subsequently, the spectral classes are matched to information classes. Information on the number of groups or classes to be defined in the analysed data along with the specification of parameters related to the separation distance among the classes and the variation within each class, may help the performance of the implemented method.

After the initial categorization to supervised and unsupervised classifiers, based on whether parameters such as mean vector and covariance matrix are used or not, the classifiers can be further discriminated into parametric and non-parametric. In parametric classifiers, the mean vector and covariance matrix are generated from the training samples while in non-parametric classifiers statistical parameters are not required for separating the classes. Classifiers are also categorised into different classes such as per pixel, sub pixel, object oriented and per field, based on the kind of pixel information used. Per pixel classifiers, combine the spectra of all training set pixels, while in sub pixel classifiers the spectral value of each pixel is used as a linear or nonlinear combination of defined classes. Object oriented classifiers perform the classification after the pixels are grouped together to form objects. Per field, classifiers are integrating raster and vector data in the classification process. Moreover, if the output of the classifier is a definitive and strong decision about each pixel in the analysed image for the class that belongs to, the classifier is defined as a hard classifier. If for each pixel a measure of the similarity for every class is calculated instead, this classifier is defined as soft classifier. Finally, if spectral, spatial or spectral and spatial information is used, the classifiers are characterised as spectral, contextual or spectral-contextual respectively [56].

Image classification has received significant interest from the computer vision learning communities. Many advanced classification approaches, such as artificial neural networks,

fuzzy-sets, expert systems, maximum likelihood, minimum distance, K-mean, nearest neighbourhood, decision tree, linear discriminating analysis and support vector machines have been widely applied for image classification [57-66].

There is no strong evidence that a classification method is inherently superior to any other. Instead, it is based on the problem set, the study area, the data sources, and the intended use of the results to determine the most appropriate, effective and efficient approach [56, 63]. Features or image descriptors such as colour, intensity, shape, texture and the selection of the most appropriate for describing the classes of the analysed images is an important factor of the designed method. Based on the selected image descriptors, the most appropriate classifier should be used. With respect to the objectives of this work support vector machines (SVM) classifiers were selected for further investigation. They are chosen because of their ability to generalize well in high dimensional spaces, a common case for the image domains discussed in this thesis [64-65]. SVMs are based on statistical theory and aim to determine the location decision boundaries that result in the optimal separation of classes. SVMs have the ability to classify non-linearly separable data in a high dimensional feature space using a kernel function. Kernel functions are applied to describe a similarity relationship between the sets to be classified. SVMs determine a decision boundary in the feature space to distinguish the classes by creating the optimal separating hyper-plane. SVM selects the linear decision boundary that minimizes the generalization error. The method finds the hyper-plane that leaves the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from it. This optimum hyper-plane is produced by maximizing the minimum margin (i.e. the sum of the distances to the hyper-plane from the closest points of the classes between the sets). Thus, the resulting hyper-plane is depended on border training patterns only, called support vectors. SVMs have been shown to generalize well on difficult image classification problems where the only available features are defined in a high-dimensional space [64]. A k-means unsupervised algorithm [66] that aims to segment *n* observations into *k* partitions, in which each observation belongs to the partition with the nearest mean, was also investigated. K-means algorithm is a simple and computational fast method that may serve the objectives of this work for analysing large data sets effectively. Towards this direction, initial experiments have been performed for identifying buildings in satellite images and abnormalities in mammographic images using SVM and K-means classification methods. Mean intensity and standard deviation have been utilized as feature descriptors. The results are presented in the following figure.

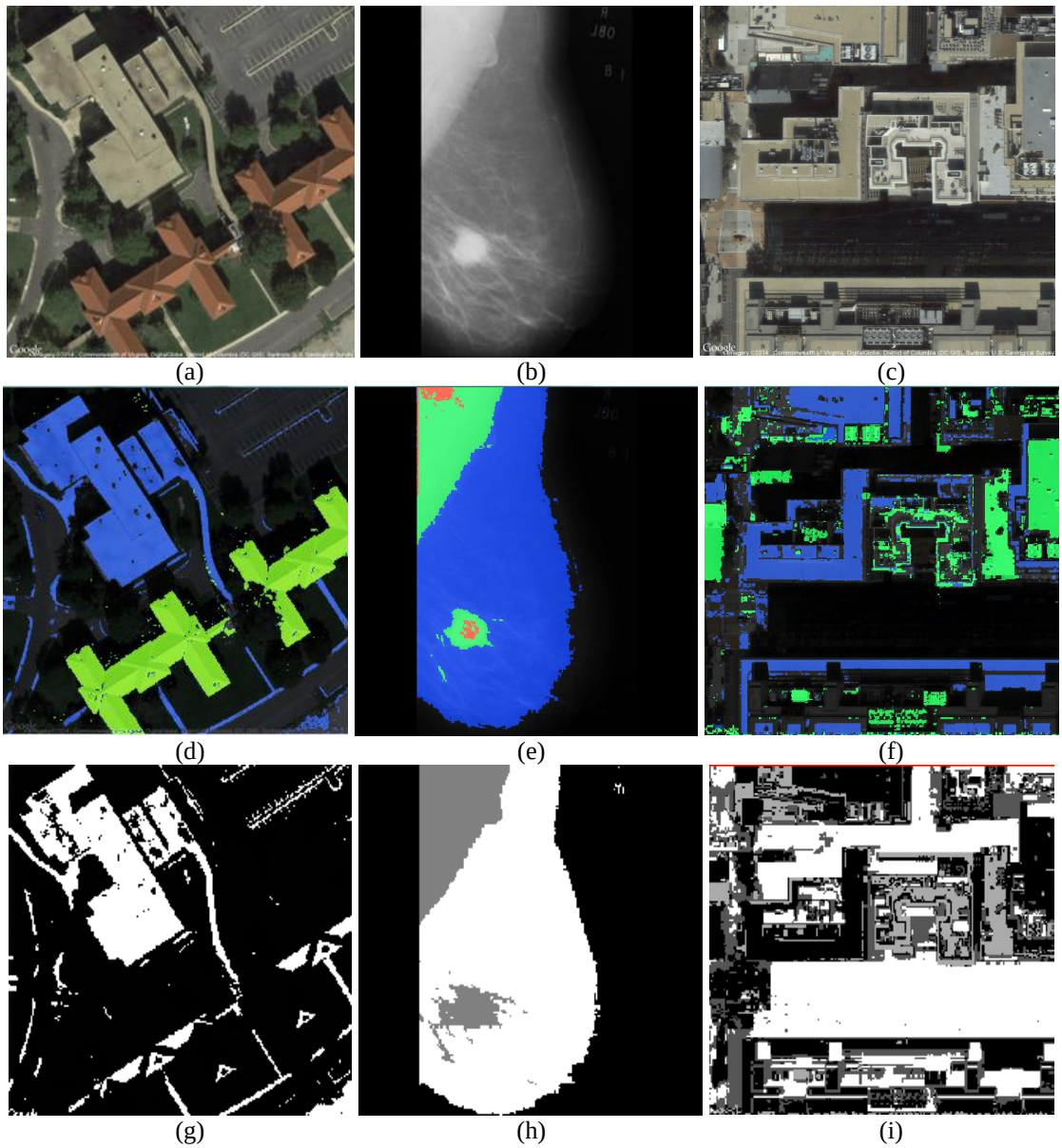


Figure 6: Image classification

(a), (b), (c) Original images, (d), (e), (f) SVM classification, (g), (h), (i) K-mean classification

The experimental results revealed that if a large number of variables are used, K-Mean's is computationally fast and produce tight clusters especially if these clusters are globular. However, the accuracy of the method is not sufficient for the evaluated domains since regions consisting from buildings were not labelled as building structures and regions presenting other structures, like roads in several cases, were classified as buildings. Similarly, in mammographic images, the area presenting the pectoral muscle and the area presenting an abnormality has been in many cases misinterpreted. Better results have been obtained using the supervised SVM classification method. This approach appears to have the potential to work well in the selected domains.

2.5 Discussion

The initial experiments performed in this work revealed that image segmentation and classification in the presence of intensity or spectral heterogeneity and feature complexity that often occurs in natural or real world images can be challenging processes. The most appropriate segmentation methods should be first selected and then combined together to achieve good results. The above initial observations lead to the conclusion that methods, using active contours and level sets combined with the appropriate imaging filters or classification tasks, are promising for selectively detecting a variety of objects. Thus, in the following chapters such methods are investigated further. Moreover, for classification processes SVMs seems to work well in the selected domains and subsequently are further evaluated in this thesis for the development of a multi-purpose classifier.

Chapter 3: Active Contour and Level Set Image Segmentation Models

3.1 Introduction

Active contours are piecewise polynomial curves that match a deformable model to an image by means of energy minimization (Kass et al. 1987). In active contour models, a parametric or non-parametric representation of an initial curve is evolved using explicit or implicit forces. Curve evolution is subject to internal and external constrains. Internal constrains are the curve constrains and external constrains are those derived from the given image. By energy minimization, active contour models are able to detect objects or obtain image segmentations.

Level set approach made active contour methods popular in image segmentation field because topological changes are treated smoothly and concavities are covered effectively [47-54, 67-70]. Level set is a numerical technique for tracking interfaces and shapes. Level set use the calculus of variations and partial differential equations (PDE) for solving the image segmentation problem. Existing active contour methods formulated on level set can be categorized into edge-based and region-based methods.

Edge-based models use edge information for guiding the motion of the active contour to delineate the regions or objects of interest [50]. These methods suffer from the general limitations of the edge-based segmentation methods. If the boundaries of the regions or objects of interest are weak, leakage problems arise very often [49].

Region-based level set models aim to partition the analysed image using region descriptors for guiding the active contour to the boundaries of the regions or objects of interest [49, 51-54, 71-72]. Region-based approaches are considered as more robust to noise and initial contour placement than the edge-based approaches [71-72]. These methods can be further discriminated into piecewise constant [73] and piecewise smooth [71-72]. Piecewise constant (PC) models are based on the assumption that there is intensity homogeneity within and across the regions of interest of the analysed scene. Piecewise smooth (PS) models are designed to delineate regions or objects of interest that their intensities vary smoothly.

3.2 Active contour models

The basic idea in active contour models is to evolve a curve, subject to constraints given for delineating the existing objects in an image [47]. Thus, given an image I and a variable closed curve C the parametric representation of the position of the snake can be defined as:

$$C(s) = \big(x(s), y(s)\big), s \in [0,1] \tag{3.1}$$

The energy functional is a combination of internal and external forces.

$$E_{snake} = a\underbrace{\int_0^1 |C'(s)|^2 ds + b\int_0^1 |C''(s)|^2 ds}_{\text{Internal Energy}} \underbrace{- \lambda \int_0^1 |\nabla I(C(s))|^2 ds}_{\text{External Energy}} \tag{3.2}$$

The weighting parameters α and b control the snake's tension and rigidity respectively. C' is the first derivative of $C(s)$ with respect to s, C'' is the second derivative of $C(s)$ with respect to s and λ is a positive fixed parameter.

The internal energy of the active contour is due to bending and it is used for regularization. It imposes piecewise smoothing constraints. The external energy is based on image forces that first place the initial contour near the desired local minimum, thus near the object to be identified and then lead the active contour towards salient image features like lines and edges. By minimizing the energy, the model locates the curve at the points where the gradient of the image is the maximum and is acting as an edge detector.

The above equation can be solved using Euler-Langrage formulation for function minimization. The classic active contour model has several limitations. It can provide an accurate location of the edges only if the initial contour is specified sufficiently near the edges, since they rely on the local information along with the contour. Estimating a proper position of the initial contour without prior knowledge is a very difficult task. Traditional active contours cannot detect more than one boundary simultaneously since the contours maintain the same topology during the evolution stage, which implies that the active contour cannot split to multiple boundaries or merge from multiple initial contours. Finally, active contours have difficulties progressing into boundary concavities.

3.3 Level set segmentation models

Level set segmentation models presented by Osher and Sethian for computing and analysing the evolution of a contour using partial differential equations [48]. The principle of an original curve built into a surface is used. The contour C is implicitly represented, via a two dimensional continuous function $\phi(x, y, t)$: $\Omega \rightarrow R$, defined in the image plane. The function $\phi(x, y, t)$, where (x, y) are the coordinates in the image plane and t is the time, is called the level set function. At any given time, the level set function defines an edge contour and a segmentation of the image. The edge contour is taken to be the zero level set $\phi(x, y, t) = 0$, and the segmentation is given by the two regions $\{\phi > 0\}$ and $\{\phi < 0\}$.

A typical example of level set function for expressing geometrical attributes and motion is given by the signed distance function to the curve. The Heaviside function $H(z)$, equal with 1 when $z > 0$ and 0 when $z < 0$ can be used to approximate the length of C and the area ω that includes.

$$|C| = \int_{\Omega} |\nabla H(\Phi)|, |\omega| = \int_{\Omega} H(\Phi)\, dx\, dy \tag{3.3}$$

Any C and regularization H_{ε} of the Heaviside function as $\varepsilon \rightarrow 0$ can approximate the length L and area function A:

$$L_{\varepsilon}(\Phi) = \int_{\Omega} \delta_{\varepsilon}(\Phi) |\nabla \Phi|\, dx\, dy\ ,\ A_{\varepsilon}(\Phi) = \int_{\Omega} H_{\varepsilon}(\Phi) |\nabla \Phi|\, dx\, dy \tag{3.4}$$

where $\delta_{\varepsilon} = H'_{\varepsilon}$ approximates the one-dimensional Dirac delta function δ_0.

The above functions can be minimized using the Euler-Lagrange equations with respect to Φ and the parameterization of the descent directions in time *t*.

$$\frac{\partial \Phi}{\partial t} = |\nabla \Phi| div\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right), \frac{\partial \Phi}{\partial t} = \delta_{\varepsilon}(\Phi) \tag{3.5}$$

In the context of level set theory, the above equations represent the motion by mean curvature and the motion with constant speed for minimizing the length and the area respectively.

3.4 Active contours without edges

Active contours without edges or the Chan-Vese (CV) model, is a region-based model formulated on level set and designed for segmenting images containing objects with or without weak boundaries [49]. This model uses a geometric representation of the contour. The optimal segmentation is found by minimizing the Mumford-Shah functional [74].

To solve the segmentation problem in computer vision, Mumford and Shah proposed the following minimization problem. In a given image I the decomposition Ω_i of Ω and the optimal piecewise smooth approximation u of I, such that u varies smoothly within each Ω_i and rapidly or discontinuously across the boundaries of Ω_i, can be found using the following equation:

$$inf_{u,c}\{E^{MS}(u,I,C) = \lambda \int_{\Omega} (I-u)^2\, dx\, dy + \nu \int_{\Omega \setminus C} |\nabla u|^2\, dx\, dy + \mu length|C|\} \tag{3.6}$$

where $\mu, \nu, \lambda > 0$ are fixed parameters responsible to weight the different terms in the energy.

A minimizer of the above energy is a piecewise smooth approximation of the analysed image where C is a smooth closed curve able to approximate edges.

A piecewise constant model can be obtained from the above equation under the assumption that u is constant inside each connected component Ω_i leading to the following transformed function:

$$E^{MS}(I,C) = \sum_i \int_{\Omega} (I-c_i)^2\, dx\, dy + \mu|C| \tag{3.7}$$

In the piecewise constant model the energy is minimized in the variables c_i by setting $ci = mean(u)$ in Ω_i for a fixed c.

Based on the above piecewise constant model Chan and Vese defined the following energy functional for partitioning a given image:

$$E(c_1,c_2,C) = \lambda_1 \int_{inside\_c} |I(x,y)-c_1|^2\, dx\, dy + \lambda_2 \int_{outside\_c} |I(x,y)-c_2|^2\, dx\, dy \tag{3.8}$$

where c_1 and c_2 are the intensity mean values of I inside and outside C respectively, and the coefficients λ_1, λ_2 are fixed parameters. In CV model, two regularization terms were incorporated for controlling the smoothness of the boundary. Therefore, the energy in the CV model is given by the following equation:

$$E_{CV}(c_1, c_2, C) = \lambda_1 \int_{\text{inside\_}c} |I(x,y) - c_1|^2 \, dx \, dy + \lambda_2 \int_{\text{outside\_}c} |I(x,y) - c_2|^2 \, dx \, dy + \mu \, Length|C| + \nu \, Area|C| \tag{3.9}$$

Considering the level set formulation with $C = \{(x,y)|\Phi(x,y=0\}$ the associate energy to be minimized is defined as follows:

$$E_{CV}(c_1, c_2, \Phi) = \lambda_1 \int_{\Omega} (I(x,y) - c_1)^2 H(\Phi) \, dx \, dy + \lambda_2 \int_{\Omega} (I(x,y) - c_2)^2 (1 - H(\Phi)) \, dx \, dy$$

$$+\mu \int_{\Omega} \delta\big(\Phi(x,y)\big)\big|\nabla\big(\Phi(x,y)\big)\big| \, dx \, dy + \nu \int_{\Omega} H(\Phi(x,y)) \, dx \, dy \tag{3.10}$$

where $H(\Phi)$ and $\delta(\Phi)$ are Heaviside and Dirac function respectively and selected as follows:

$$H_\varepsilon(z) = \frac{1}{2}(1 + \frac{2}{\pi}\arctan\left(\frac{z}{\varepsilon}\right), z \in R$$

$$\delta_\varepsilon(z) = \frac{1}{\pi}\frac{\varepsilon}{\varepsilon^2 + z^2}, z \in R \tag{3.11}$$

Keeping $\Phi(x)$ fixed and minimizing the CV energy with respect to constants c_1, c_2 then:

$$c_1(\Phi) = \frac{\int_\Omega I(x,y) H(\Phi(t,x,y)) \, dx \, dy}{\int_\Omega H(\Phi(t,x,y)) \, dx \, dy}, c_2(\Phi) = \frac{\int_\Omega I(x,y)(1 - H(\Phi(t,x,y))) \, dx \, dy}{\int_\Omega 1 - H(\Phi(t,x,y)) \, dx \, dy} \tag{3.12}$$

Subsequently, keeping c_1, c_2 fixed and minimizing the CV energy with respect to $\Phi(x)$ and

using the Euler-Lagrange equation to represent $\Phi(x)$, the following level set formulation model is derived:

$$\frac{\partial \Phi(x,y,t)}{\partial t} = \delta_\varepsilon(\Phi)[-\lambda_1(I-c_1)^2 + \lambda_2(I-c_2)^2 + \mu div\left(\frac{\nabla\Phi}{|\nabla\Phi|}\right) - \nu] \tag{3.13}$$

The above model is used for two-phase segmentations and $div\left(\frac{\nabla\Phi}{|\nabla\Phi|}\right)$ is the curvature defined using the following equation:

$$div\left(\frac{\nabla\Phi}{|\nabla\Phi|}\right) = \frac{\Phi_{xx}\Phi_y^2 - 2\Phi_{xy}\Phi_x\Phi_y + \Phi_{yy}\Phi_x^2}{(\Phi_x + \Phi_y)^{3/2}} \cdot (\Phi_x^2 + \Phi_y^2)^{1/2} \tag{3.14}$$

where Φ_x , Φ_y and Φ_{xy} are derivatives of Φ. By using the following forward finite differences scheme, the proposed model segments a given image.

$$\Phi_x = \Phi_{x+1,y} - \Phi_{x,y}, \Phi_y = \Phi_{x,1+y} - \Phi_{x,y}, \Phi_{xy} = \Phi_{x+1,y+1} - \Phi_{x,y+1} - (\Phi_{x+1,y} - \Phi_{x,y}) \tag{3.15}$$

In figure 7, the evolution of a curve in order to delineate an object based on CV model is presented.

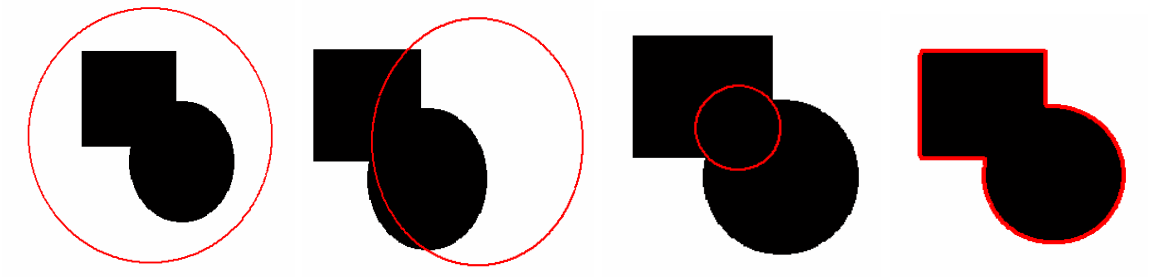


Figure 7: Chan and Vese model

The CV is an energy based minimization segmentation model that separates the image into two regions based on the mean intensity values inside and outside the evolving curve.

A generalization of the active contour model without edges, the multiphase level set framework using the Mumford and Shah model was also proposed by the same authors [72]. The model was proposed for addressing the limitation of partitioning the image into two constant regions.

Using the piecewise constant model for the segmentation of images with more than two segments and junctions, the principal is to use $\log_2 n$ level set functions to represent n

phases or segments. The segments or phases in the domain Ω are define as follows. Two pixels (x_1, y_1) and (x_2, y_2) in Ω will belong to the same phase or class, if and only if $H(\Phi(x_1, y_1)) = H(\Phi(x_2, y_2))$. $\Phi = (\Phi_1, \dots . \Phi_m)$ is called the "vector level set function" and $H(\Phi) = (H(\Phi_1) \dots . H(m))$ is called the "vector Heaviside function" where *m=log n* are the level set functions $\Phi_i : \Omega \rightarrow R$ whose components are only 1 or 0. The classes or phases are given by the level sets of the function $H(\Phi)$ and each class is formed by a constant vector. There are up to $n = 2^m$ possible vector values leading to $n = 2^m$ phases or classes in the domain Ω. The set of curves C is represented by the union of the zero level sets of the function Φ. Based on the above definition each pixel belongs to a distinct class of the decomposition of Ω, therefore vacuum and overlap among phases is avoided. The classes are labelled by $U, 1 \leq U \leq 2^m = n$ and by using a constant vector of averages $c = (c_1 \dots \dots c_n)$, where $c_I = mean(u_0)$ in the class U and the characteristic function X_U for each class U, the simplified piecewise constant Mumford and Shah model for multiphase segmentation can be written as:

$$F_n(c, \Phi) = \sum_{1 \leq U \leq n} \int_{\Omega} (I - c_I)^2 X_U dx\, dy + \nu |C| + \sum_{1 \leq U \leq m} \mu \int_{\Omega} |\nabla H(\Phi_i)| \tag{3.12}$$

The model was also extended to compute piecewise smooth approximations of a given image using the following functions:

$$I(x, y) = I^+(x, y) \quad if\ \Phi(x, y) \geq 0,$$
$$I(x, y) = I^-(x, y) \quad if\ \Phi(x, y) < 0. \tag{3.13}$$

Assuming that I^+ and I^- are C^I functions on $\Phi \geq 0$ and on $\Phi \leq 0$ respectively with continues derivatives up to the boundary $\Phi = 0$ the following minimization function is defined representing the two phase piecewise smooth model.

$$E(I^+, I^-, \Phi) = \int_{\Omega} |I^+ - u|^2 H(\Phi) dx\, dy + \int_{\Omega} |I^- - u|^2 (1 - H(\Phi)) dx\, dy$$
$$+ \mu \int_{\Omega} |\nabla I^+|^2 H(\Phi)\, dx\, dy + \mu \int_{\Omega} |\nabla I^-|^2 (1 - H(\Phi))\, dx\, dy + \nu \int_{\Omega} |\nabla H(\Phi)| \tag{3.14}$$

Multiphase segmentation piecewise smooth models can be designed by adding more than one level set functions as was described in the previous paragraphs. A customized multiphase model using two level set functions can be implemented using the following disjoint sets: $\{\Phi_1 > 0, \Phi_2 > 0\}, \{\Phi_1 > 0, \Phi_2 < 0\}, \{\Phi_1 < 0, \Phi_2 > 0\}, \{\Phi_1 < 0, \Phi_2 < 0\}$

The boundaries of the regions defining the segmentation of the given image are specified by $\{\Phi_1 = 0\}U\{\Phi_2 = 0\}$.

In this implementation, four functions are used to link *u* and Φ.

$$\begin{aligned} I(x,y) &= I^{++}(x,y) \quad if\ \Phi_1(x,y) > 0 \text{ and } \Phi_2(x,y) > 0 \\ I(x,y) &= I^{+-}(x,y) \quad if\ \Phi_1(x,y) > 0 \text{ and } \Phi_2(x,y) < 0 \\ I(x,y) &= I^{-+}(x,y) \quad if\ \Phi_1(x,y) < 0 \text{ and } \Phi_2(x,y) > 0 \\ I(x,y) &= I^{--}(x,y) \quad if\ \Phi_1(x,y) < 0 \text{ and } \Phi_2(x,y) < 0 \end{aligned} \tag{3.15}$$

Using again the Heaviside function and the notation $\Phi = (\Phi_1, \Phi_2)$ the energy in level set formulation based on the piecewise smooth model of Mumford and Shah is defined by the following equation:

$$\begin{aligned} F(u,\Phi) &= \int_{\Omega} |I^{++} - u|^2 H(\Phi_1)H(\Phi_2)dx\,dy \\ &+\mu \int_{\Omega} |\nabla I^{++}|^2 H(\Phi_1)H(\Phi_2)\,dx\,dy + \int_{\Omega} |I^{+-} - u|^2 H(\Phi_1)\big(1 - H(\Phi_2)\big)dx\,dy \\ &+\mu \int_{\Omega} |\nabla I^{+-}|^2 H(\Phi_1)\big(1 - H(\Phi_2)\big)\,dx\,dy + \int_{\Omega} |I^{-+} - u|^2 H(\Phi_2)\big(1 - H(\Phi_1)\big)dx\,dy \\ &+\mu \int_{\Omega} |\nabla I^{-+}|^2 H(\Phi_2)\big(1 - H(\Phi_1)\big)\,dx\,dy + \int_{\Omega} |I^{--} - u|^2 \big(1 - H(\Phi_2)\big)\big(1 - H(\Phi_1)\big)dx\,dy \\ &+\mu \int_{\Omega} |\nabla I^{--}|^2 (1 - H(\Phi_2))\big(1 - H(\Phi_1)\big)\,dx\,dy + \nu \int_{\Omega} |\nabla H(\Phi_1)| + \nu \int_{\Omega} |\nabla H(\Phi_2)| \end{aligned} \tag{3.16}$$

Many researchers worked and developed models for two-phase or multiphase segmentations based on the Mumford-Shah and level set theory. The "Inward and outward curve evolution using level set" model was proposed by Amadieu et al. for two-phase image segmentation [75]. The model "A variational level set approach to multiphase motion" by Zhao et al. was developed for detecting multiple boundaries devoted to motion of junctions [76]. Yezzi et al. in their model "A statistical approach to snakes for bimodal and tri-modal imagery" are using the idea of combining several level set functions, to represent more than two segments in an image [77]. They are showing numerical results with binary flows related with the models CV and Amadieu et al., but with additional choices for the segmentation criteria. Paragios and Deriche have used also multiple level set functions for image segmentation, in the probabilistic framework "Coupled geodesic active regions for image segmentation" [78]. Tsai et al. have also developed similar models with those from

the above cases, independently and contemporaneously [71]. In addition, Cohen and collaborators [79] have used a level set method for a variant of the Mumford and Shah model related with the one described in the above section. The existence of intensity or spectral heterogeneity and feature complexity in the analysed scene creates the need of incorporating many level set functions for achieving good segmentation results. This leads to the iteration of multiple partial differential equations (PDE) simultaneously for approximating the segmentation of the under evaluation image. This very time-consuming process can limit the effectiveness of the methods.

3.5 Localized Active contour image segmentation models

In the recent years the idea of using local image statistics rather than only global, to evolve an active contour for segmenting images in the presence of intensity or spectral heterogeneity became popular [52-53, 68-70, 80-82]. The utilization of local image properties made the active contour models more effective and efficient. This is because image segmentation models based on global properties typically rely on a specific region descriptor such as intensity means value in each region to be segmented and is difficult to be estimated for images with wide intensity heterogeneity. A promising method is the local binary fitting model (LBF) proposed by Li et al. [53]. Local intensity information is incorporated in the model using a Gaussian kernel function and two smooth functions to describe the local intensities inside and outside the evolving curve are defined. The energy functional of the LBF model under the level set formulation framework is defined as follows:

$$\begin{aligned} E_{LBF}(f_1, f_2, \Phi) = \lambda_1 \int \left[\int K_\sigma (x-y) |I(y) - f_1(x)|^2 H(\Phi(y))\, dy \right] dx \\ + \lambda_2 \int \left[\int K_\sigma (x-y) |I(y) - f_2(x)|^2 (1 - H(\Phi(y)))\, dy \right] dx \\ + \mu \int \delta(\Phi(x)) |\nabla(\Phi(x))|\, dx + \nu \int \frac{1}{2} (|\nabla \Phi(x)| - 1)^2\, dx \end{aligned} \tag{3.17}$$

where $\lambda_1, \lambda_2, \mu, \nu$ are fixed parameters, f_1, f_2 are the weighted intensity average values in a Gaussian window inside, outside the evolving curve respectively and K_σ is a Gaussian kernel function with standard deviation σ. Gradient descent flow is utilized for minimizing the energy functional as follows:

$$\frac{\partial \Phi}{\partial t} = \delta(\Phi)\left[\lambda_2 e_2 - \lambda_1 e_1 + \mu div\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right)\right] + \nu\left[\nabla^2 \Phi - div\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right)\right]$$

(3.18)

$$e_i(x) = \int K_\sigma\,(y - x)|I(x) - f_i|^2\, dy, \qquad for\ i = 1{,}2$$

(3.19)

$$f_1(\Phi) = \frac{K_\sigma(x) * [I(x)H(\Phi(x))]}{K_\sigma(x) * H(\Phi(x))}, \qquad f_2(\Phi) = \frac{K_\sigma(x) * [I(x)(1 - H(\Phi(x)))]}{K_\sigma(x) * (1 - H(\Phi(x)))}$$

(3.20)

A local active contour model based on Chan-Vese (LCV) was also proposed by Wang et al. where local and global region information combined in a single framework [81]. The energy functional of the LCV model under the level set formulation framework is defined as follows:

$$\begin{aligned} E_{LCV}(c_1, c_2, d_1, d_2) &= \int (\alpha\,|I(x) - c_1|^2 + \beta\,|g_k * I(x) - I(x) - d_1|^2)\, H(\Phi(x))\, dx \\ &+ \int (\alpha\,|I(x) - c_2|^2 + \beta\,|g_k * I(x) - I(x) - d_2|^2)\, (1 - H(\Phi(x)))\, dx \\ &+ \mu \int \delta(\Phi(x))|\nabla(\Phi(x))|\, dx + \nu \int \frac{1}{2}(|\nabla \Phi(x)| - 1)^2\, dx \end{aligned}$$

(3.21)

where α, β constants are for balancing the effect of local and global terms, d_1, d_2 are the intensity average values inside and outside the evolving curve and g_k is an averaging convolution operator with $k \times k$ size window.

Gradient descent method is used to minimize the energy function of the above equation leading to the following model:

$$\begin{aligned} \frac{\partial \Phi}{\partial t} = \delta(\Phi)\Big[&-(\alpha \cdot (I - c_1)^2 + \beta\,(g_k * I(x,y) - I(x,y)d_1)^2) + (\alpha \cdot (I - c_2)^2 \\ &+ \beta\,(g_k * I(x,y) - I(x,y)d_2)^2) + \mu div\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right)\Big] + \nu\left[\nabla^2 \Phi - div\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right)\right] \end{aligned}$$

(3.22)

$$d_1(\Phi) = \frac{\int (g_k(x) * I(x) - I(x)) \cdot H\big(\Phi(x)\big)\, dx}{\int H\big(\Phi(x)\big)\, dx}$$

$$d_2(\Phi) = \frac{\int (g_k(x) * I(x) - I(x)) \cdot (1 - H\big(\Phi(x)\big))\, dx}{\int 1 - H\big(\Phi(x)\big)\, dx} \tag{3.23}$$

3.6 Evaluation of globalized and localized active contour segmentation models in the presence of intensity heterogeneity and feature complexity

In this work, a large number of experiments have been performed in order to evaluate the CV, LBF and LCV active contour models for their ability to detect buildings in monocular satellite images and delineate the breast region in mammographic images.

3.6.1 Image datasets

The models are evaluated using all the mammograms from the Mammographic Image Analysis Society Digital Mammogram Database (MIAS) [83-84] for a breast region segmentation process. The MIAS database contains 322 mammograms corresponding to 161 cases. The images in MIAS are also classified in three categories fatty, fatty-glandular and dense-glandular based on the prevalence of fibro-glandular tissue as it appears on a mammogram.

All the satellite tested images acquired from Google Earth/Google maps ((DigitalGlobe, 2013-2015), contain three bands (RGB) and represent diverse building attributes such as size, shape and colour in challenging environmental conditions from different countries. The same number of images (322) as the MIAS database proposes, was used for evaluating the performance of the two models for building detection. The spatial resolution of the selected Google Earth/Google maps images varies from 1 m to 10 m. All the tested images are 8-bit images with pixel values range from 0 to 255 and contain three bands (RGB). The Google Earth/Google maps applications for image acquisitions are selected because of their free accessibility and because spectral heterogeneity and feature complexity existed in the presented images/scenes. Three different area types high, semi and low density urban regions are defined by considering the land use along with the density of the existing buildings [85]. For the classification of the different area types contextual information such as the density, building size and distance between buildings, were utilized to quantify urban morphological patterns.

3.6.2 Accuracy assessment strategy

Quantitative pixel level measures based on the work of Shufelt [86], Ok et al. [87-88] and Ghaffarian and Ghaffarian [89] are used for performance evaluation purposes. For defining the quantitative measures, the regions of interest (ROI) were delineated manually by human operators in order to create reference or ground truth (GT) images. If both the manual and the under evaluation method detected a pixel as part of the ROI, then the pixel is denoted as a true positive (TP). If both the manual and the under evaluation method detected a pixel as a non-ROI, then the pixel is denoted as a true negative (TN). If the under evaluation method incorrectly present's a pixel as ROI while the ground truth image present's it as a non-ROI, then the pixel is denoted as a false positive (FP). Finally, if the under evaluation method incorrectly detects a pixel as a non-ROI while the ground truth image present's it as a ROI, then the pixel is denoted as a false negative (FN). The following well-known pixel based statistical metrics, precision, recall and F_B score, are computed to evaluate the performance of the under evaluation methods:

$$\text{precision} = \frac{\|\text{TP}\|}{\|\text{TP}\| + \|\text{FP}\|} \times 100 \tag{3.24}$$

$$\text{recall} = \frac{\|\text{TP}\|}{\|\text{TP}\| + \|\text{FN}\|} \times 100 \tag{3.25}$$

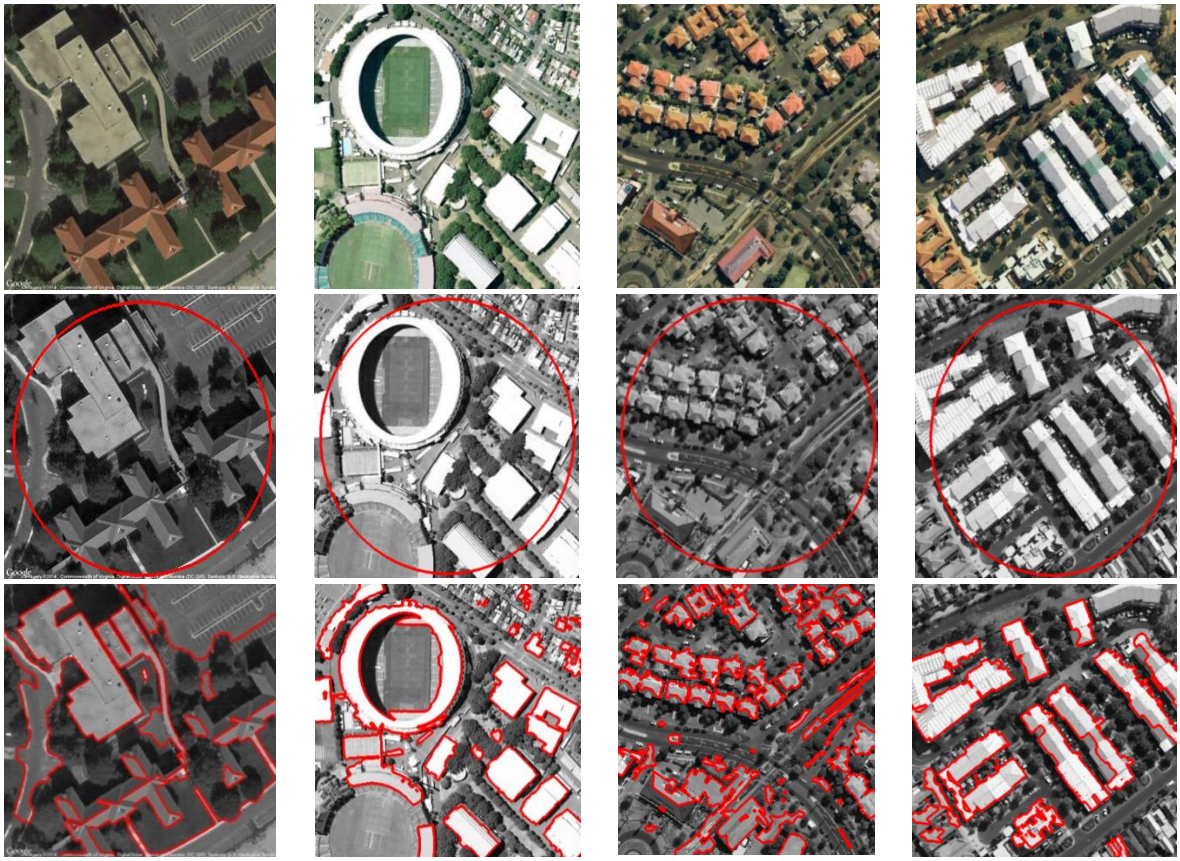
$$F_B = (1 + B^2)\frac{2 \times \text{precision} \times \text{recall}}{B^2 \times \text{precision} + \text{recall}} \times 100 \tag{3.26}$$

The precision metric gives the percentage of the true detected pixels against the true positive pixels along with the misclassified pixels as ROI pixels. The recall metric, where the denominator stands for the total number of true ROI pixels, is the ROI detection accuracy and indicates the percentage of ROI pixels correctly detected by the under evaluation algorithm. The weighted harmonic mean of precision and recall is an important measure since it combines these two metrics and is considered as a quality metric. Metrics, precision and recall are equally important in a segmentation process, thus specifying the *B* parameter equal to one, is the most appropriate.

3.6.3 Results

The models were implemented by Matlab R2013b on a computer with Intel Core i5, 2.6 GHz CPU, 4 GB RAM, and Windows 7 64 bit operating system. The common parameters of the models for the level set formulation and energy minimization such as: $\mu = 1, \text{Time step} = dt = 0.1,\ \nu = 10^{-3}255^2\ , \varepsilon = 1\ \lambda_1 = \lambda_2 = 1$ were kept the same in both models and were defined based on best practices as reported in the literature for natural images [49-54, 80-82].

In figures 7 and 8, a sample of the images that were utilized to evaluate the effectiveness of the active contour methods to segment the regions of interest are presented. Tables 1 and 2 demonstrates the pixel based quantitative measures of the overall detection performance for the under evaluation active contour models. The models were also evaluated from the aspect of computational cost. Therefore, the time needed by the three models to detect the corresponding buildings per image was recorded. The average of the computational cost of the under evaluation models are presented in Tables 3 and 4.



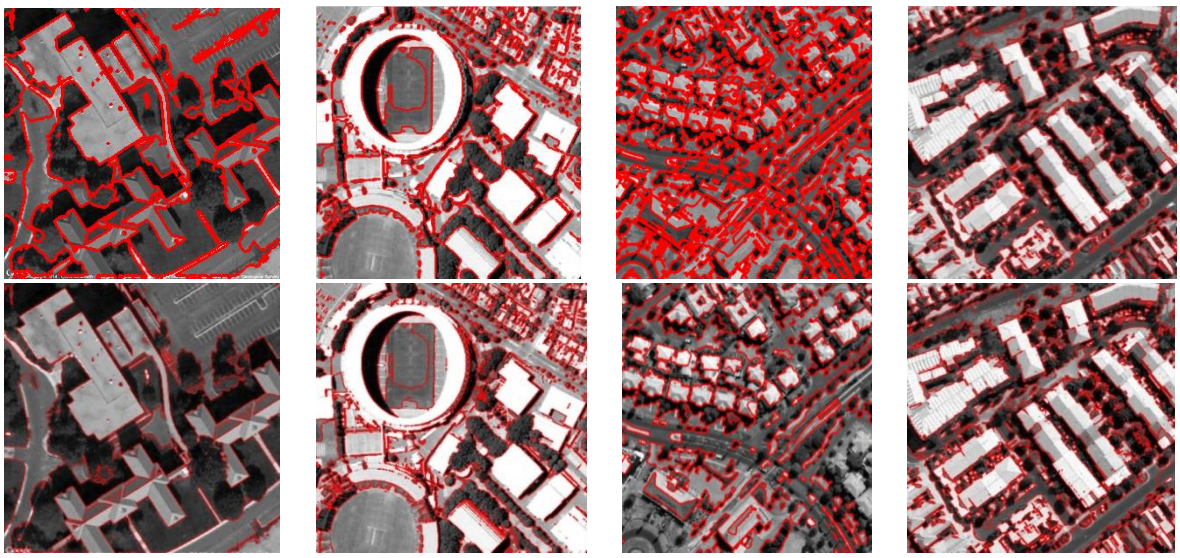
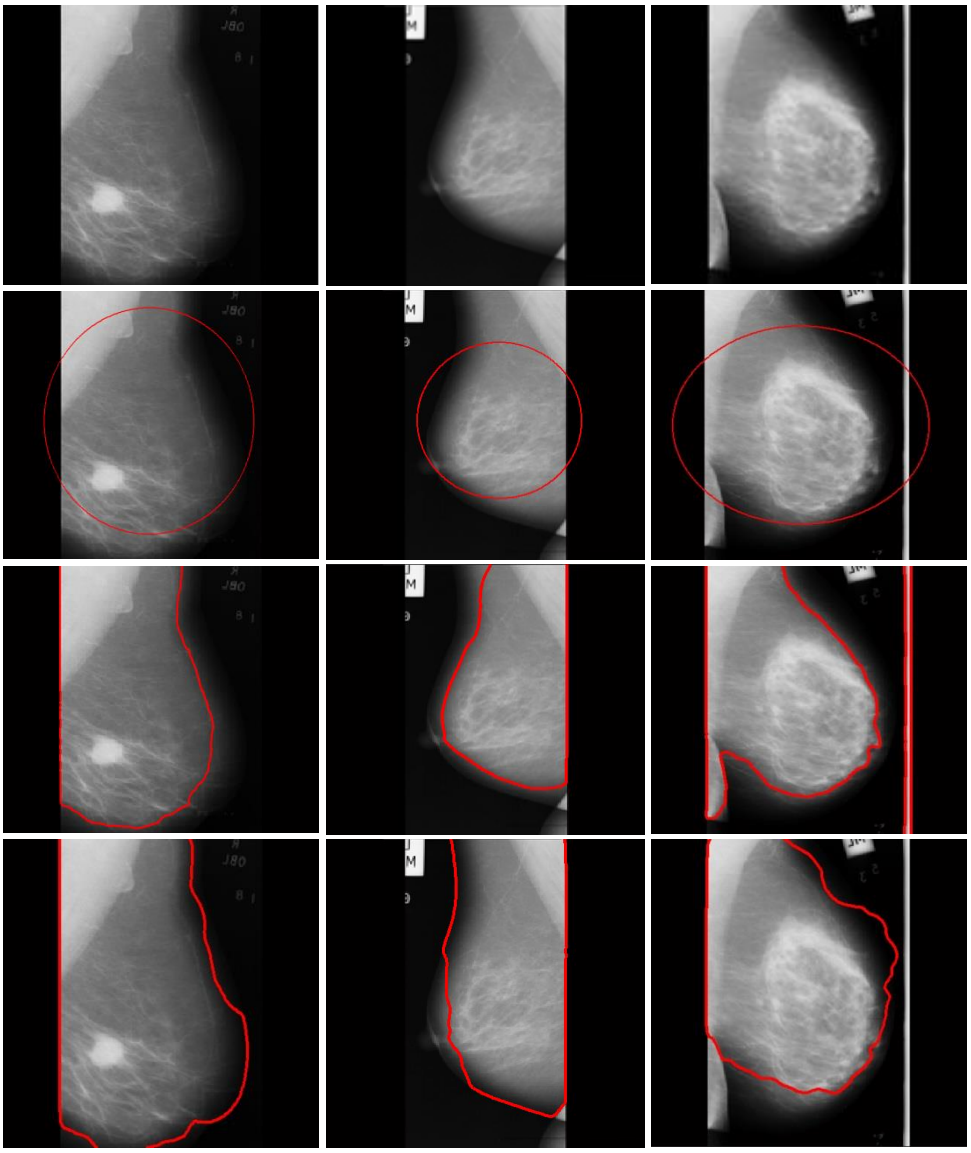


Figure 8: Applying the under evaluation active contour methods for building detection in monocular satellite images

First-row: Original images, Second-row: Initialization, Third-row: Chan-Vese segmentation, Fourth-row: Local binary fitting segmentation, Fifth-row: Local Chan-Vese segmentation.



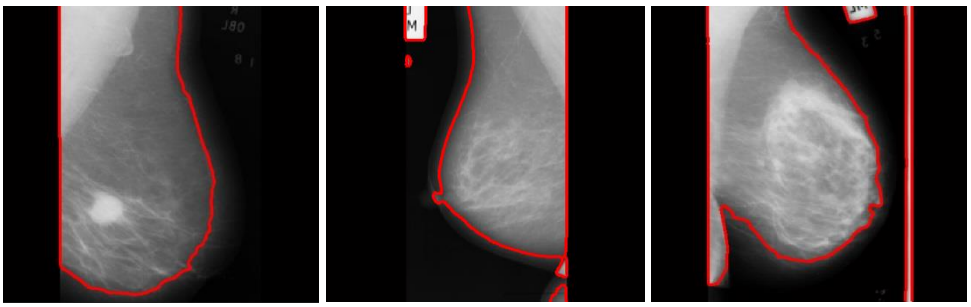


Figure 9: Applying the under evaluation active contour methods for breast boundary detection in mammographic images.

First-row: Original images, Second-row: Initialization, Third-row: Chan-Vese segmentation, Fourth-row: Local binary fitting segmentation, Fifth-row: Local Chan-Vese segmentation.

Table 1: Pixel based building structure segmentation evaluation measures using traditional active contours

Dense Urban Area Type	**Total Number of Buildings**	**Precision (%)**			**Recall (%)**			**F (%)**		
		CV	***LBF***	***LCV***	***CV***	***LBF***	***LCV***	***CV***	***LBF***	***LCV***
High density urban areas	9425	80.2	73.8	75.1	67.9	76.1	78.3	73.5	74.9	76.7
Semi density urban areas	5028	70.6	60.1	66.3	71.6	72.8	73.7	71.1	65.9	69.8
Low density urban areas	1568	69.9	53.7	60.0	70.6	70.4	76.5	70.2	61.0	67.2
Overall	16021	75.6	66.6	70.4	69.4	74.4	76.6	72.3	70.3	73.3

Table 2: Pixel based breast region segmentation evaluation measures using traditional active contours

Breast dense category		Fatty dense	Fatty-Glandular	Dense-Glandular	Overall
Total Number of Images		110	100	112	322
Precision (%)	**CV**	80.2	70.6	69.9	75.6
	LBF	73.8	60.1	53.7	66.6
	LCV	75.1	66.3	60.0	70.4
Recall (%)	**CV**	67.9	71.6	70.6	69.4
	LBF	76.1	72.8	70.4	74.4
	LCV	78.3	73.7	76.5	76.6
F (%)	**CV**	73.5	71.1	70.2	72.3
	LBF	74.9	65.9	61.0	70.3
	LCV	76.7	69.8	67.2	73.3

Table 3: Active contour traditional models computational cost evaluation

Method	**Computational Cost Average (Seconds)**	
	Satellite Images	Mammographic Images
CV	16.75	48.5
LBF	15.68	47.86
LCV	45.89	120.42

3.7 Discussion

Quantitative metrics revealed that the CV model has the lowest true detection performance or recall. LBF and LCV achieve a similar recall performance, which is better compared with the CV model. However, the false detection, which mainly influences the precision metric is

quite low when using the CV model, quite high using the LBF and slightly lower than LBF when using the LCV model.

Based on the experimental results of this work it was also found that the performance of the CV model depends on initial curve placement and number. If the initial curves do not cover effectively all the buildings that may exist in the analysed scene, some of the buildings may be missed. In LBF and LCV models due to their natural ability to split and merge faster and more efficient than CV, there is no dependency on the initial curve number and placement. However, this attribute in several cases leads to a high number of false detections. The initial curve placement and number in the CV model also influences the computational cost. Very few initial curves lead to higher computational cost.

The presence of intensity heterogeneity near the boundaries of the breast region traps the active contours in local minimum or creates leakage and affects the performance of the models. Intensity or spectral heterogeneity also leads to erroneous segmentations in satellite images affecting those buildings that part of their roofs are covered with shadows or have roof regions with lower intensity values than the average intensity value.

Several choices for the representation of the different segments and their boundaries by level sets are being used, giving promising results. However, image segmentation in the presence of intensity or spectral heterogeneity and feature complexity based on traditional active contour models is associated with several limitations, such as extensive initialization, under-detection and over-detection.

