SETTLEMENT TYPE CLASSIFICATION USING AERIAL IMAGES

by

LIZWE MDAKANE

Submitted in fulfillment of the academic requirements for the Masters degree in the School of Computer Science, Faculty of Science and Agriculture, University of KwaZulu-Natal, Durban, South Africa, December 2014

As the candidate's supervisor I have approved this dissertation for submission.

Signed:_____ Name:_____ Date:_____

ABSTRACT

In metropolitan and urban areas, the problems relating to rapid transformations that are taking place in terms of land cover and land use are now very pronounced, e.g., the rapid increase and unpredictable spread of formal and informal physical infrastructure. As a result, the availability of detailed, timely information on urban areas is of considerable importance both to the management of current urban activities and to forward planning. Remote sensing sources can make a vital contribution in this context, since they provide regular and recurring data from a single, consistent source. Pattern recognition techniques have been demonstrated to be effective in distinguishing and classifying human settlements. However, these methods are not ideal as they perform poorly when presented with imagery of the same area acquired at different dates. The poor generalization ability is mainly caused by large off-nadir viewing angles which produce image pairs with different viewing- and illumination-geometries. Classification performance is also decreased by differences in shadow length and orientation.

The objective of this research is to improve the generalisation ability of the automated classification of human settlements using only remote sensing data over urban areas. The multiresolution local binary patterns (LBPs) algorithm, extended with an orthogonal variance measure for measuring local contrast features (i.e., the extended LBP) has been shown to excel at texture classification tasks. To minimize the viewing- and illumination-geometry effects and improve settlement classification, the extended LBP was applied to high spatial resolution panchromatic aerial images. The addition of a contrast component to the LBP features does not directly affect the desired invariance to shadow orientation and length, but it is expected that the richer features will nevertheless improve settlement classification accuracy.

The extended LBP method was evaluated using a support vector machine (SVM) classifier for cross-date (training and test images of the same area acquired at different dates) and same-date analysis. For comparable results, LBPs without contrast features were also evaluated. The results showed the extended LBP to have a strong spatial and temporal generalisation ability for classifying settlements of aerial images, when compared to its counterpart. From this research, we can conclude that the extended LBP's additional contrast features can improve overall settlement type classification accuracy and generalisation ability.

PREFACE

The research work described in this dissertation was carried out in the School of Computer Science, University of KwaZulu-Natal, Durban, from February 2013 to June 2014, under the supervision of Dr. D. Moodley and Dr. F. van den Bergh

These studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others it is duly acknowledged in the text.

DECLARATION 1 - PLAGIARISM

I, Lizwe Mdakane, declare that:

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
   (a) Their words have been re-written but the general information attributed to them has been referenced
   (b) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed: _____

DECLARATION 2 - PUBLICATIONS

1. MDAKANE, L., AND VAN DEN BERGH, F. Extended local binary pattern features for improving settlement type classification of QuickBird images. In *Proceedings of the Twenty-Third Annual Symposium of the Pattern Recognition Association of South Africa, PRASA* (2012), pp. 68–74

Signed: _____

ACKNOWLEDGEMENTS

First and foremost, I thank God Almighty, for all the countless opportunities that He has given me. This thesis is dedicated to many people whom have supported me through this endeavour. I am deeply grateful to my family and my dear friends for their patience, understanding, support and for providing the foundation needed to complete the MSc over the past few years. Most importantly, none of this could have happened without my mother Nonhlanhla Mdakane and best friend Moloko Mathipa as they have experienced some ups and downs in the past years. Every time I was ready to quit, you did not let me and I am forever grateful. This dissertation stands as a testament to your unconditional love and encouragement.

I have to thank my research supervisors, Dr. Deshen Moodley and Dr. Frans van den Bergh. Without their assistance and dedicated involvement in every step throughout the process, this thesis would have never been accomplished. I would like to thank them for your support, guidance and understanding over these past years.

Finally, I would also like to show gratitude for the research support provided by the Council for Scientific and Industrial Research (CSIR) through the remote sensing and research unit (RSRU) MSc studentship program. The group has been unwavering in their personal and professional support during the time I spent with them at CSIR, Meraka in Pretoria. Special thanks to Dr Waldo Kleynhans and fellow students Vincent Seotlo, and Colin Schwegmann for their useful suggestions and advice.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

According to the Oxford dictionaries, a slum is defined as "a house or building unfit for human habitation"[1] and depending on where you are located in the world, slums are also known as shacks, squatter areas, shanty areas, or informal settlements. In 2001, about 32% of the world's population (924 million people, at that time) lived in informal settlements, with the majority found in less developed countries where up to 78.2% lived in slums [40]. Since then, there has been rapid increase of informal settlements, particularly in urban areas. The accelerated pace of urbanisation is mainly due to:

1. migration for better prospects of employment, education and access to social infrastructure,
2. poverty and unequal distribution of wealth,
3. inability of government to define clear and long-term land and housing policy [58].

Major concerns of informal settlements (IS) are that they have the highest concentrations of poor people, the worst shelter and unsafe physical environmental conditions with high population densities (i.e., building structures are unsafe, location is on hazardous land and usually overcrowded). IS also lack basic and essential services such as water supply, sewerage and waste disposal, health and education, and in addition, they are not in compliance with current planning and building regulations thus are defined as unauthorised housing [41, 40, 88, 51]. High rates of these types of settlements often results in an uncontrolled urban growth and increased consumption of natural resources. If left unchecked, these may threaten the sustainable development of urban areas in the long term, and may eventually cause environmental degradation and social tension.

In this context, an effective settlement differentiation and monitoring system of the spread of these settlements is essential. However, this is not a trivial task as informal settlements are highly condensed, dynamic (erection and removal of structures happens over short time periods), and difficult to access for surveys [82]. This makes traditional monitoring methods labour intensive, expensive, time-consuming and therefore impractical and unsatisfactory for urban management

[1] Definition found at http://www.oxforddictionaries.com/definition/english/slum?q=Slums

purposes. Thus, an automated effective and repetitive monitoring system for human settlements is of the greater importance.

In an effort to develop an automated monitoring system that can be highly effective in settlement differentiation/classification without the need for ancillary data, we considered the use of remote sensing imaging sources and pattern recognition techniques. Remote sensing (RS) imaging sources provide the opportunity to capture accurate, detailed historic data of settlement at regular intervals and high details. This enables the opportunity to monitor small-scale land use structures and dynamics in urban areas. RS and pattern recognition techniques also permit a detailed analyses of informal settlements by assisting in identifying main driving factors of informal settlements and hence the prediction of future informal patterns. Such a strategic approach is essential for an automated, systematic and replicable urban monitoring system and could be a key for future success in efficient management of informal settlements and other related environmental impacts.

1.1 Problem statement

There are numerous geo-spatial and image processing methods that can be used to map and monitor urban structures and/or objects from RS imagery. To find meaning in RS imagery, which is an unstructured array of pixels, the first step is to extract efficient visual features from the unstructured pixels [37]. This makes image feature extraction an essential step for an effective automated settlement type classification system as an appropriate feature representation can significantly improve the performance of the classification process. Using high spatial resolution imagery of urban areas, texture feature extraction methods have been shown to be successful in setting apart different settlement types [8]. Methods such as the Local Binary Patterns (LBPs) have been shown to be the most effective algorithm when compared to other known algorithms, such as, the Gray-level Co-occurrence Matrix (GLCM), Granulometrics and Discrete Wavelet Transform (DWT) [30]. In spite of the good performance, the LBP and above mentioned methods are still far from being ideal due to viewing- and illumination-geometry effects [125].

These effects cannot be easily avoided as they may not be caused by poor image quality but may be due to a pair of images of the same area acquired at different dates or time of the day. The latter may introduce large off-nadir[2] viewing angles which produces an image pair with different viewing- and illumination geometries. Illumination effects alter the orientation and length of the shadows between the image pair which results in decreased classification

[2]The nadir refers to the direction pointing directly below a particular point of observation

accuracy. In addition, the seasonality and time of day during acquisition determine the sun elevation and azimuth angles which may hinder classification performance further. Though no real change has occurred on the ground, these effects produce an image-pair that may contain a large number of spurious differences [70, 125]. A good image feature is one that is designed to have a representation that is sensitive to change in the desired variables, e.g., settlement type, while being insensitive to other types of change that may be present in the image [125]. If an image feature has these properties, then it can be expected that the feature will lead to good generalization performance in classification tasks.

1.2 Research objective

The objective of this research is to improve settlement type classification accuracy, using only remote sensing imagery, by improving the feature extraction step. The work evaluates the generalization performance particularly for cross-date images (two scenes of the same area acquired under different conditions, such as time of the year) as these contain large viewing- and illumination-geometry effects. To achieve good generalization in classification accuracy, joint distribution of gray-scale and rotational invariant LBP with rotational invariant variance measures (contrast), called extended LBPs are used. The gray-scale and rotational invariant LBP part of the algorithm is only sensitive to the structure (pattern) whilst being insensitive to rotation and contrast changes. Variance measures are invariant to pattern change and only sensitive to contrast. RS images can be viewed as 2-D texture images characterised by spatial structures and contrast thus making the extended LBP an ideal tool for the task.

In order to ascertain classification accuracy improvements and the significance of the variance measure in classifying settlement type, the extended LBP is compared with the original LBP method (which omits the variance components).

1.3 Contribution

The main contribution of the study is to investigate the significance of contrast features for the task of improving settlement type classification using cross-date aerial images.

1.4 Methodology and expected impact

To achieve the objectives mentioned above, aerial RS data was used for a semi-automated classification system to map settlement types. The methodology framework for this research was

then carried out as following:

1. RS data was acquired using a Digital Mapping Camera (DMC) panchromatic airborne sensor over the same area at different dates.

2. The process of extracting image features for distinguishing human settlements from the input data was done using LBP and extended LBP (LBP/VAR) algorithms.

3. Support Vector Machine (SVM) classifier was used for training and testing the extracted features.

4. To evaluate the general improvement and the significance of contrast measures in settlement mapping, a comparison of the classification accuracies (LBP/VAR against the LBP) was performed.

In light of the above, viewing- and illumination geometry effects have shown to be challenging problems in settlement type classification. The extended LBP is expected to aid in a development of more robust system for classifying settlements. Though contrast measures may not directly effect the desired invariance to shadow orientation and length, it is expected that the richer features will nevertheless improve settlement classification accuracy.

1.5 Thesis layout

Chapter 1 is the introduction chapter, consisting of the background, description of the problem statement, research objectives and methodology.

Chapter 2 presents a brief introduction to remote sensing history, fundamental concepts, general use and interpretation. The chapter then discusses RS systems for urban landscape mapping, in particular feature extraction and classification methods. The chapter ends by presenting the study case: classifying human settlements. This section presents the definitions of human settlements and a review of texture feature methods for settlement analysis.

Chapter 3 is the methods chapter, presenting the theoretical development of the algorithms used, namely, the extended LBP feature extraction algorithm and SVM classifier.

Chapter 4 is the experiment design chapter, presenting the dataset and implementation of the extended LBP features for training and testing the SVM classifier.

Chapter 5 provides the results from the evaluation (training and testing) of the SVM classifier and a discussion section for analysis, using different parameters.

Chapter 6 presents conclusions and recommendations for future work.

Chapter 2

Literature Review

2.1 Introduction

The chapter is divided into three sections. The first section presents the remote sensing (RS) field, discussing the field's fundamental concepts and history. The next section presents the use of remote sensing in automated or semi-automated classification tasks, particularly in urban areas. The last part of the chapter presents the use of RS for the classification of human settlements, the section also presents current work, challenges and possible solutions.

2.2 Remote Sensing: History and fundamental concepts

There are many definitions of Remote Sensing, for our purpose, it is defined as the measurement of objects on the earth's land and water surfaces using data acquired from an airborne or a spaceborne platform [14, 65, 111]. Data is acquired using electromagnetic (EM) sensors, in one or more regions of the EM spectrum, that record electromagnetic energy variations reflected and emitted by the Earth's surface features. While the output can vary, it is usually in a form of a two-dimensional spatial grid (image) representing the region of interest [65]. Data measurement is required to derive useful information from the image for the purpose of generating products (e.g., conventional maps and resources surveys) to be used in the fields of geography, geology, oceanography, urban and regional planning, agriculture, and others [110, 16, 65, 22]. Remote sensing systems such as aerial RS, satellite RS, thermal RS, radar RS and LiDAR have been used in land, agriculture, forestry, geology, hydrology, land use, land cover, oceanography, weather analysis and forecast applications [132, 66].

The use of remote sensing in mapping is the most apparent, as imaging sensors collect and render information in the inherently spatial form of a map. The precise control and geometric registration of modern imaging systems allows the geographic location of an individual image pixel to be determined within metres from sensor altitudes of hundreds of kilometres [96]. This enables accurate, reliable, timely, long-term and quality global data acquisitions that allows better prediction of natural hazards, epidemics, impacts of energy choices, climate variations and a combination of the past and present earth observations enables the detection, monitoring

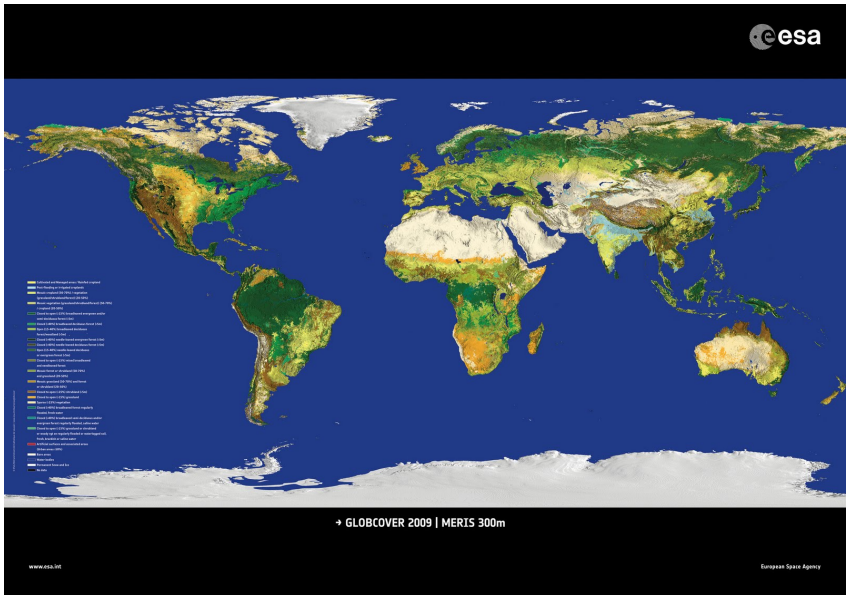


Figure 2.1: Remote Sensing global land cover product: 300-m global land cover map produced from an automated classification of MERIS FR time series for the year 2009[1].