Two different approaches are followed in addressing the above limitations. In the first approach, pre-processing and post-processing methods are implemented for optimizing the overall performance of active contour level set segmentation models applied, on spectral or intensity heterogeneous and complex images. In the second approach, an optimized localized active contour level set segmentation model is implemented. The proposed model suggests a new energy minimization active contour model for image segmentation in the presence of spectral or intensity heterogeneity which can be applied in complex scenes. The work in both approaches using real data, such as mammographic and satellite images is discussed in the following chapters.

Chapter 4: Image Classification

4.1 Introduction

Image classification has received significant interest from the computer vision and image processing communities. Standard image classification processes use image descriptors in combination with support vector machine (SVM) classifiers [63]. The use of SVM classifiers is motivated by their computational efficiency. They were initially designed for binary classification and extended for multiclass classification using various schemes. One of the simplest strategies to learn classifiers in the multi-class setting is to train one-vs-rest binary classifiers independently for each class. Most image classification approaches have adopted this strategy because of its simplicity [63-65].

Extracting suitable image representations is important for the success of image classification. A successful practice has been to extract descriptors, such as SIFT [90], LBP [91], Textons [92-93], GLCM [94-95], NLC [96] and colour histograms [97]. SIFT is an approach for detecting and extracting local feature descriptors that are reasonably invariant to changes in illumination, image noise, rotation, scaling and small changes in viewpoint. The SIFT descriptor is a coarse descriptor of the edges found in the key-points. LBPs provide a robust and computationally simple way for describing pure local binary patterns in a texture. They provide information regarding the prevalence of different edge patterns and uniformity. Textons are defined under the operational definition of clustered filter responses and provide a statistical and structural unifying approach for texture characterization. Grey-level co-occurrence matrices (GLCMs) show the relative frequency with which different combinations of different pixel intensity values at certain distances, occur in an image. The matrices are used for the evaluation of different texture features such as autocorrelation, contrast, cluster prominence, energy, entropy, homogeneity. These features are computed for the different GLCMs for different pixel distances. Normalize luminance coefficients (NLC) are Gaussian intensity distribution features that capture a broad spectrum of image statistics that often exhibit changes in tail behaviour i.e. Kurtosis. Colour features have been shown to achieve higher success rate than grayscale features in image classification because colour features contain significantly better discriminative information. Colour based image classification methods can be useful in identifying regions of interest in natural images. Recent work on colour based image classification appears in [97-98] that propose several new colour spaces and methods for object and scene classification. Fusion of colour models, colour region detection, and colour edge detection

has been investigated for the representation of colour images [99]. Some important contributions of colour, texture, and shape abstraction for image retrieval have been presented in [100]. Colour features when combined with intensity and texture descriptors are able to outperform many alternatives in image classification. In this thesis, colour, normalize luminance coefficients, texture descriptors and their combination are evaluated in mammographic and satellite image for classification purposes.

4.2 Image features

1. Gaussian intensity distribution features are based on Ruderman's research findings [96]. Ruderman observed that the removal of local mean displacements from the zero log-contrast and the normalization of their local variance have a de-correlating effect. For natural images these normalized luminance values strongly tend towards a unit normal Gaussian characteristic. This operation was used to model the contrast-gain masking process in early human vision [101-102]. The NLCs are estimated using the moment-matching based approach proposed by Sharifi and Leon-Garcia [103]. For each image, two parameters from a generalized Gaussian distribution (GGD) fit are estimated and form the first set of features that capture image statistics. The statistical relationship between neighbouring pixels is also modelled. This structure that uses the empirical distributions of pairwise products of neighbouring GGD coefficients along four orientations horizontal, vertical, main-diagonal and secondary diagonal are utilized as image representation features. An asymmetric generalized Gaussian distribution (AGGD) model is adopted in order to visualize how paired products vary. The parameters of the AGGD are estimated using the moment-matching based approach proposed by Lasmar et. al. [104]. For each paired product, 16 parameters (4 parameters / orientation × 4 orientations) are computed, yielding the next set of features. Images are naturally multi-scaled to capture better the representation feature descriptors across scales. Thus, all features are extracted at two scales, the original image scale, and at a reduced resolution scale. This process leads to a 36-feature descriptor vector extracted for each image or region of interest.

2. Scale Invariant Feature Transform (SIFT): The SIFT approach which was proposed by Lowe, transforms an image into a collection of local feature vectors which are translation, scaling and rotation invariant [90]. SIFT approach proved a robust feature descriptor process in a variety of image classification, retrieval and matching applications. It provides key-point detection through the identification of interesting points in the scale space. The SIFT descriptor is a rough descriptor of the edges found in the key-points. First, the images are convolved with Gaussians at different scales and the Differences of Gaussians images

are generated from the adjacent smoothed images. The detection stages for SIFT features include scale space extrema detection, key-point localization and orientation assignment. Key-points are identified as local extrema of differences of Gaussians and localized through interpolation of nearby data. Following, the gradient orientation histogram is computed in the neighbourhood of the key-point forming the feature descriptor. The feature descriptor used in this work, is computed as a set of orientation histograms (with 8 bins) on 4 x 4 pixel neighbourhoods. This results in a 128-dimensional feature vector for each region. A regular grid with spacing M pixels over N regular support patches is used. A SIFT key-point vocabulary is built through vector quantization of descriptors from a set of training images using k-means. Finally, each descriptor is assigned to the one closest to it in the SIFT vocabulary and the images are represented with corresponding histograms.

3. Local Binary Patterns (LBP): LBPs are evaluated by calculating the local binary difference between the grey values of a pixel i and the grey values of p pixels in a local neighbourhood of i, placed on a circle of radius r. An LBP code for a neighbourhood is produced by summing up the result of the thresholded values multiplied with weights given to the corresponding pixels. However, this representation proved rotational variant, thus Ojala et al. extended this work by introducing a rotational invariant approach [91]. This was achieved by selecting the minimum value after performing $p - 1$ bitwise shift operations on the binary pattern. The number of changes between the zeroes and ones in the pattern signify whether the corresponding local texture is uniform. This means that there are a limited number of transitions or discontinuities in the circular representation of the pattern. A pattern is defined uniform if the number of transitions between 0 and 1 of the sequence is less or equal to two. The most frequent uniform binary patterns correspond to edges and corners and can be regarded as important feature detectors triggered by the best matching pattern. Here, the rotational invariant LBP code for a neighbourhood is used, produced by multiplying the thresholded values in the central pixel's local neighbourhood with weights given to the corresponding pixels, and summing up the result. For each pixel, an 8-bit number $b_1 \ldots . b_8$ is created, where b_i is denoted as 0 if neighbour i has value less than or equal to the central pixel's value and 1 otherwise. The rotation invariant representation is achieved by circularly rotating each bit pattern to the minimum value. The histogram of all the corresponding values for all the pixels is evaluated to describe the given image.

4. Statistical Distribution of Textons: Textons, as proposed by Julesz, are the primitives of texture [92]. Structural models of texture are based on the view that texture is composed of primitives in spatial arrangements. In this work, textons are defined as clustered filter responses based on the operational definition of Leung and Malik [93]. For the computation

of the texton statistical distribution features, the following process is performed. Initially the texton dictionary must be derived. Images from the training set are filtered using the Maximum Response 8 (MR8) rotational invariant filter bank proposed by Varma and Zisserman [105]. After filtering using the filter bank, each pixel is associated with a vector that holds the filter response corresponding to each filter in the filter bank. The filter responses over all the pixels in the image are aggregated. The texton dictionary is created by clustering these aggregated filter responses over all images using the k-means classification algorithm. The texton dictionary is then used for the evaluation of the different quality image categories. Each image pixel is mapped to the texton closest to it, in the filter response space. This step provides the image's texton map on which the texton histogram, showing the relative frequency of occurrences of the textons in the texton dictionary is computed forming the texton feature.

5. Grey Level Co-occurrence Matrices: In statistical texture analysis, texture features are computed from the statistical distribution of observed combinations of intensities at specified positions relative to each other in the image. According to the number of intensity points (pixels) in each combination, statistics are classified into first-order, second-order and higher-order statistics. The Grey Level Co-occurrence Matrix (GLCM) method is a way of extracting second order statistical texture features. A GLCM is a matrix where the number of rows and columns is equal to the number of grey levels, G, in the image. The matrix element $P(i,j|d,\theta)$ contains the second order statistical probability values for changes between grey levels i and j at a particular displacement distance d and at a particular angle θ. Because a $G \times G$ matrix (or histogram array) must be accumulated for each sub-image or window and for each separation parameter set (d,θ), it is usually computationally necessary to restrict the (d,θ) values to be used to a restricted number of values such as angles of $\theta = 0$, $\theta = \pi/4$, $\theta = \pi/2$ and $\theta = 3\pi/4$ radians under the assumption of angular symmetry. For a given distance d usually four angular grey level co-occurrence matrices are created. Haralick et al. 1973 [94-95] have proposed a number of scalar texture measures $T(d,\theta)$ such as autocorrelation, contrast, cluster shade and prominence, energy, entropy, homogeneity that may be extracted from these matrices. GLCM features can be extended to support colour images. In colour images the co-occurrence matric is computed for each image region represented in the respectively colour channel.

6. Colour histogram descriptors: A colour image contains three components and each pixel of a colour image is defined in a colour space, which serves as a colour coordinate system. The commonly used colour space is the RGB colour space. Other colour spaces are

usually calculated from the RGB by means of either linear or nonlinear transformations. To reduce the sensitivity of the RGB images to luminance, surface orientation, and other photographic conditions, the RGB colour space is defined by normalizing the R, G and B components. Due to the normalization, R and G are scale-invariant and thereby invariant to light intensity changes, shadows and shading [106]. The colour histogram technique is a simple low-level method that has been found producing good results in practice [97-100]. Spatial features are lost, thus spatial relations between parts of an image cannot be used. This also ensures full translation and rotation invariance. A colour is represented by a three dimensional vector corresponding to a position in a colour space. The histogram is computed using a predefined number of bins per colour component. Usually 8 or 16 bins are used to compute the histogram from the highest spatial resolution available without losing sharp peaks.

4.3 Support Vector Machines

SVMs [64] are based on statistical theory and aim to determine the location decision boundaries that result in the optimal separation of classes. SVMs have the ability to classify linearly and non-linearly separable data in high dimensional feature spaces. In SVM models, kernel functions are applied to describe a similarity relationship between the sets to be classified. A decision boundary is determined in the feature space to distinguish the classes by creating the optimal separating hyper-plane [65]. SVMs select the linear decision boundary that minimizes the generalization error. The method finds the hyper-plane that leaves the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from it. This optimum hyper-plane is produced by maximizing the minimum margin i.e. the sum of the distances to the hyper-plane from the closest points of the classes between the sets. Thus, the resulting hyper-plane is depended on border training patterns only, called support vectors. SVMs have been shown to generalize well on difficult image classification problems where the only available features are defined in a high dimensional space [107-108]. Therefore are used in this work to investigate the statistical distributions of the different feature descriptors as presented above, for representing different quality characteristics of the satellite images, breast abnormalities and density parenchymal in mammographic images.

Given a training set of instance and label pairs: $(x_i, y_i), i = 1, \dots, N$ where $x_i \in R^n$, n being the dimension and $y_i \in \{-1, 1\}$, the aim is to divide the training set such that all the points with the same label are on the same side of the hyper-plane. This leads to the following linear classification function:

$$f(x) = w \cdot x + b$$

(4.1)

corresponding to a separating hyper-plane:

$$w \cdot x + b = 0$$

(4.2)

$f(x)$ can be normalized to satisfy $|f(x)| \geq 1$ for all x_i, so that the distance from the closest point to the hyper-plane is $1/\|w\|$.

The optimal separating hyper-plane is the one that the distance to the closest point is maximal and this can be estimated by minimizing $\|w\|$.

The objective function is:

$$\min \varphi(w) = \frac{1}{2}\|w\|^2 = \frac{1}{2} w \cdot w$$

(4.3)

subject to the constraints:

$$y_i(w \cdot x_i + b) \geq 1, i = 1, \ldots., N$$

(4.4)

The optimal separating hyper-plane can be uniquely constructed by solving a constrained quadratic programming problem using the Lagrange multipliers.

$$w = \sum_i a_i y_i x_i$$

(4.5)

where $a = (a_i, \ldots. a_N)$ are the N non negative Lagrange multipliers. Eq. (4.5) consist a subset of training patters the so-called support vectors that lie on the margin. The classification function can thus be written as:

$$f(x) = sign(\sum_i a_i y_i x_i \cdot x + b)$$

(4.6)

Loose variables ξ_i and a penalty term c are introduced if the data is not linearly separable and the objective function is modified as follows:

$$\varphi(w, c, \xi) = \frac{1}{2}(w \cdot w) + c(\sum_{i} \xi_{\iota})$$

(4.7)

A kernel function is used to map the training vectors x_i into a higher dimensional space for finding a linear separating hyperplane with the maximal margin. $C > 0$ is the penalty parameter of the error term and $K(x, y) \equiv \varphi(x) \cdot \varphi(y)$ is the kernel function that should satisfy Mercer's condition [65]. Finally, the classification function is transformed as follows:

$$f(x) = sign(\sum_{i} a_i y_i \cdot K(x_i \cdot x) + b)$$

(4.8)

The following SVM kernel functions are evaluated in this work:

Linear: $K(x_i, x_j) = x_i x_j$ (4.9)

Polynomial: $K(x_i, x_j) = (\gamma x_i x_j + r)^d, \gamma > 0$ (4.10)

Radial Basis Function (RBF): $K(x_i, x_j) = \exp\left(-\gamma \left\| x_i - x_j \right\|^2\right), \gamma > 0$ (4.11)

Sigmoid: $K(x_i, x_j) = \tanh(\gamma x_i x_j + r)$ (4.12)

where γ, r, d are the Kernel parameters.

The SVMs algorithms performance strongly depends on the above parameters that control the trade-off between margin maximization and error minimization. These parameters must be optimized for better classification results. A well-known method that can be used is the *k-fold* cross validation [107-108]. The training data are partitioned into k sets and each set is used as the validation data once, while the rest data are used as the training data. The training process repeats k times for the k validation data sets and the average validation rate is used to choose the parameter combination with the highest classification accuracy.

Since the SVMs were originally developed for two-class classification, for the multi-class image classification an ensemble of binary classifiers is used, where a binary classifier is trained for each pair of classes. The decision is then taken by combining the partial decisions of the single members of the ensemble and unknown images are classified using a majority voting strategy [108].

4.4 Discussion

The increasing use of digital imagery in many fields of science and engineering introduces a demand for accurate image analysis and classification. Image classification is attracting considerable attention and the design of an effective multi-purpose classifier to be applied on a variety of tasks in natural images is a challenging task. Application specific classifiers are usually implemented and when are applied to domains that they were not originally designed for can often be inadequate, introducing a significant barrier for automated image classification. The design of a multi-purpose image classifier and its application to a variety of image classification tasks without affecting classification accuracy includes the selection and integration of the most appropriate feature descriptor sets and kernel functions. Furthermore, feature selection or dimensionality reduction and weighting must be performed in an automated way.

In the following chapters of this thesis the automated integration of the most appropriate feature descriptor sets is investigated for the design of a multi-purpose classifier using SVMs and evaluated using a variety of tasks, using medical and satellite imagery.

Chapter 5: Estimation of the Breast Boundary in Mammograms Using Active Contour Models

5.1 Introduction

Breast segmentation plays a critical role in the extraction of useful information from mammograms. It is also important for image compression, storage and transfer. Different image processing techniques have been applied for the segmentation, classification and analysis of mammograms and segmentation of the breast region is an important pre-processing step for computer aided detection systems. Additionally to limiting the region where the algorithms must be applied, it provides important information regarding the deformation of the breasts both laterally and temporally. Moreover, identification and extraction of the effective breast region is also important in picture archiving and communication systems (PACS), image retrieval and tele-mammography systems [109].

Despite the availability of Full Field Digital Mammography (FFDM), it is of utmost importance to digitize process and analyse film mammograms because a number of screening programs only have access to screen film mammography. In addition, even in centres where FFDM is available, many mammograms from previous visits used for temporal evaluation are in film form. Yet, it is still important to keep, digitize and analyse such mammograms, as they are important for temporal evaluation for the early detection of breast cancer. The large amount of film mammograms, drives the development of computer applications that can be used to pre-process digitized mammograms for further automatic evaluation in order to assist the clinicians in the reliable and early diagnosis of cancer.

A mammogram contains two main distinctive regions: the exposed breast region and the unexposed air-background region. The breast contour or skin-air boundary is an important feature of a mammogram. The breast is a well-defined curve and the background, as imaged by the x-ray mammography, is a very low intensity, low gradient region. However, this is also true for the breast edge as it is mostly composed of adipose tissue that is radiolucent due to the compression of the breast between two plates during mammogram acquisition that produces intensity heterogeneities. Thus, it is a very difficult to distinguish the skin-air contour that separates the breast tissue area and background.

Different segmentation methods include thresholding, region growing, edge detection, random fields and deformable models [110]. As the breast appears as a well-defined curve in a mostly low intensity low gradient background, the presented segmentation uses active contours following pre-processing that incorporates anatomical information to find and

enhance the breast boundary. Based on the characteristics of mammograms, level set formulation of active contours can serve as an appropriate method for breast region extraction in a breast segmentation process. Active contours are able to track well-defined curves and in addition, the low intensity background in mammograms can be avoided by the moving interface in its search for a local minimum.

The algorithm presented in this work uses level sets to establish the corresponding skin-air boundary. An initial pre-processing procedure is performed where the image contrast is improved and the breast edge is enhanced. Artefacts, labels and noise are also removed from the mammogram. Following, a new optimized active contour (OAC) model able to selectively identify the breast boundary is developed and applied. In the proposed OAC, indices from the pre-processed image are encoded to form a new energy term that biases the contours to detect the breast boundary and avoid intensity heterogeneities that exist near the breast edges. Post-processing using spline smoothing removes line artefacts and results in a smooth breast boundary. The results are quantitatively evaluated against the radiologist's ground truth. In the following sections, the developed method is presented in detail, and different measures are used to quantify the resulting segmentations.

5.2 Literature review

Algorithms for breast region segmentation and skin-air boundary detection include use of local grey-value range on modified histogram analysis, border region search methods based on the gradient of grey values, polynomial modelling, classifiers and active contours based methods. Many of these algorithms are used for further mammogram analysis as precise segmentation of the breast region is essential in mammogram registration and in the extraction of quantitative features for classification of breast tissue and mammographic density classes [111]. The goal of intensity thresholding is to determine a brightness that will separate the image into object and background. In 1979, Hoyer et al. [112] used a simple thresholding technique in an attempt to segment the breast region from the background in mammograms. Most probably, this was the first attempt for mammogram segmentation. Simple global thresholding was also used by Lau et al. [113], Yin et al. [114] and Bying et al. [115]. Bick et al. [116] based their work on modified histogram analysis and local thresholding, including only pixels with a low range of intensities in their local neighbourhood. Using the modified histogram and a number of threshold values the pixels were classified as potential breast pixels. This procedure was followed with a region growing technique that generated the border boundary. Hein et al. [117] observed that low frequency images of a mammogram provide a better separation between breast

region and background compared to grey level intensities. By taking a reduced size wavelet expansion image at low frequencies, and applying, a simple threshold to eliminate the noise the largest object in the resulting image was identified as the breast region. Masek et al. [118] proposed a local thresholding method combined with polynomial modelling. The identification of the breast region was simplified by dividing the skin-air interface into two sets of functions, one using the rows as the x-axis, the other using the columns. This simplification led to two methods, one based on local thresholding along the estimate, the other based on a method, where the mean squared error between a polynomial and the estimated outline was minimized by changing the threshold and polynomial order in a given section of the skin-air interface. While the local-threshold technique resulted at times in a rough breast outline, the polynomial based algorithm produced an accurate skin-line in most images, with results comparing well against established algorithms. Methods using histogram grey level information are fast and usually not computationally demanding compared to other methods. However, these methods suffer from a reduced effectiveness in delineating a smooth breast contour and accurate segmentation of the breast region. Nevertheless, useful results can be obtained from automatic global and local thresholding. Depending on the application, the segmentation is accurate enough to use for further processing or as a starting point for more accurate segmentations.

Border region search methods based on the gradient of grey-level values have also been extensively utilized. Semmlow et al. [119] used spatial filters and a Sobel edge detector to obtain the breast boundary. Mendez et al. [120] used the gradient to identify the breast boundary after a two level histogram threshold technique. Morton et al. [121] and Zhou et al. [122] used similar methods where after subscripting the background using an initial threshold, an edge was found using a line gradient analysis process. The work of Karssemeijer and Brake [123] utilized a multi-resolution scheme, using a global threshold followed by the application of a Sobel operator. Abdel-Mottaleb et al. [124] identified the breast edge through the evaluation of different threshold values to create two images that were considered as candidates for identifying the breast contour. The gradient of these images and the unified gradient were utilized for detecting the breast region. Highnam and Brady's [125] algorithm initially applied linear Hough transforms in the areas where the film edges should appear to find and mark them. A spatial gradient operator was then applied to the image and local variances were calculated. The image background was found by thresholding the spatial gradient variances in the local neighbourhoods using a constant predefined threshold. Background pixels have lower gradients and this is how they were separated from the breast region. The result was post-processed using

morphological operators. Any wrongly marked pixels were removed with a small mask that performs binary morphological erosion.

Chandrasekhar et al. [126] used an initial threshold to approximate the breast region follow by a polynomial model. Fitting polynomials of different orders to the row and column interface generated from several thresholds and choosing the order and threshold that minimize the mean square error. Candidate output images were produced using polynomials of order 0, 1 and 2 and the final output image was a union of several candidates. Saha et al. [127] used a classifier where a scale-based fuzzy connectivity technique was used to segment the breast region. A fuzzy segmentation technique was also used by Wirth et al. [128].

Very few breast region segmentation methods explore the potential of active contours for breast segmentation. Ferrari et al. [129] enhanced the mammogram applying a logarithmic operator to the entire image and corrected the contrast in order to enhance the regions near the skin-air boundary. The optimal threshold was found with an iterative technique and a spline was used to approximate the boundary. Wirth et al. [130] performed dual thresholding on the original mammogram to identify all possible initial placement points and ran the initial placement points through an algorithm for a least-squares fit piecewise quadratic curve. They used points along this curve for initializing a parametric active contour. Next, they performed a right to left edge enhancement and equalization of the original mammogram to obtain the breast contour edge image. A left to right edge enhancement on the original mammogram was also performed followed by grayscale erosion in order to be used as a noise mask on the breast contour edge image. The snake's image energy measure was defined as a right to left and down to up gradient functional. Start and end point movements of the snake were limited to make an open contour. A small factor was used on left neighbours of control points to favour the snake movement to the left, since in most cases the breast was to the left of the initial placement. Finally, the breast boundary was found as the results of the snake’s energy minimization.

5.3 The development of an optimized active contour segmentation model to automatically estimate the breast boundary

A breast boundary segmentation method must result in an accurate segmentation of the breast tissue, remove any labels and digitizer-introduced wedges, provide a smooth outline of the breast border and preserve the nipple if it exists in the profile. The fact that the CV [49] model is not based on edge gradient, makes it a good choice for breast region segmentation, due to the specific characteristics of the breast edge, which is mostly

radiolucent as a result of the mammogram imaging process. Additionally, the CV model separates the image into two distinctive regions and a mammogram mainly contains two major distinctive regions, the breast region and the air-background region. Still, the enhancement of the breast edge and localization of where the breast edge lies will enable the algorithm to converge faster and achieve better edge localization. For accurate segmentation of the breast boundary, the mammogram needs to be pre-processed to remove label and other artefacts in the background region. Following, the image is enhanced to enable the CV level set to achieve the best possible segmentation.

5.3.1 Overall algorithm

Initially different thresholding, morphological and enhancement operations are applied to establish, limit and enhance the region where the breast edge lies. Mammogram images must adhere to quality standards involving optical density, contrast, image sharpness, breast compression, processing and film artefacts [131]. However, this is not the case in a significant number of cases [132]. The processing of the mammogram incorporates knowledge regarding the standard characteristics of mammograms and corresponding features. Following the breast region is identified using an optimized active contour model. Finally, spatial morphological operations using a very small disk are applied to remove any errors due to lines artefacts on the breast and/or labelling wrongly placed near or overlapping with the breast edge. Then, the breast outline is approximated using smoothing splines to get the final smooth breast boundary. Figure 10 presents the flowchart of the breast region segmentation algorithm.

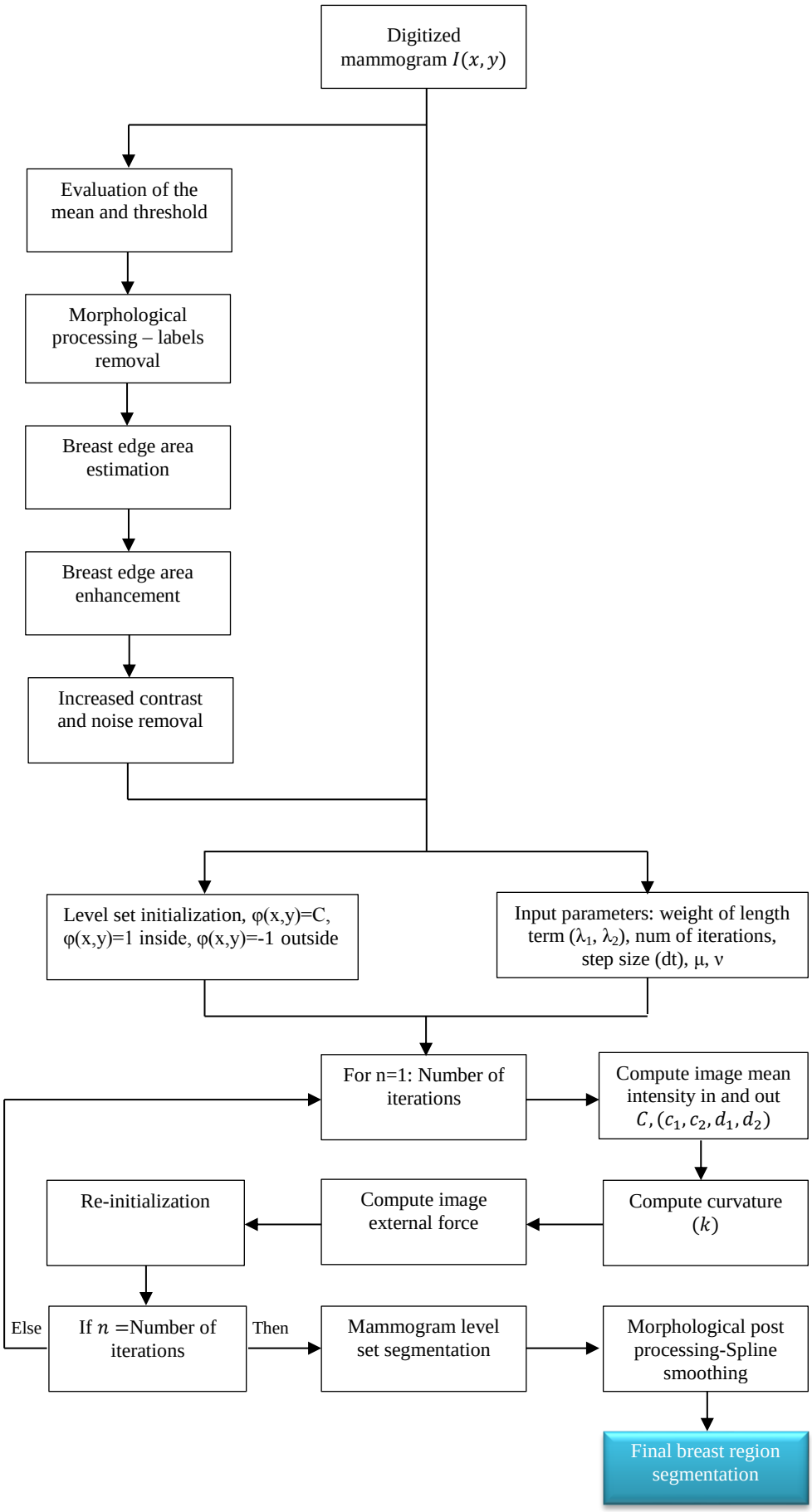


Figure 10: Breast region segmentation algorithm flowchart

5.3.2 Mammogram pre-processing

An initial pre-processing procedure is performed to improve the image contrast and enhance the breast edge to enable the CV level set to segment the true breast region. Initially, global threshodling using a multiple of the image mean is performed. Artefacts, labels and noise are removed from the mammogram using the result of the thresholding and following morphological operations. Following, a more specific region, where the breast edge may be found, is identified and enhanced in order for the level set to avoid falling into a local minimum.

The glandular tissue has relatively high x-ray absorption and the surrounding area is relatively dark. A multiple of the mean of the image provides a good threshold for constructing a mask from which a rough estimate of the breast region area is derived. The masks of the higher intensity regions resulting from the thresholding operation provide information regarding the location of the breast on the mammogram as well as the labels and other possible artefacts. The mask resulting from the breast is the largest object on the image, but still underestimates the actual breast region.

Labels and spiked edges are removed using morphological operations [1]. Closing using a disk with a relatively large radius of 1.5cm is used to remove labels. Using the size of the different masks, the breast region is identified. As this mask is an underestimation of the breast area, opening is used to enlarge the area covered by the mask. A disk element with a radius of 2.5 cm or less radius disk is big enough to capture the entire breast edge as it appears on a mammogram [125] and is used to include the breast's edge in the mask. The difference between the resulting mask and the eroded original, isolates the area where the breast edge lies. The original mask of the breast region is slightly eroded to ensure that the algorithm captures the breast edge even in cases where the mammogram is very dense. This area includes only a region along the breast's edge, with intensities below the mean and exhibits very poor contrast. To enhance the region directional median filtering is applied to enhance the pixels' intensities and remove noise.

To improve contrast, especially in this region, the corresponding intensities are mapped from zero to the maximum intensity of the original mammogram minus the maximum intensity in the localized region. In some cases, where the mammograms do not adhere to the clinical standards for mammographic imaging [131], the image contrast is really low and the background is really noisy. Thus, the mean intensity in a small region corresponding to the most convex area is evaluated and compared to the mammogram's mean intensity. Depending on the result, the pixels with the lowest intensities are mapped

to zero. The resulting image is added with the region corresponding to the original mask. The addition of the two images includes the breast region and some background area near the breast edge. More importantly, a band near and including the breast edge is enhanced. Median filtering is then applied to remove noise. This is the image where the CV level set is applied to get the breast region segmentation.

Figure 11 illustrates the output results of the different pre-processing steps, applied to the mammogram to prepare the image for the application of the CV level set.

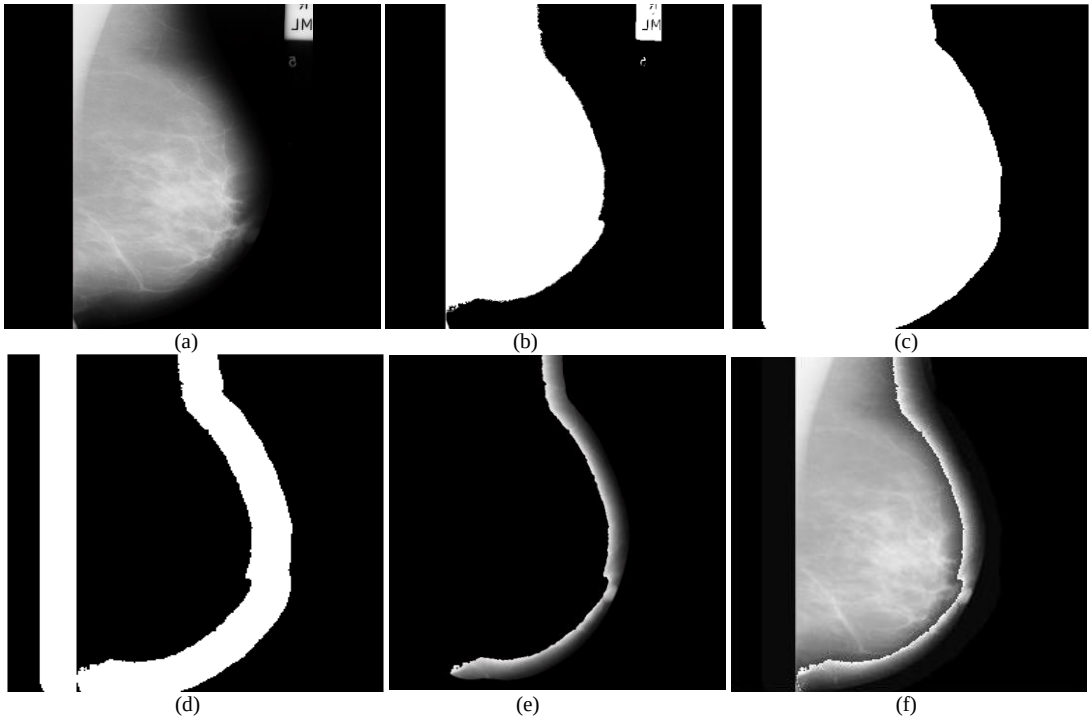


Figure 11: Mammogram image pre-processing

(a) Original image, b) Thresholded binary image, (c) Morphological operations - labels removal, (d) Breast edge area, (e) Breast edge area increased contrast, (f) Breast edge area increased contrast added to a multiple of the mask of the region corresponding to the original mask, (g) Final pre-processed mammogram.

5.3.3 Optimized active contour model for breast region segmentation

The model used is an optimal version of the formulation for the active contours without edges by Chan and Vese [49] as was defined and solved in equations 3.9-3.15. However, the proposed optimized active contour (OAC) model utilizes indices from both the original digitized mammogram and its pre-processed representation. The OAC is described as follows:

$$
\begin{aligned}
&E_{\mathrm{OAC}}(c_1, c_2, d_1, d_2, \Phi) \\
&\quad = \lambda_1 \int_{\Omega} (I(x,y) - c_1)^2 H(\Phi)\,\mathrm{dx\,dy} \\
&\quad + \lambda_2 \int_{\Omega} (I(x,y) - c_2)^2 (1 - H(\Phi))\,\mathrm{dx\,dy} + \lambda_1 \int_{\Omega} (I'(x,y) - d_1)^2 H(\Phi)\,\mathrm{dx\,dy} \\
&\quad + \lambda_2 \int_{\Omega} (I'(x,y) - d_2)^2 (1 - H(\Phi))\,\mathrm{dx\,dy} \\
&\quad + \mu \int_{\Omega} \delta\big(\Phi(x,y)\big)\big|\nabla\big(\Phi(x,y)\big)\big|\,\mathrm{dx\,dy} + \nu \int_{\Omega} H(\Phi(x,y))\,\mathrm{dx\,dy}
\end{aligned}
\tag{5.1}
$$

where $\delta(\Phi)$ is Dirac function, $c_1,\ c_2$ are the intensity mean values of the original image (I) inside and outside C respectively, d_1, d_2 are the intensity mean values of the transformed (pre-processed) image (I') inside and outside C respectively. Using the same principles and Heaviside function $H(\Phi)$ as presented in section 3.4 the following equations are used to approximate the mean intensity values:

$$
c_1(\Phi) = \frac{\int_{\Omega} I(x,y) H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}}{\int_{\Omega} H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}} \quad c_2(\Phi) = \frac{\int_{\Omega} I(x,y)\Big(1 - H\big(\Phi(t,x,y)\big)\Big)\,\mathrm{dx\,dy}}{\int_{\Omega} 1 - H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}}
$$

$$
d_1(\Phi) = \frac{\int_{\Omega} I'(x,y) H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}}{\int_{\Omega} H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}} \quad d_2(\Phi) = \frac{\int_{\Omega} I'(x,y)\Big(1 - H\big(\Phi(t,x,y)\big)\Big)\,\mathrm{dx\,dy}}{\int_{\Omega} 1 - H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}}
\tag{5.2}
$$

The level set formulation of the model where the forward finite differences scheme is applied to produce breast region segmentation is described as follows:

$$
\begin{aligned}
\frac{\partial \Phi(x,y,t)}{\partial t} &= \delta_{\varepsilon}(\Phi)[-\lambda_1((I - c_1)^2 + (I' - d_1)^2) + \lambda_2((I - c_2)^2 + (I' - d_2)^2) \\
&\quad + \mu\,\mathrm{div}\left(\frac{\nabla\Phi}{|\nabla\Phi|}\right) - \nu]
\end{aligned}
\tag{5.3}
$$

For mammogram segmentation, the initial zero level set must first be constructed. A general object model, in this case a small rectangle, away from the breast boundary is used. For the level set evolution the zero level set is described as a function of time based on equation 5.3. The algorithm steps of the active contour segmentation method are as follows:

Step 1: For the under evaluation image, the initial contour C is implemented and placed on the image by initializing the level set function as follows:

$$\Phi(x,y) = C, \Phi(x,y) = 1 \text{ inside } C, \Phi(x,y) = -1 \text{ outside } C$$

Step 2: Set the parameters for the level set formulation and energy minimization.

Num. of iterations $= 50$, Time step $=$ dt $= 0.1, \mu = 1, \nu = 10^{-3}255^2$, e $= 1, \lambda_1 = \lambda_2 = 1$.

The parameters are specified following the same principles as was discussed in chapter 3, section 3.6.3 and tested further in this work.

Step 3: Evolve the level set function Φ according to 5.3. At each time step, the constants c_1, c_2, d_1, d_2 are updated according to 5.2 and the level set function is re-initialized.

Step 4: Extract the zero level set from $\Phi_t(x, y)$ if the level set evolution terminates.

The OAC model identifies the breast edge using the mean intensities across the segmented regions.

5.3.4 Breast boundary post-processing

After the breast boundary is established artefacts such as lines that run the whole height of the image within the breast region, or labels/writing that partly cover breast tissue are suppressed using spatial morphological operations. An opening operation is used to remove lines followed by a closing operation to close any holes that may result. After the morphological operations, spline smoothing, through sampling points along the boundary and interpolating using smoothing splines, yields the final smooth breast boundary contour.

5.4 Experimental results and discussion

The algorithm is evaluated on mammograms from the Mammographic Image Analysis Society Digital Mammogram Database (MIAS) [83-84]. In this work, the performance of the algorithm is evaluated on all mammogram images of the MIAS database. Figures 11, 12 and 13 illustrate results of the presented method as well as the final segmentation of mammograms after the presented pre-processing. In a number of cases in literature the algorithms are evaluated qualitatively by experts who categorize the segmentations as accurate or acceptable [116], [123]. The segmentation algorithm presented here is quantitatively evaluated by comparing the resulting segmentation with the ground-truth (GT) segmentations of the corresponding images that are drawn by an expert clinician.

The quantitative pixel level measures as presented in chapter 3 and section 3.6.2 are used. Furthermore, a measure named mean distance (MD) is also used. It is defined as a measure of the distance between the edge of the segmented region derived by the algorithm and the true breast edge as identified by the expert clinician. This can be perceived as another, more specific measure of how accurate the method is. MD is evaluated using the following:

$$MD = \frac{\sum_{i=1}^{n} distance|actual\_skinair\_boundary - evaluated\_skinair\_boundary|}{n} \quad (5.4)$$

where n is the vertical height of breast in pixels.

The mean values of the quantitative measures for the evaluation of presented method on all the images are shown in Table 4. The algorithm can achieve good segmentation results even when the images do not adhere to the clinical standards for mammography [132]. This is demonstrated on the examples in Figures 12 and 13. The mammogram in Figure 12 contains a horizontal line across the length of the image almost halfway across the region. The vertical line can be seen in the result of the level set segmentation, but the opening and closing operations following the OAC level set segmentation remove the corresponding artefact. The same is true for the mammogram shown in Figure 13. This mammogram does not include the entire breast region and there is writing overlapping the breast tissue. Again, the post-processing removes the corresponding artefacts.

Table 4: Breast region segmentation quantification using the proposed method

Mean Recall Accuracy (%)	96.43
Mean Precision Accuracy (%)	98.71
Mean F Score Accuracy (%)	97.55
Mean Distance (mm)	0.794
Mean Distance Standard Deviation (mm)	0.723

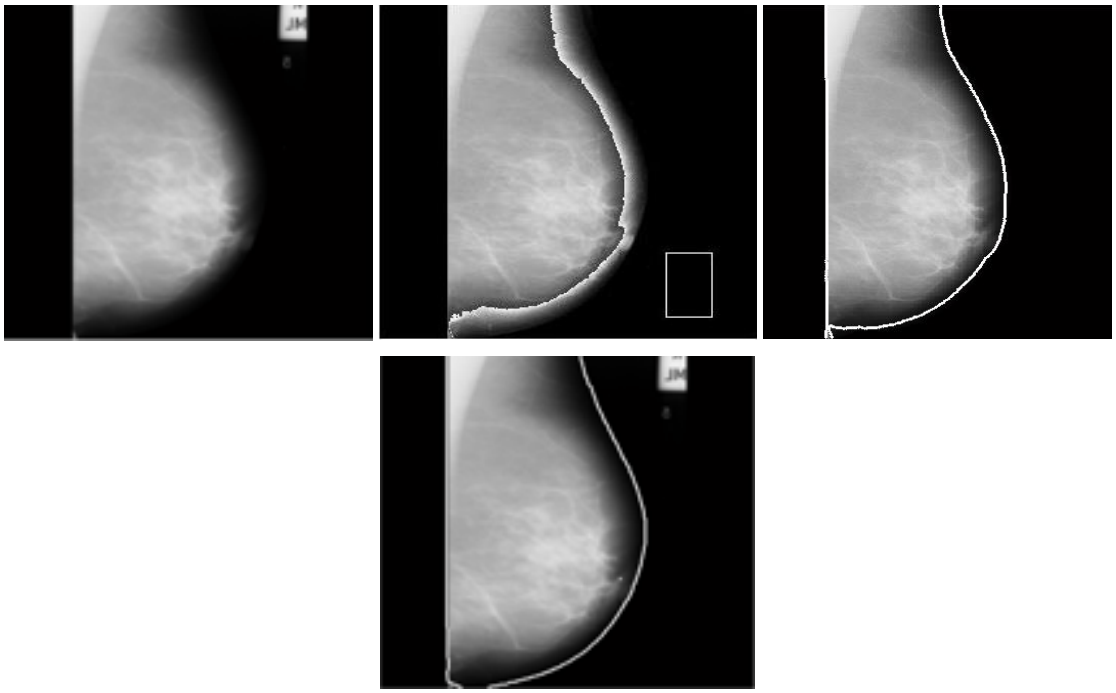


Figure 12: Mammogram image segmentation sample results

(a) original image, (b) pre-processed image with initialization, (c) OAC level set result, (d) final segmentation.

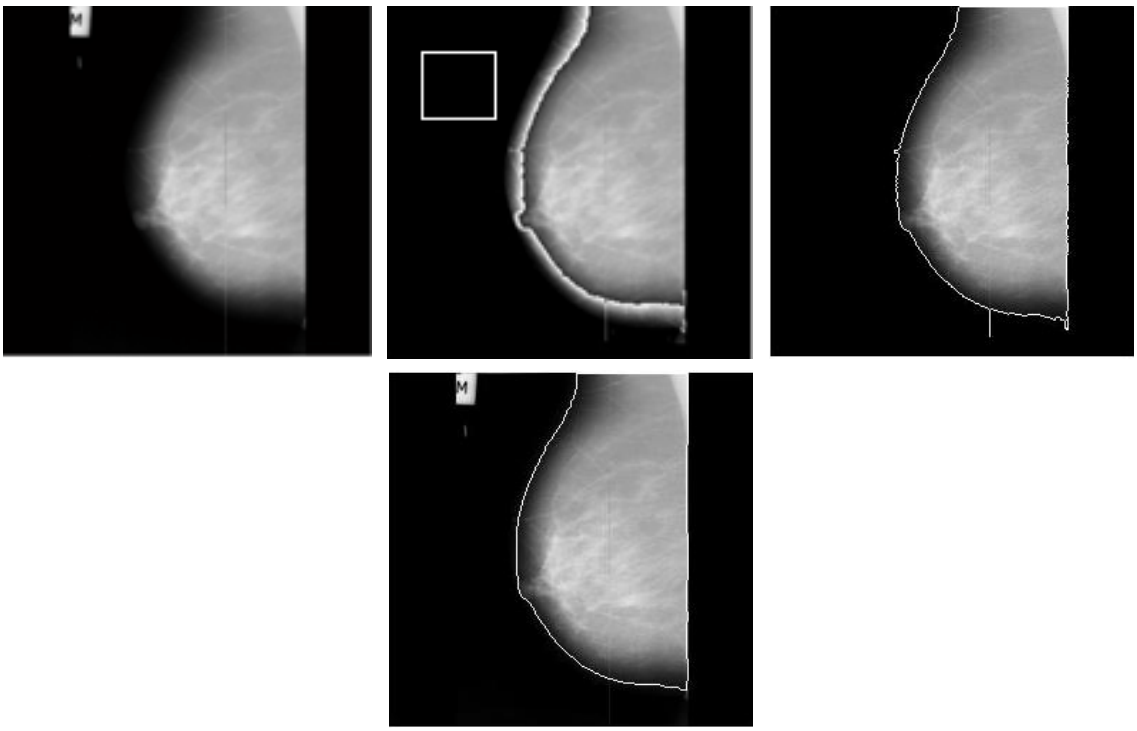


Figure 13: Mammogram image segmentation sample results-a line artefact exists vertically on the breast region

(a) original image, (b) pre-processed image with initialization, (c) OAC level set result, (d) final segmentation.

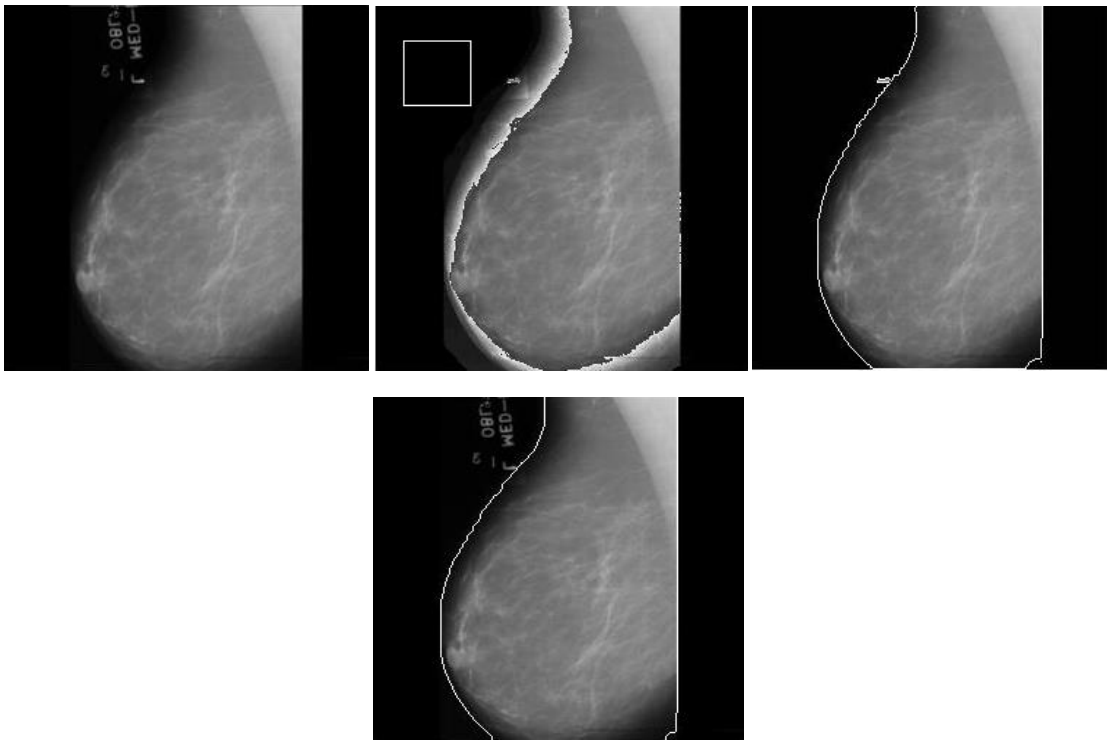


Figure 14: Mammogram image segmentation sample results-labels/writing overlaps the upper breast region

(a) original image, (b) pre-processed image with initialization, (c) OAC level set result, (d) final segmentation – the line in the upper left region of the breast is no longer part of the boundary but it is the lettering as can be seen on the original image.

The segmentation results of the presented method demonstrate that the algorithm works well with images of breasts of different sizes, qualities and densities and can detect the breast edge, even when other artefacts are present, as can be seen in Figures 11, 12 and 13. Another important advantage is that the produced outline of the breast is smooth and the nipple is preserved if it exists in the profile. The algorithm is evaluated on mammograms corresponding to different breast tissue densities: from fatty to dense, and achieves comparative segmentation results.

The algorithm achieves good segmentation results because it separately addresses different problems that arise due to different characteristics and artefacts that may appear on the mammogram. By using the initial thresholding, the enhancement of the breast edge becomes localized, resulting in higher contrast. The choice of the OAC model is based on trying to separate the image to regions based on intensities and not the breast edge gradient that can be ill-defined. The post-processing ensures that any line artefacts that may affect the breast edge are removed and the breast edge is smoothed. However, it should be noted that if a relatively large label overlaps the breast region, the opening operation might not be able to remove the artefact.

The algorithm achieves good segmentation that reaches a mean accuracy of 96% and a mean distance of 0.79 mm. These results compare favourably to other results in the literature, in the cases where quantitative results are available. Wirth et al. [130] achieve segmentation accuracy around 98%. However, their method is limited by the availability of gradients in the breast edge, especially in the bottom of the breast region and has the limitations of parametric active contours. A final important issue that the presented method overcomes is the placement of the initial contour for the level set function. As was mentioned before, the initial level set can be away from the breast region and a general object can be used.

It should also be noted that if the mammograms do not adhere to clinical quality standards [132] the resulting segmentation might suffer, as is the case with most of the breast segmentation algorithms. However, the algorithm has been shown to work well even in such cases, as is shown in Figures 12 and 13. This is because the algorithm specifically addresses some of the related issues. A disadvantage of the presented method that should be noticed is the re-initialization of the level set function that produces in some cases a delay in the segmentation process.