[2]Source: Globcover from ESA, url: $http://due.esrin.esa.int/globcover/$

and management of the Earth's surface changes (e.g., urbanisation, industrialisation and globalisation). These capabilities can assist in decision making for managing common environment problems such as water shortages, desertification, soil depletion, greenhouse gases warming the atmosphere, deforestation, elevated coastal waterway sediment and nutrient fluxes, and other troubling consequences of human activities [64]. For an example of a RS global land cover product using satellite RS, see figure (2.1).

2.2.1 History of Remote Sensing

Aerial photography : Remote sensing began with the development of photography in early 1800s by Louis Daguerre [23]. In attempts to view the Earth's surface from a vertical perspective, photography went from ground to the sky using balloons, kites and airplanes. The first aerial photographs were taken in the 1860s by Felix Tournachon, known by the pseudonym "Nadar", in France using a camera mounted on a balloon tethered 80 m above the Bievre Valley [53]. In 1909, Wilbur Wright took the first aerial photograph using a plane. The manoeuvrability of the plane provided the capability of controlling speed, altitude, and direction required for systematic use of the airborne camera [14].

Remote sensing: From then until the early 1960s, aerial photography remained the single standard tool for depicting the surface from a vertical or oblique perspective. In 1960, the

U.S. launched the world's first successful weather satellite called Television and Infrared Observation Satellite (TIROS-1) [121]. As images were now collected from space and outside the visible spectrum, the term "aerial photography" no longer described the collected data. The term "remote sensing" coined in the mid 1960s by Evelyn Pruitt (a scientist working with the U.S. Office of Naval Research) was then used to describe the collected data [14].

Satellites: The focus of remote sensing research shifted from aerial photography to the use of images acquired by Earth-orbiting satellite sensors [28]. The first satellite (known as the Earth Resources Technology Satellite ERTS-1) that was dedicated to monitoring environmental conditions on the Earth's surface was launched in 1972 [86]. It was followed by ERTS-2 (launched in 1975) and ERTS-3 (launched in 1978). Later, the names of these satellites were changed to Landsat-1, -2, and -3, respectively [86]. as Ikonos and QuickBird, were launched in the last decade and have stimulated the development of newer detailed scale applications related to urban settlements.

2.2.2 Data acquisition

Remote sensing sensors collect electromagnetic radiation from the Earth's surface by one of two ways, i.e., active or passive. Active sensors have their own energy source and emit a signal that travels through the atmosphere, reflects off the Earth's surface and returns to the sensor, which measures the signal's travel time and strength. Synthetic Aperture Radar (SAR) sensor is an example of an active sensor that uses long-wavelength signals and can penetrate clouds or adverse weather conditions. Passive sensors, also known as optical sensors, do not have their own energy source and usually record the reflected energy of electromagnetic radiation or emitted energy from Earth where sunlight is the main source. Photographic cameras and multispectral scanners are examples of passive sensors and often used in satellite remote sensing [96].

Remote sensing systems have been designed to be sensitive to the visible and other portions of the electromagnetic spectrum (see figure 2.2). This characteristic enables remote sensing analysts to see portions of the spectrum that the human eye cannot detect, thereby enhancing their ability to identify different surface materials.

For any given material, the amount of solar radiation that it reflects, absorbs, or transmits varies with wavelength. This property of matter makes it possible to identify different substances or features and separate them by their spectral signatures [65]. The spectral properties of a

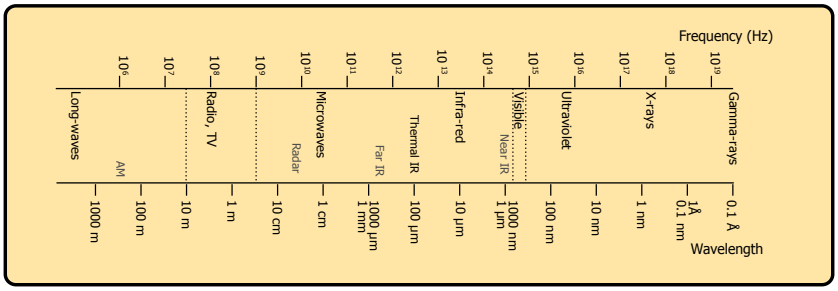


Figure 2.2: Electromagnetic spectrum.

sensor are defined by the number, placement and width of bands within the electromagnetic spectrum that it is able to record.

2.2.2.1 Sensors

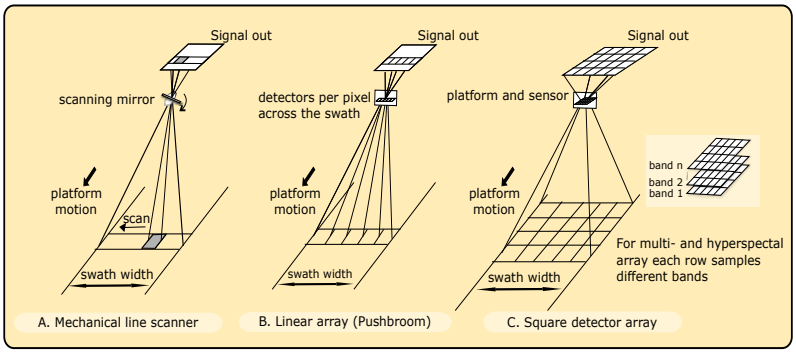


Figure 2.3: Remote Sensing devices used to collect aerial photography, multispectral and hyperspectral imagery.

Image acquisition technologies used in satellite programs have ranged from traditional cameras to line scanners. While a traditional camera is held fixed on the target of interest as it is sensed in a very brief moment, a line scanner uses a satellite/aircraft to provide motion along the track where the scanning motion *across* or *along* the target captures the scene over a time interval. There are two broad categories of line scanners, namely, mechanical line scanner and line detector array. A mechanical line scanner contains a mechanical component (e.g., a rotating mirror) that scans the surface across the swath. The forward motion of the vehicle and a rotating mirror allows an image strip to be built up from the raster-scans, see figure (2.3-A). Alternatively, a line detector array or "push-broom" provides scanning along the track and the sensed radiation moves directly through the optics onto the array detectors. The scanner carries sufficient detectors on the sensor platform such that each pixel can be recorded individually (such that, the sensor sweeps across the sensed scene), see illustration in figure (2.3-B). Other scanners include the square detector array which works like the "push broom", but instead of

the rectangular detector, a square detector array is used to capture a two dimensional image underneath the satellite, as illustrated in (figure (2.3-C). To record many spectral channels of data across track direction simultaneously (to the order of 10 bands), other dimensions are employed and the acquired data is described as multispectral or hyperspectral data.