5.5 Conclusions

Segmentation of the breast region and identification of the breast boundary constitute an important pre-processing step for mammographic image analysis and processing. Limiting the area to be analysed and processed into a well-defined region, with an accurate boundary allows for mammogram registration, accurate feature extraction and much more. Yet, segmentation of the breast region is a difficult task due to image characteristics that result from the acquisition process, as well as other artefacts such as labels, lines and other noise that may interfere with the segmentation method.

The presented optimized active contour model incorporates knowledge from mammogram characteristics and enhances a more localized region where the breast region may lie. The experimental results from this work suggest that this level set formulation is appropriate for the estimation of skin-air boundary for breast segmentation process. After utilizing indices from the proposed pre-processing of the image, the initial level set interface expands and delineates the breast edge accurately. The OAC level set achieves the segmentation of the enhanced image by evaluating the intensities and not the gradients in the segmented regions. Another important advantage of the algorithm is that the initialization of the level set can be away from the breast region and an initial general object can be used. The algorithm accounts for different artefacts that may appear in the mammogram.

Chapter 6: Building Extraction in Satellite Images using Active Contours and Colour Features

6.1 Introduction

Obtaining the segmentation of building footprints from satellite images is a complex process since building areas and their surroundings are presented with various colour intensity values and complex features. Active contour region based segmentation methods can be used to establish the corresponding boundary of building structures that exhibit a certain similarity and homogeneity. However, using the traditional active contour algorithms for building structures detection, in several cases where spectral heterogeneity exists, over-detection or under-detection drawbacks needs to be addressed. In this chapter, the RGB representation and the properties of the Hue, Saturation and Value (HSV) colour space are analysed and used to optimize the extraction of buildings from satellite images in an active contour segmentation framework. Initially, the satellite image is processed by applying a clustering technique using colour features to eliminate vegetation areas and shadows that may adversely affect the performance of the algorithm. Subsequently, the HSV representation of the image is used and a new active contour model is developed and applied for building extraction, utilizing descriptors derived from the value and saturation images. A new energy term is encoded for biasing the contours to achieve better segmentation results. An effective procedure has been designed and incorporated in the proposed model for the active contour initialization. This process enhances the performance of the model, leading to lower computational cost and higher building detection accuracy. Additionally, statistical measures are used for designing optimum morphological filters to eliminate any misleading information that may still exist. Qualitative and quantitative measures are used for evaluating the performance of the proposed method.

6.2 Literature review

The manual processing of satellite images is tedious and time-consuming, thus very expensive. This leads to the need for designing effective methods for the automated detection of buildings using remote sensing or satellite images. Designing and implementing methods for detecting building structures is an active field of research and a large number of studies have been reported in literature. The existing methods can be classified into two categories with respect to the type of the images that are used. The first category includes the detection of buildings using monocular remote sensing images and the

second category covers the detection of buildings using an additional channel of data such as height information. Detailed reviews can be found in [133-136]. Since this work deals with the detection of buildings through single monocular optical remote sensing images, the discussion of the previous studies will be focused on this domain.

Initial attempts in the monocular context mostly used edge-based methods and relied on the extraction of features such as lines and corners [137-138]. In these studies, the cast shadows of the buildings were also utilized for the detection of the corresponding buildings. Shadow information was used for estimating the location, shape and height of buildings [138]. Liow and Pavlidis used an edge detector and features such as shadows for extracting building boundaries from aerial images [139]. Furthermore, local edge information and global shape information was utilised to eliminate the segmentation errors or to enhance the resultant contours. Shufelt and McKeown combined several separated systems that relied on shadow information for the detection of a complete set of building boundaries [140]. McGlone and Shufelt have used image orientation information and shadow evidence to verify a building hypothesis based on the projective geometry and vanishing point calculations [141]. Shufelt performed a detailed evaluation and a comparison study of the edge-oriented methods [142]. The Shufelt's work revealed that with edge oriented methods it is difficult to handle the entire building complexity because these methods utilize information from only a single band and are mostly based on a simple building shape hypothesis such as a parallelogram structure. Thus, new methods are required to achieve better results.

The launch of remote sensors with the ability to acquire VHR multispectral images, inspired many researchers to use information from the additional bands in their proposed models. Classification methods were utilised in a number of research works. Lee et al. developed a supervised classification method, where Hough transformation was used in an automated building detection process from Ikonos imagery [143]. Geometric image features were extracted and a support vector machine (SVM) classification framework used for detecting artificial objects in Inglada's model [144]. Texture features were utilised in an SVM framework by Koc-San and Turker [145]. Ghaffarian and Ghafffarian collected training areas using shadow evidence and designed an automated parallelepiped supervised classification method for detecting building objects [6]. Graphs based theory was also used in implementing methods for the detection of buildings. The scale invariant feature transform (SIFT) was used in a graph cut model by Sirmacek and Unsalan [146]. A two level graph partitioning framework was developed by Ok et al. [87], for increasing the

performance of his previous proposed model, which was based on fuzzy logic and Grab Cut [88].

Markov Random Fields (MRFs) were thoroughly investigated in building detection models from satellite images. Krisshnamachari and Chellapa [147] used MRFs to group line segments and in the later years Katartzis and Sahli, extended this model by creating a stochastic framework for detecting rooftops of the presented buildings [148]. Independent component analysis (ICA) was utilised by Ghaffarian and Ghafffarian, for developing a model, named purposive fast-ICA to detect buildings from monocular high-resolution images acquired from Google Earth [89]. In this study, a colour based fusion technique was proposed to approximate shadow, vegetation, bare soil, road and building patches. This information was then used to initialize the Fast-ICA algorithm for detecting the presented buildings in a given image.

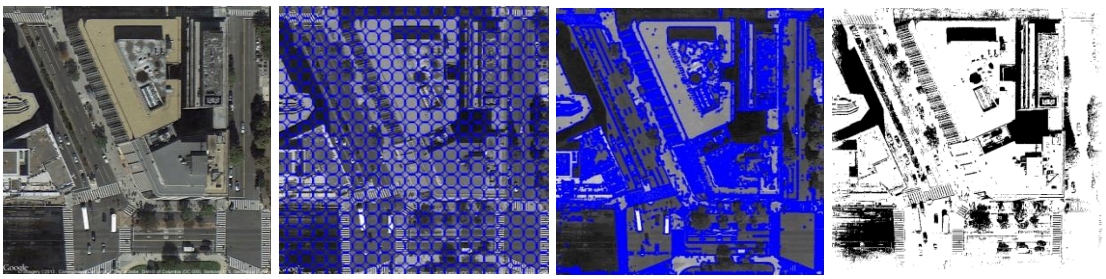
A number of studies reported in the literature used a different approach by analysing and incorporating multiple cue data to segment satellite images. Tsai analysed and evaluated the most common colour spaces for shadow detection in aerial images using a threshold technique applied to the colour spaced transformed images [149]. Sırmacek and Unsalan used invariant colour features and shadow information for building detection [150]. The RGB colour space was used and shadows were detected by applying a threshold on the blue colour invariant image. In addition, buildings with red rooftops were detected by applying a similar method on the red colour invariant image. The evaluation and subsequently the selection of the most appropriate features capable to characterize all or most of the buildings presented in the analysed scene, is important for the performance of all the above methods.

In several remote sensing applications, among the various methods already discussed, curve evolution schemes like active contours, deformable models and level set were used to segment satellite images. These methods revealed promising results due to their ability to cope with topological changes [151-158] . Niu used a geometric active contour formulation, which is based on level set methodology in a semi-automatic framework, where the initial curves are manually defined in order to detect certain objects like buildings [152]. In this formulation, the energy minimization function is based on image edges or gradients, a process that can limit the success of the method especially when the scenes are incorporating objects with weak boundaries. Cao et al. used an energy function, which was based on a modified Mumford-Shah segmentation model for artificial objects detection from aerial images [153]. They used a coarse-to-fine strategy and a fractal error metric at the evolutionary stage of the algorithm in order to detect urban or semi urban areas in aerial images. Karantzalos and Argialas also proposed a region-based level set segmentation

method for artificial object detection in aerial or satellite images [154]. A variational geometric level set function based on active contours without edge formulation was used and artificial objects such as roads and buildings were detected. These methods in some cases are vulnerable to misleading information derived from the presence of a large scale of intensities, shadows or occlusions, which is a common scenario in satellite images. Several studies have been reported to address the above issues by incorporating prior building shape information to label the segmented regions as buildings [155-156]. These schemes depend on the efficiency of the pre-defined building shape templates. A different approach can be seen in Peng et al. study, where an improved active contour model is proposed by modifying the traditional snake to be attracted from the radiometric and geometric building features by customizing the snake energy function [157]. In a similar approach, Ahmadi et al. designed an innovative model based on active contours for extraction of building boundaries from high-resolution aerial images where the radiometric properties of the building classes were predefined [158].

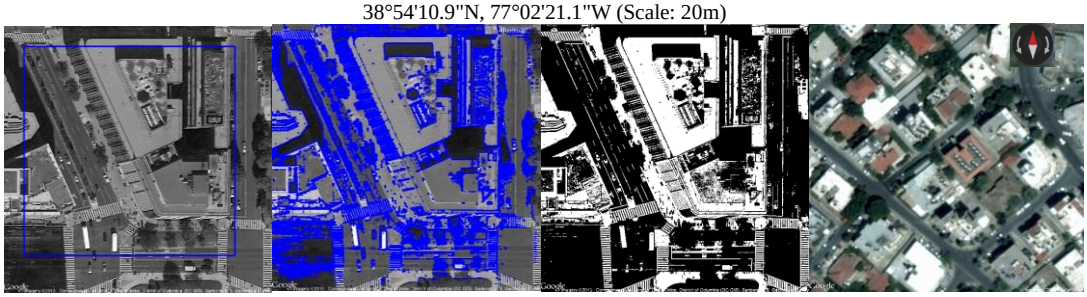
6.3 Active contours without edges model applied for building structures detection

The CV model has been used extensively in literature for the detection of buildings. Despite the good results, it is associated with several limitations such as extensive initialization, under-detection and over-detection. The model depends on initial curve placement and number. If the initial curves do not cover effectively all building structures that may exist in the analysed scene, some of them may be missed. Nevertheless, if the initial curves over-cover the whole scene including objects or regions that do not from consist building structures, then this may lead to erroneous or over-segmentations. The initial curve placement and curve number in the CV model also influences the computational cost. Very few initial curves lead to higher computational cost. Furthermore, spectral heterogeneity may also lead to misleading segmentation of building structures. A typical urban area in satellite images includes a variety of classes of buildings presented with different colour intensity values that makes it difficult for the traditional CV model to detect all buildings correctly. Several experiments have been performed in this work using the traditional CV model, with different initialization schemes and sample results of heterogeneous and complex scenes are presented in figure 15. Conclusions about the performance of the CV model are also in line with the findings of Ahmadi et al., Karantzalos and Paragios and Niu work [152, 154-156, 158].



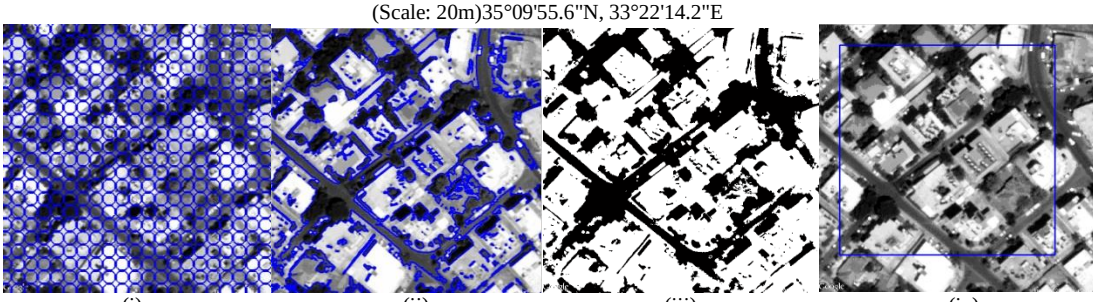
(a) (i) (ii) (iii)

38°54'10.9"N, 77°02'21.1"W (Scale: 20m)

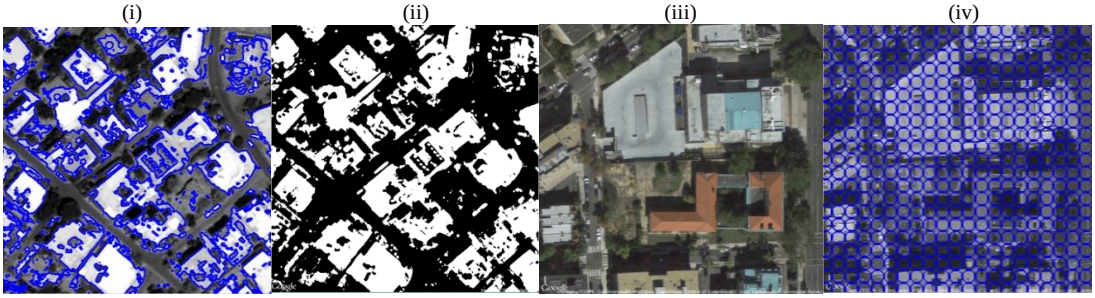


(iv) (v) (vi) (b)

(Scale: 20m)35°09'55.6"N, 33°22'14.2"E

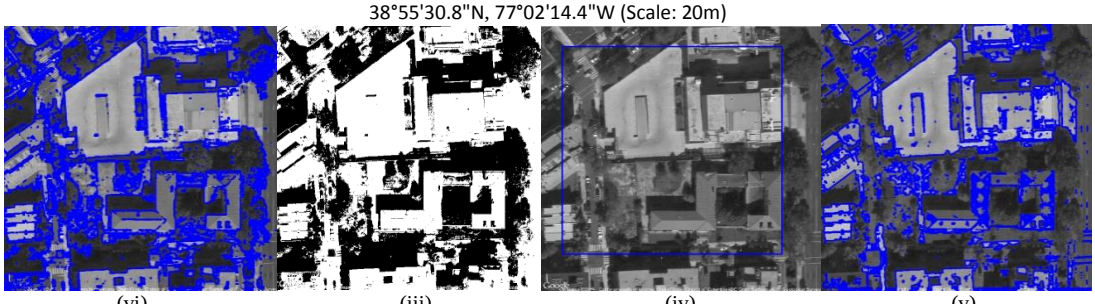


(i) (ii) (iii) (iv)

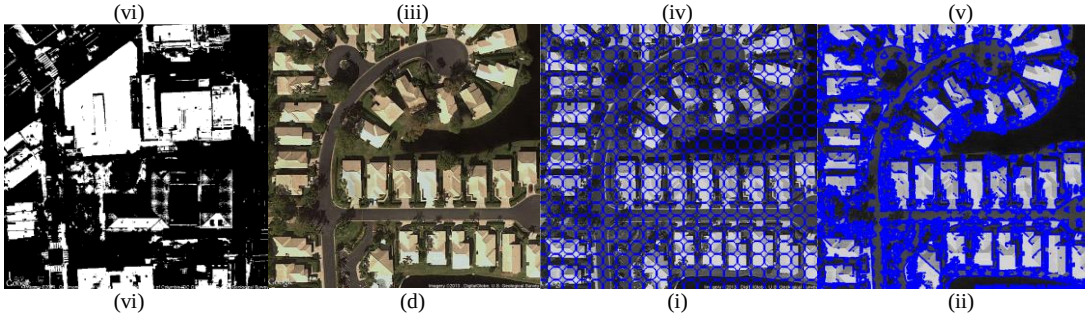


(v) (vi) (c) (i)

38°55'30.8"N, 77°02'14.4"W (Scale: 20m)



(vi) (iii) (iv) (v)



(vi) (d) (i) (ii)

27°27'24.7"N 82°36'25.9"W (Scale: 20m)

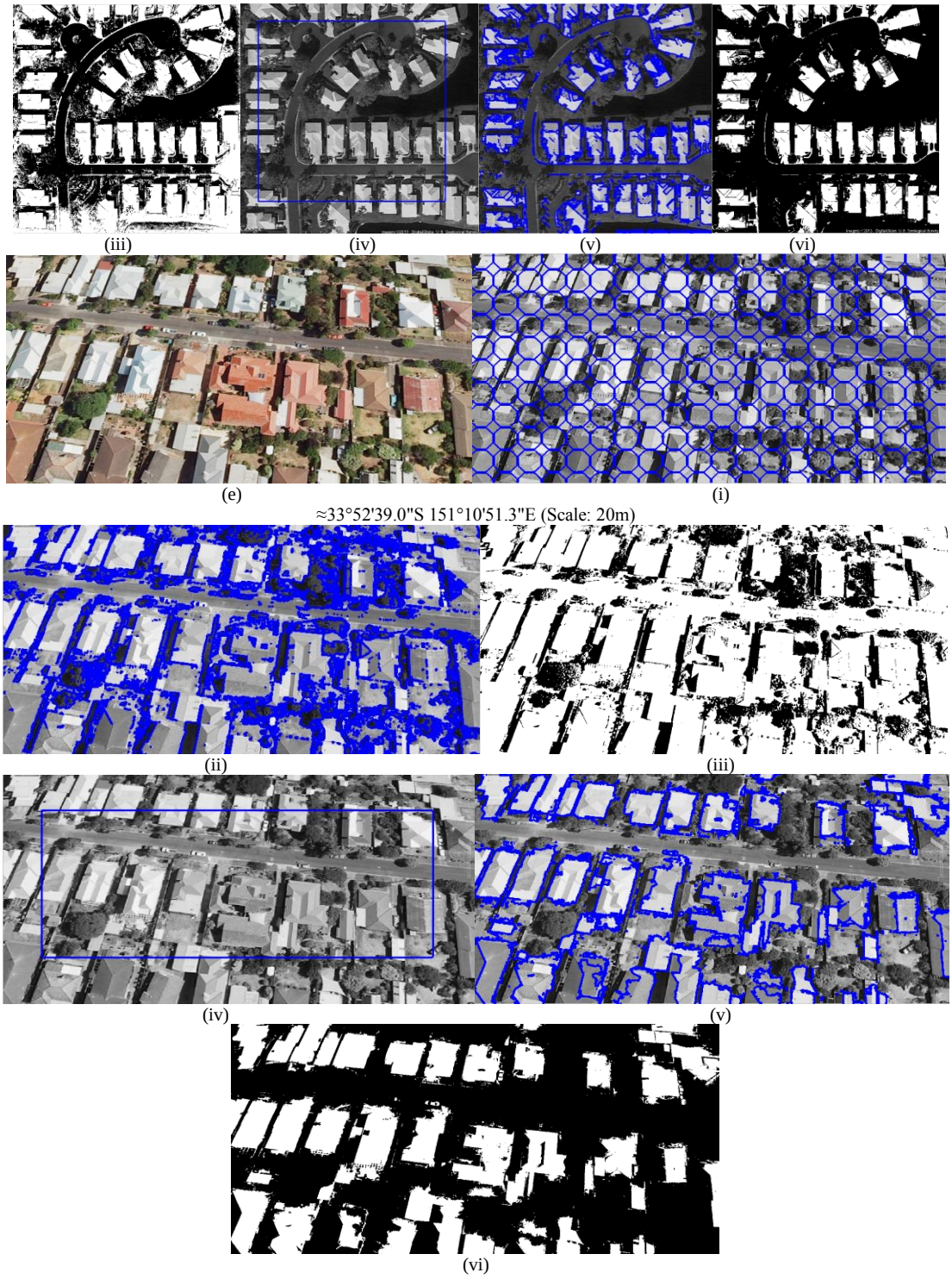


Figure 15: Traditional CV segmentation method

(a), (b), (c), (d), (e) Original RGB image, (i,) Initialization scheme using a series of regular circles all over the image, (iv) Initialization scheme using a rectangle, (ii), (v) Active contour building boundaries detection, (iii), (vi) Segmentation mask

The approach that this work follows to address the above limitations is based on designing an effective automated procedure for active contour initialization. Subsequently, the available information from colour features is utilised for enhancing the detection of building accuracy in an optimized active contour segmentation model. The proposed active

contour model can be considered an extension of the CV model. The aim was to produce balanced results between true and false detections with efficient computational cost.

6.4 The Development of an optimized active contour model to automatically detect buildings

6.4.1 Active contour initialization

The estimation of the optimal position of the initial curves can prove to be a difficult task without prior knowledge of the scene and the objects to be detected. In the urban building boundary extraction process using active contours that was proposed by Ahmadi *et al.*, the initial curves were generated automatically as a series of regular circles all over the image [158]. The number and size of the initial curves were experimentally estimated. Karantzalos and Paragios [155] used an arbitrary elliptical curve to initialize their proposed curve evolution scheme and a data term based on prior knowledge of the analysed scene was incorporated to drive the evolving curve to the building boundaries. In the above publications, it was noticed that the initialization schemes influence the models speed of execution and accuracy.

A *k*-means clustering algorithm which aims to segment *n* observations into *k* partitions, in which, each observation is assigned to the nearest partition (with respect to partition centre) can be incorporated and pre-segment the image [66]. This information can be used to develop the initial curves of the level set segmentation procedure for addressing the initialization issue. Using a *k*-means algorithm, along with morphological operations erosion and dilation for the refinement of *k*-means clustering result, the initial curves can be defined within or near the buildings that may exist in the image and the active contour model may lead to a better building segmentation result. The form and the parameters of the structuring element, to be used in erosion and dilation, are important for the success of the process. Since the aim is to detect buildings, which in most of the cases are presented as squares, rectangles or a combination of the above, a structuring element in the form of a rectangle is naturally more suitable. In this work, the length and width of the structuring element are estimated using statistical measures based on the objects that have been detected using the *k*-means algorithm. After a number of experiments, it was found that the square root of the average length and width of the detected objects are optimum values for the structuring element size.

The *k*-means algorithm calculates data features forming a vector space $s_i, i = 1 \dots k$ where k is the number of clusters and tries to find the natural clustering among them [159].

Literature suggests that this method is sensitive to initialization [160] thus a method called *K-means++* proposed by Arthur and Vassilvitskii has been incorporated for choosing initial cluster centroid positions or seeds [161]. The *k-means++* algorithm uses a heuristic to find centroid seeds for *k*-means clustering. According to Arthur and Vassilvitskii, *k-means++* improves the running time of Lloyd's algorithm, and the quality of the final solution. All the points x_j are clustered around centroids or mean points m_i ($m_i \forall = 1 \dots k$) that are obtained by minimizing the following objective function:

$$\sum_{i=1}^{k} \sum_{x_j \in s_i} (x_j - m_i)^2 \tag{6.1}$$

The algorithm steps of the proposed pre-segmentation method for designing the initial curves are as follows:

Step 1: Use the appropriate single band (saturation or value) representation of a given image and compute the intensity histogram.

Step 2: The parameters such as the number of iterations and the number of the different clusters are defined.

$Number\ of\ iterations = 10,\ \ Number\ \ of\ clusters = 2$. The number of iterations was calculated experimentally and kept static for all the images under evaluation. Two classes are defined in order to represent as foreground objects all the buildings and as background objects all the rest.

Step 3: Initialize the k centroids using the *k-means ++* algorithm.

Step 4: Repeat the following steps until the cluster labels of the image become static:

4.1 Cluster the points based on the distance of their intensities from the centroid intensities

$$c^i = \text{argmin}_j \left\| x^i - m_j \right\|^2 \tag{6.2}$$

4.2 Compute the new centroid for each cluster

$$m_i = \left(\sum_{i=1}^{k} 1(c_i = j)x^i\right) / \left(\sum_{i=1}^{k} 1(c_i = j)\right) \tag{6.3}$$

where j and i iterates over all the centroids and all the intensities respectively, m_i are the centroid intensities and k is the number of clusters.

Step 5: Apply morphological operations on foreground objects using a structuring element (SE) in the form of a rectangle.

5.1 Compute the number of connected components or objects (n), their length X_i for $i = 1 \ldots n$ and width Y_i for $i = 1 \ldots n$.

5.2 Use the length and width of the objects for the calculation of the structuring element length and width (X_{SE}, Y_{SE}) based on the following functions:

$$X_{SE} = \sqrt{\text{average}(X_i)}, Y_{SE} = \sqrt{\text{average}(Y_i)} \tag{6.4}$$

5.3 Apply erosion and then dilation using the same SE.

Step 6: Apply boundary detection for developing the final contours to be used as the initial curves for the active contour segmentation model.

Figure 16 presents sample results of initial curves implementation using the proposed method.

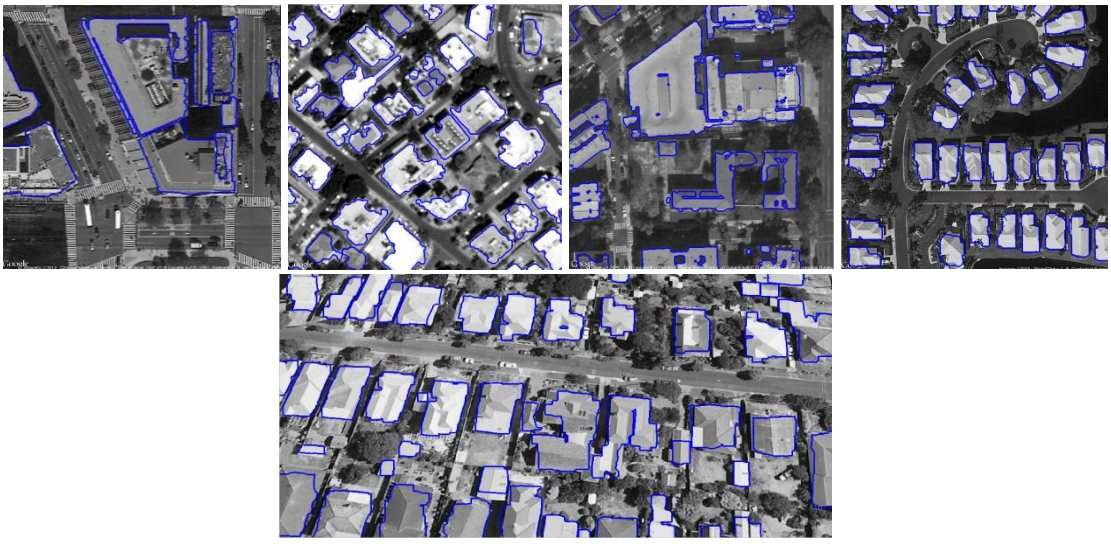


Figure 16: Active contour's initialization curves using the proposed initialization method

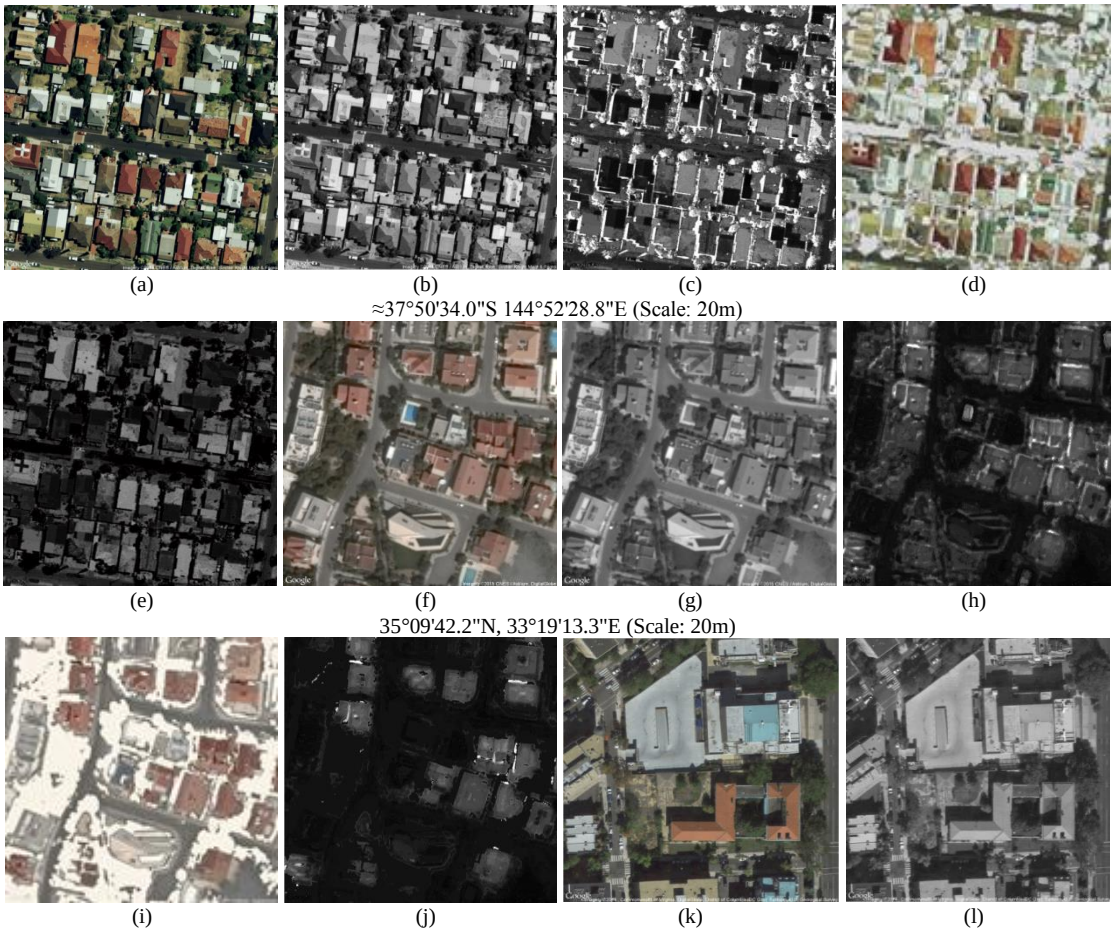
6.4.2 RGB and HSV Colour space features

In most cases, building structures in satellite images are presented with various colour intensity values and different classes of buildings are created in the grey scale representation of the image. Thus, some buildings may be missed or detected incorrectly by the active contour CV formulation segmentation model even encoded with a well-defined initialization process. During this work, it has been observed that in those cases where a large scale of colour intensity values exists, the CV model does not usually accurately detect the buildings presented with mean intensity values, lower than the majority of the existing buildings. This was also observed by Ahmadi et al. where it was noticed that the number of building and background classes, which are formed by the variety of colour intensity values, is an important factor in achieving good building detection results using the CV model [158]. Obtained segmentation results using the traditional CV method, where buildings presented with mean intensity values lower than the majority of the existing buildings, can be observed from Figure 15. This limitation can be overcome using features from the different colour spaces such as the RGB and HSV. Concrete, asphalt and white surfaces are low saturation areas while colour surfaces are high saturation areas [162]. After the initial experiments performed in this work it has been observed that the saturation image, captures well building structures presented with colours such as red, brown, blue etc. The value image seems to capture well the buildings with high colour intensity values in all three RGB bands. The above observations are utilised in this work for simplifying the capture scenes and the regions of interest in satellite images become more homogeneous. Figure 17 presents, the value and the saturation representation of Google Earth images. As it can be seen, buildings presented with red, brown or very similar colour roofs that are common in urban areas, are highlighted in saturation images. However, it should be also noticed that vegetation regions and shadows are also highlighted in the saturation image and this may affect the accuracy of the building feature detection. Thus, the applied method pre-processes the satellite image by applying a colour clustering technique to eliminate vegetation regions and shadows that are typically classified as areas with low illumination in several scenes and can adversely affect the building extraction process. To eliminate these regions, indices and information obtained from the visible bands are used. The relationships between the red, green and blue bands in these areas are utilised and the pixel colour intensity values of shadows and vegetation regions are identified and subsequently eliminated using the *k*-means and thresholding-clustering algorithm as described in section 6.4.1. Vegetated regions are classified based

on their relatively high intensity value in the green band while shadows are classified based on their low intensity values in all three bands. After the iterative procedure using equations 6.2 and 6.3 applied on the saturation component of the image, the detected regions are labelled as vegetation or shadow areas and eliminated if the following criterion functions are satisfied respectively:

$$c_{\mathrm{R}}^{i} < c_{\mathrm{G}}^{i}, c_{\mathrm{B}}^{i} < c_{\mathrm{G}}^{i} \text{ and } c_{\mathrm{G}}^{i} < T_{\mathrm{V}}$$
$$c_{\mathrm{R}}^{i} \cong c_{\mathrm{G}}^{i} \cong c_{\mathrm{B}}^{i} \text{ and } c_{\mathrm{R}}^{i}, c_{\mathrm{G}}^{i}, c_{\mathrm{B}}^{i} < T_{\mathrm{S}} \tag{6.5}$$

where $c_{\mathrm{R}}^{i}, c_{\mathrm{G}}^{i}, c_{\mathrm{B}}^{i}$ are the intensity values for the RGB bands and T_{V}, T_{S} are threshold values defined experimentally. Figure 17 shows sample results of the saturation images after the application of the clustering procedure based on *k*-means and thresholding for eliminating the shadows and vegetation areas.



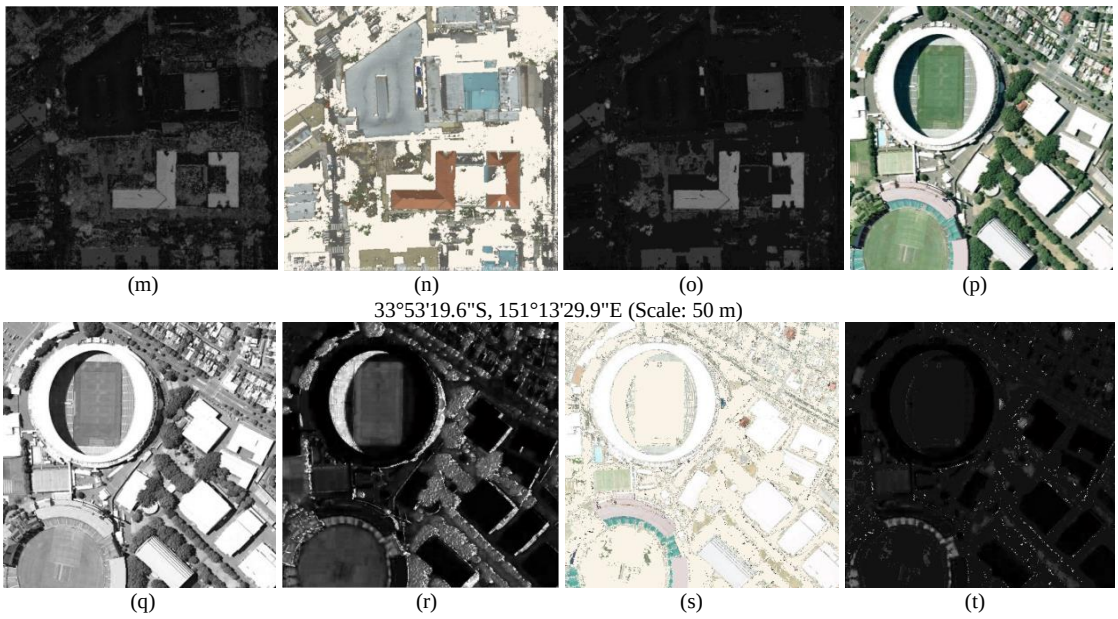


Figure 17: Clustering method using colour features

(a), (f), (k), (p) Original RGB image, (b), (g), (l), (q) Intensity value images, (c), (h), (m), (r) Saturation component of the original RGB images, (d), (i), (n), (s) Results using the proposed clustering method, (e), (j), (o), (t) Saturation components of the proposed clustering method results.

6.4.3 Optimized active contour segmentation model

The piecewise constant model of CV as was defined in equation 1 is extended in this work for partitioning a given image (I) with a variable closed curve (C) for the detection of buildings. The level set formulation $C = \{(x, y)|\Phi(x, y = 0\}$ is used to solve the energy minimization problem following the principals as presented by equations 3.10 and 3.13 in chapter 3. A new energy term is incorporated utilizing the intensity descriptors of the value (I) and saturation (S) representation of a given image after eliminating shadow and vegetation regions using the proposed scheme as presented in section 6.3.2. The proposed active contour model is described by the following equation:

$$E_{OAC}(c_1, c_2, d_1, d_2, \Phi) =$$

$$\lambda_1 \int_\Omega (I(x,y) - c_1)^2 H(\Phi)\,\mathrm{dx\,dy} + \lambda_2 \int_\Omega (I(x,y) - c_2)^2 (1 - H(\Phi))\,\mathrm{dx\,dy}$$

$$+\lambda_1 \int_\Omega (S(x,y) - d_1)^2 H(\Phi)\,\mathrm{dx\,dy} + \lambda_2 \int_\Omega (S(x,y) - d_2)^2 (1 - H(\Phi))\,\mathrm{dx\,dy}$$

$$+\mu \int_\Omega \delta\big(\Phi(x,y)\big)\big|\nabla\big(\Phi(x,y)\big)\big|\,\mathrm{dx\,dy} + \nu \int_\Omega H(\Phi(x,y))\,\mathrm{dx\,dy}$$

(6.6)

where c_1, c_2 are the intensity mean values of the value image inside and outside C respectively and d_1, d_2 are the intensity mean values of saturation image inside and outside

C respectively. $H(\Phi)$ and $\delta(\Phi)$ are the Heaviside and Dirac functions estimated by equations 3.11.

Keeping Φ(x) fixed and minimizing the CV energy with respect to the constants $c_1, c_{2,} d_1, d_2$ then these constants are estimated as follows:

$$c_1(\Phi) = \frac{\int_\Omega I(x,y)H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}}{\int_\Omega H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}} c_2(\Phi) = \frac{\int_\Omega I(x,y)\Big(1 - H\big(\Phi(t,x,y)\big)\Big)\,\mathrm{dx\,dy}}{\int_\Omega 1 - H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}}$$

$$d_1(\Phi) = \frac{\int_\Omega S(x,y)H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}}{\int_\Omega H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}} d_2(\Phi) = \frac{\int_\Omega S(x,y)\Big(1 - H\big(\Phi(t,x,y)\big)\Big)\,\mathrm{dx\,dy}}{\int_\Omega 1 - H\big(\Phi(t,x,y)\big)\,\mathrm{dx\,dy}} \tag{6.7}$$

Subsequently, keeping $c_1, c_{2,} d_1, d_2$ fixed and minimizing the CV energy with respect to $\Phi(x)$ where the Euler-Lagrange equation is used to represent $\Phi(x)$, the following level set formulation model is derived:

$$\frac{\partial \Phi(x,y,t)}{\partial t} = \delta_\varepsilon(\Phi)[-\lambda_1((I-c_1)^2 + (S-d_1)^2) + \lambda_2((I-c_2)^2 + (S-d_2)^2) + \mu \mathrm{div}\left(\frac{\nabla\Phi}{|\nabla\Phi|}\right) - \nu] \tag{6.8}$$

The algorithm steps of the active contour segmentation method are as follows:

Step 1: The initial contour C is implemented and placed on the given image by initializing the level set function using the following equations:

$$\Phi(x,y) = C, \Phi(x,y) = 1 \text{ inside } C, \Phi(x,y) = -1 \text{ outside } C$$

Step 2: Set the parameters for the level set formulation and energy minimization:

$$\text{Number of iterations} = 30 - 60, \mu = 1, \text{Time step} = dt = 0.1$$

$$\nu = 10^{-3}255^2, \lambda_1 = \lambda_2 = 1, \varepsilon = 1$$

The above settings are defined based on best practices as reported in literature for natural images and kept the same for all the images under evaluation [49, 72, 154-157, 158, 163].

Step 3: Evolve the level set function Φ according to equation 11. At each time step, the level set function is reinitialized to be the signed distance function to its zero level set curves [164], the constants c_1, c_2 and d_1, d_2 are updated according to equations in 10.

Step 4: Extract the zero level set from $\Phi_t(x, y)$ if the level set evolution terminates.

Step 5: Spatial morphological operations using a structuring element are applied on the segmentation mask. The process aims to remove small areas that are not part of a building, as well to create better-structured buildings. The same approach discussed in section 6.3.1 for designing the morphological filter and specifying its parameters is followed.

The suggested optimized active contour (OAC) segmentation algorithm for building structures detection is presented in figure 18 and all the images derived from the different processing steps are shown in figure 19.

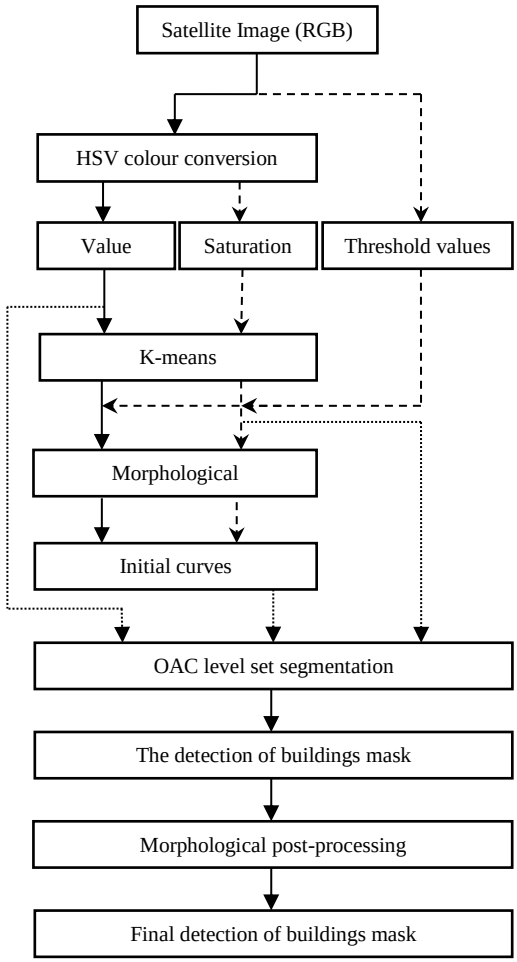


Figure 18. Detection of buildings algorithm flowchart

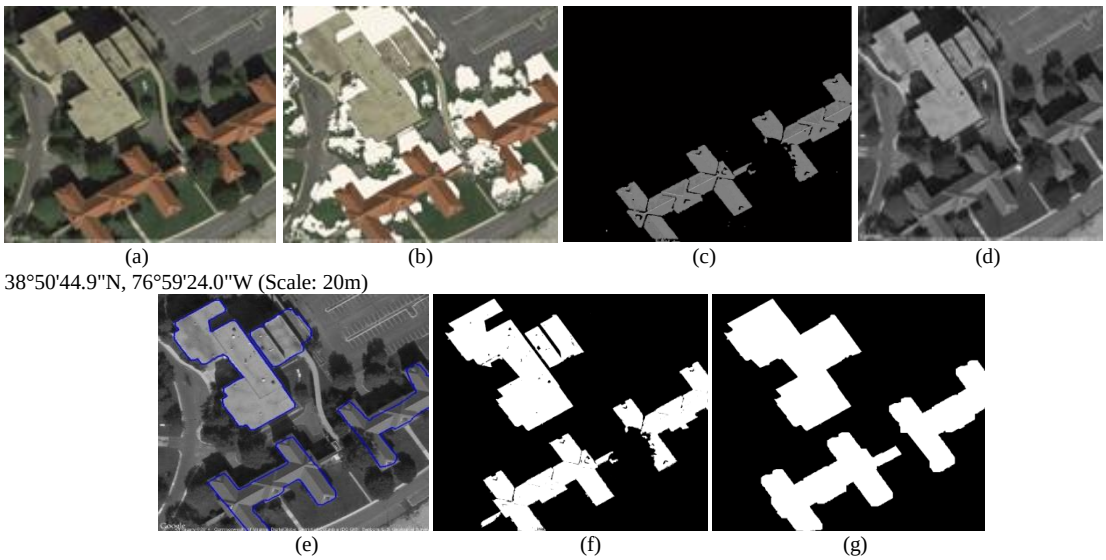


Figure 19: All stages for building detection in satellite images

(a) Original image, (b) Clustering image using colour features, (c) Saturation image, (d) Value image, (e) Active contour's initialization curves (f) Optimized active contour building segmentation mask, (g) Final building segmentation mask after morphological operations.

6.5 Experimental results and discussion

The algorithm is evaluated on 96 randomly chosen satellite Google Earth and or Google maps, images, following the same principles as was described in section 3.6.1 and 3.6.2 for the true detection and false detection of buildings. The proposed model was implemented using the same hardware and software as described in section 3.6.2. The traditional CV model has been also implemented for comparing its performance against the proposed model. The spatial resolution of the selected images varies from 1 m to 10 m. All the tested images are 8-bit images with pixel values ranging from 0 to 255 and contain three bands (RGB). They were selected to represent diverse building attributes such as size, shape and colour in challenging environmental conditions. The images that were used contained 5063 buildings in total. Table 5 presents the object-based performance of initial curves coverage. Figure 20, Figure 21, Table 6 and Table 7 illustrate the visual and quantitative object and pixel based building structure detection evaluation results. The time needed by the traditional active contour and the proposed model to detect all the corresponding buildings per image was also recorded. The average time in seconds using all the tested images is presented in Table 8.

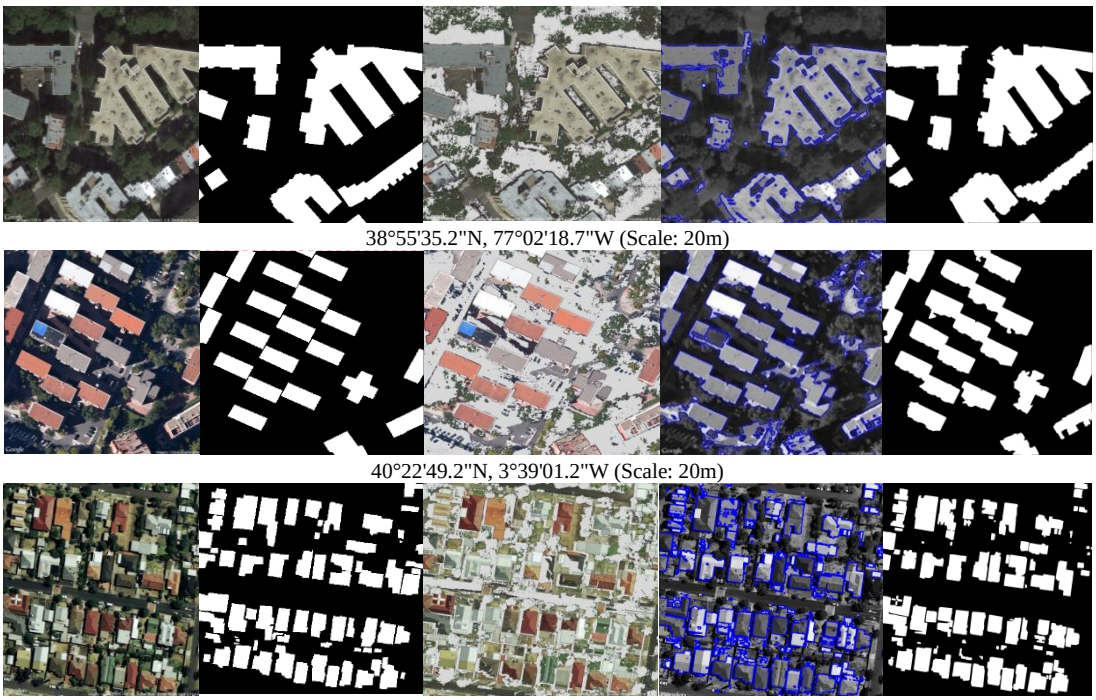
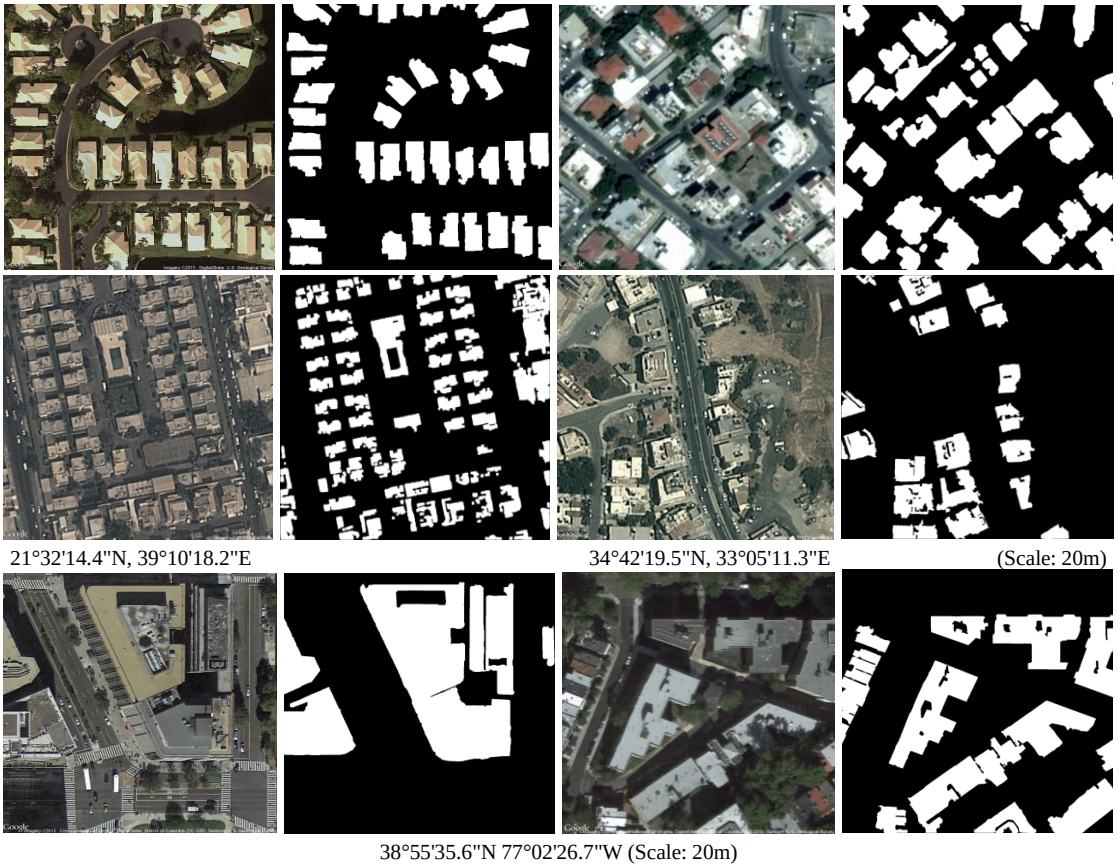


Figure 20: Building detection in satellite images

First column: Original images, Second column: Ground truth images, Third column: Clustering images using colour features, Forth column: Active contour building boundaries detection, Fifth column: Building segmentation final masks using the proposed method. First, second, third row: low, semi and high-density urban areas respectively.