2.2.2.2 Platforms

During the last two decades, there has been a rapid increase in number of platforms for remote sensing applications, these can be split into two categories: airborne and spaceborne.

- **Airborne platform**

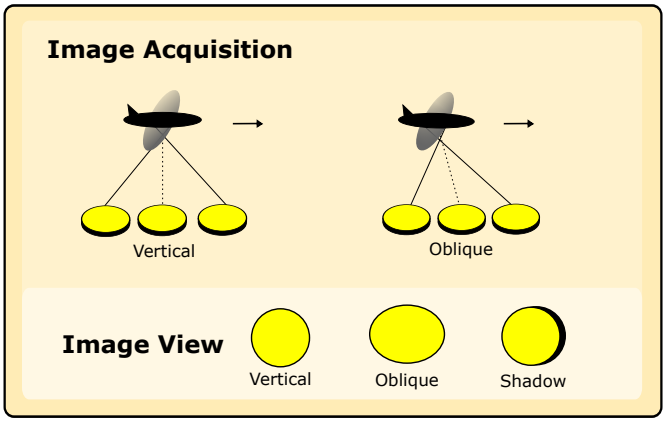


Figure 2.4: The geometry of aerial photographs (vertical or oblique) and tilt displacements on the acquired image.

Airborne platforms such as an aeroplane or helicopter mount cameras to capture an aerial view of the target of interest, the acquired images are known as aerial photographs. Aerial photography provides the longest available images history and have high spatial and radiometric resolution [81]. The images can be acquired with any type of camera, however, RS airborne cameras mainly use two types, namely, film-based and digital cameras. Both cameras are quite similar in structure, but differ in the way they store data. A digital camera records reflectance using an electronic sensor and stores it digitally while a film-based stores it on a film [81]. Film-based photographs are scanned to create digital images, which offer advantages in analysis [14]. Most optical aerial photographs commonly span panchromatic (black and white), colour, or false-colour infrared bands. However, using different emulsions and filters, photographs can be recorded in various types of electromagnetic wavelengths [65]. Aerial photography is not only limited to optical; SAR imaging can also be carried out on airborne platforms [14].

Data acquisition of aerial photographs can either be vertical or oblique, depending on the orientation of the camera's optical axis relative to the earth's surface, see figure (2.4). A vertical photograph is perpendicular to the surface with a slight tilt of no more than 3° and oblique aerial photographs are taken with an intentional inclination of the camera axis. Aerial images may have some errors due to the platform stability, weather conditions and time of acquisition (i.e., oblique objects appear longer and depending on the time of day the image was acquired shadows appear at various lengths), see figure (2.4).

- **Spaceborne platform**

  Sensors on space platforms offer broad spatial imagery, regular re-visit frequency and are often a cost effective alternative to aerial photography. An example of a spaceborne platform images is satellite imagery. RS satellites have unique characteristics which make them particularly useful for remote sensing of the Earth's surface. There are several RS satellites currently available, however, depending on the orbit they can be divided into two categories: geostationary or near polar satellites. A geostationary satellite revolves at speeds which match the Earth's rotation, making it seem to be stationary. This allows the satellite to observe and collect information continuously over specific areas. These satellites are commonly used for weather monitoring and communications. Near polar orbits follow an inclined orbit relative to line running between the North and South poles. In conjunction with the Earth's rotation (west – east), near polar satellites are able to cover most of the Earth's surface over a certain period of time. Near polar orbiting satellites are said to be sun–synchronous, i.e., at any given latitude, the position of the sun in the sky as the satellite passes overhead will be the same within the same season [15].

Acquisition of spaceborne or airborne remote sensing images has some errors due to the platform and digitisation of images. The most common are: geometric and radiometric errors. Geometric errors occur due to problems regarding:

- variations in platform altitude, stability and velocity,
- topographic/relief displacement, tilt displacement,
- the rotation of the Earth during image acquisition.

Radiometric errors can be caused by atmospheric effects (e.g., refracted light, clouds and bad weather) and the time and season of image acquisition (which can affect the clouds and the

angle of the sun) [81, 109]. Geometric distortions are corrected using ortho-rectification procedure which corrects geometric displacement errors and provide spatial reference [135]. For radiometric errors, correction procedures are specific to the nature of the distortion. The corrections usually involves histogram[2] manipulation procedures, e.g., histogram matching, histogram equalisation and histogram normalisation procedures [102, 109].

2.2.2.3 Data resolutions

Data collected by RS systems can either be in analogue or digital format and these data are primarily described by four types of resolutions:

1. **Spatial resolution:** a measure of the finest detail in an image. For digital images, this refers to the ground area captured by a single pixel. Spatial resolution can be characterised as: low (approx. 1 km or more), medium (approx. 100 m to 1 km), high (approx. 5 to 100 m) and very high spatial resolution systems (approx. 5 m or less).

2. **Spectral resolution:** represents the width of wavelength interval and/or number of spectral channels (or bands) captured by a sensor. Optical imaging systems (visible, near infrared, and shortwave infrared systems) can be classified, in terms of the spectral resolution, according to the number of spectral bands used: a) Mono-spectral or panchromatic (single wavelength band, black-and-white, gray-scale image systems), b) Multispectral, (several spectral bands), c) Super-spectral (tens of spectral bands) and, d) Hyper-spectral (hundreds of spectral bands).

3. **Temporal resolution:** the amount of time it takes a sensor to revisit a particular geographic location, possibly at a different viewing angle.

4. **Radiometric resolution:** the sensitivity of the sensor to brightness values. This metric is usually articulated in terms of binary bit-depth, which refers to the number of gray-scale levels at which data are recorded by a particular sensor. The binary bit-depth is typically expressed in the following ranges of gray-scale levels: 8-bit (0 – 255), 10-bits (0 – 1,023), 11-bits (0 – 4,095) and 16-bit (0 – 65,535).

In order to detect, distinguish between, and identify objects of interest, it is first necessary to appreciate the satellite remote sensing system's trade-off in resolution. The trade-off in resolution is dependent on the user application, i.e., depending on the application the user can

[2] the distribution of the tonal and radiometric values of the entire image

place the emphasise on the most important resolution. That is, emphasis on spectral resolution sensors results to having a medium or low spatial resolution and vice-versa, see table (2.1). Alternatively, the user may not place any particular emphasis on either high nor low resolution, but can acquire medium spatial, spectral and temporal resolution simultaneously.

Table 2.1: Remote sensing space-borne sensors in-terms of resolution characteristics.

Attribute	Resolution		
	High Spectral	High Spatial	Medium
Organisation	NASA (USA)	GeoEye (USA)	NASA (USA) and METI (Japan)
Sensor (Mission)	Hyperion (EO-1)	GeoEye-1	ASTER (EOS Terra)
Operation	2000–	2008–	1999–
Temporal Res	16 days	$<$3 days	4-16 days
Spatial Res (m)	30	1.65 (MS[a]),0.41 (PAN[b])	15(VNIR[c]), 30(SWIR[d]), 90 (TIR[e])
Swath Width (km)	7.5	15.2	60
Radiometric Res	12-bit	11-bit	8-bit(VNIR/SWIR), 12-bit(TIR)
Spectral Res (μ m)	0.353–2.577	0.45–0.92	0.52–0.86, 1.60–2.43, 8.125–11.65
Pan Band Res (m)	N/A	0.45–0.80	N/A
Bands (Total)	220	5	14

[a]MS: Multispectral
[b]PAN: Panchromatic
[c]VNIR: Visible and Near-Infrared
[d]SWIR: Short Wave Infrared
[e]TIR: Thermal Infrared

2.3 Urban analysis RS systems

Remotely sensed data can be a useful source of data for mapping the composition of urban settings and analysing changes over time that cannot be obtained from ground-level observations. It provide a more spatially complete representation of urban areas, using sensors onboard airborne/spaceborne platforms that provide a unique overhead perspective on the diversity of urban environments over a wide range of spatial and temporal scales. The geometric precision of the images combined with repeated revisits provided by a satellite orbit extends spatial mapping into the time dimension and makes it possible to monitor subtle changes in urban areas. For example, recently launched high-resolution sensors, like Quickbird, PLEIADES and WorldView, combined with archives of moderate-resolution Landsat and SPOT imagery, can provide detailed multitemporal observations of every city on earth including cities in developing countries [96]. Using RS sensors at various wavelengths, e.g., thermal and microwave radar sensors, RS data can reveal urban characteristics that are not visible to the eye.

2.3.1 RS data selection

A wide variety of passive and active remote sensing systems with various spatial resolutions have been useful for urban studies. As early as the 1970s, when the first Landsat was launched, medium-resolution RS data has been used to examine urban phenomena or processes over large areas. In an urban environment two major classes can be remotely sensed: land-cover types (e.g., bare soil, water, vegetation) and land-uses (e.g., residential, commercial, industrial and roads). Using RS multispectral sensors, diverse land-cover types can be accurately separated to derive accurate thematic land-cover maps. For land-uses, however, as object identification often responds in a strongly correlated manner to spatial resolution of the imagery, analysing multispectral imagery with relatively coarse spatial resolution may not distinguish these urban categories as accurate. Another problem with spectral data is due to the highly similar spectral response of these areas of the urban land-cover mix, e.g., asphalt roads versus black tiled roofs [46]. In this context, to effectively use RS data for urban landscape monitoring, the choice of data should meet certain conditions in terms of spatial, spectral, radiometric, and temporal characteristics [57].

To discern the inner-city structures, one requires remote sensing data with fine detail (very high spatial resolutions) [57]. The availability of very high spatial resolution satellite imagery and aerial photography (1 m and below) provides valuable information in various forms to detect and differentiate inner-city structures, e.g., settlements in the urban environment [131, 57, 119]. However, there are disadvantages in using very high resolution imagery, i.e., they need more storage and processing time. However, given the rapid and continuing improvements in computer technology, the latter appears to be of secondary importance [28]. A notable disadvantage about high spatial resolution sensors is that presently the majority of highest resolution images are recorded in panchromatic mode only (multispectral images covering wavelengths from visible to near-infrared) with high geometric accuracies, are usually at lower spatial resolution than the panchromatic band, see Table (2.1) on page 13. Regardless of this trade-off, the usefulness of satellite remote sensing for distinguishing inner-city structures from its surrounding neighbourhoods has been addressed in the last decade [7, 48, 62, 108, 118, 130]. The degree of "slumness" of neighbourhoods and places within a city has also been estimated using remotely sensed data from very high spatial resolution platforms such as QuickBird, IKONOS and SPOT 5 [96].

To understand temporal resolution requirements for monitoring urban areas, three urban temporal scales must be considered [57]:

1. The temporal development cycle of an urban phenomenon, i.e., how long formal infrastructure is developed from non-built to a fully developed residential area.

2. The revisit time for the remote sensor system to acquire data of the urban landscape (this knowledge is critical for an up-to-date urban monitoring system).

3. How often the users (urban managers/planners) need a specific type of information.

Aerial or satellite-based sensor systems are sufficient, in terms of spatial and temporal resolution, to monitor urban areas as the imagery can be detailed and is acquired at regular time intervals. This ensures an up-to-date system and combined with the archived imagery, RS systems can provide detailed multi-temporal observations of urban phenomena. Airborne sensor's flexible revisit time makes them suitable for capturing specific type of information, e.g., urgent information in emergency situations such as flooding and earthquakes.

2.3.2 Feature selection and extraction

To successfully monitor/classify urban objects, the choice of distinguishing features is a critical task. Prior knowledge about the object features plays a major role in the design of the feature extractor, where the knowledge may be about the form of the underlying categories or attributes of the patterns. Using this prior knowledge, the object features/properties (i.e., the values of the chosen features) can be are extracted/measured and passed for classification. However, due to the composition of urban areas having many small objects composed of many different materials in a spatial arrangement that does not produce many homogeneous pixels, extracting urban feature presents more challenging problems [79]. Further complicating the classification of urban environments is the diversity of its elements (e.g., buildings, bare soil and vegetation), creating a spectral diversity that far exceeds natural environments. This complexity, along with three-dimensional surface heterogeneity, creates a particularly challenging mapping environment for urban areas [46, 92, 57]. However, with recent improvements in RS imagery resolution, feature extraction methods and computational power, a detailed physical analysis of characterisation of urban landscape such as human settlements and other different urban materials is now possible.

For an effective classification system, the first and key step is having a powerful image feature extraction technique as it can significantly improve the accuracy of the classification process [37]. Image feature extraction techniques extract characteristics/features that capture certain properties of an image either globally for the entire image or locally for regions or objects. There are two types of approaches for identifying and extracting features of interest in

remotely sensed images, i.e, traditional/manual and semi/fully-automated approaches. Traditionally, information has been obtained from aerial photography through manual interpretation. Analogue aerial photographs are usually analysed visually by experienced analysts whilst digital images can be analysed and interpreted with the aid of a computer. In aerial photography (analogue or digital), non-geometric image characteristics such as tone or colour, texture, pattern, shape, shadow, size and situation normally give clues in recognition, identification and interpretation [110, 81]. Manual interpretation can be useful as it requires limited image preparation, can be fairly accurate and is a well-developed discipline. Although this is still the predominant approach, it is not efficient mainly because of the laborious, inconsistent, expensive and time-consuming nature of manual feature identification process [96, 81].