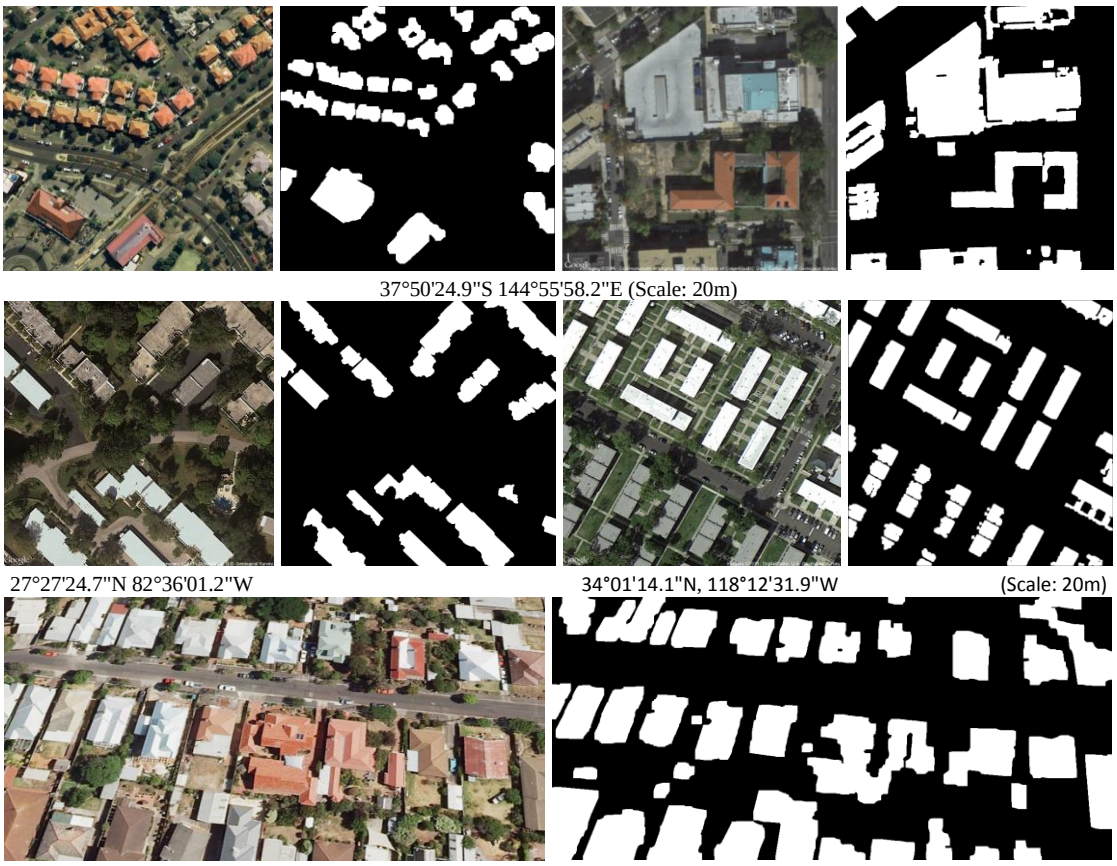


Figure 21: Building detection in satellite images

First, Third column: Original images, Second, Fourth column: Building segmentation masks using the propose method.

Table 5: Object-based initial curves coverage evaluation

Urban Area Type	**Total Number of Buildings**	**Precision (%)**	**Recall (%)**	**F (%)**
High density	2898	95.31	98.39	96.83
Semi density	1728	93.22	94.83	94.02
Low density	437	91.43	94.12	92.75
Overall	5063	93.86	96.40	95.11

Table 6: Object-based building structure detection evaluation using 60% overlap threshold

Urban Area Type	**Total Number of Buildings**	**Precision (%)**		**Recall (%)**		**F (%)**	
		CV	***OAC***	***CV***	***OAC***	***CV***	***OAC***
High density	2898	78.72	95.06	71.60	90.37	74.99	92.66
Semi density	1728	68.60	90.39	69.27	89.81	68.93	90.10
Low density	437	67.56	88.24	69.11	85.81	68.33	87.01
Overall	5063	74.03	92.83	70.59	89.79	72.27	91.29

Table 7: Pixel-based building structure detection evaluation

Dense Urban Area Type	**Total Number of Images**	**Precision (%)**	**Recall (%)**	**F (%)**
		OAC	***OAC***	***OAC***
High density	32	92.94	88.12	90.47
Semi density	32	88.05	86.02	87.02
Low density	32	82.33	84.42	83.36
Average		87.77	86.19	86.95

Table 8: Computational cost evaluation

Method	**Computation Cost Average (Seconds)**
CV	17.42
OAC	12.96

In several cases, as can be seen in figure 1, where multiple buildings with different colour intensity values exist in satellite images, darker than the average building roofs, are not usually detected using the traditional level set CV segmentation method applied in the grey-scale representation of the image. In this work, it was observed that a saturation image is able to highlight building features that are presented with at least a high intensity colour value in one of the RGB bands. This was expected since the pure colour surfaces have high saturation representation [149, 162]. However, in the case of vegetation regions and/or shadows, these were also highlighted in several scenes on the saturation representation. By utilizing the intensity properties of the saturation representation of the image after eliminating the vegetation areas and shadows in the proposed active contour segmentation model that also encompass a well-defined initialization process, several buildings that were missed by the traditional CV active contour model were detected and extracted in most of the cases. In this work, along with the general initialization *k-means++* scheme that has been utilised, additional steps such as threshold and morphological filters have been incorporated for refining further the results of the k-means clustering method . The threshold values T_V and T_S of the two criterion functions as presented in equations (6.5) have been estimated using a training process. Samples of shadow and vegetation regions were used to define the range of the colour intensity values. It has been observed that these values vary for different captured areas. However, the estimation of the most appropriate values is not a very difficult task since in Google Earth images, large areas are presented with similar radiometric characteristics. Thus, the range of the threshold values can be estimated only once using a small number of representative images for each area. The median value of the observed range for each area is calculated and used as threshold value in the corresponding area. As can be seen in figure 17 the

proposed clustering method for eliminating shadows and vegetation regions works well in various scenes. Furthermore, buildings with green or variations of green roofs are not affected when the elimination of vegetation and shadow regions process is applied (figure 17, image a). This can be explained since the radiometric characteristics of green artificial objects as presented in satellite images are different from the radiometric characteristics of vegetation regions. Nevertheless, some buildings are not detected when they have very similar radiometric characteristics with shadows. This type of building structures are expected to be eliminated as shadows. However, this drawback is limited since only those buildings that are eliminated as shadows, are finally missed. The visual interpolation of all figures suggests that this type of building roofs is very rare in the evaluation scenes. As can be seen in figure six row three, building structures with similar radiometric characteristics with shadows are successfully detected if a small region of the corresponding building roofs still exists in the cluster image. The proposed active contour model manages to detect such building structures by utilising more iterations in the curve propagation process ($N \approx 60$), since initial curves are created for these buildings by the proposed initialization scheme.

The segmentation results of the presented method demonstrate that the building detection algorithm works well with satellite images of urban areas. Results presented in all figures suggest that the proposed methodology can detect almost all the buildings and their basic structure and shape, even when these buildings have irregular shapes, interact with environmental objects and occlude each other. In addition, the proposed method discards most of the non-building areas with similar building characteristics and any misleading artefacts. Applying optimized morphological post-processing functions ensures that the building shapes are retained.

The algorithm's overall per object segmentation results, reach a building detection accuracy or recall of 89.79%, precision 92.83 % and a radio F equal to 91.29% as can be seen in Table 6. The values in Table 6 correspond to 60% overlapping threshold. This value has been selected following best practices used in other studies in literature [6,87-89]. The algorithm performs well in all the predefined density urban area types. The pixel-based evaluation results in Table 7, are in line with the object-based results. The recall metric estimated at 86.19%, the precision at 87.77% and the F score at 86.95%. All the results reveal that the performance of the algorithm decreases as building density decreases. This was expected since active contour models based on CV formulation detect regions of interest that exhibit a certain similarity and homogeneity. In high dense urban areas, the building structures are presented with greater homogeneity and similarity. In

addition, the results revealed that the performance of the proposed method is superior when compared against the traditional CV model. The computational cost is an important factor in practical applications of the method and it was found that for proposed method it is significantly faster than the CV model (Table 8). After the application of the proposed initialization scheme, less iteration compared to the traditional CV model are required to detect the building structures. As it can be seen in figure 16, the initial contours are near the boundaries of the building structures. The proposed initialization scheme also enforces the active contours to avoid several misleading regions during the propagation process, leading to better segmentation. The coverage of the buildings as presented in the analysed scenes using the proposed initialization method has been thoroughly assessed. The aim of the initialization development curves process is to create curves of any form and structure, within or near the building objects. Thus, in this assessment if any part of the initial curve coincides with a building in the ground truth image, is considered as a TP. An initial curve is denoted as a FP if does not coincide with any of the building objects as presented in the ground truth image. A FN object is denoted a building object presented in the ground truth image where an initial curve has not been developed. As it can be seen from Table 5, the overall recall accuracy has been calculated at 96.40%, the precision at 93.86% and the F score at 95.11%.

Finally, it should also be noticed that if the image does not adhere to quality standards the resulting segmentation might suffer, as is the case with most of the building segmentation algorithms. However, the suggested algorithm has been shown to work well even in such cases, as is shown in Figure 21 and the presented results for the images in row two.

A limitation of the proposed method that should be noticed is that in high-density urban environments, some non-building objects such as bridges and parking places or even parts of the roads are classified as buildings. This behaviour is expected and is noticed in other schemes proposed in literature [87-89, 146, 154-158], because of the similarity between these objects and the buildings. Finally, the results of this work revealed that in some cases where two or more buildings are extremely close to each other, they are classified as a single building. This influences the object based results as reported in Table II for the TPs buildings where two or more merged buildings based on the 60% coverage threshold are classified as two or more TPs even though one of them might has less than 60% TP pixels. This has been observed mainly in high-density areas where very few buildings, less than 1.5% of all the existing buildings in all the tested images, were affected. This percentage can be thought as the highest margin of error of the object-based results.

In this work, extensive analysis of those parameters that might affect the building detection results was performed. One parameter, which is important for the object-based measures, is the object-overlapping threshold of 60%. Figure 22 presents this relationship and as it can be seen, even for a high overlap threshold, such as 80% and more, the proposed method achieves satisfactory detection performance where the accuracy of precision, recall and F score estimated near 90%, 85% and 87% respectively.

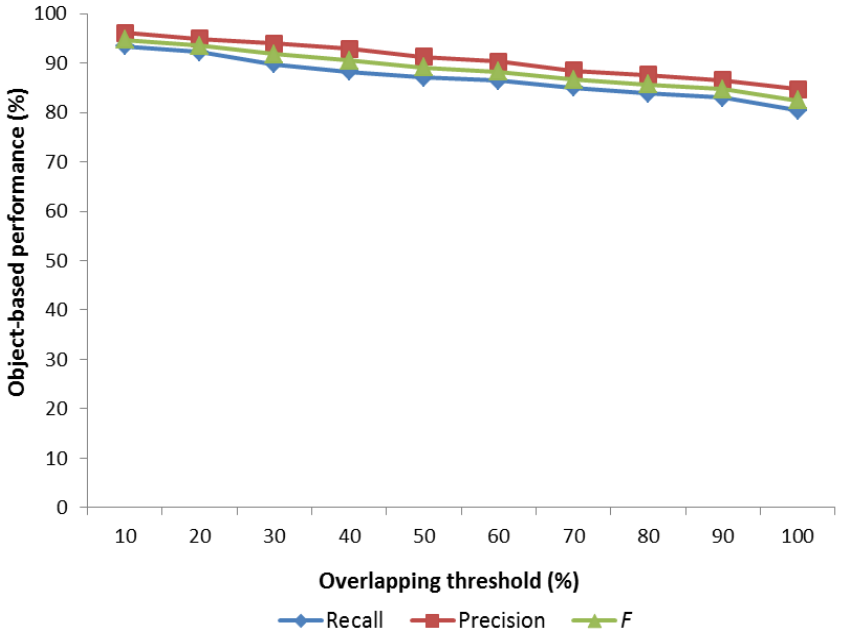


Figure 22: Object-based performance of the presented method for different overlapping threshold values.

Another important factor for the overall performance of the proposed method is the automated estimation of the structuring element size used in erosion and dilation morphological operations. It is expected that if the structuring element is too large or too small the results will be affected, leading to under-segmentation or over-segmentation respectively. Figure 23 presents how the pixel based performance metrics are affected when the size of the structuring element is increased or decreased by a percentage in both directions height and width. The 0% corresponds to the size of the SE as has been estimated in a given image by the proposed method and 100% increase indicates a double size SE in both directions while -100% indicates a half size SE in both directions. As can be seen in the corresponding figure, even for high increase or decrease on the structuring element estimated size, the results are satisfactory. This leads to the conclusion that the method used for calculating the structuring element parameters, works well in the building detection process.

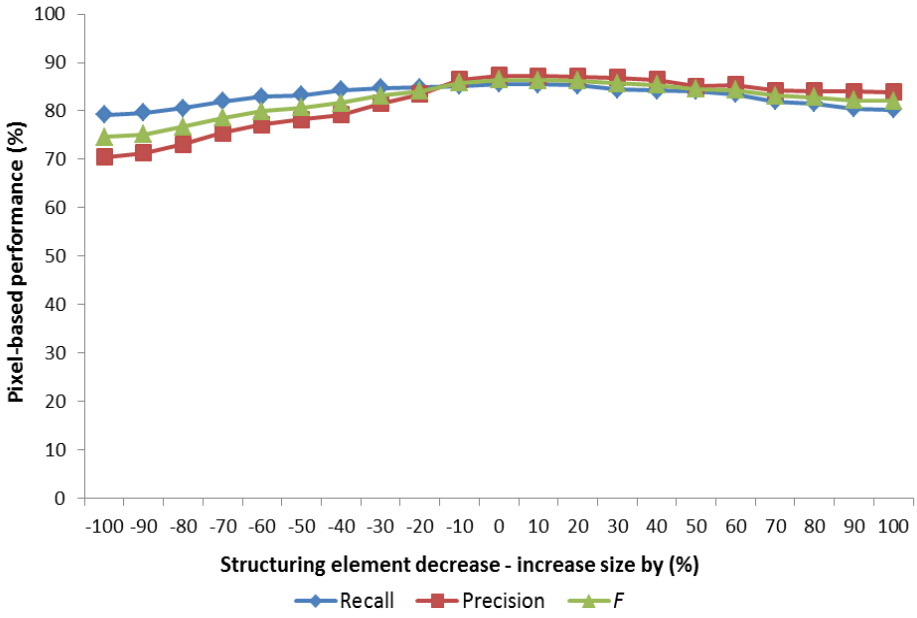


Figure 23: Pixel-based performance of the presented method if the structuring element size is increased or decreased.

6.6 Conclusions

The detection of buildings from monocular satellite images using an optimized active contour model is proposed. Active contour models are highly dependent on the spectral homogeneity of the objects in an image and may often lead to erroneous segmentation of satellite images where spectral heterogeneity and complexity exist. The proposed method works efficiently and has the ability to detect buildings with arbitrary shapes, sizes and even with significantly different coloured rooftops. It addresses limitations of the traditional active contour without edges model, for building detection and optimizes the overall segmentation procedure. A colour clustering method has been developed and used to remove the vegetation areas and shadows resulting in less complex scenes. The image is decomposed in HSV colour components and the intensity properties of the saturation and value representations are used to develop the optimized active contour model. An additional energy term has been encoded for biasing the active contours to delineate several buildings that were represented with lower intensity values than the average of the buildings in the grey-scale representation of the image, and often missed by the traditional active contour models. The proposed automated active contour initialization process is simple and drives the contours to a better and faster segmentation of the image avoiding misleading regions. Post-processing operations for retaining the building's structure are also applied. The estimation of the parameters used in the morphological erosion and dilation operations proved to be appropriate for retaining the shape of the detected buildings and remove any small misleading artefacts. Extensive experiments were

performed on a large number of images and the produced segmentation results for heterogeneous and complex satellite images verify the efficiency of the suggested method. The proposed method achieves better accuracy when compared to traditional active contour without edges segmentation model. Furthermore, the method is time efficient, there is no dependency on any additional data and user interaction is not required.

Chapter 7: Satellite Images Analysis for Shadow Detection and Building Height Estimation

7.1 Introduction

Building shadows can be used for estimating the structure and height of the corresponding buildings. This information can be utilized in many telecommunication or remote sensing applications such as natural disaster monitoring, urban change detection, urban scene reconstruction, cartography update, urban inventory and wireless network planning. For a successful satellite image analysis and height estimation of building structures, the complexity and spectral heterogeneity that often exist in urban landscape scenes has to be handled properly and shadows should be accurately detected and isolated.

Shadows are usually cast when objects totally or partially occlude direct light from a source of illumination. In remote sensing and telecommunication literature, various methods have been used for shadow detection using satellite images. These are often classified into two categories, the property-based and the model-based [4, 162]. In property-based methods, spectral and spatial features of shadow regions are utilized [165-170]. Model-based methods make use of additional metadata or scene information, such as sensor localization data, light source direction and objects geometry. This is typically priori knowledge [171-173].

The work presented in this chapter involves with building shadows detection through single monocular images in the presence of spectral heterogeneity and feature complexity. An automated segmentation method for delineating building shadows and estimating the corresponding building heights as presented in Google Earth satellite images using active contours is proposed. Applying a region-based active contour model for selectively detecting building shadows is difficult since several regions or objects are presented with similar intensities or characteristics as building shadows. Furthermore, spectral or intensity heterogeneity that often exists in shadow regions leads to erroneous segmentations. Thus, a pre-processing method of the satellite image is applied and a customized filter is developed for enhancing the shadows and reducing intensity heterogeneity. To achieve this, spectral and spatial analysis of the satellite image is performed. An automated scheme for selecting the most appropriate parameters and designing the optimized filter is proposed. The radiometric property of shadows is automatically estimated and encoded with the proposed filter in an active contour model. A new energy term able to bias the contours to detect shadow regions is formed. Morphological and variance threshold operations are applied for

eliminating misleading non-building shadow regions from the shadow segmentation mask. Finally, the detected building shadows are analyzed further for estimating the heights of the corresponding buildings. The shadow length and the predefined solar elevation angle are used for calculating the height of the presented buildings by applying simple trigonometric functions. In those cases where the solar elevation angle is unknown, indices from the building shadow segmentation mask are used for estimating it.

7.2 Related work

Since this work attempts to detect building shadows using a single monocular image, where additional metadata information is often unavailable, the discussion of the previous studies will be focused on property-based models. Furthermore, in this study the detected building shadows are used for estimating the corresponding building heights, thus, a review of methods estimating the building heights is also conducted.

Methods based on histogram threshold techniques have been extensively investigated for shadow detection because of their simplicity. Nagao and Matsuyama used a simple linear thresholding technique to guide the detection of shadow regions as presented in colour infrared aerial images [165]. Dare used a bimodal histogram splitting method for selecting the most appropriate threshold value to separate satellite images into shadow and non-shadow regions [166]. Chen et al. analysed the number of peaks and valleys of image histograms to estimate the threshold value for able to segment the existing shadows in the analysed scene [167]. Sırmacek and Unsalan after applying a median filter the local minimum in the grayscale histogram was utilised as the threshold value to detect shadow segments automatically [150]. A drawback of thresholding methods that needs to be resolved for successful implementations is the selection of the most suitable threshold value able to distinguish building shadow regions from other dark regions that may exist in the analysed scene. Thus, several researchers utilized invariant colour models towards this direction. A shadow detection method based on the image grey characteristics and threshold segmentation has been applied by Ye and Xu. Shadows on urban aerial true colour and colour infrared images were detected by using the HSI (H-hue, S-saturation, I-intensity) representation of the image [168]. A Region-based algorithm has been proposed by Arévalo et al. to locate shadow areas in satellite images. Shadows appear with several similar properties. Region-based algorithms are able to explicitly delineate shadows according to these similarities. Arévalo et al., used colour invariants and edges to identify shadows in QuickBird images. The image was pre-processed and the HSV representation was utilized for detecting the edges of the image, and for identifying the initial seed points

in a semi-automatic region growing technique [162]. Tsai performed a detailed comparative study on shadow compensation of colour aerial images in invariant colour models, concluding that shadows have high value in hue channel, high saturation value in blue or violet channels and low intensity. Thus, a spectral ratio between hue and intensity has been developed and used to detect shadow regions [149]. Chung et al. extended and improved the spectral ratio of Tsai by changing the calculation of Hue in the original HSI colour model. Based on the modified ratio map a connected component process was performed to separate the pixels of the input image into shadow and non-shadow pixels [169]. Recently Elbakary and Iftekharuddin proposed a regional based active contour model for detecting the shadows of artificial buildings in high-resolution panchromatic satellite images [170]. A predefined radiometric value for describing the shadow regions has been encoded in the active contour model for forcing the propagating curve to delineate the dark regions of the image. Many of the studies as reported in literature have considered the misclassification of sunlit dark objects classified as shadows and shadowed bright objects classified as non-shadow regions, as an important drawback. Texture information has been utilized by Dare et al. [166] and Chen et al. [167] to overcome the dark regions misclassification limitation.

A few attempts have been reported in literature where shadow information is used for estimating building heights. Irvin and McKeown used shadows along with information about the solar elevation angle in aerial photographs to recover the shape and height of buildings [138]. Shettigara and Sumerling described a method for determining building heights based on their extended features and the shadows they cast [174]. The building shadows were also used by Cheng and Thiel [175], Turker and San [176], to calculate the height of buildings that collapsed in an earthquake.

7.3 The Development of an optimized segmentation model to automatically detect the shadow of buildings and estimate their height

Automatic shadow detection is a very important process for many remote sensing applications, particularly for images acquired with high spatial resolution. This work is motivated by the important advantages and wide application of the active contour models using regional descriptors for satellite image segmentation. In the following paragraphs the proposed method, which includes four major processes is discussed. In the first process a filter for shadow enhancements is developed. Subsequently an optimized active contour model for selectively detecting building shadows is formed. Post-processing operations for distinguishing the shadows cast by buildings and other dark regions are also applied.

Finally, building heights are estimated using information derived from the shadow segmentation mask and solar elevation angle.

7.3.1 Spectral and spatial image analysis for shadow enhancement

The aim of spectral and spatial analysis of the satellite images is to develop a filter able to enhance shadow regions and subsequently improve the overall performance of a building shadow detection process using an active contour segmentation model. With special consideration to the typical types of satellite images, a shadow enhancement method should result in improving the clarity and homogeneity of the shadow and non-shadow regions as presented in landscape urban scenes.

The spatial frequency transform is a widely used tool for image analysis. The image signal is represented in terms of magnitude and phase. The importance of spectral analysis in images has been noticed in a variety of applications. Usually the content of image signal is not stationary, thus the localized frequency analysis has become an important and powerful image processing approach. Moreover, it is advantageous to analyse the signal frequency and spatial information simultaneously using various tools such as Short Time Fourier Transform (STFT), Gabor Transform (GT) and Wavelet Transform [1-2]. Gabor filters were originally introduced by Dennis Gabor [177] and extended into two dimensions by Daugman [178-179]. Gabor filters are used extensively in image processing and computer vision because of their optimal localization properties in both spatial and frequency domains [180-184]. This work utilises both the local phase and magnitude response of a Gabor filter for enhancing shadows displayed in satellite images.

The Gabor filter is a two-dimensional filter formed by the combination of a cosine with a two-dimensional Gaussian function and it has the following general form:

$$g(x, y, \theta, f, \sigma_x, \sigma_y) = exp\left\{-\frac{1}{2}\left[\frac{{x_\theta}^2}{{\sigma_x}^2} + \frac{{y_\theta}^2}{{\sigma_y}^2}\right]\right\} \cdot \cos(2\pi f x_\theta),$$

$$x_\theta = x \cdot \cos\theta + y \cdot sin\theta, y_\theta = -x \cdot \ sin\theta + y \cdot cos\theta \tag{7.1}$$

where θ denotes the rotation of the filter related to the x-axis and f denotes the local frequency. The Gabor filter as presented above is centred at the origin and the standard deviations of the Gaussian function along the x-axis and y-axis are presented as σ_x and σ_y, respectively. Gabor filters are band-pass filters that have both frequency-selective and orientation-selective properties [179]. This means the filters can be effectively tuned to specific frequency and orientation values for enhancing the building shadows and reducing

intensity heterogeneity. Therefore, the regularity of spatial structure can be utilised by applying a Gabor filter that is tuned to match the local ridge orientation and frequency as presented in satellite images. Based on the local orientation and ridge frequency around each pixel, the Gabor filter is applied to each pixel location in the image. The effect is the increase of the ridges oriented in the local orientation direction and the decrease of anything oriented differently. Hence, the filter enhances the contrast between the foreground ridges and the background. Furthermore, it effectively reduces any existing noise.

A Gabor filter for enhancing the shadows as presented in satellite images can be designed by thoroughly investigating and subsequently selecting the most appropriate parameters of the corresponding function. The implementation of Gabor filters, follow mainly two directions, the filter design approach and the filter bank approach [185]. In the filter design methods the parameters are chosen by considering the available data and the expected result. In the filter bank methods the filter parameters are chosen in a data independent way and subsequently a subset of filters is selected for a particular application. In this work an automated optimization scheme is incorporated for selecting the most appropriate parameters in a filter design approach. The proposed procedure follows a scheme based on evolutionary computing and Genetic Algorithms (GAs). Genetic algorithms are robust optimization and search methods, based on natural selection principles. GAs were inspired by Darwin's theory about natural biological evolution and firstly introduced by John Holland in 1975 [186]. Since then, GA's have been extensively used for finding optimal solutions to various problems in computer vision [187]. The idea is to define a search space with all the feasible solutions for the investigated optimization problem and estimate the one that fits better, through an evolved process. The Genetic algorithm starts with a set of solutions represented by chromosomes called a population. Solutions from one population are taken and used to form a new mutated population called offspring through recombination or crossover in an effort to produce a better one. A fitness function is defined and the new solutions are selected according to their fitness in an iterative process until a predefined condition is satisfied.

A two region image with approximately piecewise constant distinct intensities, presenting shadow and non-shadow regions is required as a result, of the convolution of the implemented filter and the satellite image. To achieve this, the most appropriate values for the parameters of the Gabor filter function such as $\sigma_x, \sigma_y, \theta$ and f should be selected.

The development of the automated procedure for selecting filter parameters using the proposed scheme begins with the formulation of the objective function. For the creation of

a two region image, with approximately piecewise constant distinct intensities the distance between two points needs to be maximized. This is a differentiable and continuous function that can serve as an appropriated objective function and is implemented using the following equation:

$$d_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2]^{1/2} \; for \; i \neq j \tag{7.2}$$

where d_{ij} is the Euclidean distance between feature points $(x_i, y_i), (x_j, y_j)$ and the Gabor filter parameters will be optimized by maximizing the above objective function.

In image enhancement methods based on Gabor filters as reported in literature, quadratic masks of size ~10×10 pixels and choices of the standard deviation of the Gaussian $\sigma_x = \sigma_y \approx \text{mask size}/2$ or very similar are often used [1, 177-180]. The selection of these parameters involves the trade-off between an ineffective filter in the case of small values and the risk of creating artefacts or the loss of important edge information in the enhanced image if large values are specified. In this work the same principles are followed by specifying $\sigma_x = \sigma_y \approx 5$ without affecting the overall performance of the method. With the above settings only the parameters f and θ needs to be optimized with GA for maximizing the objective function. Following the best practice reported in literature, f and θ are defined as integers from the set of numbers 0, 2, 4, 8, 16, 32, 64 and 0, π/6, π/3, π/2, 3π/4, π respectively [179-185]. For the parameter sets coding, f and θ are represented as bit strings where initially a random population is taken and the genetic algorithm rules and operators are applied. The fittest solution is chosen down the generations and the objective function is maximized in each generation. Finally a stopping term is specified based on the highest intensity differences relative to the mean of the image. For each set of the generated parameters f_i, θ_i this difference is calculated and the process is terminated when the difference remains constant. Gabor filter is implemented using the equation (7.1) and utilised for shadow detection in an optimized region-based active contour segmentation model.

7.3.2 Active contour segmentation model

Based on the principle that the response of the convolution of the satellite image with the customized Gabor filter is an image that includes approximately two distinct regions presenting the shadows and non-shadows, the response of the filter is utilised in an active contour model for driving the evolving curves to delineate shadows. This idea derived by analysing the radiometric properties of the acquired satellite images following the

application of the customized filter. In this process, it was observed that the distinct shadow regions are approximately homogeneous and their radiometric property can be represented by a global constant. The proposed optimized active contour (OAC) model for building shadow detection using image I is defined as follows:

$$\begin{aligned} E(c_1, c_2, d_1, d_2, \Phi) &= \lambda_1 \int_{\Omega} (I(x,y) - rc_1)^2 H(\Phi)\, dx\, dy + \lambda_2 \int_{\Omega} (I(x,y) - c_2)^2 H(\Phi)\, dx\, dy \\ &+ \lambda_1 \int_{\Omega} (S(x,y) - d_1)^2 H(\Phi)\, dx\, dy + \lambda_2 \int_{\Omega} (S(x,y) - c_2)^2 H(\Phi)\, dx\, dy \\ &+ \mu \int_{\Omega} \delta\big(\Phi(x,y)\big)\big|\nabla\big(\Phi(x,y)\big)\big| \,\mathrm{dx\,dy} + \nu \int_{\Omega} H(\Phi(x,y))\,\mathrm{dx\,dy} \end{aligned} \tag{7.3}$$

where $H(\Phi)$ and $\delta(\Phi)$are the Heaviside and Dirac function respectively. The unknown constant responsible to characterize the radiometric property of the shadows is represented as r. The complementary image $S(x,y)$derived from the convolution of the optimal filter G with the original image is defined as follows:

$$S(x,y) = 1 - (G * I(x,y)) \tag{7.4}$$

Using the response of the proposed filter and the constant r, the active contours are bias to detect homogeneous dark regions such as the shadow of the buildings. In this work, the constant r is automatically estimated as the optimal threshold value after analysing the histogram of $S(x,y)$ using Otsu's threshold method [39].

The energy minimization of the model as presented in equation (7.3) with respect to Φ and C can be achieved using the traditional CV model. The level set, partial differential equation for the segmentation of a given image I based on the proposed model using the equation (7.3) and the principles as have been described in equations (3.11-3.13) becomes:

$$\begin{aligned} \frac{\partial \Phi(x,y,t)}{\partial t} &= \delta(\Phi)\left[-\lambda_1(E_1 + E_3) + \lambda_2(E_2 + E_4) + \nu\, div\left(\frac{\nabla\Phi}{|\nabla\Phi|}\right)\right] \\ &+ \mu\left[\nabla^2\Phi - div\left(\frac{\nabla\Phi}{|\nabla\Phi|}\right)\right] \end{aligned}$$

$$E_1 = \int |I(x,y) - rc_1|^2\, dx\, dy = (I)^2 - 2rc_1 I + (rc_1)^2$$

$$E_2 = \int |I(x,y) - c_2|^2\, dx\, dy = (I)^2 - 2c_2 I + (c_2)^2$$

$$E_3 = \int |\mathrm{S}(x,y) - d_1|^2 \, dx \, dy = (\mathrm{S})^2 - 2d_1 S + (d_1)^2$$

$$E_4 = \int |\mathrm{S}(x,y) - d_2|^2 \, dx \, dy = (S)^2 - 2d_2 S + (d_2)^2$$

(7.5)

The unknown constants c_1, c_2, d_1, d_2 are calculated as follows:

$$c_1(\Phi) = \frac{\int_\Omega I(x,y) H\big(\Phi(t,x,y)\big) \, dx \, dy}{\int_\Omega H\big(\Phi(t,x,y)\big) \, dx \, dy}, c_2(\Phi) = \frac{\int_\Omega I(x,y) \left(1 - H\big(\Phi(t,x,y)\big)\right) dx \, dy}{\int_\Omega 1 - H\big(\Phi(t,x,y)\big) \, dx \, dy}$$

$$d_1(\Phi) = \frac{\int \mathrm{S}(x,y) H\big(\Phi(t,x,y)\big) \, dx \, dy}{\int H\big(\Phi(t,x,y)\big) \, dx \, dy}, d_2(\Phi) = \frac{\int \mathrm{S}(x,y)(1 - H\big(\Phi(t,x,y)\big)) \, dx \, dy}{\int 1 - H\big(\Phi(x)\big) \, dx \, dy}$$

(7.6)

Finally, r is a global distinct constant and is estimated as:

$$r = 1 - (\max\{\sigma^2(t)\}), 1 \le t < L$$

(7.7)

where σ^2 as proposed in [39] is the between-class variance represented by:

$$\sigma^2 = q_1(\mu_1 - \mu_\tau)^2 + q_2(\mu_2 - \mu_\tau)^2$$

(7.8)

where $[1, L]$ is the range of grayscale levels of image $S(x,y)$. The intensity means for dividing the image into two classes are represented as μ_1, μ_2 and calculated using the following probability functions:

$$\mu_1 = \sum_{i=1}^{t} i \cdot p_i / \sum_{i=1}^{t} p_i, \qquad \mu_2 = \sum_{i=t+1}^{L} i \cdot p_i / \sum_{i=t+1}^{L} p_i$$

(7.9)

μ_τ is the global mean intensity of the image $S(x,y)$ and q_1, q_2 are the class probabilities, estimated as:

$$q_1 = \sum_{i=1}^{t} p_i, q_2 = \sum_{i=t+1}^{L} p_i$$

(7.10)

The parameters for the level set formulation of the energy minimization model are defined following the same practice as has been described in chapter 3 and are kept constant for all the images under evaluation ($Time\ step = dt = 0.1, \mu = 1, \nu = 10^{-3}255^2, \lambda_1 = \lambda_2 = 1, \varepsilon = 1$) [49, 72, 154-157, 158, 163]. The proposed model segments the given image using the forward finite differences scheme as has been described in equations (3.14-3.15).

7.3.3 Morphological and variance threshold operations

A difficulty in building shadow detection methods that should be resolved is water bodies that are often obtained as shadow regions since their radiometric characteristics are very similar with that of shadows [4, 166]. Thus, further steps are incorporated to remove these misleading regions from the shadow segmentation mask. A simple however effective method is the application of a region filter based on the variance of groups of pixels representing the true shadows and water bodies [166]. Many features are visible within the shadowed regions mainly because of secondary illumination thus the variances of shadowed regions are higher than the variances of water regions. A variance threshold value can be selected as a distinguishing factor in the comparison of shadowed regions and water body surfaces by examining the satellite image under evaluation. In this work a sample of shadow and water body regions has been examined by calculating their variance. It was found that the variances of water body regions are approximately within the range of 32-43 and the variances of the shadow regions are within the range of 68-130. Thus a refinement procedure is incorporated as a final step using a global variance threshold value ~ 45 to filter out the water body regions.

Moreover, the shadows cast by buildings and threes needs to be separated. Spatial morphological operations erosion and dilation are utilised for eliminating tree shadows that are considered as artefacts. The form and the parameters of the structuring element (SE), to be used in erosion and dilation, are automatically estimated following the same principles as presented in equation (6.4).

7.3.4 Overall shadow of buildings detection method

A building shadow segmentation method must result in an accurate delineation of the shadows, avoid misleading dark regions and provide a smooth outline for the boundaries of the detected regions. The suggested optimized segmentation model for building shadow detection is presented in Figure 24 and all the images derived from the different processing steps are shown in Figure 25.

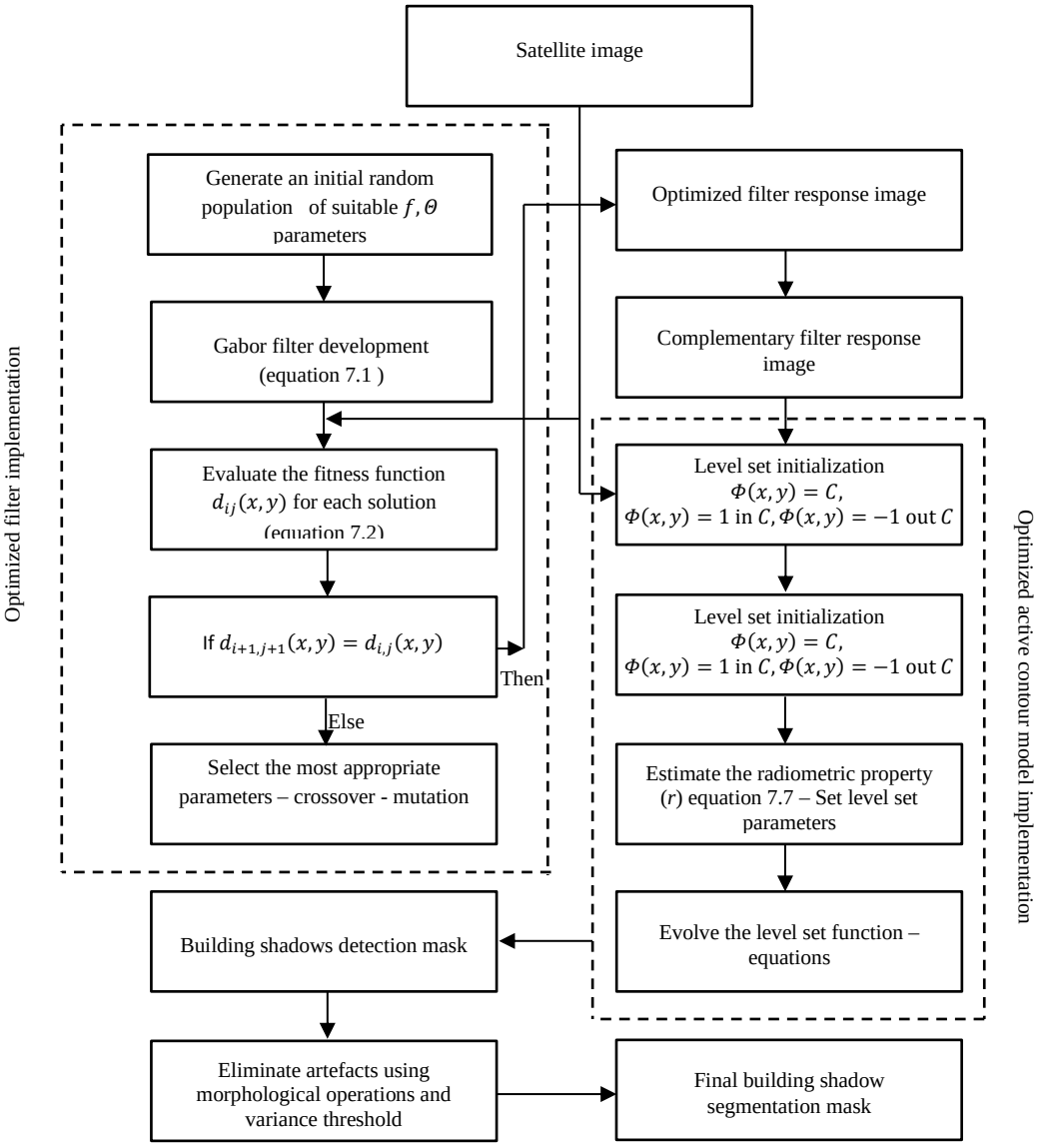


Figure 24: Optimized segmentation model for building shadow detection

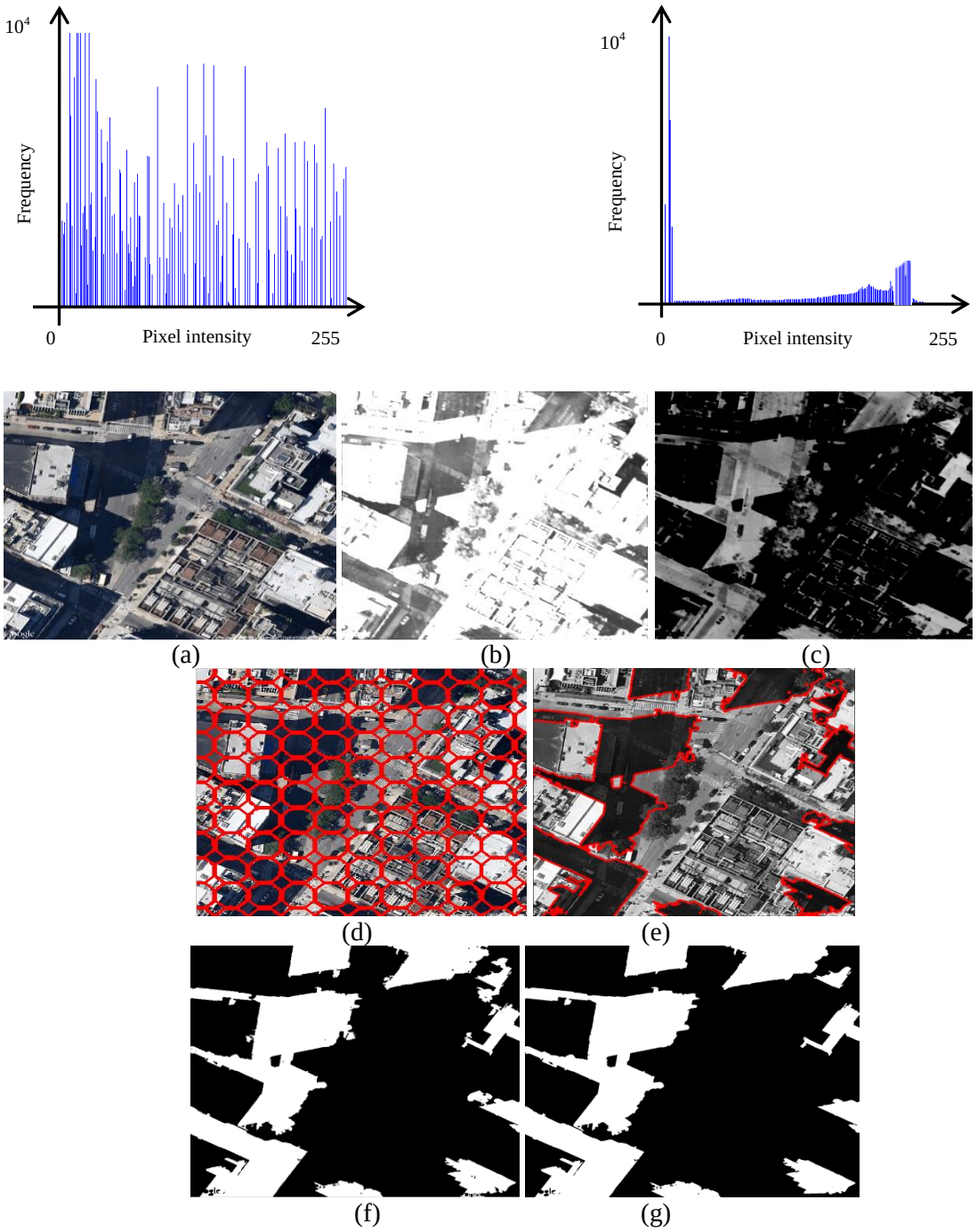


Figure 25: All stages for building shadow detection in satellite images

All stages for building shadow detection in satellite images (a) Original satellite image and intensity histogram of its grayscale representation, (b) Optimized filter response image, (c) Complementary filter response image and its intensity histogram, (d) Level set initialization, (e) Building shadows delineation using OAC, (f) Segmentation mask, (g) Post-processing final segmentation mask

7.3.5 Estimating the height of buildings

The building shadow mask is used to estimate the height of the corresponding building as presented in a landscape urban scene. To achieve this, information about the geometric relationship between the building and its shadow is required. The relationship is defined by the sun and satellite elevation angles against the shadow length. When a building is

illuminated by the sun, the whole or part of its shadow can be seen in the image based on the location of the sun and the location of the satellite. Figure 26 shows that there is a linear relationship between the shadow length and building height, which can be expressed by the following equation if the sun and satellite are on the same side:

$$h = sl \times \tan(a) \tag{7.11}$$

where h is the building height, sl is the shadow length and a is solar elevation angle [138, 172-173].

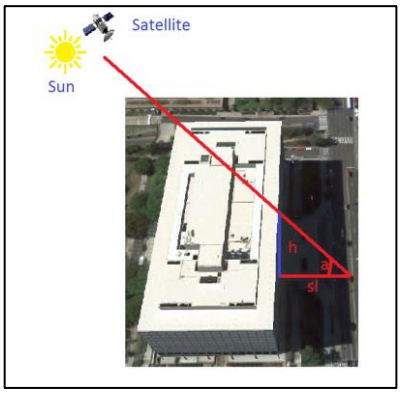


Figure 26: Sun, satellite and building geometry relationship

In this work, it was found that the buildings located among a limited area have the same projection characteristics if the satellite image was acquired at the same date and time. This means that the solar elevation angle is constant for a given satellite or remote sensing image within the specific limited area. As a result, the building height can be computed directly from equation (7.11) which has the critical variables a and sl. The shadow length sl is calculated using the shadow mask. The solar elevation angle a can be found from the direction of the sun, according to its azimuth at the date, time and place that the satellite image is acquired.

The overall proposed algorithm for building height estimation can be seen in Figure 27.

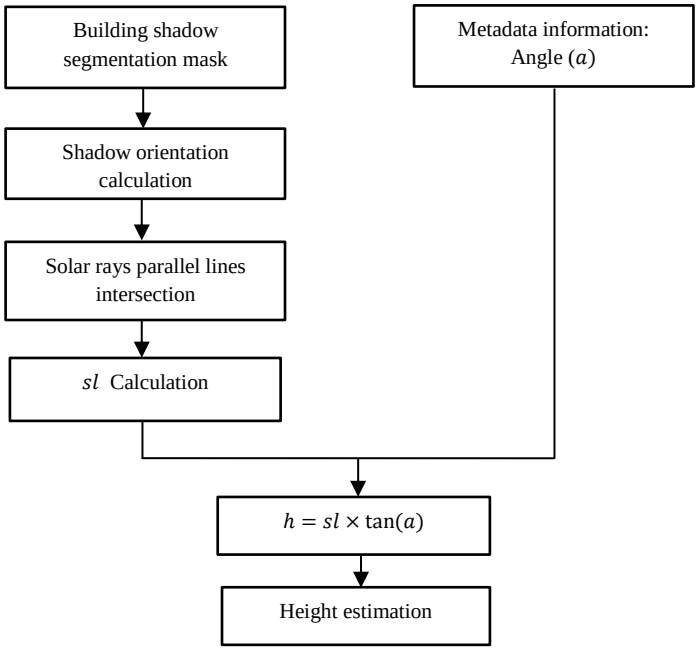
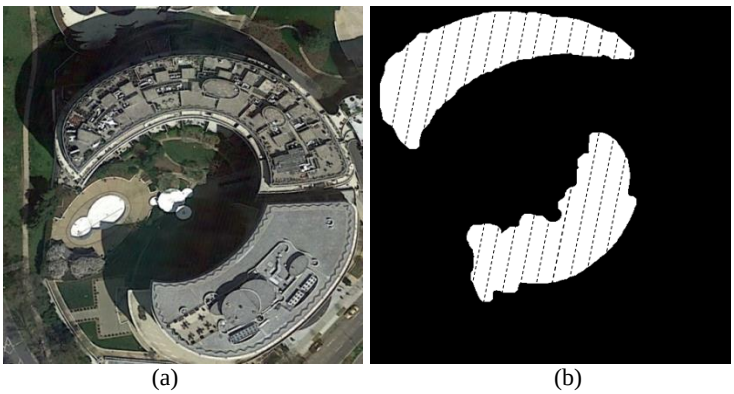


Figure 27: Building height estimation algorithm flowchart

The algorithm receives as input the shadow segmentation mask derived from the building shadow detection model and the solar elevation angle that can be found from the metadata information (date, time, location) of the acquired image. Information about the bounding box, the volume and the orientation of the shadow areas can be obtained and utilised in order to calculate the shadow length. The boundaries of shadow areas, as can be seen in figure 25, are quite irregular and this makes shadow length estimation more complicated. In this work, it was found that the shadow length sl can be defined if the shadow is intersected with solar rays. For the intersection of the solar rays with the shadows, the orientation of the individual shadows after extracting the corresponding bounding box is utilised. A series of parallel lines (figure 28) with a relatively small interval are designed and used to estimate the shadow length by obtaining their median length. Solar elevation angle (a) is also needed for calculating the height of buildings. In the cases where the date, time and place that the satellite image was acquired are known from the metadata properties of the image, the solar elevation angle is found according to this information. However, if the time of the acquired image is unknown, the orientation of the shadows can be used to estimate the sun azimuth and subsequently define the solar elevation angle. In these cases, some error is expected due to the difficulty to establish the exact relationship between the orientations of the various shadows presented in the processed scene with the azimuth of the sun. The expected error is evaluated in this work. Finally, by using solar elevation angle a and the shadow length sl estimation, equation 7.11 is applied and the height of the individual buildings is calculated.



(a) (b)

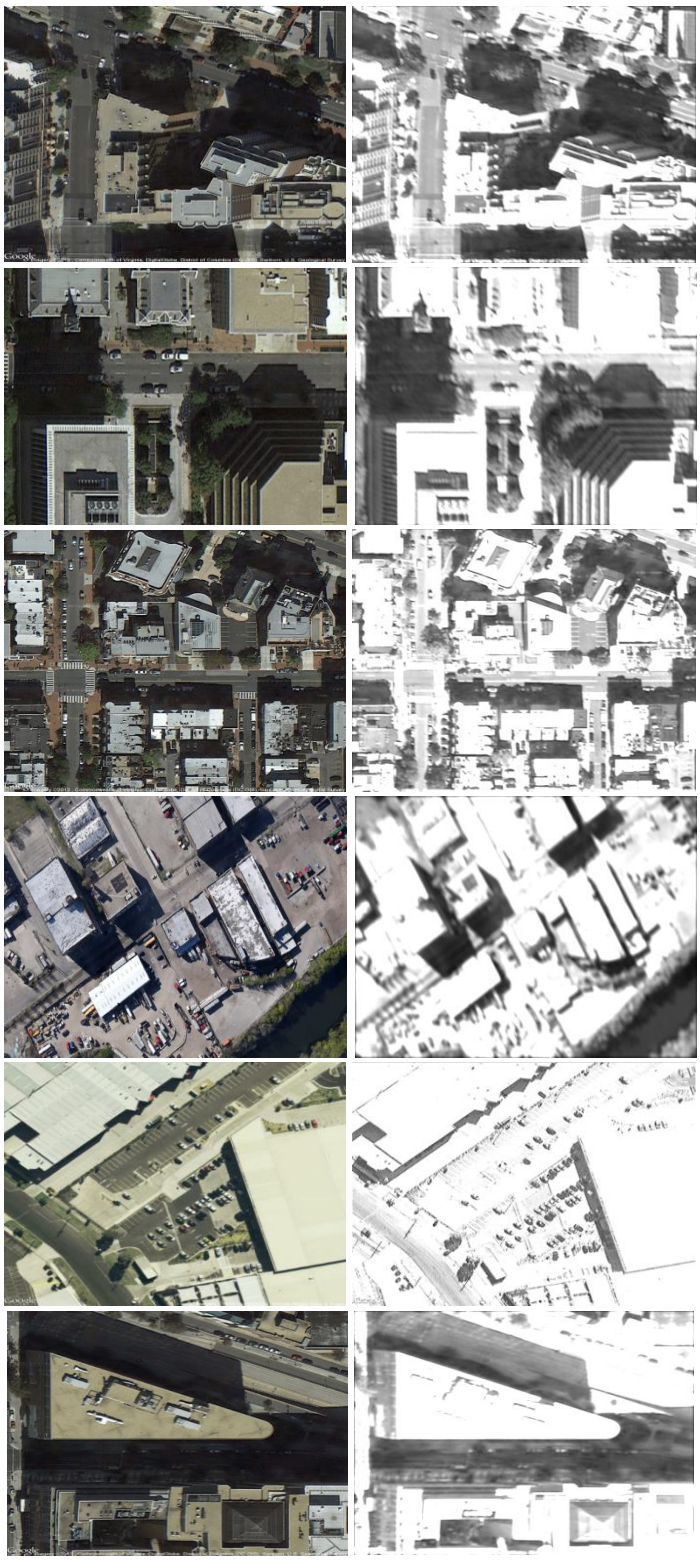
Figure 28: Solar ray intersection

(a) Original satellite RGB image, (b) Solar ray parallel line intersection using the shadow segmentation mask

7.4 Results and discussion

The algorithm is evaluated on 120 randomly chosen satellite Google Earth and/or Google maps, images (DigitalGlobe, 2013-2015) from different countries and different area types, using the same dataset as described in section 3.6.1 and following the same principles as presented in section 3.6.2 for the true detection and false detection of regions of interest.

Gabor filters can be used to enhance the shadows in satellite images. The optimization scheme that is utilised for finding the best suited Gabor filter parameters in order to enhance shadows and reduce the intensity heterogeneity proved to be very effective for the analysed images. This can be seen from the intensity histograms as presented in figure 25. The distribution of the intensity values as displayed in histogram figures highlight the effectiveness of the application of the proposed filter since the large distribution of the grayscale representation of the original satellite image is reduced in the complementary image. Moreover, the histogram of the complementary image clearly suggests that two major features exist after the application of the proposed filter, the shadow and non-shadow regions. Figure 29 presents more results of the convolution of satellite images with the customized Gabor filter, using the proposed scheme. The results noticeably suggest that the shadow areas in satellite images are effectively highlighted, without the loss of important information. Utilising the response of the customized Gabor filter in the proposed active contour segmentation model also proved to be very effective for shadow detection. The segmentation results demonstrate that the presented method works well in satellite images of urban landscape areas. Figure 30 illustrates the results of the presented method for detecting the building shadows for various scenes. Furthermore, in this study Dare's histogram threshold model [7] has been implemented and evaluated against the proposed model. Table I presents the quantitative pixel based evaluation results.



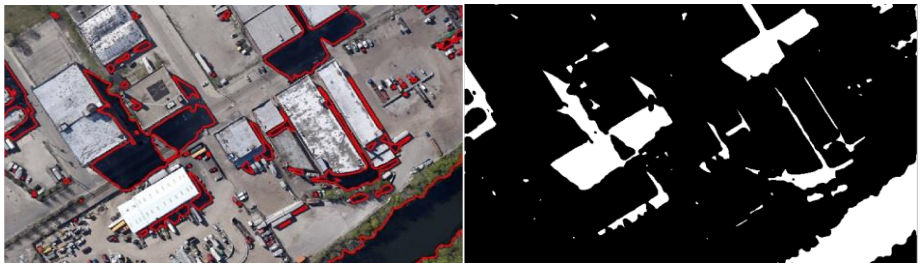
(a) (b)

Figure 29: Optimized Gabor filter response

(a) Original satellite RGB image, (b) Gabor feature image

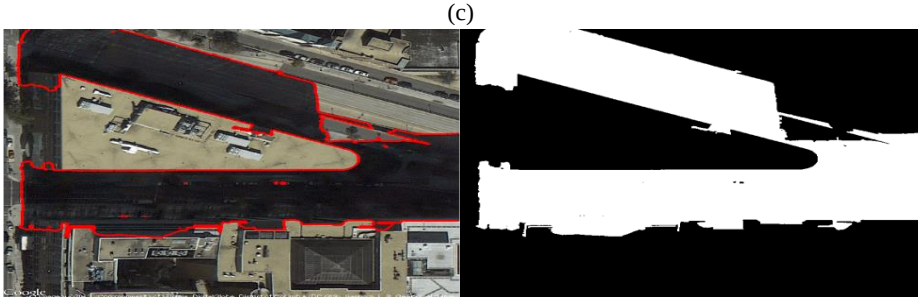
(a) (b)

(c)



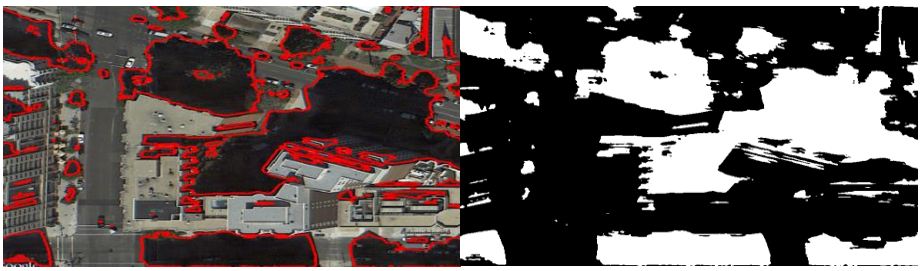
(a) (b)

(c)



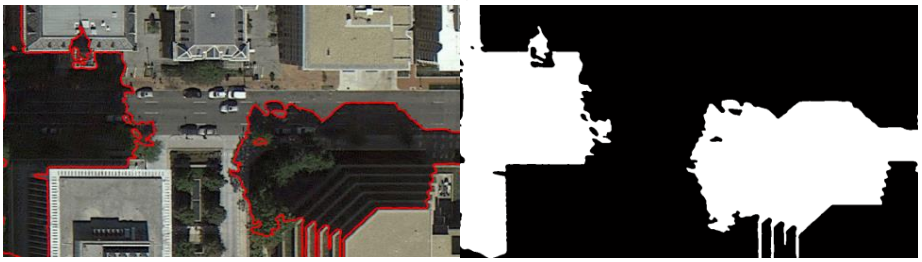
(a) (b)

(c)



(a) (b)

(c)

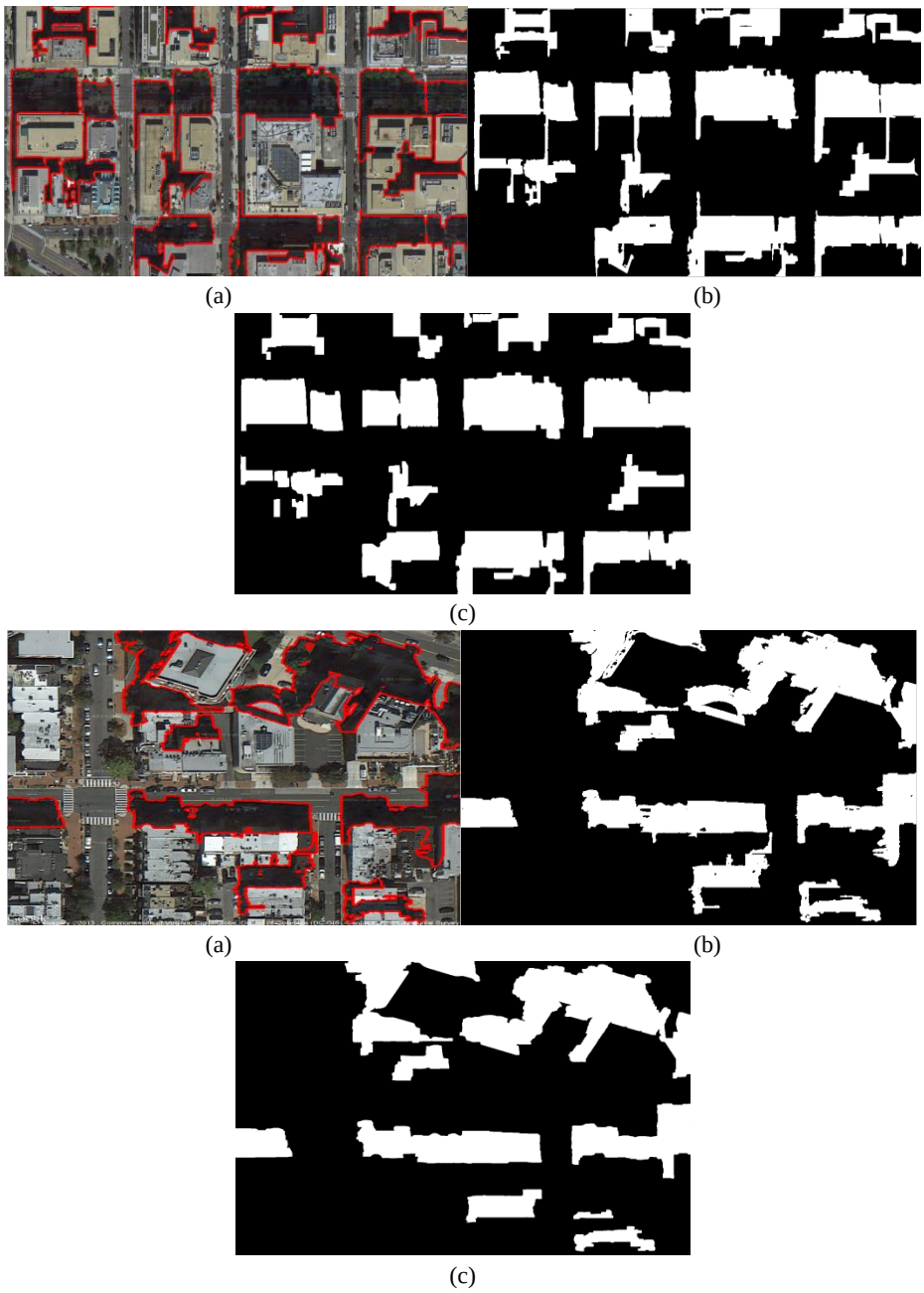


(a) (b)

(c)

(a) (b)

(c)



(a) (b)

(c)

(a) (b)

(c)

Figure 30: Buildings shadow detection in satellite images

(a) The shadow of building boundary delineation using the proposed active contour model, (b) The shadow of the buildings segmentation mask, (c) The shadow of buildings final segmentation mask

Table 9: Pixel based building shadow detection evaluation

Total number of images	**Precision (%)**		**Recall (%)**		**F (%)**	
	Dare's model	Proposed model	Dare's model	Proposed model	Dare's model	Proposed model
120	83.64	95.43	81.22	93.78	82.44	94.60

Results presented in all figures suggest that the proposed methodology can detect almost all of the building shadows, even when buildings have irregular shapes and interact with other objects from the surrounding environment. The proposed method also discards most of the non-building shadow areas with similar characteristics and any misleading artefacts.

The segmentation results reach a building shadow recall accuracy of 93.78%, a precision accuracy of 95.43% and F score of 94.60%. All the acquired results compare favourably to other results in literature, in the cases where similar quantitative results are available [4]. The highest recall and precision accuracy as presented in [170] was 91.29% and 86.64% respectively. The highest recall accuracy as reported by Arevalo et al. in [162] was 90.41%, while the value of the highest precision accuracy was 90.94%. Furthermore, the results revealed that the performance of the proposed method is superior when compared against Dare's histogram bimodal threshold model.

In this work extensive analysis of the radiometric property value, as automatically estimated using the proposed threshold scheme and its effect to the building shadow detection results has been performed. Figure 31 presents this relationship after a set of trials that were conducted involving the application of the proposed method for a range of values near the estimated radiometric property.

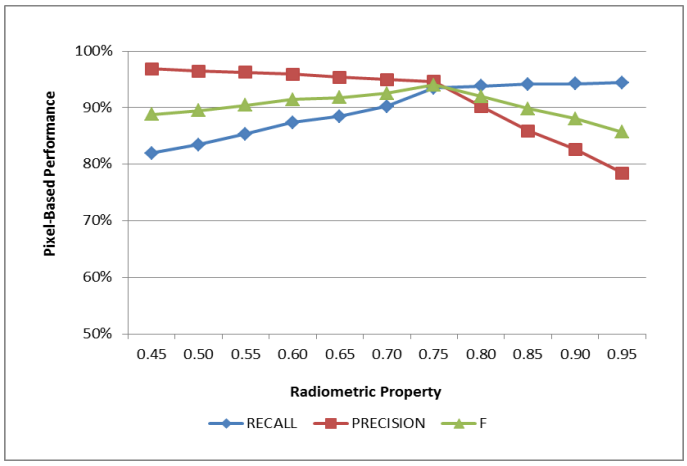


Figure 31: Radiometric property sensitivity analysis

As can be seen from the above figure assigning a high value for the radiometric property leads to the presence of artefacts and cause over segmentation in the detection process while a low value decrease the recall metric and thus the detection accuracy of the method. However, in all cases the results are still satisfactory which proves that the proposed scheme for estimating the radiometric property is efficient.

The proposed method for building heights estimation is evaluated by selecting a number of buildings in different areas and countries where their actual height is known. Three different building types based on function, residential, office and mixture (residential/office) are investigated. The actual height of the buildings in each of the three categories has been compared to the height determined by the proposed algorithm. The variance along with the aggregated variance for each category was also calculated. A sample of the results using

known buildings is presented in Table 10. Table 11 presents the overall aggregate variance for the three different building types.

Table 10: Estimated building height evaluation for known buildings using the proposed method

Name or Address	Location	Function	Actual Building Height (m)	Estimated Height. (m)	Variation (%)
Alm. Tow.	Dubai	Office	360	343.20	- 4.67 %
Mazaya	Dubai	Office	180	174.00	- 3.33 %
HorizonT	Dubai	Office	190	187.50	- 1.31 %
City Center	Sidney	Office	206	193.40	- 6.11 %
1101 NY	Washing.	Resident	50	46.20	- 7.60 %
1625 Eye	Washing.	Resident	49	44.75	- 8.67 %
Hilton Ca.	Washing.	Resident	46	48.30	+ 5.00 %
Capit. V.	Washing.	Mixture	52	48.90	- 6.00 %
Tower 42	London	Mixture	183	192.30	+ 5.10 %
Church. Pl.	London	Office	125	121.50	- 2.25 %
Wo. Tr. C.	Genoa	Office	102	109.20	+ 7.06 %
Piacentini	Genoa	Mixture	108	110.30	+ 2.13 %
1010 Mass	Washing.	Mixture	48	47.70	- 0.70 %
NRA	Washing.	Office	51	53.00	+ 3.90 %
Watergate	Washing.	Office	46	44.50	- 3.26 %
	Aggregated variance:				**4.47 %**

Table 11: Estimated building height overall evaluation using the proposed method

Number of buildings	Function	Aggregate Variance
64	Office	3.95%
72	Resident	4.46 %
58	Mixture	5.22 %
Total variance:		**4.84 %**

In those cases where all the information was available, an aggregated height variance less than 4.5% has been achieved. The overall height estimation aggregate variance after analysing 198 buildings using the proposed algorithm was 4.84% for all the different building categories. The lack of information about the time of the acquired image was treated by using the orientation of the shadows as presented on the shadow mask for calculating the solar elevation angle. Metadata information is a critical aspect for the success of the proposed algorithm for estimating building heights. If the time when the analysed image was acquired is not available, then the azimuth of the sun and subsequently

the solar elevation angle are unknown parameters. The shadow orientations are closely related to the above values [173]. Using building shadow orientations calculated directly from the building shadow segmentation mask for defining the azimuth of the sun and then the solar elevation angle is expected to lead to some inaccuracies. In this work, experiments were performed in order to estimate this expected error. The under evaluation buildings were analysed and their height was estimated for different orientation values near the calculated value. After evaluating the results against the actual height of the buildings, it was found that there is a linear relationship between the expected building height error and the orientation error. One radian error during orientation calculation is expected to give ±3 meter error on building height estimation. Figure 32 presents graphically this relationship.

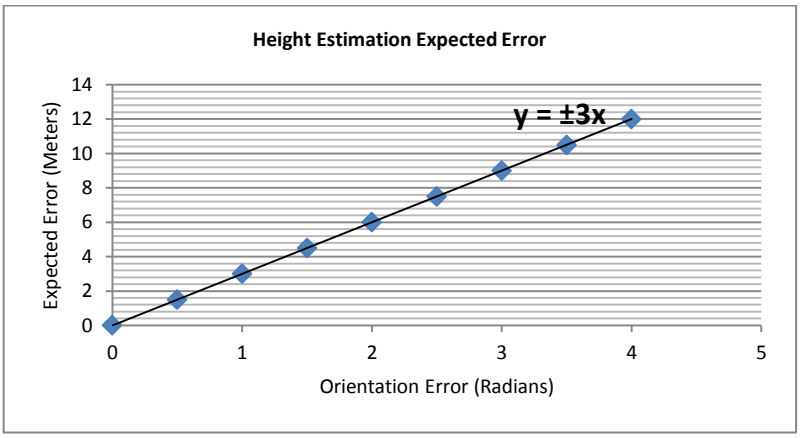


Figure 32. Height estimation, expected error

The proposed building height estimation approach is similar to the commonly used technique as initially presented by Irvin and McKeowen applied in aerial imagery and utilised by Shettigara and Sumerling in multispectral panchromatic images [138, 174]. However, in this work, high-resolution monocular Google Earth and/or Google map satellite images presenting diverse features from different areas or countries are evaluated and the automated estimation of building heights is emphasized. This can be thought as an important advantage when the analysis involves a variety of large study areas and large numbers of buildings. All the results revealed that the incorporated proposed scheme for the automated shadow length estimation is efficient. The scheme works well also in those cases where parts of the building shadows are covered by the corresponding buildings. In these cases, an overestimation of shadow length is expected as noticed in literature since the bases of the analysed buildings are unknown [173]. However, using the simulation of the solar rays intersected with the shadows and calculating the median value of the intersected lines for estimating the shadow length produced balanced results. In most of the examine cases the overestimation of the building shadow length is minimal. Furthermore, in this work a

thoroughly analysis has been performed for estimating inaccuracies derived from the lack of information on image acquisition time.

7.5 Conclusions

Building shadows in monocular satellite images can provide valuable information about the displayed landscape urban scene. Active contour segmentation methods can be used to detect these shadows. However, the performance of the traditional region-based active contour models is highly dependent on the intensity homogeneity or feature similarity of the regions of interest. Using traditional active contour models in the presence of intensity heterogeneity and feature complexity often leads to erroneous segmentation of shadows as presented in satellite images. The proposed algorithm optimizes the building shadow segmentation procedure by implementing a new active contour model, after utilising information derived from the spectral and spatial analysis of the satellite image. A filter able to reduce the effect of intensity heterogeneity and highlight the shadows has been designed. The most appropriate parameter settings of the filter are automatically estimated. The radiometric feature of the shadows is approximated using the response of the proposed filter, applied to the under evaluation image without any user interaction. The response of the filter along with the radiometric feature of the shadows is then utilised to form an optimum active contour model able to selectively delineate the existing shadows. Post-processing using morphological and variance thresholding operations is performed for eliminating any remaining artefacts. Extensive experiments were done and the produced segmentation results of urban satellite images verify that the suggested method is robust. The proposed approach deals very well, with complex-shaped buildings and avoids the possible disturbance of misleading information such as water-bodies, tree shadows and a number of other dark regions. Furthermore, by using the shadow mask for calculating the shadow length and by incorporating or estimating the solar elevation angle in the suggested algorithm, building height estimations were achieved with an aggregated variance less than 5%.

Chapter 8: An Optimized Localized Active Contour Segmentation Model with Application to Intensity Heterogeneous and Complex Images

8.1 Introduction

The last years, several methods have been proposed in literature to improve the image segmentation performance in the presence of intensity heterogeneity and feature complexity using active contour models [68-70, 80, 188-192]. Results produced by above models when applied on various images have been very promising. However, these attempts seem to produce good results usually in cases where intensity heterogeneity is a slowly varying field [193]. When there is arbitrary or sharp, intensity and colour variations, the above models can still yield erroneous segmentations [193]. The work presented in this chapter involves with the evaluation and optimization of the localized active contour models for image segmentation in the presence of arbitrary intensity heterogeneity and scene complexity. More specifically, an optimized localized active contour segmentation model (OLAC), addressing effectively the intensity heterogeneity of complex scenes that often occurs in real world applications such as mammography and remote sensing or satellite images is proposed. Developing a segmentation model that can be used effectively in both these areas is challenging since images acquired from these domains vary. The optimized localized active contour model proposes a new energy minimization model for image segmentation. The proposed model decomposes the image into two intrinsic components, the intensity heterogeneity and intensity homogeneity and a low pass filter is applied in local regions to transform the image and reduce the intensity heterogeneity. A kernel function is integrated to approximate the remaining heterogeneity artefact of the transformed image and produce the image segmentation. The OLAC segmentation model is evaluated on mammography images for breast boundary detection, and on satellite, Google Earth images for building boundaries detection. Furthermore, LIC [188, 194] and LSACM [190, 195] segmentation models were also applied on the same datasets for comparison purposes. To the best of the authors' knowledge, these models have not been extensively evaluated for breast boundary detection and/or building boundaries delineation.

8.2 Review of related work

Intensity heterogeneity can be thought as an attribute of the image. It can be defined as a smooth and spatially varying field that multiplies the constant representation of the original signal [196-197]. Thus, any given image in the presence of intensity heterogeneity and with zero mean Gaussian noise can be modelled as follows:

$$I(x) = b(x)I'(x) \tag{8.1}$$

where I represents the original signal, I' is the homogeneous representation of the original image, b represents the intensity heterogeneity. Intensity heterogeneity is a global feature of the image, but in smaller regions, it can be thought as a local constant [80-81, 188-190, 198, 225]. Therefore, the intensity homogeneous image I' is expected to approximately take n distinct constant values $d_i, \dots, d_n$ in n separated regions $\Omega_i, \dots, \Omega_n$ where the intensity heterogeneity field b varies slowly. In these separated regions, intensity heterogeneity can be approximated by a constant. The intensity heterogeneity field b and the constants d_i can be estimated using intensity criterion functions and their localized properties in terms of the regions Ω_i with respect to the neighbourhood denoted as y. These criterion functions can be incorporated in an active contour level set segmentation model for defining the different parts of the image domain and estimating the field that accounts for the intensity heterogeneity [80-81, 188-190, 197-198]. The above principles were used by Li et al. [188] for the development of the so-called local intensity clustering active contour model (LIC) as an extension of the LBF active contour model [53] as presented in equation 3.17 for overcoming difficulties derived from the presence of intensity heterogeneity. The image model as already described in equation (8.1), has been also utilized by Wang et al. [189] and Zhang et al. [190] to extend the CV model and improve its segmentation performance for images with intensity heterogeneity. An average filter in a local circular region has been used for estimating the intensity heterogeneity in Wang et al. active contour model [189]. The global term of the proposed model has been defined based on similarity measures using the Bhattacharrya coefficient. In a similar approach, Zhang et al. modelled the heterogeneous regions of interest as Gaussian distributions of different means and variances for the development of a localized statistical active contour model (LSACM). A sliding window is used to map the original image into another domain, where the intensity distribution of each object is still Gaussian but better separated. The means of the Gaussian distributions in the transformed domain is adaptively estimated by multiplying a bias field

with the original signal within the window. Maximum likelihood energy functional is then defined on the whole image region, which combines the bias field, the level set function and a piecewise constant function which approximates the true image signal [190].

8.3 The development of an optimized localized active contour level set segmentation model

Intensity heterogeneity in complex scenes where the number and placement order of elements in the visual stimulus makes their description difficult is defined as an inherent artefact caused by uneven illuminations. Since, the intensity heterogeneity is considered as a slowly varying field in the image domain, its spectrum in the frequency domain is expected to be concentrated in the low frequency area [196]. This work proposes the incorporation of a low pass convolution operator applied locally for correcting the presence of intensity heterogeneity by eliminating the low frequency artefact using a localized active contour, level set formulation framework. The proposed optimized localized active contour (OLAC) model based on the above principles, image representation (8.1) and the CV model as presented in equations 3.9-3.13 can be written as follows:

$$
\begin{aligned}
&E_{OLAC}(d_1, d_2, b, \Phi) \\
&\quad = \lambda_1 \int \int k_s(x,y)\,|(I(x) - l_s * I(x)) - b(y)d_1|^2\, H\big(\Phi(x)\big)\, dx\, dy \\
&\quad + \lambda_2 \int \int k_s(x,y)\,|(I(x) - l_s * (I(x)) - b(y)d_2|^2\, (1 - H\big(\Phi(x)\big))\, dx\, dy \\
&\quad + \nu \int \delta\big(\Phi(x)\big)\big|\nabla\big(\Phi(x)\big)\big|\, dx + \mu \int \frac{1}{2}(|\nabla\Phi(x)| - 1)^2\, dx
\end{aligned}
\tag{8.2}
$$

where l_s represents a low pass convolution filter in a window of a size $s \times s$ and $\nu, \mu, \lambda_1, \lambda_2$ are constants. The local intensity averages are denoted as d_1, d_2 and $b(y)$ represents the intensity heterogeneity, which may still exist in the transformed image $(I(x) - l_s * (I(x))$. The last two terms of the proposed equation are used for preserving the regularity of the level set function and the smoothing of the zero level set contours respectively [95]. The choice of the kernel function K_s is open and must satisfy the following criteria: $K_s(x,y) = 1,\ |y - x| \leq s,\ K_s(x,y) = 0\ |y - x| > s$. It is regarded as the indicator function of the local regions. A constant kernel function or a Gaussian kernel function can be used.

The performance of the model is expected to be highly depended on the ability of the convolution function l_s to eliminate or reduce the intensity heterogeneity, thus the selection and the optimization of this function is very important. The proposed model does not depend on the assumption that the intensity heterogeneity is expected to be well approximated by a constant parameter in its neighbourhood working with the original signal, but can be well approximated by a constant parameter in its neighbourhood after the application of the low pass filter convolution function in local regions. This is a distinct principle, if it is compared with the LIC the models presented in [188-190]. It should be also mentioned that the transformation of the image into the new domain is not a pre-processing step applied to image under evaluation, since it is performed in the local regions as defined by the window of a size $s \times s$ during the localized energy minimization process.

In this work, the well-known Butterworth and Gaussian low pass filters are utilized in correcting the presence of intensity heterogeneity in natural images such as satellite and mammography. Various kernels could also be examined. The low pass Butterworth function $h_B(u)$ and the low pass Gaussian function $h_G(u)$ are represented as follows:

$$h_B(u) = \frac{1}{1 + \left(\frac{u}{u_c}\right)^{2n}}, \qquad h_G(u) = e^{\frac{-\frac{1}{2}u^2}{u_c}} \tag{8.3}$$

where n is the filter order of the Butterworth function responsible for sharpening the result and u_c is the cut off frequency. Low pass filters can be constructed with arbitrary bandwidth and the bandwidth can be optimized to produce a filter with minimal spatial extent. The above parameters, along with the size of local regions are thoroughly investigated in this work. Initially the most appropriate filter for the image domains of interest must be selected. Subsequently, the most appropriate filter settings for implementing a filter able to estimate and correct the intensity heterogeneity under the proposed energy minimization model as presented in equation (8.2) must be defined.

The energy minimization with respect to Φ, C and b can be achieved using the traditional CV model and the level set methodology as has been described in chapter 3. The level set partial differential equation for the segmentation of a given image based on the proposed model becomes:

$$\frac{\partial \Phi(x,y,t)}{\partial t} = \delta(\Phi)\left[\lambda_2 E_2 - \lambda_1 E_1 + \nu\, div\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right)\right] + \mu\left[\nabla^2 \Phi - div\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right)\right]$$

$$E_1 = \int\int K_s(x,y)|(I(x) - l_sI(x)) - b(y)d_1|^2\,dx\,dy$$
$$= (I - l_sI)^2 * K_s - 2d_1(I - l_sI)(b * K_s) + (b^2 * K_s){d_1}^2$$
$$E_2 = \int\int K_s(x,y)|(I(x) - l_sI(x)) - b(y)d_2|^2\,dx\,dy$$
$$= (I - l_sI)^2 * K_s - 2d_2(I - l_sI)(b * K_s) + (b^2 * K_s){d_2}^2$$

(8.4)

where I is the given image, δ is the Delta function, $H$is the Heaviside function, $\mu \geq 0, \nu \geq 0, \lambda_1, \lambda_2 > 0$ are fixed parameters responsible to weight the different terms in the energy and d_1 , d_2 , b are unknown constants that are calculated as follows:

$$d_1(\Phi) = \frac{\int (b * K_s)(I - l_sI)H\big(\Phi(y)\big)\,dy}{\int (b^2 * K_s)H\big(\Phi(y)\big)\,dy}, d_2(\Phi) = \frac{\int (b * K_s)(I - l_sI)(1 - H\big(\Phi(y)\big))\,dy}{\int (b^2 * K_s)(1 - H\big(\Phi(y)\big))\,dy}$$

$$b = \frac{\big((I - l_sI)d_1H + (I - l_sI)d_2(1 - H)\big) * K_s}{(d_1^2H + d_2^2(1 - H)) * K_s}$$

(8.5)

8.3.1 Overall algorithm

A segmentation method must result in an accurate delineation of the regions of interest, avoid misleading artefacts, and provide a smooth outline for the boundaries of the detected objects. The algorithm steps of the proposed segmentation method are as follows:

Step 1: For a given image, the initial contour C is implemented and placed on the given image by initializing the level set function as follows:

$$\Phi(x,y) = C, \Phi(x,y) = 1\ inside\ C, \Phi(x,y) = -1\ outside\ C$$

Step 2: Select the low pass convolution function (l_s), set its parameters and develop the filter according to (8.6). Three parameters are defined as variables ($u_c = 0.1{:}\,0.5,\ n = 4{:}\,10,\ s = |17 \times 17|{:}\,|50 \times 50|$) and used for evaluating the performance of the function in correcting the presence of the intensity heterogeneity.

Step 3: Define the indicator function of the local regions K_s. A constant kernel is used.

Step 4: The following parameters for the level set formulation and energy minimization are set: $\mu = 1, Time\ step = dt = 0.1,\ \nu = 10^{-3}255^2$, $\varepsilon = 1,\ \lambda_1 = \lambda_2 = 1$. All the above parameters are kept constant following the same practice as has been described in chapter 3.

Step 5: Evolve the level set function Φ according to (8.4). At each time step, the level set function is reinitialized to be the signed distance function to its zero level set curves [53, 68], and the constants d_1 , d_2 and b are updated according to equations (8.5).

Step 6: Extract the zero level set from $\Phi_t(x, y)$ if the level set evolution terminates.

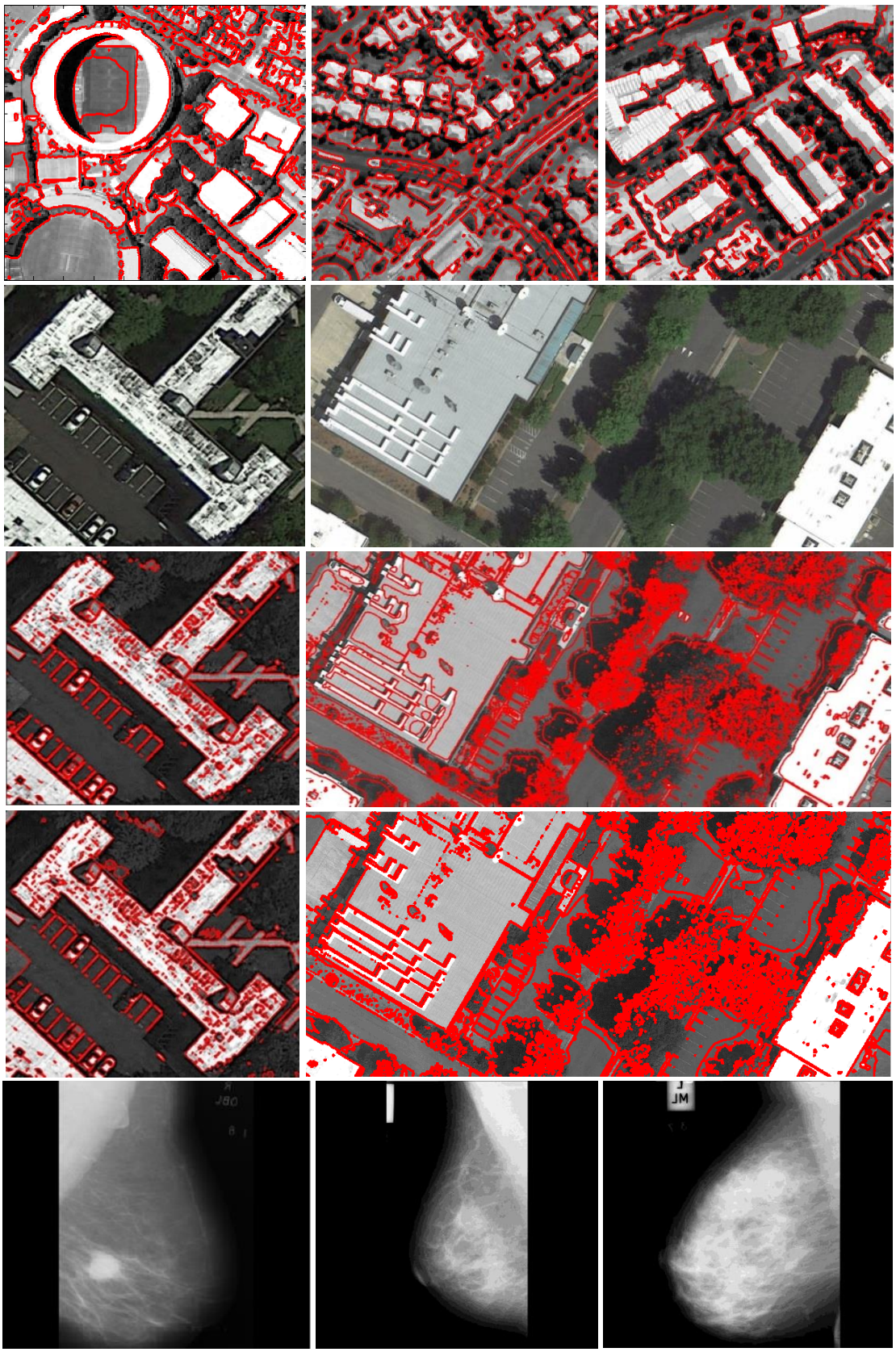
Step 7: Export the segmentation mask.

Step 8: Asses the overall performance of the algorithm using the ground truth images by calculating the recall, precision and F score.

8.4 Experimental results and discussion

The algorithm is evaluated for its ability to detect buildings and delineate the breast region using the same dataset as presented in chapter 3. The level set method for image segmentation in the presence of intensity heterogeneity (LIC) proposed by Li et al. and the locally statistical active contour model for image segmentation with intensity heterogeneity (LSACM) proposed by Zhang et al. have been also applied and evaluated on the same images using the same principles and parameters. The original source codes as have been published by the corresponding authors were downloaded and used for this study [194-195].

Figure 33 and Table 12 presents the results of the LIC and LSACM models. Figure 34, Figure 35 and Table 13 illustrates the results of the proposed method. Table 14 illustrates the computational cost of all models.



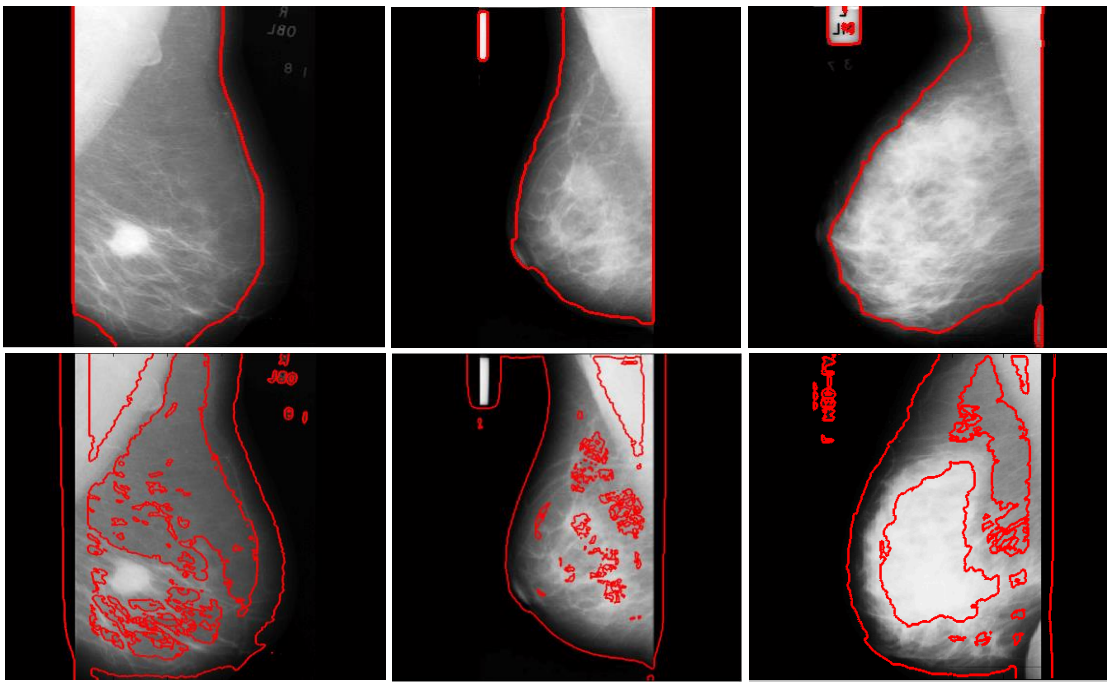


Figure 33: Google earth and Mammographic test images and obtained results by applying the under evaluation active contour methods for building detection and breast region segmentation.

First-Fourth-Seventh row: Original images, Second-Fifth-Eighth row: LIC segmentation, Third-Sixth-Ninth row: LSACM segmentation.

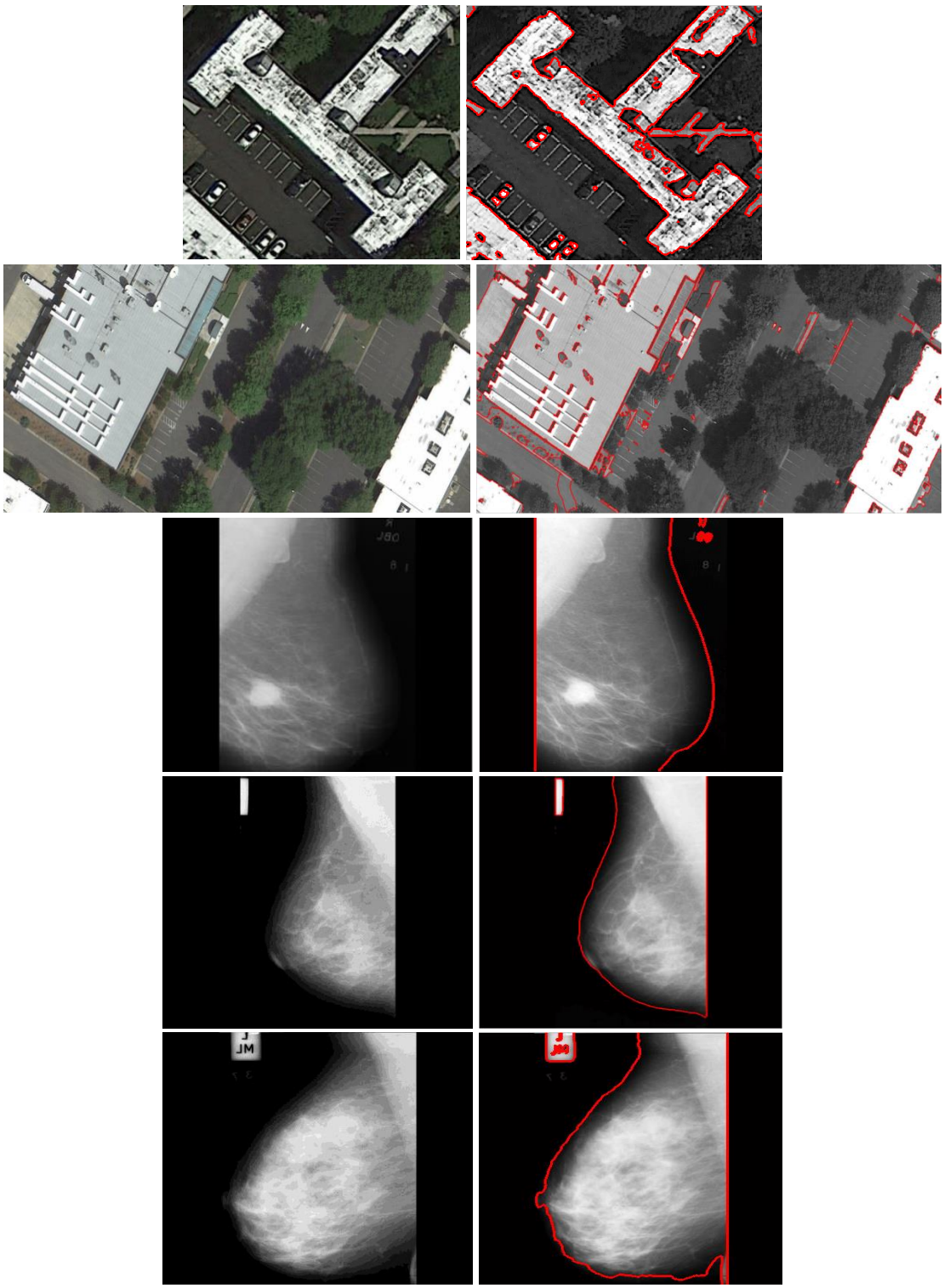
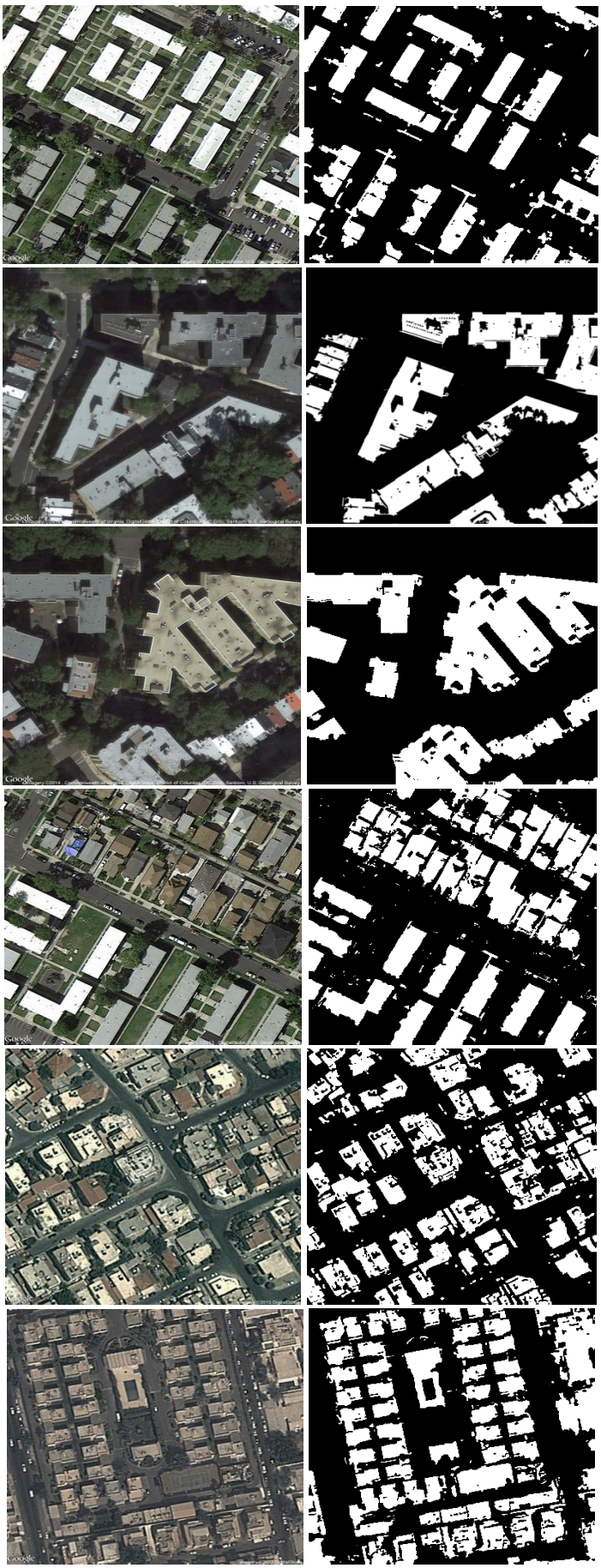


Figure 34: Google earth and Mammographic test images and obtained results by applying the proposed active contour model (OLAC) for building detection and breast region segmentation.

First column: Original image, Second column: Optimized localized active contour representation.



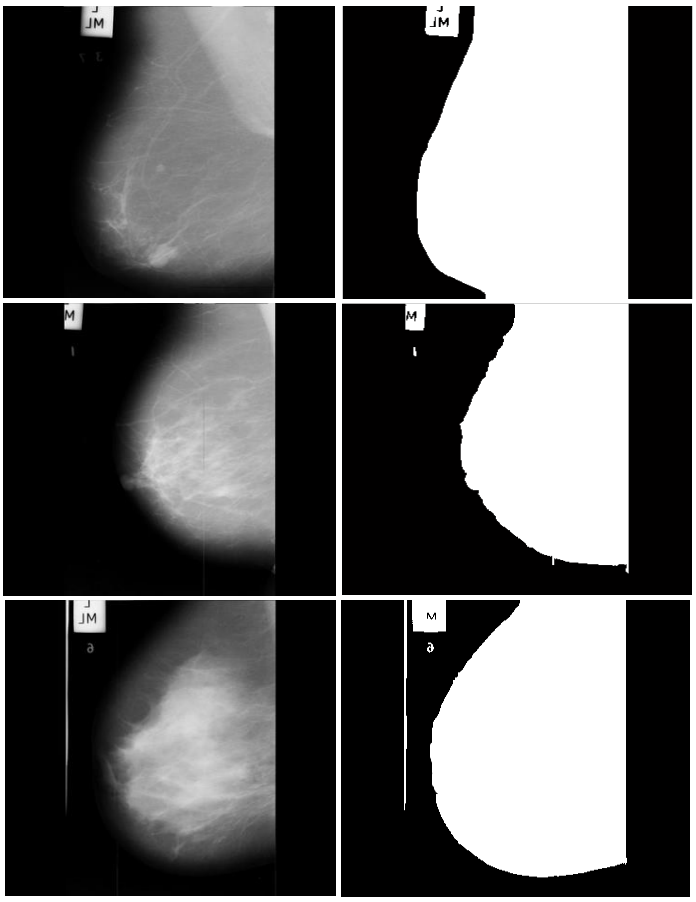


Figure 35: Google earth and Mammographic test images and obtained results by applying the under evaluation active contour methods for building detection and breast region segmentation.

First-column: Original images, Second column: OLAC segmentation masks.

Table 12: Pixel based segmentation evaluation of the LIC and LSACM models

Image Domain	**Total Number of Images**	**Average Precision (%)**		**Average Recall (%)**		**Average F (%)**	
		LIC	***LSACM***	***LIC***	***LSACM***	***LIC***	***LSACM***
Satellite	322	65.8	64.4	87.6	86.1	75.2	73.7
Mammograms	322	85.4	66.7	86.3	84.8	85.8	74.7

Table 13: Pixel based segmentation evaluation of the proposed model using the Butterworth and Gaussian low pass kernel

Image Domain	**Total Number of Images**	**Average Precision (%)**		**Average Recall (%)**		**Average F (%)**	
		Butterworth	***Gaussian***	***Butterworth***	***Gaussian***	***Butterworth***	***Gaussian***
Satellite	322	88.5	82.6	93.2	90.4	90.8	86.3
Mammograms	322	93.2	97.4	95.6	98.2	94.7	97.8

Table 14: Computational cost evaluation

Method	Mammography Domain Computational Cost Average (Seconds)	Satellite Domain Computational Cost Average (Seconds)
OLAC (Gaussian)	56.5	6.8
OLAC (Butterworth)	57.8	5.4
LIC	93.2	19.5
LSACM	165.4	47.2

In this work, an optimized localized active contour model (OLAC) is proposed. The model has been designed to enhance the performance of the localized active contour models in the presence of intensity heterogeneity and feature complexity. A localized filter function such as a low pass Gaussian or a low pass Butterworth is incorporated for transforming the image into a new representation, which is considered as less heterogeneous. It is expected that the remaining intensity heterogeneity artefact can be better estimated and eliminated in the local regions of the transformed image defined by a spatially weighted kernel function, using active contour and level set methodologies.