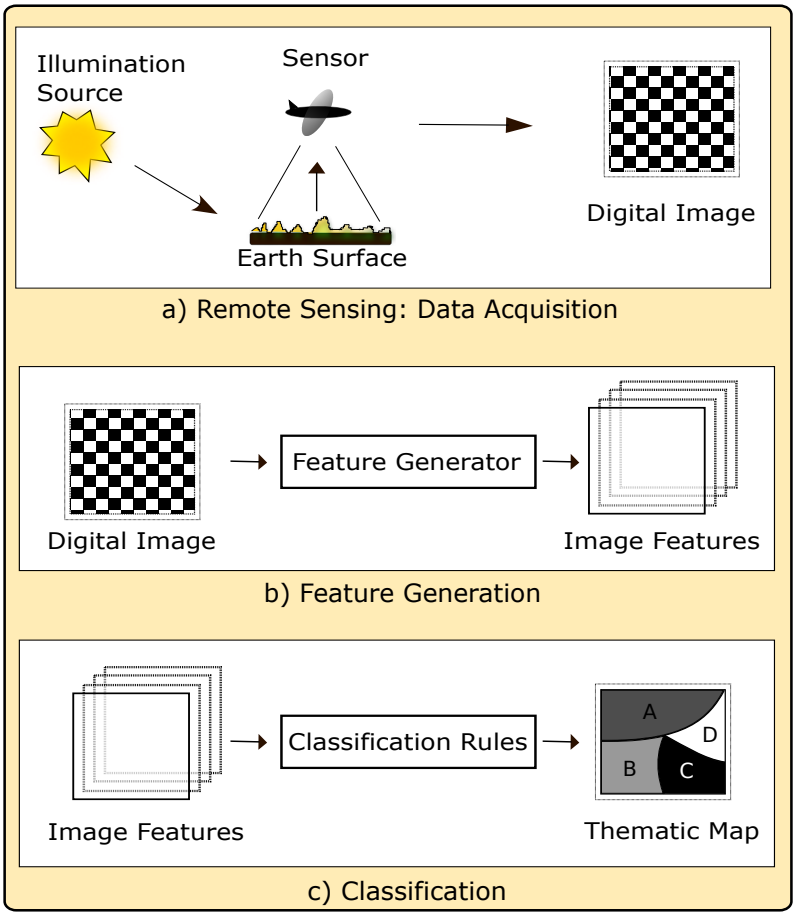


Figure 2.5: An overview of a classification system.

Alternatively, automated systems use unique characteristics (e.g., spatial, structural and contextual) to automatically detect and extract non-geometric image characteristics mentioned earlier (i.e., tone, texture, size, ...) to identify and classify objects of interest. The system makes use of digital RS data, where image interpretation is basically a classification process. This entails the repetition of a number of different activities: data collection, feature selection or

extraction, training/learning and evaluation, see illustrated in figure (2.5). Automated object recognition and feature extraction approaches from imagery have been studied since the 1970s. Depending on the task/information being extracted, three major approaches have been studied, namely spectral-, object- and textural-based methods. Spectral-based methods have been studied to extract spectral informal, object-based method for extracting objects of interest (individual buildings for example) and textural method for extracting textural information. These methods can be used individually or can also be combined depending on what information is being extracted.

2.3.2.1 Spectral-based methods

Classifying urban landscapes, such as buildings and road networks, purely on spectral features is a difficult task as some classes have a constant low reflectance over the whole spectral range with no or only minor distinct absorption features, e.g., gray composite shingle and dark gray tar roofs, asphalt roads and parking lots, etc. These features are important for land cover type separability and should be included to obtain higher classification accuracy. Though spectral methods have been criticised, there are a few studies that have and still focus on the spectral properties of urban materials, for example, spectral techniques have been used in the mapping of impervious urban surfaces and for predicting population [68, 69, 5].

Using spectral methods for urban classification can be useful, however, these methods do not achieve high accuracies [46]. To successfully classify urban surface materials purely on spectral analysis, RS hyperspectral data or advanced evaluation techniques are used. Heiden *et al.* [45] successfully created a database for an area-wide identification of urban surface materials using hyperspectral (HyMap) data to systematically measure, analyse and store urban surface material categories in a spectral library (for wavelengths between $0.35 - 2.5\mu$m in 2151 channels). Their work discussed the importance of different spectral regions in mapping urban areas and found that urban objects do hold significant spectral fingerprints, in some spectral regions. Even though the work showed reasonable spectral recognition, advanced techniques and hyperspectral data are difficult to work with due to the dimensionality of the data sets. For a simpler and effective classification system, authors have suggested the use of object-oriented or other classification techniques using spatial, textural or contextual information might provide a further significant improvement of land cover mapping accuracy and help to overcome spectral similarities between specific classes.

2.3.2.2 Object-based methods

Object-based image analysis (OBIA) analysis has also been shown to be a better approach when compared to spectral based approaches in identifying urban land-cover classes using high resolution imagery [85]. The discriminant analysis of spectra information received an overall accuracy of 63.33% whilst the object-based method produced a significantly higher overall accuracy of 90.40%.

OBIA methods extract dwelling footprints and make estimates based on shape, size, and spacing [9, 44, 67]. Some approaches require the analyst to proceed from the object centroid. For example, Mayunga *et al.* [76] presented a new semi-automated approach for extracting buildings, for informal settlement mapping, from high-resolution QuickBird imagery. Using snake models to effectively extract building, a user is required to measure a single point at the approximate centre of the building. The results obtained in the study were satisfactory in the context of building extraction in a complex environment, but had some shortcomings. Problems were not only experienced in identifying some small buildings, but also the proposed system was not able to accurately delineate building corners. That is, using IKONOS imagery and incorporating digital surface models from specialised LiDAR data, Sohn *et al.* [115] was able to automatically detect and delineate building objects and their boundaries. The technique used consisted of a two step procedure: building detection and building description. The building detection method reduced scene complexity in urban areas and simplified the building description process. The results proved that terrain information extracted from LiDAR data and chromatic cues provided by multispectral bands of IKONOS imagery is important information detecting buildings. The results showed that OBIA can successfully classify an urban scene. However, OBIA methods have several limitations in identifying informal settlements. Examples include: continuous roof lines, improper pixelation of building outlines, and incorrect dwelling separation due to diverse materials on a single roof and they can be computationally expensive when dealing with finer resolution data for a relatively large area [92].

2.3.2.3 Texture-based methods

Texture is one of the most important characteristics used in identifying objects or regions of interest in an image as it can be used to differentiate features that may otherwise have similar reflectance and dimensional characteristics. Despite its importance, a formal definition of texture does not exist. A collection of definitions compiled by [19] demonstrated that the "definition" of texture varies (it is dependent on a person and on the particular application) and as a result

there is no generally agreed upon definition [122]. Tuceryan *et al.* [122] defined image texture as a function of the spatial variation in pixel intensities.

Table 2.2: Categories of texture feature extraction methods.