The visual interpretation along with the quantitative segmentation results of the presented method revealed that the proposed algorithm works well with satellite images of urban areas for detecting building boundaries. Results presented in all figures suggest that the proposed methodology can detect almost all of the buildings and their basic structure and shape, even when the buildings have irregular shapes, interact with environmental objects and occlude each other. In addition, the proposed method discards most of the non-building areas with similar building characteristics and many misleading artefacts. The experimental results suggest also that the proposed model is appropriate for the estimation of skin-air boundary for a breast segmentation process. It works well with images of breasts of various sizes and densities and can detect the breast edge, even when other artefacts are present.

In this work, when the algorithm was applied with a Butterworth low pass filter on satellite images for the detection of buildings, reached a building detection accuracy or recall of 93.2%, precision 88.5 % and a ratio F score equal to 90.8% as can be seen in Table III. Using the Gaussian low pass filter a recall of 90.4%, a precision of 82.6 % and a ratio F score equal to 86.3% has been achieved. The algorithm achieves a recall of 98.2%, precision 97.4 % and a ratio F score equal to 97.8% if is applied on mammogram images for delineating the breast region using the Gaussian filter. Using the Butterworth filter, a recall of 95.6%, precision 93.2% and a ratio F score equal to 94.7% is achieved. It must be noticed that despite the good results in both domains for both low pass filters, Gaussian works superior to mammography images while Butterworth works more efficient on satellite images. These results compare favorably to the results of both localized active contour

models LIC and LSACM as can be seen in the presented tables. The computational cost is also an important factor for practical applications of the proposed method, where large datasets are involved, and has been proved that is significantly lower than LIC and LSACM. In several cases especially, in satellite images, very few iterations such as 30 to 50 are required for detecting almost all the buildings that exist in the analyzed scene.

In addition, the results of the proposed model are compared favourably to other results in literature, in the cases where quantitative results are available. However, it must be noticed that in the satellite imagery domain, common databases are not available and researchers are using their own distinctive data for evaluation purposes. Ok et al. [87] achieved a pixel based recall accuracy of 83.3%, precision 80.3% and F score 81.7% in a building detection process. Furthermore, it has to be noted that Ok et al. work is very different, since shadow evidence was utilized along with near-infrared (NIR) information in a Grab-Cut partitioning methodology. In the proposed method, only the information of the three true colour bands (RGB) has been used. It has to be noted that the NIR band or any other metadata information is unavailable for the acquired Google Earth images. Ghaffarian and Ghaffarian [89] in their proposed method for building detection used monocular high resolution Google Earth images in a purposive fast independent component analysis framework. The results were as follows: pixel based recall accuracy 73.6%, precision 88.6% and F score 80.1%. Based on the results of both methods, i.e. the presented method and the Ghaffarians method, where distinct schemes were used one can conclude that monocular Google Earth or Google maps images can be used for building boundaries detection. It has to be noted that the proposed method achieves superior performance in all quantitative metrics.

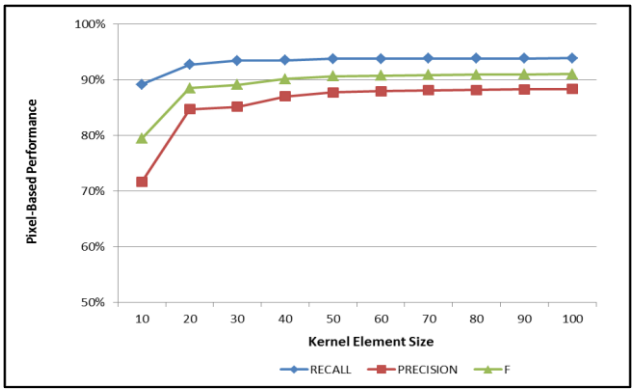
In the mammographic image domain, several researchers such as Wirth et al. [130] and Tzikopoulos et al. [199] utilised MIAS database for evaluation purposes. The results of this work compare favourably to the results achieved by the above models. Wirth et al. achieved a segmentation accuracy of around 98%. However, their method is limited by the availability of gradients in the breast edge - especially in the bottom of the breast region - and has the limitations of parametric active contours. Tzikopoulos et al. achieve segmentation accuracy also around 98% following Masek's proposed scheme [118]. The algorithm relies on the idea that the skin-air interface is the smoothest section of identical pixels near the breast boundary. Based on that, a specific threshold is defined for breast region segmentation. Naturally, the algorithm highly depends on the correct estimation of this threshold, which in some cases is a very difficult task to define and the scheme may suffer by the general limitations of the segmentation methods that are based on thresholds.

The proposed method achieves segmentation accuracy results near 98.5%, which is very similar with the above results using a simple and time efficient method.

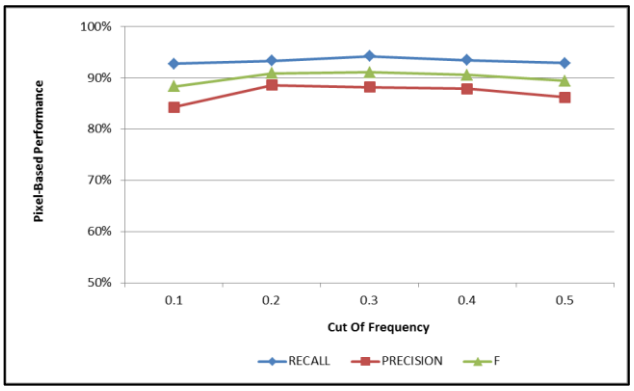
Finally, it should also be noted that if the images in both domains do not adhere to quality standards (figure 35, row fifth, sixth and ninth), the resulting segmentation might suffer, as is the case with several segmentation algorithms [89, 146, 200]. However, the proposed algorithm has been shown to work well even in such cases.

In high-density complex urban environments, the proposed method still cannot classify correctly some non-building objects such as bridges and parking places or even parts of the roads, which are classified as buildings. However, this behaviour was expected because of the feature similarity between these objects and the buildings. This is also noticed in other schemes proposed in the literature [89, 121-123, 146]. Also in certain cases where two or more buildings are very close to each other, these are classified as a single building.

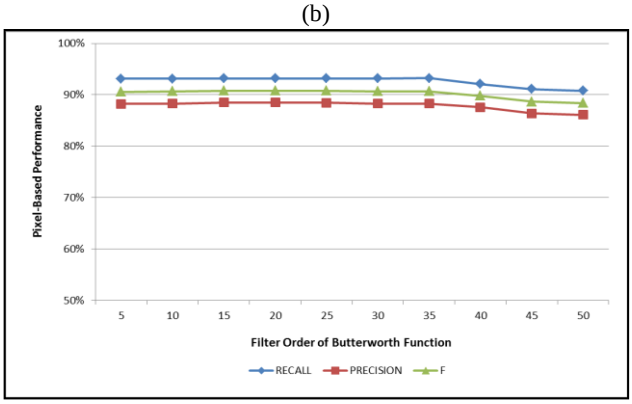
In this work an extensive analysis of the parameters that might affect the building detection results or the breast region extraction of the proposed method was performed. One parameter, which is investigated thoroughly for its effect to the final segmentation result, is the size of the low pass kernel function. It is expected that if the size is too large or too small the results will be affected, leading to under-segmentation or over-segmentation.The diagrams *a* and *d* in figure 34 present the relationship of the kernel size with the quantitative metrics and as it can be seen, a kernel size between 20-30 pixels leads to satisfactory results. However, the kernel size is in close relationship with the size of the under evaluation images and the size of the regions of interest thus must be adjusted accordingly. Cut-off frequency is also an important factor in the overall performance of the proposed method since it controls the process of eliminating the intensity heterogeneity. In figure 36, the diagrams b and e present the relationship of the cut-off frequency and the segmentation results. The results appear optimized for cut off frequency values between 0.15 and 0.35. Finally, the filter order was also examined and presented in the diagrams c and f of figure 36. The most appropriate order value has been estimated between 5 and 15.



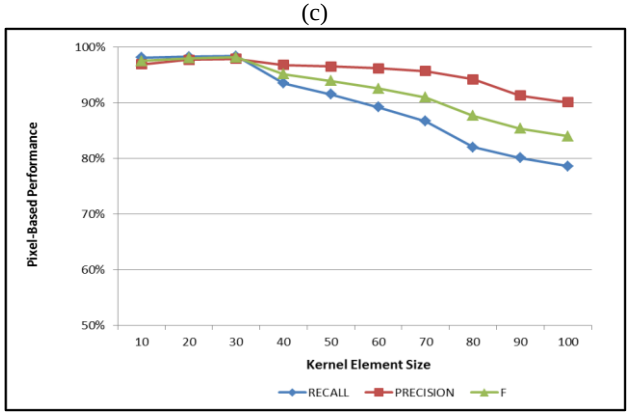
(a)


100%
90%
80%
70%
60%
50%
Pixel-Based Performance
0.1
0.2
0.3
0.4
0.5
Cut Of Frequency
RECALL
PRECISION
F

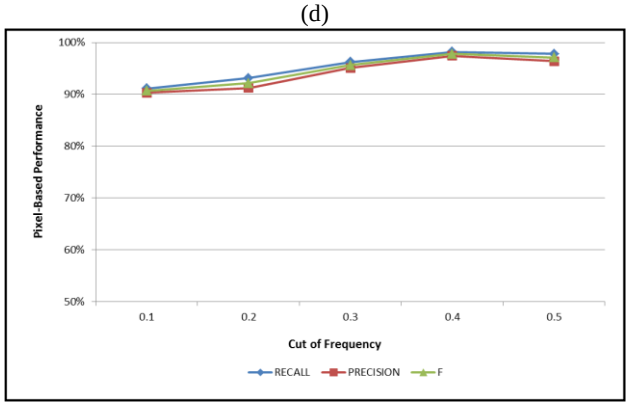
(b)


100%
90%
80%
70%
60%
50%
Pixel-Based Performance
5
10
15
20
25
30
35
40
45
50
Filter Order of Butterworth Function
RECALL
PRECISION
F

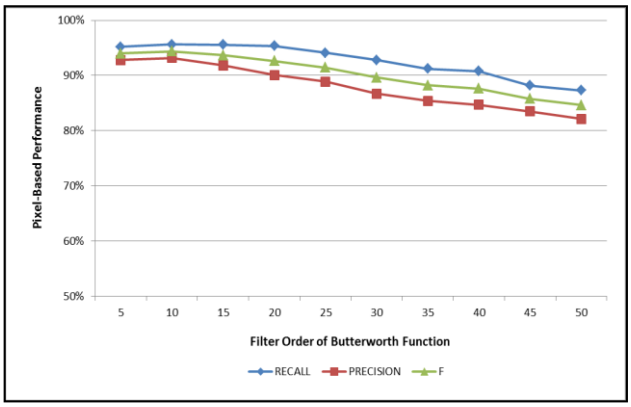
(c)


100%
90%
80%
70%
60%
50%
Pixel-Based Performance
10
20
30
40
50
60
70
80
90
100
Kernel Element Size
RECALL
PRECISION
F

(d)


100%
90%
80%
70%
60%
50%
Pixel-Based Performance
0.1
0.2
0.3
0.4
0.5
Cut of Frequency
RECALL
PRECISION
F

(e)



(f)

Figure 36: Parameters analysis: low pass kernel size, cut off frequency and filter order of Butterworth function

(a), (b), (c): Parameter analysis for the detection of buildings, (d), (e), (f): Parameter analysis for breast region segmentation in mammographic images

8.5 Conclusions

Localized active contour models were evaluated in this work for their ability to detect buildings in monocular satellite images and delineate the breast region in mammography images in the presence of intensity heterogeneity. Even though these models achieve good results, extensive initialization, under detection and over detection can still be noted. In addition, in several cases the computational cost was high. An optimized algorithm was proposed to address some of these issues, which works efficiently in the presence of intensity heterogeneity, and appears able to detect objects with arbitrary shapes, sizes even when the intensities significantly vary. Proposed algorithm addresses the extensive initialization, under detection and over detection limitations of the localized active contour models, and optimizes the overall segmentation procedure in breast region segmentation using mammography images, as well as the detection of building boundaries in satellite images. The intensity heterogeneity is considered as a low pass artefact, which can be well approximated with a constant parameter in its local neighbourhood after the application of a suitable low pass convolution filter. The proposed active contour model encompasses a low pass kernel function for transforming the image into a new representation where the regions of interest are presented with less intensity heterogeneity. During the localized curve propagation process, the remaining intensity heterogeneity artefacts are estimated and eliminated. The proposed optimized localized active contour model is not sensitive to the initialization process and drives the curve evolution faster leading to a better image segmentation by avoiding several misleading regions. Experiments were carried out on a large number of scenes and the produced segmentation results from heterogeneous and complex images from a diversity of fields, which include satellite, and mammography

images, verify that the suggested method is efficient. It was found that the proposed method achieves better accuracy when compared with the original LIC and LSACM segmentation methods. Results are also compared favourably with other studies proposed in literature, where similar pixel based performance measures exist. Furthermore, the method appears to be more time efficient while no dependency on any additional data or user interaction is required.

Chapter 9: Classification Model in the Presence of Intensity Heterogeneity and Feature Complexity

9.1 Introduction

Different set of descriptors can add important information in classification processes. Integrating different sets of image descriptors is fundamental for achieving better performance in classification models evaluating images where heterogeneity and feature complexity exists. To have the maximum advantage from feature integration, each feature descriptor set should represent diverse information and an integration scheme should be applied. Statistical distributions of different feature descriptors such as normalize luminance coefficients, second level statistics, scale invariant feature transforms (SIFT), local binary patterns (LBP), texton histograms as presented in chapter 4, can be used for the development of an objective classification model. In the proposed classification model a robust integration scheme is developed where the different feature sets are automatically and optimized integrated into a support vector machines (SVMs) classification framework.

9.2 Related work

Integration schemes can be divided into two main categories, low-level and high-level as presented in literature in various domains [201]. In low-level strategies, the feature descriptor sets are integrated at feature level to form a new representation. The integration does not directly involve the classification process. Feature descriptor set concatenation is used where two or more feature vectors are combined into a single feature vector that is normalized and is then used for classification. The information related to each feature descriptor set is combined without a weighting factor that allows the control of the influence of information derived from the individual descriptor sets to the final decision result [202]. A low-level integration method has been also investigated and presented in chapters 9 and 10 with promising results. Despite the good performances for specific tasks, there are several drawbacks of the low-level methods. First, if a feature descriptor set gives misleading information, it is expected that the new feature vector will be adversely affected influencing the whole performance. Second, the dimension of the feature vector increases, as more feature sets are added. This implies longer learning and recognition time, greater memory requirements, and a possible curse of dimensionality effect.

High-level integration methods start from the output of two or more classifiers using two or more feature descriptor sets producing individual decisions about the object to be classified. All these decisions are then combined together to achieve a consensus decision. In high-level integration methods, the strategy is to keep the feature sets separated, train individual classifiers and integrate the outputs of the individual classifiers [201-203]. These techniques are more robust, allow the use of different classifiers adapted to the characteristics of each single feature set and decide on the number of feature sets that should be extracted and used for each particular classification task [202]. The success of these methods is related to the selected rules for defining the different weights of the different feature descriptor sets to the final decision. This task can be very difficult without a prior knowledge of the importance of the individual descriptors to the classification process.

9.3 Classification model and optimal feature descriptor set integration method

Inspired from the advantages and disadvantages of high and low level integration methods, a so-called mid-level integration method using an additional SVM classifier able to automatically weight the information derived from the different feature descriptor sets has been developed. In the proposed method for each set of features, an SVM classifier that outputs a set of decision values corresponding to the various classes of each sample image is initially applied. A kernel function is then used to integrate the outputs generated by the single feature descriptor classifiers to form a new feature descriptor set. The parameters of the integration function are learned by applying an optimization process and the weights of the different feature descriptor sets are automatically estimated. More specifically, the proposed feature descriptor sets integration model can be described as follows:

1. For a training set containing samples belonging to all different quality classes (*M*), extract feature vectors *T* each relative to a single descriptor set ($t = 1, \dots, T$).

2. Apply SVM [64] multiclass extension training process one versus the rest, including k-fold cross validation [65] for estimating the optimal SVM parameters and define the training model f.

3. For the same sample *s* for each of *T* apply the SVM training model and calculate the output decision values using the following equation:

$$V_{t,u}(T_t(s) = f_{t,u}\big(T_t(s)\big), u = 1, \dots, M \tag{9.1}$$

4. Apply the following integration function to form the new feature descriptor set:

$$V_u(s) = \sum_{n=1}^{N} a_{u,n}\gamma_n K(v_n, v) + \beta_u, u = 1, \dots, M \tag{9.2}$$

where v is a vector containing all the outputs for all descriptors T:

$$v = \left[\left\{V_{1,v}\left(T_1(s)\right)\right\}_{v=1}^{V_T}, \dots, \left\{V_{T,v}\left(T_T(s)\right)\right\}_{v=1}^{V_T}\right] \tag{9.3}$$

The choice of the kernel function K is open and any linear or nonlinear kernel can be used. The parameters $a_{u,n}\gamma_n, \beta_u,$ and the support vectors v_n are estimated during the training optimization process. This new feature descriptor set is used as an input to an additional SVM classifier to form the training model for the image quality assessment. The new SVM training model can be used to assess unknown images by classifying them into one of the predefined quality categories. Figure 37 presents the flow chart of the proposed optimal feature descriptor set integration method and classification model.

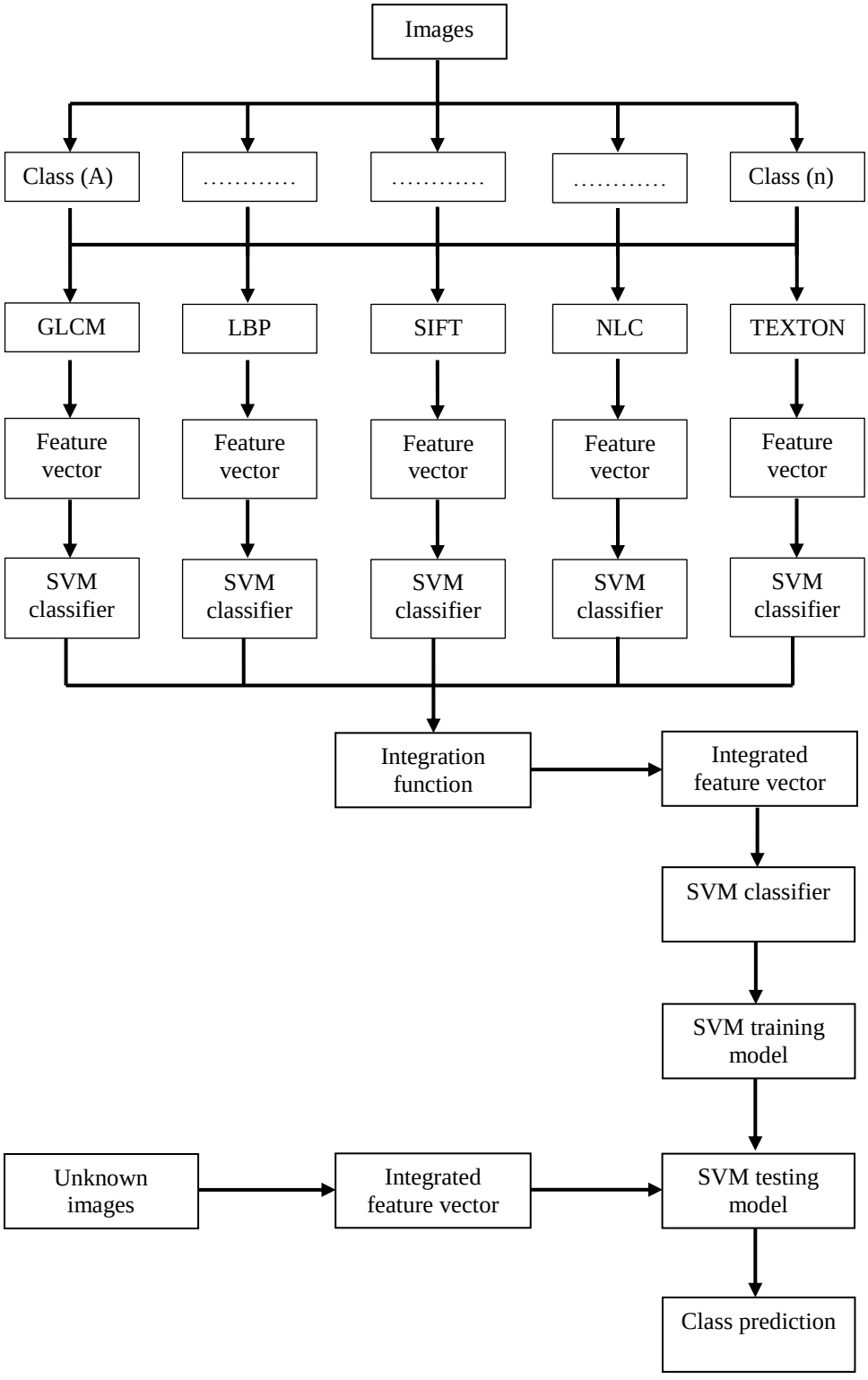


Figure 37: Feature integration and classification model

9.4 Discussion

The performance of the proposed feature integration and classification model should be assessed using practical applications and real data where heterogeneity and feature complexity often occurs such as satellite and mammographic imagery. Initially the ability of the feature descriptors using different SVM kernel functions should be assessed and the

most appropriate should be selected for further evaluation using the proposed integration scheme. Furthermore, the performance should be assessed against high and low level integration methods. In the following chapters a thoroughly assessment of the proposed method is performed for breast density and breast abnormality classification. Finally, a no-reference quality assessment model is developed for satellite images.

Chapter 10: Integration of Different Texture Features for Mammographic Breast Density and Abnormalities Classification

10.1 Breast density classification

Mammographic breast density refers to the prevalence of fibro-glandular tissue as it appears on a mammogram. Breast density is not only an important risk for developing breast cancer but can also mask abnormalities. Breast density information can be used for planning individualized screening and treatment. In this work, statistical distributions of different texture descriptors such as scale invariant feature transforms (SIFT), local binary patterns (LBP), texton histograms and their combination are investigated with Support Vector Machines (SVMs) for objective breast density classification. SIFT is an approach for detecting and extracting local feature descriptors that are reasonably invariant to changes in illumination, image noise, rotation, scaling and small changes in viewpoint. The SIFT descriptor is a coarse descriptor of the edges found in the key-points. LBPs provide a robust and computationally simple way for describing pure local binary patterns in a texture. They provide information regarding the prevalence of different edge patterns and uniformity. Textons are defined under the operational definition of clustered filter responses and provide a statistical and structural unifying approach for texture characterization. The different set of descriptors can add different however important information in breast density classification. Thus, an algorithm for breast density classification using texture features automatically integrated at feature level in an SVM classifier framework is developed.

10.1.1 Related work

Mammographic breast density has been shown to be one of the most important risks for developing breast cancer first reported by Wolfe in 1976 [204]. Breast density is associated with both lower sensitivity and an increased rate of interval cancers [204-207]. It is an inherent trait, but can be altered by endogenous and exogenous hormonal factors, elective estrogenic receptor modulators, diet and generally declines with age [207]. Better understanding of breast density and how it corresponds to a significant increase of breast cancer risk as well as the case that cancer in dense breasts are more often mammographically occurred [208] have led to the need of breast density assessment and

use of the information for supplementary screening using other imaging techniques such as whole-breast ultrasound (US) [209]. Information regarding breast density has also been shown to be a cost-effective strategy for sharing risk information that may become useful in prevention decisions. Women given such information reported being very likely to have an annual clinical breast examination in addition to their screening mammogram [210]. X-ray mammography has been the method of choice for breast cancer population screening. With the development of novel breast imaging techniques and the endeavour to move towards individualized screening and prevention, mammographic breast density can be used to establish appropriate plans. In the case of dense breasts use of whole breast ultrasound has been shown to significantly increase breast cancer detection despite a corresponding increase in false positives [211]. Magnetic Resonance Imaging (MRI) can also be used for screening in dense breasts. In Addition to US, tomosynthesis may be useful in addition to mammography for reducing false positive recalls especially in dense breasts [212]. Thus, breast density and change thereof may be used for risk assessment, for reducing screening intervals, for the development of Computer Aided Detection (CAD) systems with higher sensitivity and specificity, but most importantly for signalling the necessity for developing individualized risk evaluation, screening, and treatment for achieving the earliest possible diagnosis for the best prognosis. Mammographic breast density is a powerful breast cancer risk factor that has considerable potential in risk stratification and in monitoring the effects of interventions in risk alteration. Yet, the need still exists to develop objective methods that provide precise, simple and reproducible density measures to achieve this [208]. Following Wolfe's mammographic breast parenchymal density categorization [209], the American College of Radiology proposed the Breast Imaging Reporting and Data System (BIRADS) [213] mammographic parenchymal density classification as follows:

(i) the breast is almost entirely fat,
(ii) there are scattered fibro-glandular densities,
(iii) the breast is heterogeneously dense which may lower the sensitivity of mammography, and
(iv) the breast tissue is extremely dense, which could obscure a lesion in mammography.

The first two classes correspond to low density, low risk mammograms whilst BIRADS mammographic breast density classes III and IV correspond to high density, higher breast cancer risk classes. The four BIRADS mammographic breast density classes can be seen in Figure 38.

Different computer aided techniques and methodologies have been developed for more objective and reproducible mammographic breast parenchymal density evaluation and classification, which fall in two main categories: qualitative and quantitative [208]. These include physics based models and evaluation of different parameters based on intensity and/or texture. Byng et al. [214] proposed a semi-automatic area based quantitative computer measure. Interactive thresholding and the percentage of the segmented dense tissue over the segmented breast area provide a relative quantitative evaluation of breast parenchymal density. The method that is implemented in CumulusTM software is widely used for density classification. The Standard Mammogram Form (SMF) provides the height of interesting tissue i.e. volume of non-fat tissue in the breast [125] and more recently VolparaTM which uses a relative - rather than an absolute - physics model which reduces the need for accurate imaging physics data [215]. As breast density is evaluated on the 2D projection of the breast on the mammogram, a number of breast density classification methods have used different texture based features for density classification. Miller and Astley [216] investigated texture-based discrimination between fatty and dense breast types applying granulometric techniques and Laws texture masks. Petroudi et al. [217] proposed a scheme that uses statistical based texton models to capture the mammographic appearance within the breast area. Musta [218] et al. presented an overview of the accuracy of different breast density classification methods using different features, and achieved a maximum classification accuracy of 73.3% through a selection of Haralick and Soh texture features, genetic search and wrappers. Kallenberg et al. [219] developed a breast density segmentation algorithm using a set of different features with information about location, intensity, texture and global context with a neural network for pixel classification intensity. Keller et al. [220] used adaptive Fuzzy Mean Clustering combined with Support Vector Machines (SVMs) for cluster tissue classification. He et al. [221] calculated different texture features on the grey-level intensity histograms to characterize different mammogram blocks to different tissue types and used a binary model based Bayes classifier for BIRADS mammogram classification.

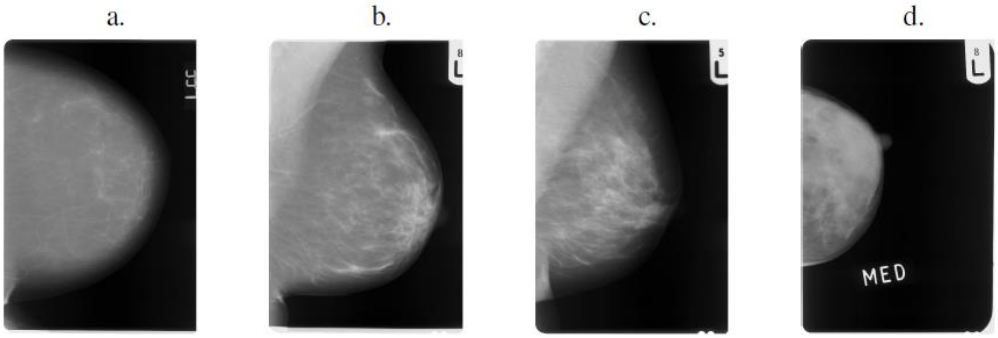


Figure 38: BIRADS density classification system

10.1.2 Proposed method

For breast density characterization the mammograms are pre-processed and normalized as in [217] and the mammographic breast region is segmented. Following, the different features are evaluated on the segmented breast region. Each training image is represented by a vector(s)/histogram(s), and SVMs [64] are trained to discriminate vectors corresponding to positive and negative training images. Finally, the trained SVMs are applied to testing images.

The methodology for the integration of the different texture statistical distribution vectors and the development of the SVM classification model follows the principles as presented in chapter 9 and is shown in Figure 39. Different kernel functions were also investigated and the Gaussian Radial Basis Function (RBF) was chosen as it resulted in the best separation between the different density classes. Ten-fold cross validation is used for training and evaluating the corresponding SVMs.

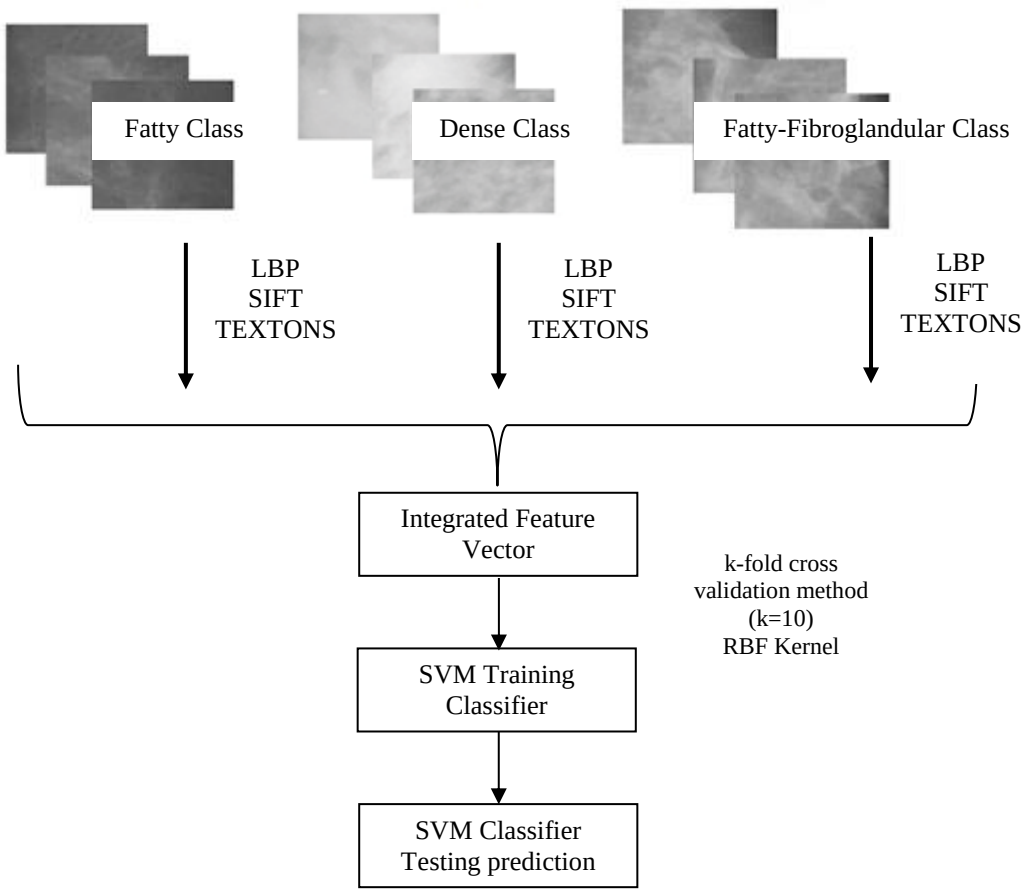


Figure 39: Proposed classification diagram

10.1.3 Results and discussion

The algorithm is evaluated on a set of mammograms from the MIAS mammographic database [83-84]. The MIAS database contains 322 mammograms corresponding to 161 cases. The mammograms in the database are classified to three different density classes, fatty, fatty-glandular and dense glandular. Thus for the purposes of this work the mammograms are automatically classified to one of these three classes. Since there is different representation of the three classes in the database, only 62 mammograms are chosen per density class. The performance of the SVMs evaluated for breast density classification using tenfold cross validation modelled on each of the histograms of the corresponding texture features is shown in Table 15. Accuracy is calculated as the percentage of correctly classified mammograms in a breast parenchymal density category over the ground truth total number of mammograms in that category. Figure 40 shows the classification for each of the mammograms in every category.

Table 15: Classification accuracy results for the MIAS breast density characterization using SVMs

MIAS Density Classification SVM trained on	Accuracy %
SIFT	74.7312
LBP	82.7957
TEXTONS	75.8065
LOW-LEVEL INTEGRATION	91.1836
HIGH-LEVEL INTEGRATION	93.5484
PROPOSED INTEGRATION MODEL	95.8674

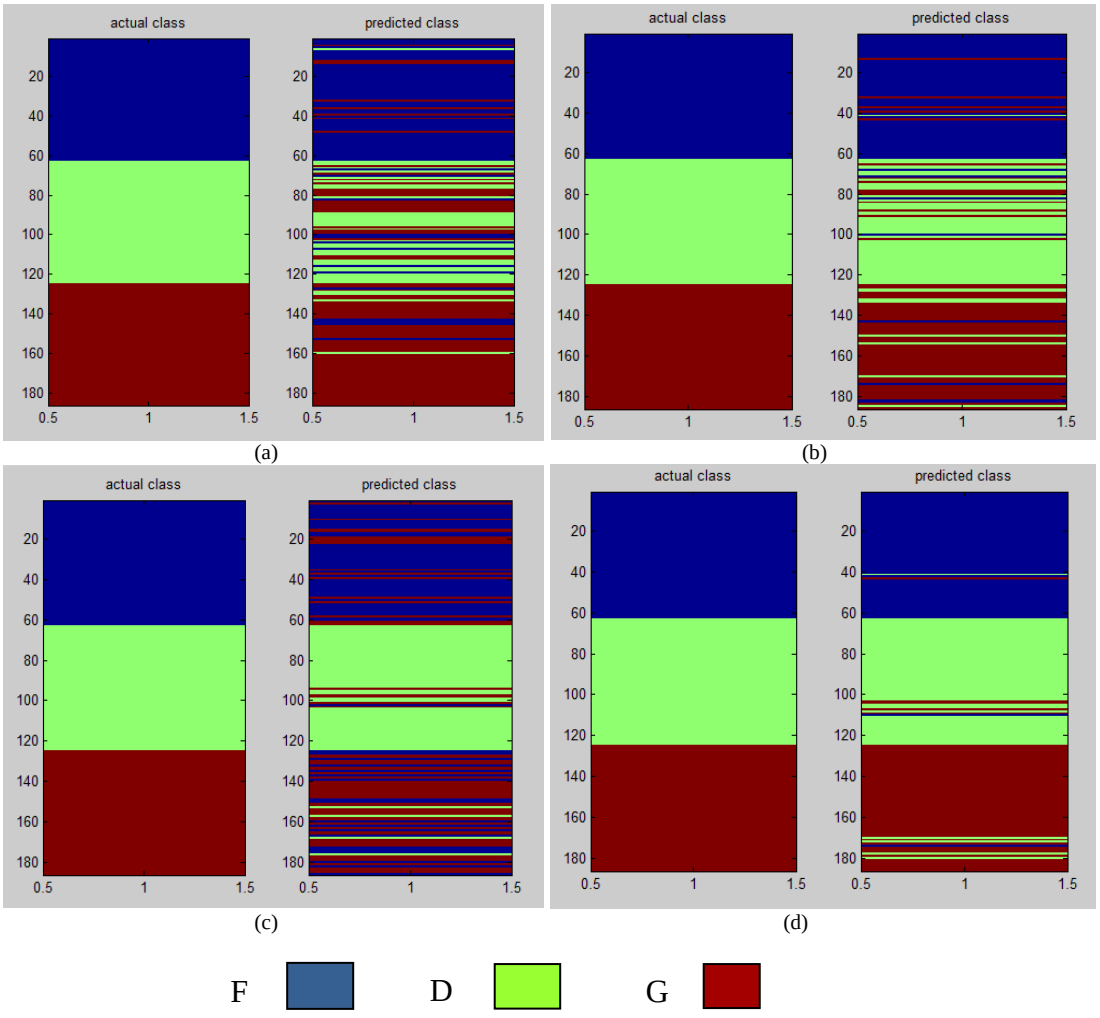


Figure 40: Accuracy classification results graphical representation

(a) Sift texture, (b) LBP texture, (c) Texton Texture, (d) SIFT-LBP-Texton Texture

The classification accuracy for the SVM when all texture feature distributions are combined at feature level reaches over 93%. SVMs achieve high classification when they are trained on all the different texture features distributions. However, these distributions have different sizes. Different combinations of different features with different classifiers may result in better classification accuracy. Thus, the best agreement of 95.8674% with the density annotations provided with the MIAS database was achieved using the proposed feature integration and classification model. The different texture features used seem to capture different breast region characteristics. The SIFT captures local oriented features

over different visual scales. LBP provide information whether the region of interest (ROI) is relatively uniform. Textons can be thought as the building blocks of the present texture the same way phonemes make up speech [126]. The presented method compares favourably with other methods presented in the literature. Petroudi et al. [217] evaluated texton histograms using chi-square distance achieved classification accuracy of 76% but on a different database. Blot and Zwiggelaar [108] using Grey Level Co-occurrence matrices and a method by Karssemeijer [222] achieved 65% accuracy for classifying the MIAS images to the MIAS given 3-class density categorization [83]. Bosch et al. [223] used a combination of SIFT and textons histograms resulting from local image patches achieving a classification accuracy of 91.4%. Oliver et al. [224] extracted morphological, texture features from the segmented breast areas, and used a Bayesian combination of a number of classifiers achieving 84% accuracy. Furthermore, from the evaluation of the SVMs with the histograms of the different texture feature one can see that the LBPs achieve the best classification accuracy when used alone, whilst training the SVM with the SIFT features' histogram achieves the worst classification result. Further investigation of the different parameters for the evaluation of the different SIFT features is required.

10.2 Breast abnormalities classification

Breast cancer is considered a major health problem amongst women. A breast mass is a generic term to indicate a localised inflammation or lump in the breast. Mammography remains the key screening tool for the detection of breast abnormalities. Masses are characterised by their location, size, shape and margin. The proportion of breast tumours that are detected using screening programs, increased significantly in the resent years. However, it is well known that expert radiologists can miss a significant proportion of abnormalities. Furthermore, several mammographic abnormalities are classified as benign after biopsy. Thus, the development of effective computer aided detection systems to assist radiologists in the detection and/or classification of mammographic abnormalities is utmost importance. This section investigates the use of different texture features evaluated from grey-level co-occurrence matrix (GLCM) [94-95] for classifying abnormalities. For characterization of the abnormalities, GLCM features such as autocorrelation, contrast, cluster shade and prominence, energy, entropy, homogeneity etc. are evaluated using SVMs on ground truth segmentation provided by expert radiologists. The different features are evaluated not only on the entire training set, but also on subsets that hold mammograms corresponding to different breast density classes. This work will present the ability of these features to classify normal and abnormal tissue for different tissue density classes. For the

feature selection, the importance of each feature and its effect on the classification accuracy is investigated. The aim is to select the most prominent features that provided the best discrimination for abnormality characterization and develop an optimized classification model. The selection of the most appropriate features is not only important for the performance of the classification model, is also important for the computational cost and the generalization of the model. The reduction in features in classification frameworks is vital for developing models able to generalize sufficiently when the number of training samples in a class is limited which often happens with mammographic data sets [225]. In this work, it was found that abnormal and normal regions have significant differences in terms of their texture uniformity. Normal images contained smooth texture while abnormal images were heterogeneous thus the utilization of only three features such as entropy, energy and cluster shade were used for abnormality classification.

10.2.1 Related work

Many computer vision techniques have been proposed in literature for the detection, segmentation and classification of masses using digital or digitized mammograms. Detection is defined as the identification of potential masses within the entire breast region. Usually, these methods define a suspicious region in a mammogram. Segmentation is defined as a method able to detect the precise outline of the potential mass. There are algorithms that combine both procedures. Classification is defined as the process of characterizing a suspicious region as mass or no mass and subsequently as benign or malignant.

Mammographic computer aided detection and classification algorithms can be initially categorized into those that use multi-view images and those that utilize single-view images. Several reviews for mammographic detection, segmentation and classification have been published in literature [22, 226-227]. Cheng et al. cover the general enhancement of mammographic images, the detection and classification of masses [226]. Rangayyan et al. provided a review covering the full range of mammographic abnormalities [227].Oliver et al. focused their review only on mass detection and segmentation [22]. This work is involved with the detection of a suspicious region and the classification of this region as mass or no mass using single-view mammograms and utilizing classifiers with texture features. Thus, the following review of related work is concentrated on mammographic mass detection and classification methods that utilized texture classifiers.

Bovis and Singh proposed bilateral subtraction of both mammograms (right and left) in order to find asymmetries in either mammogram and texture features for classifying these

asymmetries as abnormalities were used [228]. Wang et al. use a single view of the mammogram and extracted six textural features from the suspicious region that were detected after a pre-processing procedure using exponential transformation [229]. Mohamed and Kadah proposed a three-step system, namely the ROI extraction, features extraction and classification. A set of 88 features were extracted and it was found that 78 of those features are capable of discriminating normal and abnormal breast tissues [230]. A few researchers have also considered automatic detection of masses using artificial neural network [231-232]. Weidong et al. performed an iterative thresholding procedure to locate masses on fatty tissues while in dense tissues black hole registration based on discrete wavelet transform was used instead. Finally, the segmented suspicious masses were filtrated using 10 selected features via a multilayer perceptions classifier [231]. A classification approach for detection of breast abnormalities in digital mammograms using particle swarm optimized wavelet neural network developed by Dheeba et al. [233]. The proposed abnormality detection algorithm is based on extracting laws texture energy measures from the mammograms and classifying the suspicious regions by applying a pattern classifier. Polakowski [225] used shape, size, contrast, and texture features for the classification of masses. A difference of Gaussians filter to highlight suspicious regions has been utilized and derivative-based feature saliency techniques were used to determine the best features for the classification of abnormalities.

Most methods use a pre-processing procedure for breast region segmentation. This is important because computer aided systems can avoid useless computational and processing time. It is also of utmost important to pre-process mammograms because mammogram labels, artefacts and the pectoral muscle if exist should be avoided in the search of abnormalities. In an automate breast abnormality detection process these objects might affect the performance of the method.

10.2.2 Proposed method

Grey-level co-occurrence matrices (GLCMs) show the relative frequency with which different combinations of different pixel intensity values at certain distances occur in an image. The matrices are used for the evaluation of different texture features such as autocorrelation, contrast, cluster shade and prominence, energy, entropy, homogeneity etc. (Haralick et al. 1973) [234]. These features are computed for the different GLCMs for different pixel distances. In this implementation, the matrices have been constructed at a distance of $d = 1$ and for direction of θ given as $0°$,$45°$,$90°$ and $135°$. The various GLCM features were evaluated on ground truth segmentation provided by expert radiologists taking

into consideration the different breast density classes as presented in the BIRADS system [213]. Figure 41 presents samples of abnormal and normal regions as characterized by the radiologists.

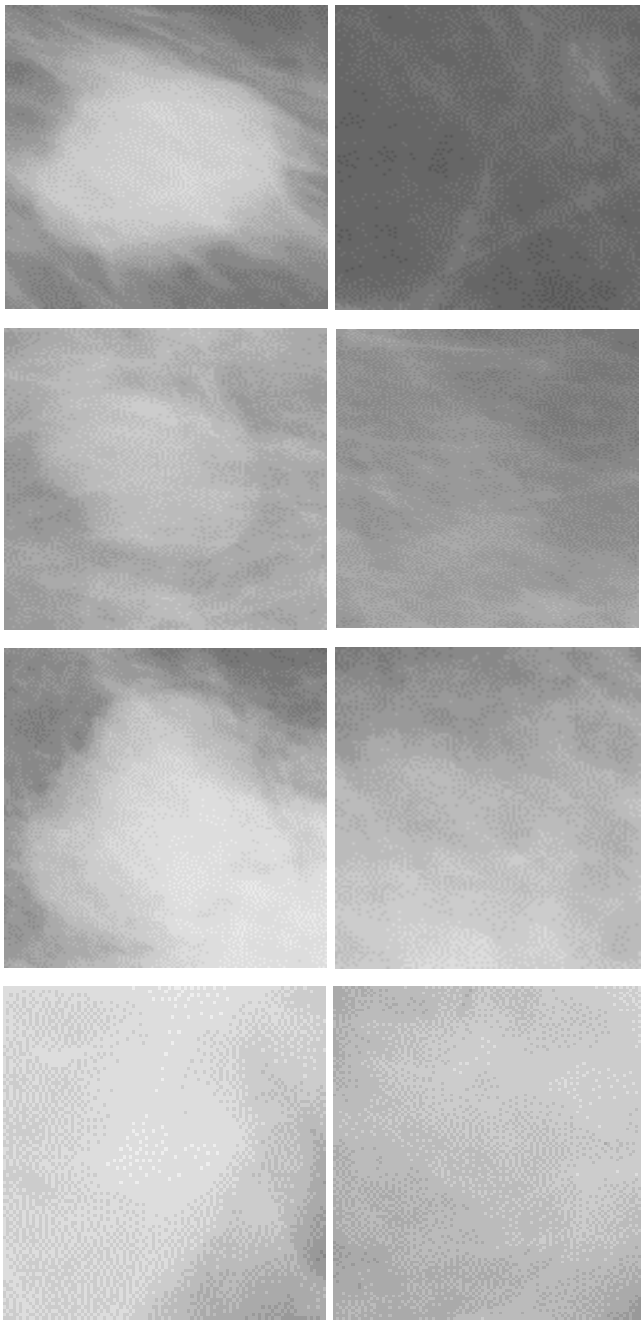


Figure 41: Sample with Abnormality – Normal Sample

Column (a) Region contains an abnormality, Column (b) Normal region
Row (a), (b), (c), (d): Sample from BIRADS Class a, b, c, d respectively

The ability of the corresponding features to characterize distinctly the abnormal and normal regions is evaluated using SVMs and RBF kernel function following the same principles as presented in chapter 4 and section 10.1.2. The most appropriate features are selected and integrated to form the new feature vector able to optimize the performance of

the classification process as proposed in chapter 9. It has been found that cluster shade, energy and entropy provide the best discrimination between normal and abnormal regions. Cluster shade measures the lack of symmetry. When cluster shade is high, the image is not symmetric. Entropy is a measure of non-uniformity or complexity of the texture. Energy is a measure of the uniformity. In this work, cluster shade, entropy and energy have been combined together to form a single feature vector, which is evaluated for its ability to classify abnormalities. Training models for the individual BIRADS classes and a training model for the entire set, are developed and subsequently utilized for abnormality detection in unseen mammograms.

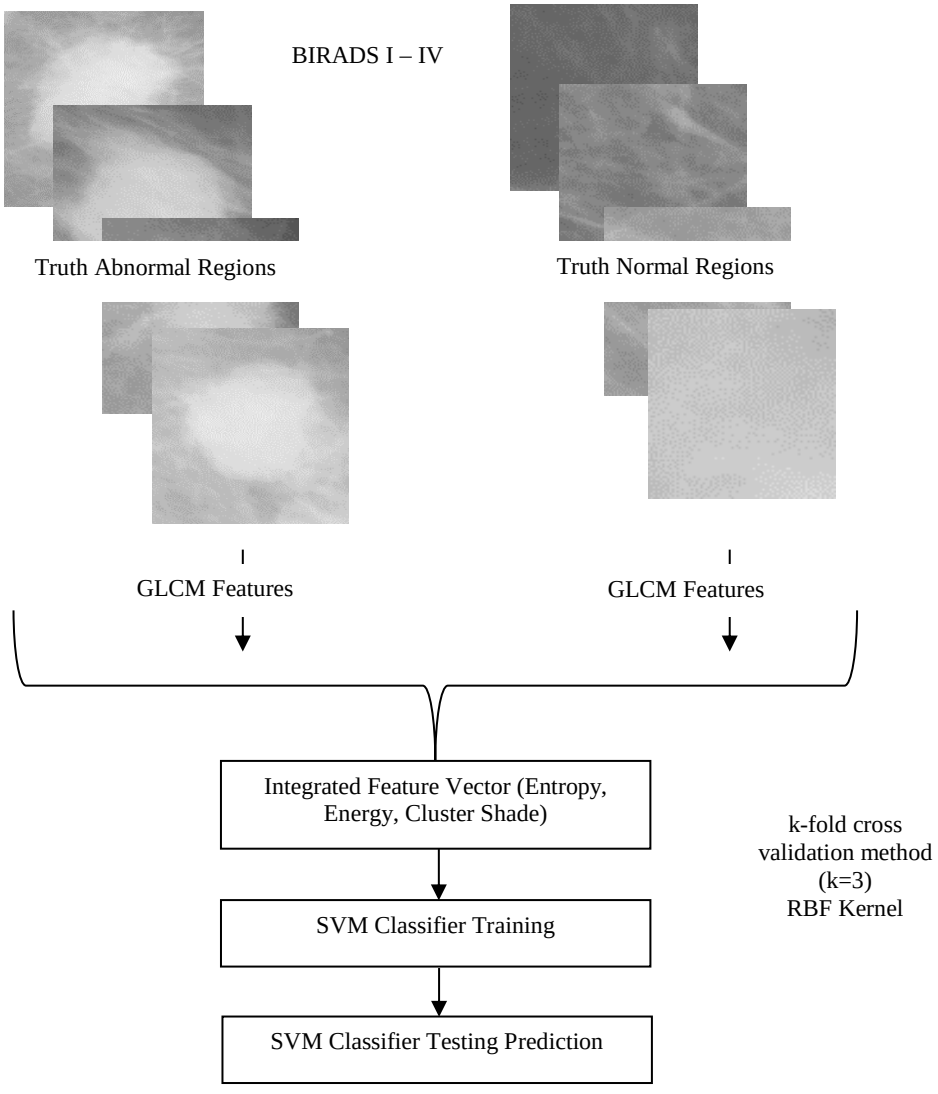


Figure 42: Feature integration and classification model for abnormality detection

10.2.3 Experimental results and discussion

The algorithm is evaluated on the Mammographic Image Analysis Society Digital Mammogram Database (MIAS) by comparing the resulting abnormality detection ability with the given experts ground truth. An expert has classified mammograms from MIAS

database according to the BIRADS annotations, obtaining 24 abnormal and 47 normal BIRADS I, 27 abnormal and 54 normal BIRADS II, 21 abnormal and 49 normal BIRADS III and 12 abnormal and 21 normal BIRADS IV samples. Half of the samples are used for developing the training models and the other half for testing the proposed model.