Category	Description and examples
Statistical methods	First-order, second-order and the higher-order statistics are used for texture discrimination. Texture is described by a collection of statistics of selected features [133]. Generally it is not possible to reconstruct the texture from the features, so these types of methods are usually only used for classification purposes [93]. There have been many surveys on various statistical approaches in the image processing literature [42]. Spatial gray level co-occurrence estimates image properties related to second-order statistics. The grey level co-occurrence matrix (GLCM) [43] is one of the most well known statistical measure in RS literature.
Structural methods	Structural methods are texture analysis methods that characterize texture as being composed of "texture elements" or primitives, arranged according to some placement rules [122, 75]. The method of analysis usually depends upon the structural properties of these texture elements. Structural approaches to texture analysis aim to discern the textural primitive and to determine the underlying structure of the texture. These elements (texels) are organized into a string descriptor, and syntactical pattern recognition techniques are used to measure the similarity of two descriptors [133].
Model-based	In model based techniques, image texture is modelled as a probability model or as a linear combination of a set of basis functions [134]. The model parameters captures the underlying texture property using stochastic or generative models [133]. Model based methods include: Gaussian Markov random fields [21], Gibbs random fields [113], wavelet models [123], Fractal dimension [75], and many other techniques. The Markov Random Field (MRF) model, or variations of it, that is most often used in modelling texture [93].
Signal processing	Various methods of texture analysis rely on signal processing techniques. Signal processing based techniques try to compute certain texture features from filtered images with specific filters which are then used in classification tasks [122]. In textured surfaces, different frequencies have their own textural properties. Both spatial and frequency domain approaches can be used for filtering images and capturing relevant information. Spatial domain filters [72], Fourier domain filtering [20] and Gabor filters [32] are among the popular measures.

Different texture measures have been proposed to address the recognition of different spatial textures. The analysis tends to be more driven by the desired application rather than any pure fundamentals. A summary of possible approaches, generalised in terms of the domain from which the texture feature is extracted, can be found in the following literature in [1, 42, 122, 134, 133, 80, 93]. These can be broadly divided into the following categories: statistical methods, structural methods, model based methods and signal processing methods [122]. A description of the variety of different texture categories are presented on table (2.2 pg. 19).

The most common applications found in the literature are the detection of urban deprivation hot spots, quality of life index assessment, urban growth analysis, house value estimation, urban

population estimation and urban social vulnerability assessment [96]. Examples include, urban population estimation [2, 68, 69, 5], damage assessment [56], assessment of socio-economic status by area [87], growth models and predictions [114, 129], extraction of informal enclaves within the larger settlement fabric [48], and impacts of informal settlements on diminishing natural resources [84].

2.3.3 Classification

Once the appropriate features for representing the input patterns have been measured and extracted from RS imagery, the next step is classification. Classification involves some general model (classifier), using training patterns to classify an unknown/unlabelled pattern to a class/-category. The classifier uses the feature vector provided by the feature extractor as training patterns. A classification task is mainly accomplished in one of two ways: supervised or unsupervised classification. Supervised classification/learning is the process of learning a set of rules from examples (i.e., labelled instances in training data) defined by the system designer, to create a classifier that can be used to generalize to new unlabelled instances. The training samples are labelled to represent the category/class in which the sample belongs.

By contrast, unsupervised classification does not rely on a set of labelled examples. In supervised classification then, the aim is to use training examples to design a classifier which generalises well to new examples. Using the non-labelled examples, the classifier determines the most appropriate group/cluster directly from the image data without referring to any predefined classes. Clustering algorithms are examples of unsupervised classification as they seek to identify groups of unlabelled examples directly within the overall body of data and features which enable them to distinguish one group from another [116]. Since number of classes are learnt along with the structure of each class, the user will likely have to make certain decisions to guide the clustering algorithm, such as the number of desired output clusters or the number of training iterations [59].

2.3.3.1 Choosing a classifier

Given a training set of patterns of known classes (labelled patterns), we seek to design a classifier that is optimal for the classification of unknown pattern classes under consideration based on the measured features. Numerous classification algorithms have been developed using different techniques. However, they all aim to find the decision boundaries that can be used to separate out different classes. That is, given the extracted features the classifier assigns the features to

one of the classes under consideration based on the measured features. The choice of a classifier is a difficult problem and is often based on which classifier(s) happen to be available, or best known, to the user [54].

A large number of classifiers have been developed based on Artificial Intelligence (Logical/Symbolic techniques), Perceptron-based techniques and Statistics (Bayesian Networks, Instance-based techniques) [63]. In remote sensing applications, the most popular learning techniques are the maximum likelihood (ML) classifier, artificial neural network (ANN) classifiers [95] and decision tree classifiers [106]. An alternative classification technique that is not as popular as the above-mentioned is the support vector machines (SVM) classifier. The use of SVM classifier has significantly increased in solving remote sensing problems (see, e.g., [50, 136, 10, 11, 128, 78, 83]). The SVMs are not only becoming popular in RS, they may possibly provide the best classification performance when compared to other known classifiers [17].

2.4 Classifying human settlements

Individual urban objects (e.g., buildings, cars, streets, and vegetation) do not provide much added information when identifying settlement types. Rather, human neighbourhoods are defined by a combination of features of homogeneous zones containing a multitude of urban objects, such as formal or informal settlements. To achieve an ideal settlement classification and monitoring system, a clear definition that defines human neighbourhoods such as formal or informal settlements is critical. This may be difficult as a clear definition of what constitutes informal settlements does not exist as they can vary from country to country (e.g, they can be shacks, slums, squatter areas and shanty areas). According to the UN, informal settlements are defined as unplanned and unauthorised housing having inadequate basic services [41]. This definition of informal settlements is not adequate as it only refers to the legality (authorised or unauthorised) of settlement types. Owen *et al.* [92] established a qualitative description of formal and informal settlements, i.e., the author referred to "Informal" as poor, unplanned neighbourhoods of low economic value while "formal" referred to higher economic value characteristics. These definitions are to be used to differentiate between formal and informal settlements. Visually it is easy to differentiate between formal and informal settlements as they often share unique spatial, structural, and contextual characteristics. For instance, informal settlements characteristics typically include, but are not limited to: 1) high heterogeneity in building orientation, 2) high variance in building materials and density, 3) small building size, 4) irregular and narrow streets, and 5) close spatial proximity to hazardous zones such as landfills, airports, rail roads, and steep

slopes [88, 55]. We can use these characteristics to distinguish informal settlements from other types of urban structures, e.g., formal residential, industrial, and commercial buildings. Unique characteristics such as vegetation, roads, texture, and geomorphology have also shown to play a role in describing the differences between informal and formal settlements [91]. Despite the unique characteristics mentioned above, distinguishing settlement type from other types of urban infrastructure is not a straightforward process due to the diversity of land features, mixed-use settlements, terrain, and heterogeneity of building materials, and neighbourhood structure in informal settlements and the heterogeneous landscape of urban areas [37]. However, using remote sensing imagery, several image classification techniques have shown some success in mapping several cities around the world.