To evaluate the performance of the proposed method the average accuracy and the sensitivity of the classification process are calculated. The accuracy is calculated as the percentage of correctly classified regions as normal or abnormal. Sensitivity is a term that show the significance of a test related to the presence or absence of the disease. Sensitivity is the percentage of pathological or healthy subjects correctly detected. These metrics are calculated using the following equations:

$$sensitivity = \frac{TP}{TP + FN}, \qquad accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{10.1}$$

where TP are the regions that present abnormal regions on the ground truth images and classified by the proposed method as abnormal regions, FN are abnormal regions classified as normal regions, TN are normal regions classified as normal and FP are normal regions classified as abnormal. Table 16 presents the evaluation results of the four different BIRADS classes.

Table 16: Breast abnormality classification results

BIRADS I	**Accuracy %**	**Sensitivity %**	
		Abnormal regions	**Normal regions**
Cluster shade	83.10	75.00	87.23
Cluster prominence	57.26	18.17	78.15
Entropy	78.23	54.83	87.24
Energy	76.61	52.10	89.36
Contrast	66.20	49.52	81.23
Dissimilarity	59.15	16.67	80.85
Homogeneity	53.45	17.38	74.26
Autocorrelation	67.61	37.50	82.98
BIRADS II	**Accuracy %**	**Sensitivity %**	
		Abnormal regions	**Normal regions**
Cluster shade	91.36	88.89	92.59
Cluster prominence	57.26	18.17	78.15
Entropy	71.60	53.15	83.33
Energy	68.43	51.40	96.30
Contrast	58.20	42.52	73.23
Dissimilarity	49.15	16.67	65.85
Homogeneity	53.45	27.68	71.26
Autocorrelation	56.79	25.93	72.22
BIRADS III	**Accuracy %**	**Sensitivity %**	
		Abnormal regions	**Normal regions**
Cluster shade	81.43	76.19	83.67
Cluster prominence	53.45	19.17	68.15
Entropy	70.00	52.75	81.60
Energy	64.57	50.29	87.96
Contrast	67.14	16.52	75.92
Dissimilarity	49.15	26.67	59.85
Homogeneity	48.57	17.85	97.96
Autocorrelation	68.86	19.05	75.92
BIRADS IV	**Accuracy %**	**Sensitivity %**	
		Abnormal regions	**Normal regions**
Cluster shade	88.00	60.10	82.24
Cluster prominence	57.26	18.17	78.15
Entropy	76.00	49.25	88.48
Energy	68.00	48.36	80.95
Contrast	66.00	15.23	80.48
Dissimilarity	59.15	16.67	80.85
Homogeneity	53.45	17.38	74.26
Autocorrelation	67.61	37.50	82.98
ALL BIRADS CLASSES	**Accuracy %**	**Sensitivity %**	
		Abnormal regions	**Normal regions**
Cluster shade	95.14	91.80	97.30
Cluster prominence	81.43	64.75	91.63
Entropy	74.29	52.79	91.28
Energy	75.24	53.51	93.96
Contrast	70.95	19.59	95.25
Dissimilarity	59.15	16.67	80.85
Homogeneity	53.45	17.38	74.26
Autocorrelation	82.86	57.38	93.29

The features that provided the best discrimination for abnormality characterization as it can be observed from Table 16 are the cluster shade, entropy and energy. These features also achieve balanced results between the abnormal and normal sensitivity accuracy. Table 17 presents the results using the proposed feature integration and classification model.

Table 17: Breast abnormality classification results using combined features: energy, entropy and cluster Shade

BIRADS CLASSES	Accuracy %	Sensitivity %	
		Abnormal regions	Normal regions
BIRADS I	98.28	96.30	98.78
BIRADS II	97.26	96.15	97.45
BIRADS III	97.36	94.30	96.78
BIRADS IV	93.76	91.30	92.79
BIRADS ALL	99.12	99.06	99.42

As it can be observed from Table 17, the proposed method achieves high accuracy and high sensitivity in all BIRADS classes. A good generalization requires that the minimum number of training samples in a class must be at least three times the number of features [225]. Thus, the reduction in features down to three works well even in the case of the density class BIRADS IV where the abnormal samples were only 12.

The presented method compares favourably with other methods presented in literature. The proposed method by Younesi et al. has achieved a sensitivity of 90.73% using a linear transformation for enhancing the evaluation images and a fuzzy entropy feature to characterize the regions of interest [235]. This method was tested on 78 mammograms (30 normal & 48 cancerous) from the BIRADS and local databases. The model by Polakowski et al. has achieved 92% accuracy using the laws of textures for describing abnormal and normal regions. The method has been evaluated using 36 mammograms for the training and 236 for the testing phase [225]. Kom et al. utilized a local adaptive thresholding technique and obtained a sensitivity of 95.91% using 61 mammograms, 48 of them containing masses [236]. Eltonsy et al. implemented a concentric morphology model and indicated 92% sensitivity at testing 155 mammograms with masses and 82 without masses and using 164 training mammograms [236]. Karssemeijer and Te Brake used a combination of 9 stellated lesions, 10 architectural distortions and 31 normal mammograms to reach a sensitivity of 80% by utilizing second-order Gaussian derivative operators [237]. Petrick et al. utilized adaptive contrast enhancement and texture classification. A true positive rate of 90% has been indicated after a false positive reduction process where 15 texture features were used derived from the spatial grey-level dependence matrix [238].

In the above models, the evaluation has been performed on the entire data set and without taking into consideration the different density classes. In Olivers et al. review it is noticed that when testing the MIAS database the abnormal mammograms belonging to BIRADS I tend to show improved detection over abnormal mammograms belonging to other BIRADS category. This is expected since it is related to the increase in parenchymal tissue especially in BIRADS IV, which is mistaken for abnormal regions [22]. This has

been observed also in this work however, the obtained results are satisfactory for all breast density classes.

10.3 Conclusions

Different texture features and their combination are used with SVMs for breast density classification. Training the SVM on the integrated statistical distributions of all the texture features results in the best classification accuracy, when evaluated on the MIAS database on the corresponding 3-category classification provided.

The proposed method for classifying regions containing abnormalities in mammographic images by combining energy, entropy and cluster shade GLCM features into a single vector to train the SVM classification model proved to be sufficient for characterizing abnormalities in all BIRADS classes.

Chapter 11: Integration of Different Feature Descriptors for No-Reference Quality Assessment of Satellite Images

11.1 Introduction

In this work, a no-reference satellite image quality assessment model using feature appearance descriptors in a classification framework is proposed. Image feature descriptors such as grey-level co-occurrence matrices, normalized luminance coefficients, scale invariant feature transforms, local binary patterns and texton histograms are utilized in a support vector machine classifier for the development of the proposed model. Integrating different sets of image descriptors is fundamental for achieving good performance in quality assessment models. To have the maximum advantage from feature integration, each feature descriptor set should represent diverse information and a robust integration as proposed in chapter 9 should be applied. The obtained results reveal the robustness of the method for automatically perceiving the quality of the images under evaluation.

11.2 Related work

Satellite imagery is used in several scientific and commercial applications. A fundamental and critical objective for the use of satellite imagery is their quality. Image quality assessment (IQA) can be defined as a process to automatically predict the quality of distorted images. There are three major objective image quality assessment methods, no-reference image quality assessment (NR-IQA), full-reference image quality assessment (FR-IQA) and reduced reference image quality assessment (RR-IQA). In NR-IQA the only information that is used includes the image under evaluation. In FR-IQA methods both the image under evaluation and a pristine reference image is required. Finally, RR-IQA methods require some information regarding the reference image and the image under evaluation [239-240].

NR-IQA algorithms in literature are based on the prior knowledge of the distortion type used in the assessment and can be further classified as distortion-specific (DS) and non-distortion-specific (NDS). Most existing NR-IQA approaches are distortion specific using the distortion type as prior knowledge. Moorthy and Bovik develop a two-stage framework involving distortion identification, followed by a distortion-specific quality assessment, capable of assessing the quality of a distorted image across multiple distortion categories [241]. Saad et. al. in their general-purpose, non-distortion specific NR-IQA algorithm, used

natural scene statistical models of discrete cosine transform coefficients to perform distortion-agnostic quality assessment [242].

Very few researchers used texture analysis and texture features to approach the problem. Ye and Doermann used the visual codebook method and Gabor filter based features as appearance descriptors to capture image statistics and developed a NR-IQA model [243]. Recently Mittal et al. published two no-reference IQA models using measurable deviations from statistical regularities observed in natural images, with or without training on human-rated distorted images with good results. These models are based on the construction of a classification model using a collection of low-level statistical features [244-245].

In satellite imagery archives, information about the distortion type is not available in most of the cases and it is very difficult and time consuming to assess the quality of the images. A NR-IQA model that assesses image quality without knowledge of anticipated distortions, or human opinions of these distortions, should be developed in order to automatically evaluate satellite images. In these models it is important to select the most appropriate 'quality aware' features. An important process of the development of an image quality assessment model using image descriptors and SVM classifiers is the initial objective classification of the satellite images into categories, representing different qualities. Images can be effectively divided into quality categories by using human subjective judgments of quality. To achieve this, large-scale human studies should be performed, where human observers should rate a large number of images. The individual opinions could then average across and the mean opinion score could be obtained for each of the image. Based on the mean scores the different quality image categories can be defined. Subsequently, image descriptors derived from grey-level co-occurrence matrices, normalized luminance coefficients, scale invariant feature transforms, local binary patterns and texton histograms are extracted for each category and each colour band. An SVM classifier is applied to estimate the decision values for the individual feature descriptors. The most appropriate decision values are then integrated to form a distinct feature descriptor set. The optimized feature descriptor set is utilized to train an SVM classifier able to automatically predict the quality category of unknown satellite images. More details about the descriptors used can be found in chapter 4 of this thesis.

11.3 Proposed method

The initial objective classification of the satellite images into different quality categories was performed based on subjective experiments that were conducted in two separate sessions with 32 and 28 observers, respectively. Observers provided their quality ratings on

a continuous scale from 1 (higher quality) to 100 (lowest quality) and the mean opinion scores (MOS) were calculated to develop five subjective classification categories.

Figure 44 presents samples from the different categories. Class A includes images where the average quality score was less than 20, Class B includes images with quality score on the interval 21-40, Class C 41-60, Class D 61-80 and Class E higher than 80.

Figure 43: Satellite image samples from the five different quality categories

(a) Class A, (b) Class B, (c) Class C, (d) Class D, (e) Class E

Subsequently the different feature descriptors are extracted and evaluated using four different kernel functions (Linear, RBF, Polynomial and Sigmoid). The most appropriate are selected and integrated. The integration of the different feature descriptors and the development of the SVM classification model follows the principles as presented in chapter 9 and are shown in Figure 44.

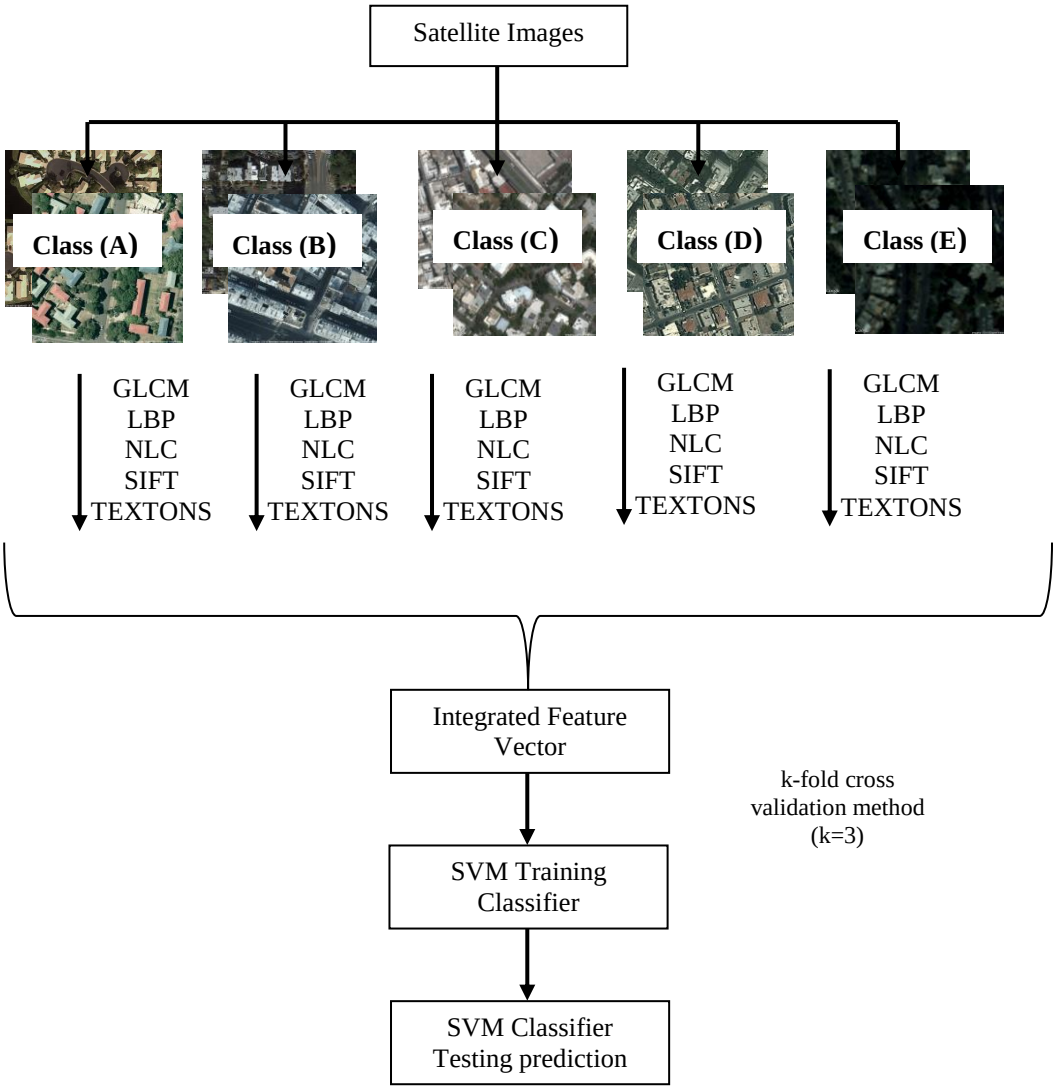


Figure 44: Proposed scheme for optimal feature descriptors integration

11.4 Experimental results and discussion

Table 18 presents the classification sensitivity and accuracy results using the individual features, their low-level integration and the proposed integration method as presented in chapter 9 and section 11.3 using different kernel functions. For the evaluation of the algorithm, 1250 satellite images were used for training and 682 images for testing, acquired using Google Earth application. The dataset as presented in chapter 3 and section 3.6 was utilized where the satellite images were divided into 6 patches and used to train and test the proposed classification model. The application of the integration function 11.6 leads to a new feature vector which include the decision value scores for the five different quality classes (M=5) using the five descriptor sets (T=5). Experiments were performed using four popular kernel functions (linear, RBF, polynomial, sigmoid) for both the individual feature vectors and the integrated feature vectors.

Table 18: Satellite image quality assessment results

LINEAR KERNEL SVM TRAINED ON:	**Sensitivity %**						**Accuracy %**
	CLASS A	**CLASS B**	**CLASS C**	**CLASS D**	**CLASS E**	**Average**	
GLCM	92.00	80.00	30.67	14.67	45.33	52.53	81.01
LBP	84.00	57.33	100.00	96.00	62.67	80.00	92.00
NLC	70.67	100.00	100.00	97.33	64.00	86.40	92.85
SIFT	100.00	92.00	100.00	100.00	62.67	90.93	94.37
TEXTONS	82.67	76.00	94.67	78.67	58.67	78.13	91.25
INTEGRATED (LOW-LEVEL)	98.67	100.00	100.00	98.67	76.00	94.67	97.78
INTEGRATED (PROPOSED)	94.77	100.00	100.00	100.00	100.00	98.94	99.67
POLYNOMIAL KERNEL SVM TRAINED ON:							
GLCM	85.33	64.00	29.33	13.13	45.33	47.46	78.99
LBP	98.67	66.67	100.00	93.33	37.33	79.2	91.68
NLC	61.33	98.67	94.67	98.67	57.33	82.13	92.56
SIFT	98.67	93.33	100.00	100.00	52.00	88.80	95.52
TEXTONS	73.33	80.00	93.33	86.67	49.33	76.53	90.61
INTEGRATED (LOW-LEVEL)	98.67	100.00	100.00	98.67	72.00	93.86	97.55
INTEGRATED (PROPOSED)	85.33	93.33	73.33	53.33	100.00	81.07	92.43
RBF KERNEL SVM TRAINED ON:							
GLCM	73.33	72.00	32.00	16.00	46.67	48.01	79.20
LBP	96.00	57.33	100.00	97.33	69.33	84.00	93.60
NLC	76.00	88.00	98.67	60.00	82.67	81.06	92.43
SIFT	98.67	94.67	100.00	100.00	66.67	92.00	96.80
TEXTONS	86.67	82.67	96.00	84.00	57.33	81.33	92.53
INTEGRATED (LOW-LEVEL)	98.67	100.00	100.00	98.67	73.33	94.13	97.65
INTEGRATED (PROPOSED)	92.00	100.00	100.00	74.67	100.00	93.33	97.33
SIGMOID KERNEL SVM TRAINED ON:							
GLCM	82.67	64.00	32.00	13.33	46.67	47.73	79.09
LBP	68.00	57.33	100.00	93.33	41.33	72.00	88.80
NLC	54.67	100.00	98.67	64.00	82.67	80.00	92.00
SIFT	100.00	92.00	100.00	100.00	61.33	90.66	96.27
TEXTONS	82.67	73.33	94.67	77.33	52.00	76.00	90.40
INTEGRATED (LOW-LEVEL)	98.67	100.00	100.00	98.67	76.00	94.66	97.87
INTEGRATED (PROPOSED)	96.00	100.00	100.00	97.33	100.00	98.67	99.47

The use of different features seems to capture different region characteristics. The SIFT captures local oriented features over different visual scales. LBP provide information whether the region of interest (ROI) is relatively uniform. Textons can be thought of as the building blocks of the present texture, the same way phonemes make up speech. Normalized luminance coefficient features describe how paired products vary in the presence of distortion. The co-occurrence matrix presents the repeated occurrence of configurations of pixels intensity values in the image. These configurations vary rabidly in

flat texture and gentler in rough texture. After the thoroughly investigation of the different feature descriptor sets evaluated using the four different kernel functions (Table 18) the results vary, thus the most appropriate combination should be selected for developing the optimal integrated feature vector. GLCMs achieve the best average sensitivity of 52.53% and accuracy of 81.01% when the linear kernel was utilized. LBPs achieve best average sensitivity of 84.00% and accuracy of 93.60% when the RBF kernel was used. The results using the NLC feature descriptor are similar for all the different kernels. However, using the linear kernel where the average sensitivity is 86.40% and the accuracy is 92.85%, the results are more balanced for all different classes. Promising results were derived using the SIFT descriptor where the average sensitivity using the RBF kernel is 92.00% and the accuracy 96.80%. Finally, textons trained and evaluated on SVM RBF kernel achieve average sensitivity of 81.33% and accuracy of 92.53%. Based on the above observations the most appropriate integrated feature descriptor set is derived using GLCM and NLC features, trained on SVM linear kernel combined with LBP, SIFT and TEXTONS trained on SVM RBF kernel.

The proposed method is a general integration scheme that achieves balanced results for all classes and all kernels. When trained and evaluated on linear kernel it reaches the best average sensitivity of 98.94% and accuracy of 99.67%. In addition, as can be seen from Table 19 it has the ability to produce superior results compared to a low-level integration scheme where all the features were combined together at feature level. Furthermore, the integrated feature descriptor set depends on all the outputs of the single feature descriptor classifiers, and the coefficients of the proposed training model are learned automatically. Therefore, there is no need to include any time-consuming procedures for estimating the appropriate weights to combine different optimized feature descriptor sets, as is usually recommended in high-level integration schemes. Finally, it should be noted that the outputs could be derived from a combination of different large margin classifiers, and not only from SVMs.

11.5 Conclusions

A NR-IQA model for satellite images is proposed. The quality can affect several image processing applications for extracting useful information. This work suggests that level-of-detail metrics such as edge-based contrast and region-based segmentation can be utilized to initially classify the satellite images into five different quality categories. The feature descriptor sets and their combination derived from the different categories can be used to train SVMs for the development of an image quality assessment model. Subsequently the

model can be applied to classify satellite images with unknown quality into these five different categories. The proposed model can be considered as a no reference quality assessment model since no additional information is used except the under evaluation image. The new developed scheme for the optimal integration of the various image feature sets at the SVM level, clearly shows better performance. Experimental results suggest that when the proposed integration method is used the accuracy of the classification model is increased.

Chapter 12: Conclusions

12.1 Introduction

The research presented in this thesis is involved with image segmentation and classification in the presence of intensity and/or spectral heterogeneity in complex scenes as presented in natural images. It has been inspired by the difficulty to develop methods with the ability to work effectively in a diversity of fields such as mammographic and satellite images. A series of algorithms have been presented, which enable both visual and quantitative assessment of images, and which can be incorporated in real world applications such as CAD systems and urban planning or monitoring. These algorithms utilize novel concepts from active contours, texture analysis, computer vision and image processing. This chapter summarises the main contributions, conclusions and limitations of this thesis. Suggestions of further research directions are also presented.

12.2 Summary and contributions

The image segmentation and classification methods that have been developed include: mammogram segmentation, breast density classification, breast abnormality detection, building structures and the corresponding building shadows detection, building structures height estimation and satellite images no-reference quality assessment. The thesis also suggests how the different methods are combined and utilized without or very little customization in a diversity of fields for quantitative analysis. It must be noted that for all the developed methods in the mammographic imagery domain, significant attention was given for achieving balanced results in all the different mammogram classes. In satellite imagery domain, emphasis was given in dealing with big data and work with images that do or do not ad-here high quality standards.

12.2.1 Image segmentation in the presence of intensity heterogeneity and feature complexity.

Image segmentation is the initial process in image processing and analysis. It is used to detect the regions of interest, and an accurate, reliable segmentation is a prerequisite for effective image understanding. In particular, it restricts the area where further processing is applied. Intensity or spectral heterogeneity and feature complexity needs to be addressed for effectively producing region of interest segmentations in a variety of fields. Two different

approaches have been proposed in this thesis for overcoming the erroneous segmentations derived by the presence of intensity or spectral heterogeneity and feature complexity.

In the first approach, pre-process methods have been designed to reduce intensity or spectral heterogeneity and feature complexity in mammographic and satellite images. Subsequently, optimized active contour models encoding indices derived from both the original and pre-processed images have been developed and applied for selectively delineating regions of interest, in the analysed scenes. In mammographic images, the breast segmentation algorithm was developed to ensure an accurate processing region for breast density quantification and abnormalities detection. The breast segmentation algorithm removes background artefacts and restricts the search region. Local enhancement based on local and global intensity statistics is initially performed. The breast region becomes more homogeneous without losing important details and an optimized active contour model formulated on level sets is applied to identify the skin-air boundary. In the proposed segmentation of building structures method, the spectral heterogeneity and feature complexity that often occurs has been addressed using indices derived from the HSV and RGB colour systems. A colour clustering method has been developed and used to pre-process the satellite images for eliminating the vegetation areas and shadows resulting in less heterogeneous and complex scenes. Furthermore, *k*-means is applied at the pre-process stage, for designing optimum initial curves to drive the active contour model for detecting faster and more accurately building structures. The image is decomposed in HSV colour components and the intensity properties of the saturation and value representations are used to develop an optimized active contour model to selectively detect building structures. The encoded energy terms associated with saturation and value components of the image, biased active contours to delineate several buildings, which are represented with lower intensity values than the average and often missed by the traditional active contour models. In the proposed method of building shadow detection, the image is pre-processed with a filter able to reduce the effect of intensity heterogeneity and highlight the shadows. The filter is designed by utilizing information derived from the frequency and spatial analysis of the satellite image signal. The most appropriate parameter settings of the filter are automatically estimated. Furthermore, the radiometric feature of the shadows is estimated automatically and is encoded in the active contour model to attract the contours to delineate the existing building shadows.

The second approach involves the development of an optimized localized active contour model to address extensive initialization, under-detection and over-detection that is noticed when traditional localized active contour models evaluated on mammographic images for

breast region segmentation and on satellite images for building structures detection. A kernel function is implemented and used to form a new localized energy term in an active contour model for reducing the intensity heterogeneity during the curve propagation process. The proposed optimized localized active contour model worked effectively in the presence of spectral or intensity heterogeneity. Regions or objects of interest with arbitrary shapes, sizes or even presented with intensities that significantly vary, are successfully detected. The proposed algorithm addressed the extensive initialization, under-detection and over-detection limitations of the localized active contour models and optimizes the overall segmentation procedure in breast region segmentation using mammography images, as well as the detection of building boundaries in satellite images.

12.2.2 Image classification in the presence of intensity heterogeneity and feature complexity.

Classification is an important process in image analysis and understanding. In complex scenes, as the number of object classes and the degree of within-class heterogeneity increases, classification rules become difficult to be formed, leading to varying classification accuracies. Key to the success of a classification algorithm in the presence of intensity or spectral heterogeneity and/or feature complexity is the selection and integration of the most appropriate feature descriptor sets. In this thesis, the proposed classification method utilized an effective scheme of integrating multiple feature descriptors in a support vector machine framework. Initially, an integration scheme of the most appropriate features at feature level for tasks such as breast density and breast abnormality classification was developed. Breast density is a risk factor for breast cancer. Breast pattern classification potentially has many applications in clinical practice such as evaluation of screening intervals and breast density quantification. Texture is used as the basis for breast density classification and characterization. Therefore, a variety of texture features such as scale invariant feature transform, statistical distribution of textons and local binary patterns were integrated to form a new feature vector to train an SVM classifier to classify breast density. The development of effective methods to assist in the early detection and/or classification of mammographic abnormalities is of utmost importance. Second order statistical features such as cluster shade, entropy and energy have been combined together at feature level to form a single feature vector and used to train a classifier to identify breast abnormalities.

The integration method by selecting only the most appropriate features and kernel functions at feature level achieved promising results thus was extended by incorporating additional feature descriptor sets for the development of a multipurpose classifier. A so-

called mid-level integration method was developed, where the decision values corresponding to the various classes of each sample image are integrated to form a new feature descriptor. An additional SVM classifier responsible to automatically weight the information derived from the different feature descriptor sets was applied. The parameters of the model, learned by applying an optimization process and the weights of the different feature descriptor sets, were automatically estimated. The model was evaluated in a variety of classification tasks in mammographic and satellite images domains and proved to be effective, robust and efficient.

12.2.3 Integrating image segmentation and classification methods for quantitative and qualitative image analysis.

The information derived from the proposed methods can be combined and utilized for acquiring quantitative and qualitative information for further understanding the analysed scenes.

In mammographic image domain using the suggested optimized localized active contour model, breast region can be identified and pectoral muscle if exists can be removed. Subsequently, the proposed classification method can be applied for breast density characterization. Differential of Gaussian filter (DoG) is considered and recorded in literature as a filter able to highlight those frequency ranges where an abnormality exists [225]. Thus, using the OLAC after applying the DoG filter, regions that may include abnormalities could be identified. Customized morphological operations can be also applied for removing misleading artefacts. Subsequently, a bounding box centred on the centroid of the detected masks, able to include the potential abnormality region can be applied for defining the regions of interest. Abnormalities are expected to have various sizes ranging from 0.5 cm to 4.0 cm as reported in literature [22, 225, 236] and thus a bounding box of size $S = 280 \times 280$ pixel, can be used for defining the regions of interest of a potential abnormality. Finally, the regions of interest are evaluated using the proposed abnormality classification model to label these regions as normal or abnormal. A flow chart of the algorithm that combines the proposed methods for qualitative and quantitative information acquisition from mammographic images is presented in Figure 45. More results are presented in Figure 46.

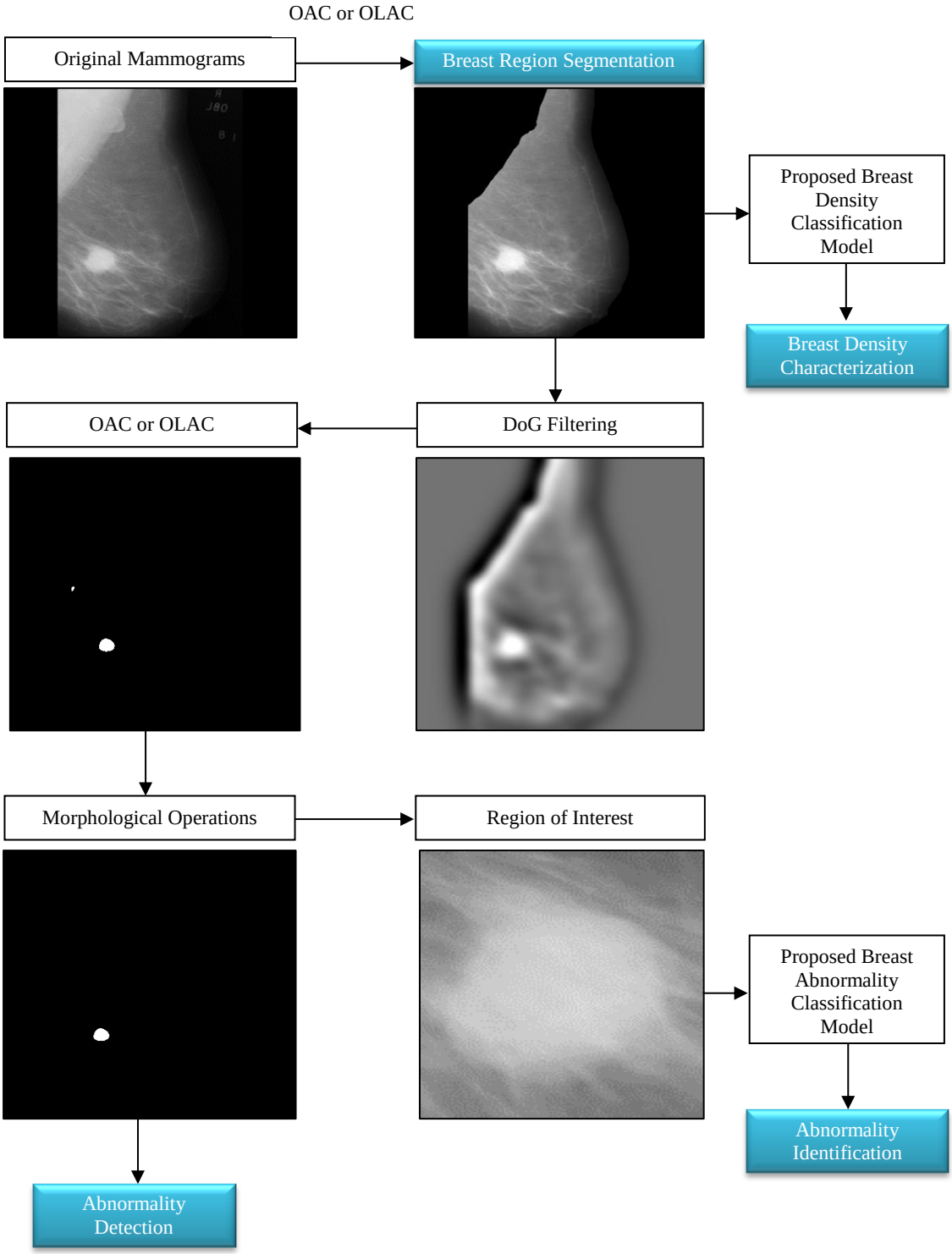


Figure 45: Integration of the proposed methods model for mammographic images analysis

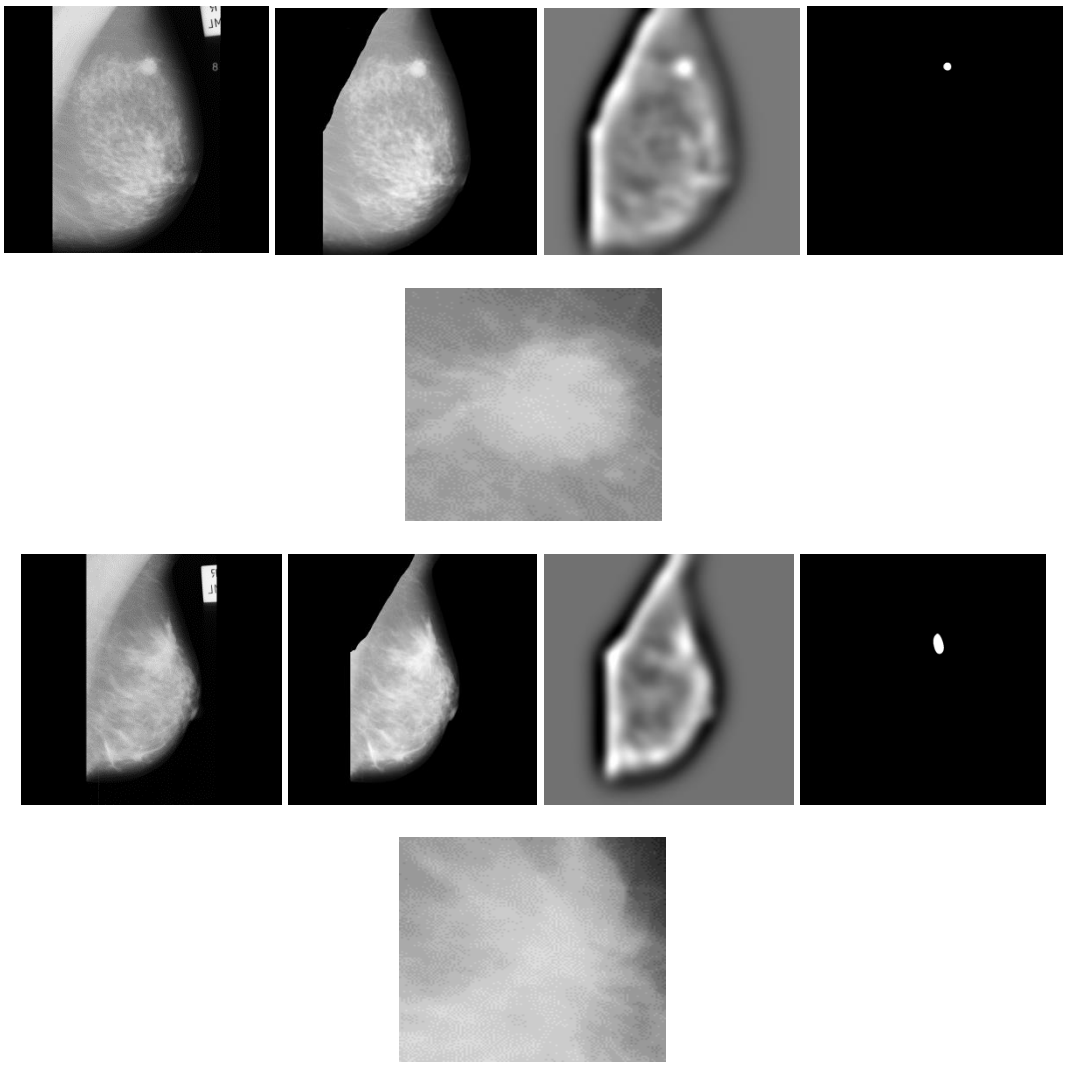


Figure 46: Abnormality detection and identification

The detection of building structures method along with the shadow of the corresponding buildings detection and the estimation of their height can be utilized for urban modelling. Furthermore, the no-reference image quality assessment model can be used to assist the overall performance of segmentation and classification methods by assigning a quality score to the under evaluation scenes or images and help the further customization and optimization of the proposed models. A flow chart of the algorithm that combines the proposed methods for qualitative and quantitative information acquisition from satellite images is presented in Figure 47. More results are presented in Figure 48.

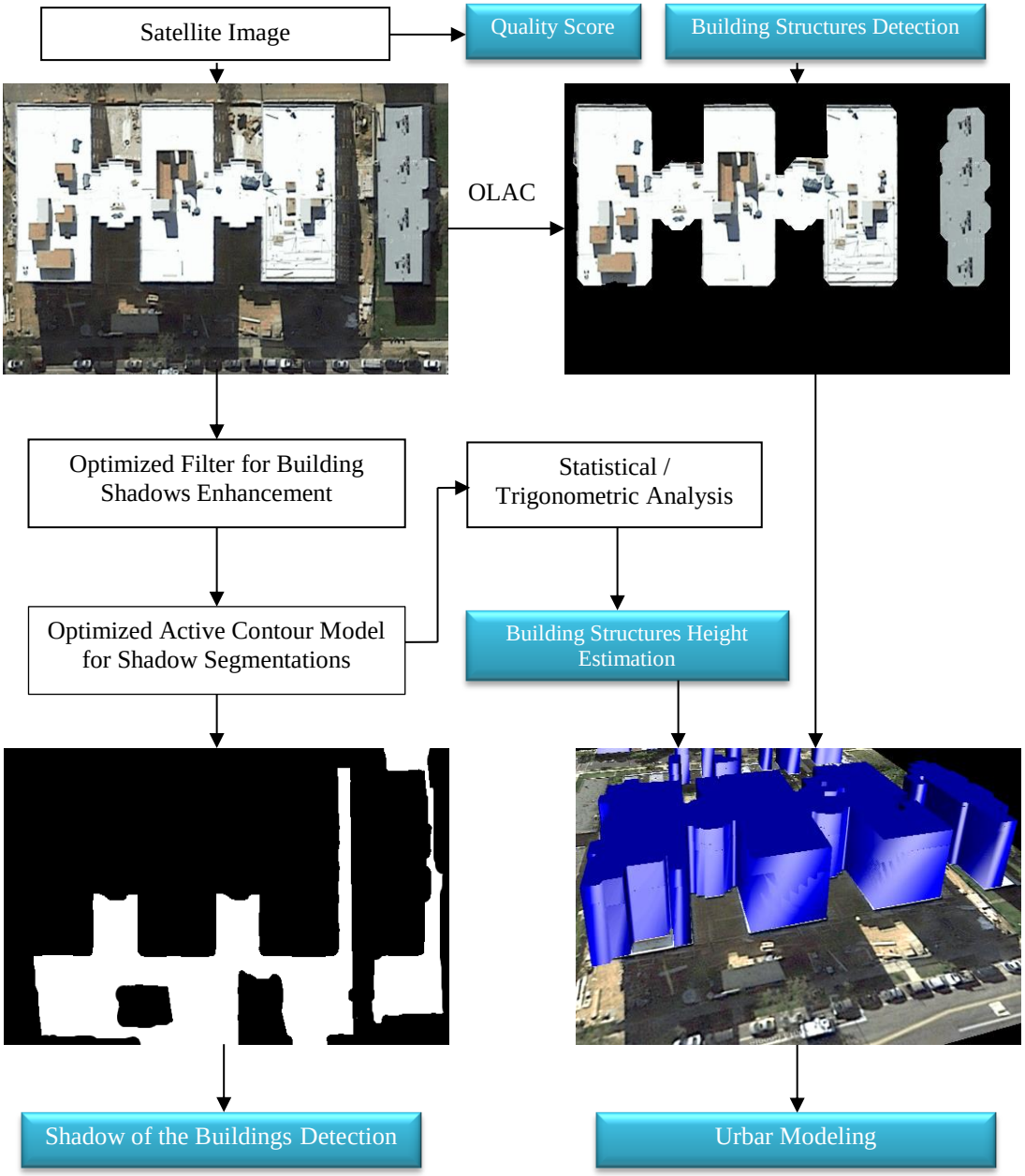


Figure 47: Integration of the proposed methods model for satellite images analysis

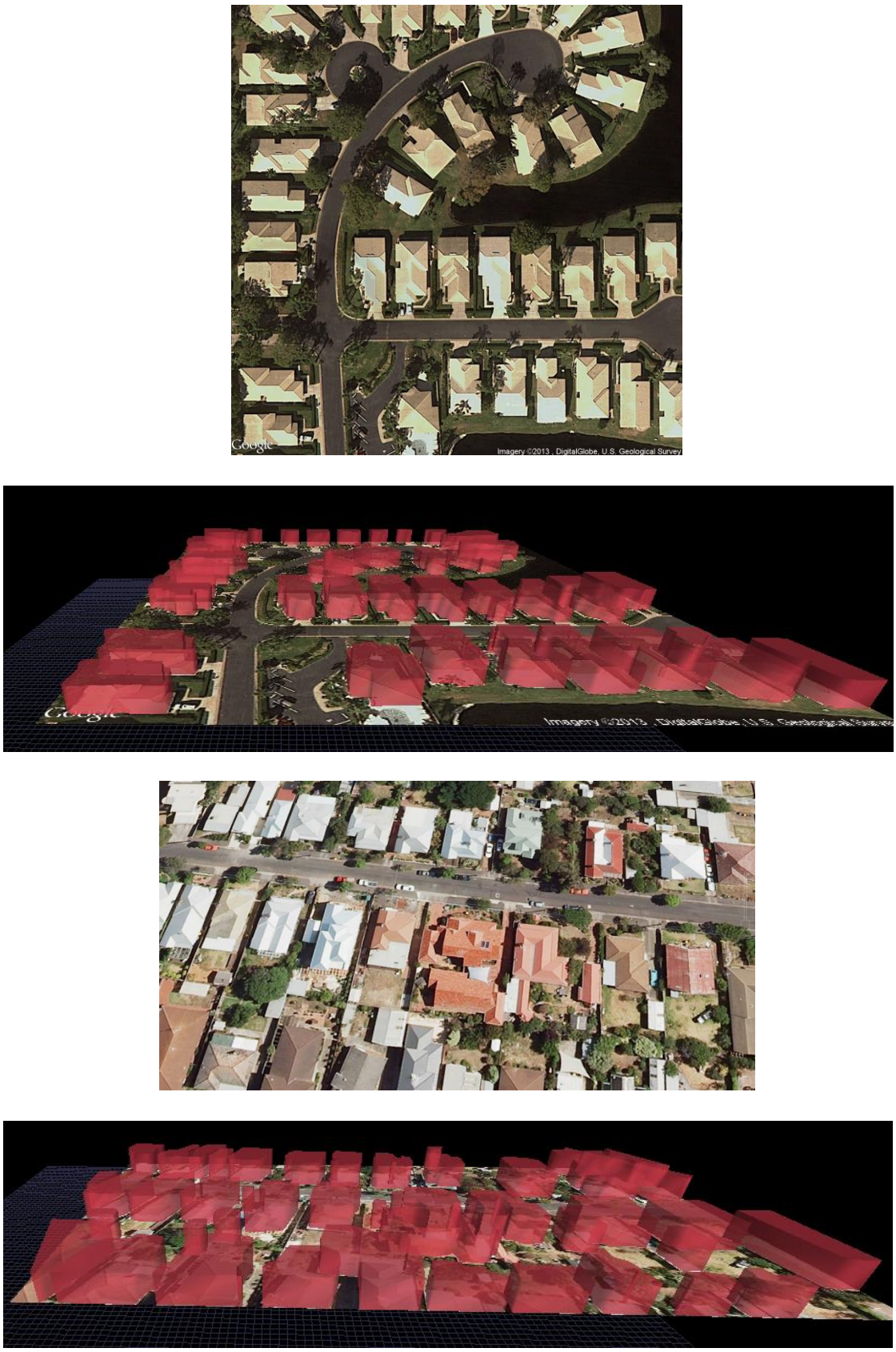


Figure 48: Initial urban modelling

12.3 Future work

The proposed segmentation and classification models can be extended for further optimization and evaluation in different fields and sectors. Prior to such future work, a

number of other issues could be investigated. In segmentation domain in the proposed optimized active contour segmentation models different image descriptors could be incorporated and evaluated for their ability to perform better in the presence of intensity or spectral heterogeneity. In classification domain since the integration of classifiers can enhance the performance of any one classifier, various strategies such as committee machine classifier, co-training and deep learning should be examined for the design and implementation of more robust integration schemes [246-249].

12.3.1 Mammography screening

A Computer-Aided System (CAD) is a set of automatic or semiautomatic tools developed to assist radiologists in the early abnormality detection and evaluation of mammographic images. CAD systems can aid in the interpretation and early detection of breast cancer. Breast region segmentation, breast abnormality detection and characterization along with information on breast density and features can be most important in the development of Computer Aided Detection (CAD) systems with higher sensitivity and specificity for use in clinical settings. The results of this thesis in mammography screening suggest that the proposed segmentation and classification methods can be incorporated in the development and application of CAD systems in the screening process. However, future work should include further evaluation and optimization of the proposed methods using different datasets to achieve the sensitivity and specificity required in the clinical setting. It is of utmost importance to evaluate the proposed methods on full-field digital mammography since in the recent years it is gaining ground compared to the typical film-based mammography investigated in this thesis. Furthermore, the proposed methods should be evaluated on different modalities such as MRI of breast. Finally, texture analysis can be extended and evaluated for density classification using the BIRADS annotations. New methods can be developed using the definition of textons and texton spatial dependence matrices, incorporating statistical and structural information that is not based on intensity. The methods will be evaluated using different distances, at different orientations and with different feature descriptors. The corresponding results can be used to build a complete CAD system that acknowledges the breast density.

12.3.2 Satellite imagery

Urban model design is a subject that is concerned with the shape, the surface and the physical arrangement of all kinds of urban elements. Satellite remote sensing images provide a valuable and irreplaceable source for urban monitoring. However, creation of

three-dimensional digital urban models using conventional satellite images is cumbersome for many applications. The work of this thesis can be extended for the design of complete deterministic three-dimensional urban models. Since the building footprints and their corresponding heights are identified using the proposed building segmentation and height estimation method, the next process should involve the identification of other urban elements such as roads. A similar approach based on an optimized active contour model can be incorporated for delineating the existing roads in the analyzed scenes. Initial experiments can be seen in Figure 49.

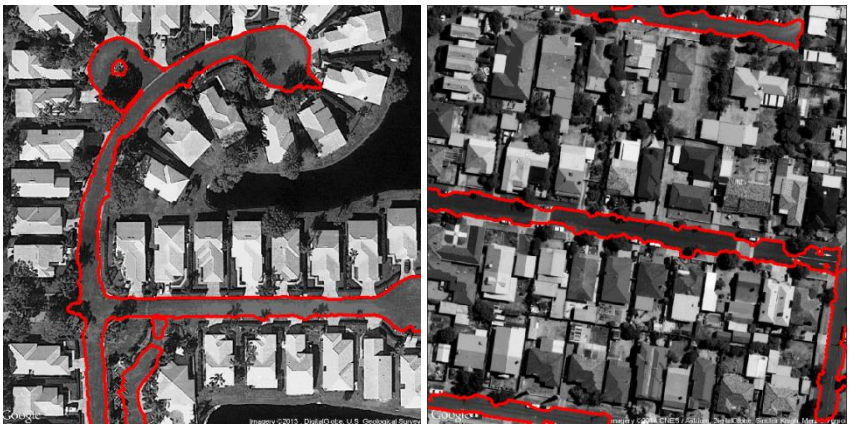


Figure 49: Roads extraction

Furthermore, for a photorealistic result, the reconstruction of all urban elements should include a close representation of the actual materials and elements used for the building roofs, building exteriors and terrain. The identification of building materials and structural objects in landscape urban scenes can serve wireless communications since the various materials and objects used in building constructions affect signal distribution and propagation. In wireless communication and urban monitoring applications is often essential to identify the facade elements such as walls, windows, doors, balconies and their shape or structure along with their construction material. This task is not only difficult because of the vast diversity of buildings, but also because of shadows, occlusions and reflections [250].

Existing work mainly utilized datasets such as the “Ecole Centrale Paris Facades Database” [251] and the eTRIMS database [252]. However, the methods for identifying the facade elements as presented in literature using the above datasets mainly characterize the different elements as roof, building or wall, balcony, door, window, car, pavement, road, sky and vegetation. Thus, novel methods are needed to identify not only the element but also the construction material type.

The proposed classification model using texture feature descriptors as described in this work can be extended and utilized for identifying facade elements and materials. The availability of facade images of building structures in applications such as Google Earth can be utilized for the development of a texture library to represent all the different materials and structural or non-structural objects. Subsequently, this library along with spectral information can be utilized for the development of an optimized classification model for labelling the building materials and objects. This information can be used to produce a more photorealistic, three-dimension visualization of the analyzed scenes.

References

[1] R. C. Gonzalez and R. E. Woods, “Digital Image Processing,” Third edition, Prentice Hall, 2008.

[2] R. Szeliski, “Computer Vision: Algorithms and Applications,” ISBN: 978-1-84882-934-3 (Print), 978-1-84882-935-0 (Online), Texts in Computer Science, SpringerLink, 2011.

[3] S. Pedroudi, “Texture in Mammographic Image Analysis,” PhD thesis, Wolfson Medical Vision Laboratory Department of Engineering Science, University of Oxford, 2005.

[4] K.R.M. Adeline, M.Chen, X. Briotte, S.K. Pang, N. Paparoditis, “Shadow detection in very high spatial resolution aerial images: A comparative study,” Journal of Photogrammetry and Remote Sensing, vol. 80, pp. 21-38, 2013.

[5] A. Oliver, “Automatic Mass Segmentation in Mammographic Images,” PhD thesis, University of Girona, 2007.

[6] S. Ghaffarian and S. Ghaffarian, “Automatic building detection based on supervised classification using high resolution Google Earth images,” Journal of Photogrammetry and Remote Sensing XL-3, pp. 101–106, 2014.

[7] G. Liasis and S. Petroudi, “Estimation of the Breast Boundary in Mammograms Using Level Sets,” Journal of Medical Imaging and Health Informatics, Vol. 1, pp. 1–8, 2011.

[8] G. Liasis and S. Petroudi, “Estimation of the Skin-Air Interface in Mammograms Using Level Sets,” In IEEE 1st Middle East Conference on Biomedical Engineering (MECBME’11), Sharjah, UAE, Feb. 22-25, 2011.

[9] G. Liasis and S. Stavrou, “Optimizing Level Set Initialization for Satellite Image Segmentation,” In IEEE 20th International Conference on Telecommunications (ICT), May 6-8, Casablanca, Morocco, 2013.

[10] G. Liasis and S. Stavrou, “Building Extraction in Satellite Images using Active Contours and Colour Features,” International Journal of Remote Sensing, vol. 37 (5), pp. 1127-1153, 2016. DOI**:** 10.1080/01431161.2016.1148283.

[11] G. Liasis and S. Stavrou, “Satellite Images Analysis for Shadow Detection and Building Height Estimation,” Accepted for publication in ISPRS Journal of Photogrammetry and Remote Sensing, 2016.

[12] G. Liasis and S. Stavrou, “An Optimized Localized Active Contour Segmentation Model with Application to Intensity Heterogeneous and Complex Images,” Submitted for publication in IEEE Transactions on Image Processing, 2016.

[13] G. Liasis and S. Petroudi, “Evaluation of Grey-Level Co-Occurrence Features for Abnormality Detection,” In IEEE, 4th Cyprus Workshop on Signal Processing and Informatics, University of Cyprus, Nicosia, Cyprus, July 14, 2011.

[14] G. Liasis and S. Petroudi, “Evaluation of Different Texture Features for Breast Density Classification,” In IEEE, 5th Cyprus Workshop on Signal Processing and Informatics, University of Cyprus, Nicosia, Cyprus, July 10, 2012.

[15] G. Liasis, C. Pattichis and S. Petroudi, “Combination of Different Texture Features for Mammographic Breast Density Classification,” In IEEE 12th International Conference on Bioinformatics and Bioengineering (BIBE), Cyprus, 2012.

[16] G. M. Farinella, S. Battiato and R. Cipolla “Advances in Computer Vision and Pattern Recognition,” ISSN 2191-6586 ISSN 2191-6594 (electronic), ISBN 978-1-4471-5519-5 ISBN 978-1-4471-5520-1 (eBook), DOI 10.1007/978-1-4471-5520-1, Springer, 2013.

[17] E. Krupinski, “The future of image perception in radiology: Synergy between humans and computers,” Acad Radiologist, vol. 10, pp. 1-3, 2003.

[18] M. P. Sampat, A. C. Bovik, G. J. Whitman and M. K. Markey, “A model-based framework for the detection of spiculated masses on mammography,” Medical Physics, vol. 35, pp. 2110-2123, 2008.

[19] J. J. Fenton, S. H. Taplin, P. A. Carney, L. Abraham, E. A. Sickles, C. D’Orsi, E.A. Berns, G. Cutter, R. E. Hendrick, W. E. Barlow and J. G. Elmore, “Influence of computer-aided detection on performance of screening mammography,” New England, Journal of Medicine vol. 536 (14), pp. 1399-1409, 2007.

[20] H. Zhang and S. W. Foo, “Computer aided detection of breast masses from digitized mammograms,” IEICE Transactions Information Systems, E89D (6), 1955-1961, 2006.

[21] P. Taylor, J. Champness, R. Given-Wilson, K. Johnston and H. Potts, “Impact of computer-aided detection prompts on the sensitivity and specificity of screening mammography,” Health Technology Assessment vol. 9 (6), pp. 1-58, 2005.

[22] A. Oliver, J. Freixenet, J. Marti a, E. Perez, J. Pont, E.R.E. Denton and R. Zwiggelaar, “A review of automatic mass detection and segmentation in mammographic images,” Medical Image Analysis Journal, vol. 14, pp. 87–110, 2010.

[23] K. Karantzalos and N. Paragios, “Large-Scale Building Reconstruction through Information Fusion and 3D Priors,” IEEE Transactions on Geoscience and Remote Sensing, vol. 48, No.5, pp. 2283-2296, 2010.

[24] J. Hu, S. You and U. Neumann, “Approaches to large-scale urban modeling,” IEEE Computer Graphics Application, vol. 23, no. 6, pp. 62–69, 2003.

[25] Z. Chen, Y. Zhang, B. Guindon, T. Esch, A. Roth and J. Shang, “Urban land use mapping using high resolution SAR data based on density analysis and contextual information,” Canadian Journal of Remote Sensing, 38(06), pp. 738-749, 2012.

[26] A. K. Jain, “Fundamentals of Digital Image Processing,” Englewood Cliffs, NJ: Prentice-Hall, 1989.

[27] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. H. Romeny, J. B. Zimmerman, and K. Zuiderveld, “Adaptive histogram equalization and its variations,” Computer Vision Graphics and Image Processing, vol. 39, (3), pp. 355–368, 1987.

[28] J. A. Stark, “Adaptive image contrast enhancement using generalizations of histogram equalization,” IEEE Transactions on Image Processing, vol. 9 (5), pp. 889–896, 2000.

[29] J. B. Zimmerman, S. M. Pizer, E. V. Staab, J. R. Perry, W. McCartney, and B. C. Brenton, “An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement,” IEEE Transactions on Medical Imaging, vol. 7 (4), pp. 304–312, 1988.

[30] D.C. Chang and W.R. Wu, “Image contrast enhancement based on a histogram transformation of local standard deviation,” IEEE Transactions on Medical Imaging, vol. 17 (4), pp. 518–531, 1998.

[31] I. M. Fuks, “Radar contrast polarization dependence on subsurface sensing,” In IEEE International Geoscience and Remote Sensing Symposium, vol. 3, pp. 1455–1459, 1998.

[32] P.C. Wu, F. C. Cheng, and Y. k. Chen, “A Weighting Mean-Separated Sub-Histogram Equalization for Contrast Enhancement,” IEEE Transactions on Biomedical Engineering and Computer Science, 2010.

[33] M. Kaur, Jas. Kaur and Jap. Kaur, “Survey of Contrast Enhancement Techniques based on Histogram Equalization,” International Journal of Advanced Computer Science and Applications, vol. 2, (7), 2011.

[34] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3d transform-domain collaborative filtering,” IEEE Transactions on Image Processing, vol. 16 (8), pp. 2080–2095, 2007.

[35] L. Sendur and I. W. Selesnick, “Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency,” IEEE Transactions on Signal Processing, vol. 50, (11), pp. 2744–2756, 2002.

[36] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, “A joint interand intrascale statistical model for Bayesian wavelet based image denoising,” IEEE Trans. Image Process., vol. 11, no. 5, pp. 545–557, May 2002.

[37] J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli, “Image denoising using a scale mixture of Gaussians in the wavelet domain,” IEEE Trans. Image Process., vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

[38] J. Guerrero-Colon and J. Portilla, “Two-level adaptive denoising using Gaussian scale mixtures in overcomplete oriented pyramids,” presented at the IEEE Int. Conf. Image Process., Genova, Italy, Sep. 2005.

[39] N. Otsu, “A Threshold Selection Method from Grey-level Histograms,” IEEE Transactions on Systems Man and Cybernetics, vol. 9 (1), pp. 62-66, 1979.

[40] X. Yang, X. Shen, J. Long and H. Chen, “An Improved Median-based Otsu Image Thresholding Algorithm,” Conference on Modelling, Identification and Control, AASRI Procedia, vol. 3, pp. 468–473, 2012.

[41] K. Sharma, C. Kamargaonkar and M. Sharma, “An Improved Image Segmentation Algorithm Based on Otsu Method,” International Journal of Engineering Research and Technology, vol. 1 (6), pp. 2278-0181, 2012.

[42] R. Kumar and A. Arthanariee, “A Comparative Study of Image Segmentation Using Edge-Based Approach,” World Academy of Science, Engineering and Technology International Journal of Mathematical, Computational, Physical and Quantum Engineering vol. 7 (3), 2013.

[43] P. F. Felzenszwalb and D. P. Huttenlocher, “Efficient graph-based image segmentation,” International Journal of Computer Vision, vol. 59 (2), pp. 167–181, 2004.

[44] J. Shi and J. Malik, “Normalized cuts and image segmentation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8 (22), pp. 888–905, 2000.

[45] D. Comaniciu and P. Meer, “Mean shift: a robust approach toward feature space Analysis,” IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 24 (5), pp. 603-619, 2002.

[46] P. Arbelaez, M. Maire, C. Fowlkes and J. Malik, “Contour Detection and Hierarchical Image Segmentation,” Technical Report UCB/EECS-2010-17, EECS Department, University of California, Berkeley, PAMI, 2010.

[47] M. Kass, A. Witkin and D. Terzopoulos, “Snakes-active contour models,” International Journal of Computer Vision, vol. 1, pp. 321-331, 1987.

[48] S. Osher, and J. Sethian, “Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations,” Journal of Computational Physics, no. 79, pp. 12-49, 1988.

[49] T. Chan, and L. Vese, “Active contours without edges,” IEEE Transactions on Image Processing, vol. 10, (2), pp. 266-277, 2001.

[50] V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic active contours,” International Journal of Computer Vision, vol. 22, pp.61-79, 1997.

[51] O. Bernard, D. Friboulet, P. Thevenaz, and M. Unser, “Variational B-Spline level-set: A linear filtering approach for fast deformable model evolution,” IEEE Transactions on Image Processing, vol. 18 (6), pp.1179-1191, 2009.

[52] S. Lankton and A. Tannenbaum, “Localizing region-based active contours,” IEEE Transactions on Image Processing, vol. 17 (11), pp.2029-2039, 2008.

[53] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, “Minimization of region- scalable fitting energy for image segmentation,” IEEE Transactions on Image Processing, vol. 17 (10), pp.1940-1949, 2008.

[54] Y. Shi and W. C. Karl, “A real-time algorithm for the approximation of level-set based curve evolution,” IEEE Transactions on Image Processing, vol. 17 (5), pp.645-656, 2008.

[55] R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification, 2nd Edition, Wiley-Interscience, San Diego, 2001.

[56] D. Lu and Q. Weng, “A survey of image classification methods and techniques for improving classification performance,” In International Journal of Remote Sensing, vol. 28 (5), 2007.

[57] Bosch, A. Zisserman, and X. Munoz, “Representing shape with a spatial pyramid kernel,” In CIVR, 2007.

[58] M. Varma and D. Ray, “Learning the discriminative power-invariance trade-off,” In ICCV, 2007.

[59] G. Griffin, A. Holub, and P. Perona, “Caltech-256 object category dataset,” Technical report, CalTech, 2007.

[60] J.-L. Yuan and T. L. Fine, “Neural-network design for small training sets of high dimension,” IEEE Trans. Neural Networks, vol. 9, pp. 266–280, 1998.

[61] H. Ishibuchi, T. Nakashima, Improving the performance of fuzzy classifier systems for pattern classification problems with continuous attributes, IEEE Transactions Industrial Electronics, vol. 46 (6), pp. 1057–1068, 1999.

[62] V. Malik, A. Gautam, A. Sahai, A. Jha and A. Singh, “Satellite Image Classification Using Fuzzy Logic,” International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, vol. 2 (2), 2013.

[63] J. Demsar, “Statistical Comparisons of Classifiers over Multiple Data Sets,” The Journal of Machine Learning Research,vol. 7, pp. 1-30, 2006.

[64] C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, pp. 273–297, 1995.

[65] O. Chapelle, P. Haffner, and V. Vapnik, “Support vector machines for histogram-based image classification,” IEEE Transactions on Neural Networks, vol. 10, no. 5, pp. 1055 –1064, 1999.

[66] J. A. Hartigan and M. A. Wong, “A k-means clustering algorithm,” Applied Statistics, vol. 28, pp. 100–108, 1979.

[67] D. Cremers, “A multiphase levelset framework for variational motion segmentation,” in Proc. Scale Space Meth. Comput. Vis., Isle of Skye, U.K., pp. 599–614, 2003.

[68] L. Wang, L. He, A. Mishra and C. Li, “Active contours driven by local Gaussian distribution fitting energy,” Signal Processing, vol. 89 (12), pp. 2435–2447, 2009.

[69] C. Li, R. Huang, Z. Ding, C. Gatenby, D. Metaxas, and J. Gore, “A variational level set approach to segmentation and bias correction of medical images with intensity heterogeneity,” In Proc. Med. Image Comput. Comput. Aided Intervention , vol. LNCS 5242, pp. 1083–1091, Part II, 2008.

[70] C. Li, C. Xu, C. Gui, and M. D. Fox, “Distance regularized level set evolution and its application to image segmentation,” IEEE Transactions on Image Processing, vol. 19 (12), pp. 3243–3254, 2010.

[71] A. Tsai, A. Yezzi, and A. S.Willsky, “Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification,” IEEE Trans. Image Process., vol. 10, no. 8, pp. 1169–1186, 2001.

[72] L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," Int. J. Comput. Vis., vol. 50, no. 3, pp. 271–293, 2002.

[73] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," International Journal of Computer Vision, vol. 46, no. 3, pp. 223–247, 2002.

[74] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and variational problems," Communication Pure Applied Math, vol. XLII (5), pp. 577-685, 1989.

[75] O. Amadieu, E. Debreuve, M Barlaud and G. Aubert, "Inward and outward curve evolution using level set method," In Proceedings ICIP, Japan, pp. 188–192, 1999.

[76] H. K. Zhao, T. Chan, B. Merriman and S. Osher, "A variational level set approach to multiphase motion," In JCP 127:179–195, 1996.

[77] A. Yezzi, A. Tsai and A. Willsky, "A statistical approach to snakes for bimodal and trimodal imagery," In Proceedings ICCV, pp. 898–903, 1999.

[78] N. Paragios and R. Deriche, "Coupled Geodesic active regions for image segmentation: A level set approach," In Proceedings ECCV, Dublin, vol. II, pp. 224–240, 2000.

[79] L.D Cohen, "Avoiding local minima for deformable curves in image analysis," In Curves and Surfaces with Applications in CAGD," A. Le M´ehaut´e, C. Rabut, and L.L. Schumaker (Eds.), pp. 77–84, 1997.

[80] K.H. Zhang, H.H. Song, L. Zhang, "Active contours driven by local image fitting energy," Pattern Recognition vol. 43 (4), pp. 1199-1206, 2010.

[81] X.F. Wang, D.S. Huang, H. Xu, "An efficient local Chan–Vese model for image segmentation," Pattern Recognition vol. 43 (3), pp. 603-618, 2010.

[82] Y. Wang, S.M. Xiang, C.H. Pan, L.F. Wang, G.F. Meng, "Level set evolution with locally linear classification for image segmentation," Pattern Recognition vol. 46 (6), pp. 1734-1746, 2013.

[83] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, et al., The mammographic image analysis society digital mammogram database, In Exerpta Medical International Congress Series, pp. 375-378, 1994.

[84] M. Wirth, MIAS Mask Database, University of Guelph, Canada, 2005.

[85] E. Baltsavias," Object extraction and revision by image analysis using existing geodata and knowledge: Current status and steps towards operational systems," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 58 (3/4), pp. 129-151, 2004.

[86] J. A. Shufelt, "Performance evaluation and analysis of monocular building extraction from aerial imagery," IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 4, pp. 311-326, Apr. 1999.

[87] A. Ok, C. Senaras and B. Yuksel, "Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery," IEEE Transactions Geoscience Remote Sensing, vol. 51 (3), pp. 1701-1717, 2013.

[88] A. Ok, "Automated detection of buildings from single VHR multispectral images using shadow information and graph cuts," ISPRS Journal of Photogrametric and Remote Sensing 86, pp. 21-40, 2013.

[89] S. Ghaffarian and S. Ghaffarian, "Automatic building detection based on Purposive FastICA (PFICA) algorithm using monocular high resolution Google Earth images," J. Photogr. Remote Sens. vol. 97, pp. 152–159, 2014.

[90] D. Lowe, "Object recognition from local scale-invariant features," In Computer Vision, The Proceedings of the Seventh IEEE International Conference on, vol. 2, pp. 1150 -1157, 1999.

[91] T. Ojala, M. Pietikinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognition, vol. 29 (1), pp. 51-59, 1996.

[92] B. Julesz, "Textons, the elements of texture, perception, and their interactions," Nature, vol. 290, (5802), pp. 91-97, 1981.

[93] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials seing three-dimensional textons," International Journal of Computer Vision, vol. 43 (1), pp. 29-44, 2001.

[94] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques. Computer Vision, Graphics, and Image Processing," vol. 29 (1), pp. 100-132, 1985.

[95] R. M. Haralick and L. G. Shapiro, "Computer and Robot Vision," Addison-Wesley, Reading, MA, 1992.

[96] D. L. Ruderman,"The statistics of natural images," Network Computer Neural Systems, vol. 5 (4), pp. 517-548, 1994.

[97] C. Liu and J. Yang, "ICA color space for pattern recognition," IEEE Transactions on Neural Networks, vol. 20 (2), pp. 248-257, 2009.

[98] A. Verma, S. Banerji and C. Liu, “A new color SIFT descriptor and methods for image category classification,” In International Congress on Computer Applications and Computational Science, Singapore, pp. 819-822, 2010.

[99] H. Stokman and T. Gevers, “Selection and fusion of color models for image feature detection.,” IEEE Transactions on Pattern Analysis and Machine Intelligence vol. 29 (3), pp. 371-381, 2007.

[100] R. Datta, D. Joshi, J. Li and J. Wang, “Image retrieval: Ideas, influences, and trends of the new age,” ACM Computing Surveys vol. 40 (2), pp. 509-522, 2008.

[101] M. Carandini, D. J. Heeger, and J. A. Movshon, “Linearity and normalization in simple cells of the macaque primary visual cortex,” J. Neurosci., vol. 17 (21), pp. 8621-8644, 1997.

[102] M. J. Wainwright, O. Schwartz, and E. P. Simoncelli, “Natural image statistics and divisive normalization: Modeling nonlinearities and adaptation in cortical neurons,” in Statistical Theories of the Brain. Cambridge, MA: MIT Press, pp. 203-222, 2002.

[103] K. Sharifi and A. Leon-Garcia, “Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video,” IEEE Transactions on Circuits Systems and Video Technolodgy, vol. 5 (1), pp. 52-56, 1995.

[104] N. E. Lasmar, Y. Stitou, and Y. Berthoumieu, “Multiscale skewed heavy tailed model for texture analysis,” In Proceedings IEEE International Conference Image Processing, pp. 2281-2284, 2009.

[105] M. Varma and A. Zisserman, “Classifying images of materials: Achieving viewpoint and illumination independence,” In Proceedings of the European Conference on Computer Vision, Copenhagen, Denmark, pp. 255-271, 2002.

[106] T. Gevers, J. van de Weijer and H. Stokman, “Color feature detection: An overview. In: Lukac, R., Plataniotis, K.N. (eds.) Color Image Processing: Methods and Applications,” CRC Press, University of Toronto, Ontario, Canada, 2006.

[107] F. Melgani and L. Bruzzone, “Classification of hyperspectral remote sensing images with support vector machines,” Geoscience and Remote Sensing, IEEE Transactions on, vol. 42 (8), pp. 1778-1790, 2004.

[108] L. Blot and R. Zwiggelaar, “Background texture extraction for the classification of mammographic parenchymal patterns,” In International Workshop on Digital Mammography, 2001.

[109] S. Lou, H. Lin, K. Lin, and D. Hoogstrate, “Automatic breast region extraction from digital mammograms for pacs and telemammography applications,” Computerized Medical Imaging and Graphics, vol. 24, pp. 205-220, 2000.

[110] D. L. Pham, C. Xu, and J. L. Prince, “Current methods in medical image segmentation,” Annu. Rev. Biomed. Eng. 2, pp. 315-337, 2000.

[111] M. Goodsitt, and H. Chan, “Classification of compressed breast shapes for the design of equalisation filters in x-ray mammography,” Medical Physics, vol. 25, pp. 937-947, 1998.

[112] A. Hoyer, W. Spiesberg, “Computerized mammogram processing,” Phillips Technical Review, vol. 38, pp. 347-355, 1979.

[113] T. Lau and W. Bischoff, “Automated detection of breast tumors using the asymmetry approach,” Computers and Biomedical Research, vol. 24, pp. 273-295, 1991.

[114] F. Yin and M. Giger, “Computerized detection of masses in digital mammogram: analysis of bilateral substraction images,” Medical Physics, vol. 28, pp. 955-963, 1991.

[115] J. Byng, and N. Boyd, “Automated analysis of mammographic densities,” Medical Physics, vol. 41, pp. 909-923, 1996.

[116] U. Bick, M.L. Giger, R.A. Schmidt, R.M. Nishikawa, D.E. Wolverton, and K Doi, “Automated segmentation of digitized mammograms,” Acad Radiology, 2 (1), pp. 1-9, 1995.

[117] J. Hein and M. Kallargi, “Multiresolution wavelet approach for separating the breast region from the backround in high resolution digital mammography,” Digital Mammography, Nijmegen, Kluwer Academic Publishers, pp. 295-298, 1998.

[118] M. Masek, and Y. Attikiouzel, “Skin-air interface extraction from mammograms using an automatic local thresholding algorithm,” ICB, Brno, CR, pp. 204-206, 2000.

[119] J. Semmlow and A. Shadagopappan, “A fully automated system for screening xeromammograms,” Computers and Biomedical Reseach, vol 13, pp. 350-362, 1980.

[120] A. Mendez and P. Tahoces, “Automatic detection of breast border and nipple in digital mammograms,” Computer Methods and Programs in Biomedicine, vol 49, pp. 253-262, 1996.

[121] A. Morton, H. Chan and M. Goodsitt, “Automated model-guided breast segmentation algorithm,” Medical Physics, pp. 1107-1108, 1996.

[122] C. Zhou and H. Chan, “Computerized image analysis: Estimation of breast density on mammograms,” Medical Physics vol. 28, pp. 1056-1069, 2001

[123] N. Karssemeijer and G. N. te Brake, “Combining single view features and asymmetry for detection of mass lesions,” IWDM, pp. 95-102, 1998

[124] M. Abdel-Mottaleb and C. Carman, “Locating the boundary between the breast skin edge and the background in digitized mammograms,” Digital Mammography, pp. 467-470, 1996.

[125] R. Highnam, and M. Brady, Mammographic image analy J. Parker sis. Kluwer Academic Publishers, 1999.

[126] R. Chandrasekhar and Y. Attikiouzel, “Gross segmentation of mammograms using a polynomial model,” International Conference of the IEEE in Medicine and Biology Society, vol 3. pp. 1056-1058, 1996.

[127] P. Saha and J. Udupa, “Breast tissue density quantification via digitized mammograms,” IEEE Transactions on Medical Imaging, vol. 20, pp. 792-803, 2001.

[128] M.Wirth, J. Lyon and D. Nikitenko, “A fuzzy approach to segmenting the breast region in mammograms,” IEEE FI., vol. 1, pp. 474-479, 2004

[129] R.J. Ferrari, R.M. Rangayyan, J.E.L. Desautels, and A.F. Frere, “Segmentation of mammograms: Identification of the skin-air boundary, pectoral muscle, and fibroglandular disc,” International Workshop on Digital Mammography, pp. 573-579, 2000.

[130] A. M. Wirth, and A. Stapinski, “Segmentation of the breast region in mammograms using snakes,” In 1st Canadian Conference on Computer and Robot Vision, pp. 385-392, 2004.

[131] L. Lee, V. Stickland, R. Wilson, E. Roebuck, G.W. Eklund, G. Cardenosa and W. Parsons, “Fundamentals of Assessing adequacy of mammographic image quality,” Radiology, pp. 190:297–307, 1994.

[132] L.W. Bassett , D.M. Farria , S. Bansal , M.A. Farquhar , P.A. Wilcox and S.A. Feig, “Reasons for failure of a mammography unit at clinical image review in the American College of Radiology Mammography Accreditation Program,” Radiology, pp.215(3):698-702, 2000.

[133] C. Ünsalan and K. L. Boyer, “A system to detect houses and residential street networks in multispectral satellite images,” Comput. Vis. Image Understand., vol. 98, no. 3, pp. 423–461, Jun. 2005.

[134] N. Haala, M. Kada, “An update on automatic 3D building reconstruction,” ISPRS J. Photogr. Remote Sens., vol 65 (6), pp. 570–580, 2010.

[135] Z. Chen, Y. Zhang, B. Guindon, T. Esch, A. Roth and J. Shang, “Urban land use mapping using high resolution SAR data based on density analysis and contextual information,” Canadian Journal of Remote Sensing, vol. 38 (06), pp. 738-749, 2012.

[136] L. Chen, S. Zhao, W. Han and Y. Li, “Building detection in an urban area using lidar data and QuickBird imagery,” International Journal of Remote Sensing, vol. 33 (16), pp. 5135-5148, 2012.

[137] A. Huertas and R. Nevatia, “Detecting buildings in aerial images,” Comput. Vis., Graph., Image Process., vol. 41, no. 2, pp. 131–152, Feb. 1988.

[138] R. Irvin, D. Mckeown, “Methods for exploiting the relationship between buildings and their shadows in aerial imagery,” IEEE Trans. Syst. Man. Cybernet, Vol. 19 pp. 1564–1575, 1989.

[139] Y. T. Liow and T. Pavlidis, “Use of shadows for extracting buildings in aerial images,” Comput. Vis., Graph., Image Process., vol. 49, no. 2, pp. 242–277, Feb. 1990.

[140] J. Shufelt and D. Mckeown, “Fusion of monocular cues to detect manmade structures in aerial imagery,” CVGIP: Image Understand., vol. 57, no. 3, pp. 307–330, May 1993.

[141] J. C. McGlone and J. A. Shufelt, “Projective and object space geometry for monocular building extraction,” in Proc. IEEE Comput. Soc. Conf. CVPR, 1994, pp. 54–61.

[142] J. A. Shufelt, "Performance evaluation and analysis of monocular building extraction from aerial imagery,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 21 (4), pp. 311-326, 1999.

[143] D. S. Lee, J. Shan, and J. S. Bethel, “Class-guided building extraction from Ikonos imagery,” Photogramm. Eng. Remote Sens., vol. 69, no. 2, pp. 143–150, Feb. 2003.

[144] J. Inglada, “Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features,” ISPRS J. Photogramm. Remote Sens., vol. 62, no. 3, pp. 236–248, Aug. 2007.

[145] D. Koc-San and M. Turker, “Support vector machines classification for finding building patches from IKONOS imagery: the effect of additional bands,” J. Appl. Remote Sens. 8 (1), 2014.

[146] B. Sirmacek and C. Unsalan, “Urban-area and building detection using SIFT keypoints and graph theory,” IEEE Trans. Geosci. Remote Sens. 47 (4), pp. 1156–1167, 2009.

[147] S. Krishnamachari and R. Chellappa, “Delineating buildings by grouping lines with MRFs,” IEEE Trans. Image Process. 5 (1), pp. 164–168, 1996.

[148] Katartzis and H. Sahli, “A stochastic framework for the identification of building rooftops using a single remote sensing image,” IEEE Trans. Geosci. Remote Sens. 46 (1), pp. 259–271, 2008.

[149] V. J. D Tsai, “A comparative study on shadow compensation of color aerial images in invariant color models,” IEEE Transactions on Geoscience and Remote Sensing, vol. 44, no. 6, pp. 1661–1671, 2006.

[150] B. Sirmacek and C. Unsalan, “Building Detection from Aerial Images Using Invariant Color Features and Shadow Information,” In: 23rd International Symposium on Computer and Information Sciences, Istanbul, Turkey, 2008, pp. 105-110.

[151] O. Besbes, Z. Belhadj and N. Boujemaa, “Adaptive satellite images segmentation by level set multiregion competition,” INRIA, Research Report, 5855, 2006.

[152] X. Niu, “A semi-automatic framework for highway extraction and vehicle detection based on a geometric deformable model,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 61, pp. 170–186, 2006.

[153] G. Cao, X. Yang and D. Zhou, “Mumford-Shah model based man-made objects detection from aerial images,” in Proceedings of the International Conference on Scale Space Theories in Computer Vision, 2005, pp. 386-395.

[154] K. Karantzalos and D. Argialas, “A region-based level set segmentation for automatic detection of man-made objects from aerial and satellite images,” Photogrammetric Engineering and Remote Sensing Journal, vol. 75, 6, pp. 667-677, 2008.

[155] K. Karantzalos and N. Paragios, “Recognition-Driven Two-Dimensional Competing Priors Toward Automatic and Accurate Building Detection” IEEE Transactions on Geoscience and Remote Sensing, vol. 47, no. 1, 2009.

[156] K. Karantzalos and N. Paragios, “Large-Scale Building Reconstruction through Information Fusion and 3D Priors,” IEEE Transactions on Geoscience and Remote Sensing, vol. 48, No.5, pp. 2283-2296, 2010.

[157] J. Peng, D. Zhang, and Y. Liu, “An improved snake model for building detection from urban aerial images,” Pattern Recognition Letters, vol. 26, (5), pp. 587-595, 2005.

[158] S. Ahmadi, M. Zoej, H. Ebadi, H. A. Moghaddam, and A. Mohammadzadeh, “Automatic urban building boundary extraction from high resolution aerial images using an innovative model of active contours,” Int. J. Appl. Earth Observ. Geoinf., vol. 12, no. 3, pp. 150–157, Jun. 2010.

[159] S. P. Lloyd, “Least Squares Quantization in PCM,” IEEE Transactions on Information Theory, 28, pp. 129–137, 1982.

[160] J. M. Pena, J. A., Lozano and P. Larranaga, “An empirical comparison of four initialization methods for the k-means algorithm,” Pattern recognition letters, 20 (10), pp. 1027-1040, 1999.

[161] D. Arthur and s. Vassilvitskii, “k-means++: the advantages of careful seeding,” In Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics Philadelphia, PA, USA. pp. 1027–1035, 2007.

[162] V. Arevalo, J. González and G. Ambrosio, “Shadow detection in color high-resolution satellite images,” International Journal of Remote Sensing, 29 (7), pp. 1945–1963, 2008.

[163] T. Chan and W. Zhu, “Level set based shape prior segmentation,” Technical Report 03-66, Computational Applied Mathematics, UCLA, Los Angeles, 2003.

[164] M. Sussman, P. Smereka and S. Osher, “A level set approach for computing solutions to incompressible two-phase flow,” UCLA CAM Report 93-18, 1993.

[165] M. Nagao, T. Matsuyama and Y. Ikeda, “Region Extraction and Shape Analysis in Aerial Photographs,” Computer Graphics and Image Processing, 10 (3), pp. 195-223, 1979.

[166] P.M. Dare, “Shadow analysis in high-resolution satellite imagery of urban areas,” Journal of Photogrammetric Engineering and Remote Sensing 71 (2), pp. 169–177, 2005.

[167] Y. Chen, D. Wen, L. Jing, P. Shi, “Shadow information recovery in urban areas from very high resolution satellite imagery,” International Journal of Remote Sensing 28 (15), pp. 3249–3254, 2007.

[168] Q. Ye, H. Xie and Q. Xu, “Removing shadows from high-resolution urban aerial images based on color constanct,” International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B3, 2012 XXII ISPRS Congress, Melbourne, 2012.

[169] K.L. Chung, Y.R. Lin, Y.H. Huang, “Efficient shadow detection of color aerial images based on successive thresholding scheme,” IEEE Transactions on Geoscience and Remote Sensing 47 (2), pp. 671–682, 2009.

[170] M. Elbakary and K. Iftekharuddin, “Shadow detection of Man-Made buildings in high-resolution panchromatic satellite images,” IEEE Transactions on Geoscience and Remote Sensing, 52 (9), pp. 5374–5386, 2014.

[171] Y. Li, P. Gong, T. Sasagawa, “Integrated shadow removal based on photogrammetry and image analysis,” International Journal of Remote Sensing 26 (18), pp. 3911–3929, 2005.

[172] L. Lorenzi, F. Melgani, and G. Mercier, “A complete processing chain for shadow detection and reconstruction in VHR images,” IEEE Transactions on Geoscience and Remote Sensing, vol. 50, no. 9, pp. 3440–3452, Sep. 2012.

[173] Y. Shao, G. Taff and S. Walsh, “Shadow detection and building-height estimation using IKONOS data,” International Journal of Remote Sensing, 32 (22), pp. 6929–6944, 2011.

[174] V. Shettigara and G. Sumerling, “Height determination of extended objects using shadows in SPOT images,” Photogrammetric Engineering & Remote Sensing, 64 (1), pp. 35–44, 1998.

[175] F. Cheng and K. Thiel, “Delimiting the building heights in a city from the shadow in a panchromatic SPOT image-Part 1 – Test of forty two buildings,” International Journal of Remote Sensing 16 (3), pp. 409–415, 1995.

[176] M. Turker and E. Sumer, “Building-based damage detection due to earthquake using the watershed segmentation of the post-event aerial images,” International Journal of Remote Sensing 29 (11), pp. 3073–3089, 2008.

[177] D. Gabor, “Theory of communication,” Journal of the Institute of Electrical Engineers, vol. 93, pp. 429–457, 1946.

[178] J. Daugman, “Two-dimensional analysis of cortical receptive field profiles,” Vision Research, vol. 20, pp. 846–856, 1980.

[179] J. Daugman, “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters,” Journal of the Optical Society of America-A, 2 (7), pp. 1160–1169, 1985.

[180] A. Jain, N. Ratha, and S. Lakshmanan, “Object detection using gabor filters,” Pattern Recognition, vol. 30, pp. 295-309, 1997.

[181] J. Yuan, D. Wang and R. Li, “Remote Sensing Image Segmentation by Combining Spectral and Texture Features” IEEE Transactions on Geoscience and Remote Sensing, vol. 52 (1), pp. 16-24, 2014.

[182] A. Tzotsos, K. Karantzalos, and D. Argialas, “Object-based image analysis through nonlinear scale-space filtering,” Journal of Photogrammetry and Remote Sensing, vol. 66 (1), pp. 2–16, 2011.

[183] B. Sirmacek and C. Unsalan “A probabilistic framework to detect buildings in aerial and satellite images,” IEEE Transactions on Geoscience and Remote Sensing, vol. 49 (1), pp. 211–221, 2011.

[184] P. Xiao, X. Feng, R. An and S. Zhao, “Segmentation of multispectral high-resolution satellite imagery using log Gabor filters,” International Journal of Remote Sensing, vol. 31 (6), pp. 1427-1439, 2010.

[185] F. Bianconi and A. Fernándezb, “Evaluation of the effects of Gabor filter parameters on texture classification,” Pattern Recognition, Vol. 40 (12), pp. 3325–3335, 2007.

[186] J. Holland, “Adaptation In Natural and Artificial Systems” University of Michigan Press, 1975.

[187] Goldberg, E. David, “Genetic Algorithms in Search, Optimization, and Machine Learning,” Addison-Wesley, 1989.

[188] C. Li, R. Huang, Z. Ding, J. C. Gatenby and D. N. Metaxas, “A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI,” IEEE Transactions on Image Processing vol. 20 no. 7, pp. 2007-2016, 2011.

[189] X. F. Wang, H. Min, L. Zou, Y. G. Zhang, “A novel level set method for image segmentation by incorporating local statistical analysis and global similarity measurement,” Pattern Recognition 48 pp. 189-204, 2015.

[190] K. Zhang, K. Lam, L. Zhang and D. Zhang, “A level Set Aproach to Image Segmentation With Intensity Inhomogeinity,” IEEE Transactions on Cybernetics 43 (4), pp. 1199-1206, 2015.

[191] K. Zhang, L. Zhang and H. Song, “Active contours with selective local or global segmentation: A new formulation and level set method,” Image and Vision Computing vol. 28 (4), pp. 668–676, 2010.

[192] K. Zhang, Q. Liu, H. Song and X. Li, “A Variational approach to simultaneous image segmentation and bias correction,” IEEE Transactions on Cybernetics, vol. 45 (8), pp. 1426-1437, 2015.

[193] L. Dai, J. Ding and J. Yang, “Inhomogeneity-embedded active contour for natural image segmentation,” Pattern Recognition, vol. 48, pp. 2513–2529, 2015.

[194] C. Li, code, “level set segmentation and bias correction,” available from: http://www.engr.uconn.edu/~cmli/, accessed 2014.

[195] L. Zhang, code, “A locally statistical active contour model for image segmentation with intensity inhomogeneity,” available from: http://www4.comp.polyu.edu.hk/~cslzhang/LSACM/, accessed 2014.

[196] R Guillemaud, “Uniformity Correction with Homomor-phic filtering on Region of Interest,” IEEE International Conference on Image Processing 2**,** 872–875, 1998.

[197] C. Li, R. Huang, Z. Ding, C. Gatenby, D. N. Metaxas, and J. C. Gore, "A Variational Level Set Approach to Segmentation and Bias Correction of Images with Intensity Inhomogeneity", MICCAI 2008.

[198] C. Li, J.C. Gore, and C. Davatzikos, "Multiplicative intrinsic component optimization (MICO) for MRI bias field estimation and tissue segmentation", Magnetic Resonance Imaging , vol. 32 (7), pp. 913-923, 2014

[199] S. D. Tzikopoulos, M.E. Mavroforakis, H.V. Georgiou, N. Dimitropoulos and S. Theodoridis, “A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry,” Computer Methods and Programs in Biomedicine vol. I (2), pp. 478-63, 2011.

[200] M. Masek, Hierarchical Segmentation of Mammograms Based on Pixel Intensity, Ph.D. Thesis, University of Western Australia School of Electrical, Electronic and Computer Engineering and University of Western Australia Centre for Intelligent Information Processing Systems, 2004.

[201] A. Pronobis, O. M. Mozos, B. Caputo and P. Jensfelt, “Multi-modal Semantic Place Classification,” The International Journal of Robotics Research, vol. 29, pp. 298-320, 2010

[202] A. Pronobis, O. M. Mozos and B. Caputo, “SVM based discriminative accumulation scheme for place recognition,” Proceedings of the International Conference on Robotics and Automation (ICRA), Pasadena, CA, 2008.

[203] M. E. Nilsback and B.Caputo, “Cue integration through discriminative accumulation,” Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Washington, DC, 2004.

[204] J. Wolfe, “Breast patterns as an index of risk for developing breast cancer,” American Journal of Roentgenology, vol. 126, pp. 1130–1139, 1976.

[205] N. Boyd, J. Byng, R. Jong, E. Fishell, L. Little, A. Miller, G. Lockwood, D. Tritchler, and M. Yaffe, “Quantitative classification of mammographic densities and breast cancer risk: Results from the canadian national breast screening study,” Journal of the National Cancer Institute, vol. 87, no. 9, pp. 670–675, 1995.

[206] J. Byng, N. Boyd, E. Fishell, R. Jong, and M. Yaffe, “Automated analysis of mammographic densities,” Physics in Medicine and Biology, vol. 41, pp. 909–923, 1996.

[207] V. Assi, J. Warwick, J. Cuzick, and S. W. Duffy, “Clinical and epidemiological issues in mammographic density,” Nature Reviews Clinical Oncology, vol. 9, pp. 33–40, 2012.

[208] K. Kerlikowske, “The mammogram that cried wolfe,” New England Journal of Medicine, vol. 356, pp. 297–300, 2007.

[209] N. Arora, T. King, L. Jacks, M. Stempel, S. Patil, E. Morris, and M. Morrow, “Impact of breast density on the presenting features of malignancy,” Annals of Surgical Oncology, vol. 17, pp. 211–218, 2010.

[210] R. Highnam, N. Sauber, S. Destounis, J. Harvey, and D. McDonald, “Breast density into clinical practice,” in Breast Imaging, ser. Lecture Notes in Computer Science, A. Maidment, P. Bakic, and S. Gavenonis, Eds. Springer Berlin / Heidelberg, 2012, vol. 7361, pp. 466–473.

[211] J. Bottorff, P. Ratner, J. Johnson, J. Buxton, G. Hislop, W. Chen, and C. Zeisser, “The effects of providing information about mammographic breast density in a mammography screening program: a controlled trial,” in Oncology Nursing Society Congress 2005, 2005.

[212] W. Berg, J. Blume, J. Cormack, E. Mendelson, D. Lehrer, M. Bhm-Vlez, E. Pisano, R. Jong, W. Evans, M. Morton, M. Mahoney, L. Larsen, R. Barr, D. Farria, H. Marques, and Boparai, “Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer,” JAMA: The Journal of the American Medical Association, vol. 299, no. 18, pp. 2151–2163, 2008.

[213] Illustrated Breast Imaging Reporting and Data System (BI-RADS), 3rd ed. American College of Radiology, 1998.

[214] J. W. Byng, N. F. Boyd, E. Fishell, R. A. Jong, and M. J. Yaffe, “The quantitative analysis of mammographic densities,” Physics in Medicine and Biology, vol. 39, no. 10, p. 1629, 1994.

[215] R. Highnam, S. Brady, M. Yaffe, N. Karssemeijer, and J. Harvey, “Robust breast composition measurement - volparaTM,” in Digital Mammography, ser. Lecture Notes in Computer Science, J. Mart, A. Oliver, J. Freixenet, and R. Mart, Eds. Springer Berlin / Heidelberg, vol. 6136, pp. 342–349, 2010.

[216] P. Miller and S. Astley, “Classification of breast tissue by texture and analysis,” Image and Vision Computing, vol. 10, pp. 277–282, 1992.

[217] S. Petroudi, T. Kadir, and M. Brady, “Automatic classification of mammographic parenchymal patterns: A statistical approach,” in Proceedings of EMBC, International Conference on Engineering in Medicine and Biology. IEEE, 2003, pp. 798–801.

[218] M. Mustra, M. Grgic, and K. Delac, “Feature selection for automatic breast density classification,” in ELMAR, Proceedings, pp. 9 –16, 2010.

[219] M. G. J. Kallenberg, M. Lokate, C. H. van Gils, and N. Karssemeijer, “Automatic breast density segmentation: an integration of different approaches,” Physics in Medicine and Biology, vol. 56, no. 9, pp. 2715–2729, 2011.

[220] B. M. Keller, D. L. Nathan, Y. Wang, Y. Zheng, J. C. Gee, E. F. Conant, and D. Kontos, “Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation,” Medical Physics, vol. 39, no. 8, pp. 4903–4917, 2012.

[221] W. He, E. Denton, and R. Zwiggelaar, “Mammographic segmentation and risk classification using a novel binary model based bayes classifier,” in Breast Imaging, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2012, vol. 7361, pp. 40–47.

[222] N. Karssemeijer, “Automated classification of parenchymal patterns in mammograms,” Physics in Medicine and Biology, vol. 43, pp. 365–378, 1998.

[223] A. Bosch, X. Munoz, A. Oliver, and J. Marti, “Modeling and classifying breast tissue density in mammograms,” in Computer Vision and Pattern Recognition, 2006.

[224] A. Oliver, J. Freixenet, R. Marti, J. Pont, E. Perez, E. Denton, and R. Zwiggelaar, “A novel breast tissue density classification methodology,” Information Technology in Biomedicine, IEEE Transactions on, vol. 12, pp. 55 –65, 2008.

[225] W. Polakowski, D. Cournoyer, S. Rogers, M. DeSimio, and D. Ruck, “Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency,” IEEE Transactions on Medical Imaging, vol. 16, no. 6, pp. 811 – 819, 1997.

[226] H.D. Cheng, X.J. Shi, R. Min, L.M. Hu, X.P. Cai, and H.N. Du, “Approaches for automated detection and classification of masses in mammograms,” Pattern Recognition vol. 39 (4), pp. 646‑668, 2006.

[227] R.M. Rangayyan, F.J. Ayres and J.E.L Desautels, “A review of computer-aided diagnosis of breast cancer: toward the detection of subtle signs,” J. Frankl. Inst. 344 (3‑4), pp. 312‑348, 2007.

[228] K. Bovis, and S. Singh, “Detection of Masses in Mammograms Using Texture Features”, Proc.15th International Conference on Pattern Recognition, Barcelona, IEEE Press 2 pp. 267-270, 2000.

[229] He Wang, Lin-Lin Huang and Xiao-Jie Zhao, “Automated detection of masses in digital mammograms based on pyramid”, International Conference on Wavelet Analysis and Pattern Recognition, pp 1-183, 7187, 2007.

[230] W. Mohamed and Y. Kadah, “Computer aided diagnosis of digital mammograms”, International Conference on Computer Engineering & Systems, 299-303, 2007.

[231] X. Weidong , L. Li and P. Xu, “A new ANN-based detection algorithm of the masses in digital mammograms”, IEEE International Conference on Integration Technology, pp 26-30, 2007.

[232] G. Zhang, P. Yan, H. Zhao, “A computer aided diagnosis system in mammography using artificial neural networks”, International Conference on BioMedical Engineering and Informatics, 2, 823-826, 2008.

[233] J. Dheeba, N. Albert Singh and S. Tamil Selvi, “Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach,” Journal of Biomedical Informatics, Vol. 49, pp. 45–52, 2014.

[234] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification”, IEEE Transactions on Systems, Man, and Cybernetics, 3, 610-621, 1973.

[235] F. Younesi, N. Riyahi Alam, R. A. Zoroofi, A. Ahmadian and M. Guiti, “Computer-Aided Mass Detection on Digitized Mammograms Using Adaptive Thresholding and Fuzzy Entropy,” Proceedings of the 29th Annual International Conference of the IEEE EMBS, Lyon, France, 2007.

[236] G. Kom, A. Tiedeu, M. Kom, “Automated detection of masses in mammograms by local adaptive thresholding,” Comput. Biol. Med., vol. 37 (1), pp. 37‑48, 2007.

[237] N. Karssemeijer, G.M. te Brake, “Detection of stellate distortions in mammograms,” IEEE Trans. Med. Imaging, vol 15 (5), pp. 611‑619, 1996.

[238] N. Petrick, H.P. Chan, D. Wei, B. Sahiner, M.A. Helvie, D.D Adler, “Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification,” Med. Phys., vol 23 (10), pp. 1685‑1696, 1996.

[239] A. C. Bovik, “Handbook of Image and Video Processing,” New York, Academic, 2005.

[240] Z. Wang and A. C. Bovik, “Modern Image Quality Assessment,” San Rafael, CA: Morgan & Claypool, 2006.

[241] A. K. Moorthy and A. C. Bovik, “Blind image quality assessment: From natural scene statistics to perceptual quality,” IEEE Transactions on Image Processing, vol. 20, no. 12, pp. 3350–3364, 2011.

[242] M. A. Saad, A. C. Bovik and C. Charrier, “Model-Based Blind Image Quality Assessment: A natural scene statistics approach in the DCT domain,” IEEE Transactions on Image Processing, pp. 3339-3352, vol. 21, no.8,2012.

[243] P. Ye and D. Doermann, “No-reference image quality assessment using visual codebooks,” IEEE Transactions on Image Processing, vol. 21, no. 7, pp. 3129-3138, Jul. 2012.

[244] A. Mittal, R. Soundararajan and A. C. Bovik, “Making a Completely Blind Image Quality Analyzer ”, IEEE Signal processing Letters, vol.22, no.3, pp. 209-212, 2013.

[245] A. Mittal, A. K. Moorthy, and A. C. Bovik, “No-reference image quality assessment in the spatial domain,” IEEE Transactions on Image Processing, vol. 21 no. 12, pp. 4695-4708, 2012.

[246] W. Wang and Z. H. Zhou, “Analyzing Co-training Style Algorithms,” Machine Learning, Lecture Notes in Computer Science, vol. 4701, pp 454-465, Springer, 2007.

[247] Y. Kokkinos and K. G. Margaritis, “A distributed privacy-preserving regularization network committee machine of isolated Peer classifiers for P2P data mining,” Artificial Intelligence Review, vol. 42, no. 3, pp 385-402, Springer, 2014.

[248] C. L. P. Chen, C. Y. Zhang, L. Chen and Min Gan, “Fuzzy Restricted Boltzmann Machine for the Enhancement of Deep Learning,” IEEE Transactions on Fuzzy Systems, vol. 23, no. 6, pp. 2163 – 2173, 2015.

[249] Y. Chen, Z. Lin, X. Zhao, G. Wang, Y. Gu, “Deep Learning-Based Classification of Hyperspectral Data,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 6, pp. 2094 - 2107,2014.

[250] P. Musialski, P. Wonka, D. Aliaga, M. Wimmer, L. Van Gool, and W. Purgathofer, “A survey of urban reconstruction,” Computer Graphics Forum, vol. 32, no. 6, 2013.

[251] O. Teboul, “Ecole centrale paris facades database,” 2010.

[252] F. Korč and W. Forstner, “eTRIMS Image Database for interpreting images of manmade scenes,” Technical Report, TR-IGG-P-2009-01, 2009.

ISBN 978-9963-695-45-4