For example, texture based classification has been extensively used in remotely sensed imagery. Texture features have been used to map and understand land-cover patterns in an urban region and urbanisation in multiple cities and for other urban related studies, i.e., [124, 52, 21, 4, 3, 74, 107, 6, 61]. Importantly, they have been demonstrated to be an effective means of describing the extreme heterogeneity of urban surface materials at both inter-pixel and intra-pixel scales that sets apart different urban settlement classes [97, 8, 30, 60, 125, 98, 117, 47].

2.4.1 Human settlements: Formal and informal settlements

As presented in section (2.3.2.3 pg. 18), texture classification involves deciding what texture category an observed object/region of interest, in an image, belongs to. To achieve this, knowledge of the classes/categories needs to be established, thereafter, texture features can be extracted according to the classes. While other feature extraction methods are usually pixel based, texture can only be measured from a group of pixels which makes them ideal for detecting human neighbourhoods. Different texture measures have been proposed to address the recognition of different spatial textures. A summary of possible approaches, generalised in terms of the domain from which the texture feature is extracted, can be found in the following literature in [1, 42, 122, 134, 133, 80, 93]. These can be broadly divided into the following categories, as illustrated in table (2.2 pg. 19).

Using just mono-spectral data and gray-level co-occurrence matrix (GLCM) textural features, Pesaresi [98] and Benediktsson *et al.* [8] demonstrated empirically that it is possible to obtain very good results in terms of automatic discrimination accuracy, even in applications requiring detailed mapping of different built-up surfaces. Since then, numerous works have attempted to improve urban land cover classification using GLCM algorithm to measure textural

information [127, 119, 101, 37]. Using SPOT satellite imagery, Stasolla *et al.* [117] proposed a semi-automatic procedure to detect human settlements (with emphasis on informal settlements) in arid environments. GLCM homogeneity co-occurrence features with the K-Means unsupervised algorithm were used repeatedly to discriminate formal from informal settlements. Their findings showed that the introduction of textures can considerably improve the results of differentiating settlements from arid areas. Based on the GLCM, Pesaresi *et al.* [100] proposed the so-called anisotropic rotation invariant built-up presence index. The method is based on the idea that built-up structures have a certain spatial dimension and can be discriminated from the background by their known spatial relationships. The method calculates a compact built-up area presence using rotational-invariant isotropic GLCM textural analysis on panchromatic satellite data. The texture-derived index was tested under a realistic scenario including degraded and non-calibrated data input and extensive validation exercise and obtained an overall accuracy of 86.7%. The method overestimated the built-up areas in cases where scattered vegetation had the same spatial pattern as settlements, to reduce the problem the method incorporated the normalised difference vegetation index (NDVI), gaining +20.76% improvement over the basic procedure [99].

Other than optical imagery, GLCM textural features have been used with other RS sensors for urban analyses. For example, using SAR data, Dell *et al.* [27] used the GLCM texture features to measure information on the difference in building densities inside a town structure. The results showed that it is possible to extract some information on urban environments from current satellite SAR images and classify them with respect to building density. SAR data has also shown the potential for mapping human settlement extents in various parts of the world [35, 36, 34].

An alternative approach from the statistical approach (i.e., GLCM), is to measure texture elements in urban areas using either structural, model-based and signal processing methods or a combination. Using high resolution QuickBird imagery, Khumalo *et al.* [60] utilised rotation invariant Gabor Filters (signal processing approach) along with the GLCM to perform the identification of the different textural regions in urban areas. In order to select the most appropriate texture algorithm for an automated informal settlement classification system, Abeigne Ella *et al.* [30] performed an experiment to compare the performance of the GLCM with that of nine other texture features. Some of the texture features were able to separate the different urban settlement classes very well, these include, Lacunarity measures [105], Discrete Wavelet Transform (DWT) [112, 13], Granulometric [18] , Local Binary Pattern (LBP) [90] and GLCM

methods. In addition, it appeared that the LBP features are more powerful than the commonly used GLCM features for this particular problem.

2.4.2 Challenges

Although some texture features, such as those generated by the LBP algorithm have been applied effectively in the settlement classification task, subsequent experiments have demonstrated that the generalization performance of classifiers using these features are not ideal [125]. The poor generalization performance is mainly caused by two dominant factors: viewing- and illumination geometry. Data of the same area acquired under different dates (cross-date images) may introduce large off-nadir viewing angles which produce two images with different viewing- and illumination geometries. This also introduces spurious differences between the two images, and is referred to as illumination geometry differences. Illumination geometry differences alters the orientation and length of the shadows between the image pair, this results in decreased classification accuracy. In addition, the seasonality and time of day during acquisition determine the sun elevation and azimuth angles of the landscape which may also hinder classification performance. Even though no real change has occurred on the ground, these effects produce an image-pair that may contain a large number of spurious differences [70, 125]. For multitemporal aerial imagery, the flight strips/paths can also have a negative impact in generalization performance. This can be caused by the difference in the flight paths and altitude between the different acquisition dates. A good image feature is one that is designed to have a representation that is sensitive to change in the desired variables, e.g., settlement type, whilst being insensitive to other types of change that may be present in the image [125]. If an image feature has these properties, then it can be expected that the feature will lead to good generalization performance in classification tasks.

2.4.3 Proposed solutions

A 2-dimensional surface texture has two properties, spatial structure (pattern) and contrast (intensity of texture). The purpose of texture description is to derive some measurements that can be used to classify a particular texture. Texture measurements are required to be in invariant with respect to position, scale and rotation, and that texture extraction can apply equally. The multiresolution LBP algorithm is invariant with respect to position, scale and rotation and by definition is invariant to any monotonic changes in gray-scale. The multiresolution LBP algorithm is thus an ideal measure for spatial structures such as human settlements [30]. However,

due to viewing- and illumination-geometry effects, the LBP algorithm was shown to offer less than ideal generalization performance [125]. Ojala *et al.* have demonstrated that combining spatial structure with the gray-level contrast can improve discrimination ability of texture features [89]. To improve performance, the multiresolution LBP algorithm was extended with a contrast extension (here denoted as the extended LBP) as a joint distribution of gray-scale and rotational invariant LBP with the rotational invariant Variance measure [90]. The addition of contrast features in the extended LBP (LBP/VAR) algorithm is expected to improve settlement classification accuracy.

2.5 Summary

We reviewed the fundamental concepts and history of the remote sensing field for a detailed understanding of the acquisition of RS sources and available applications for monitoring the Earth's surface. We reviewed the pros and cons of either using satellite or aerial imagery or the trade-offs in using low, medium or high resolutions. Aerial RS sources have a flexibility advantage while satellites provide a more stable and repeatable RS source. Both airborne and spaceborne RS sources are capable of achieving high temporal, spatial and spectral resolutions. This makes the choice of RS data source for automated or semi-automated classification tasks, particularly in urban areas, a matter of availability and cost. However, the literature showed that the use of high spatial resolution data is more appropriate for urban and suburban classification tasks. Following data acquisition, data analysis was reviewed.

Three major feature extraction methods were reviewed: Spectral-based, Object-based and Texture-based methods. Spectral-based methods showed to be reasonable for distinguishing urban features. However, in order to achieve good results, advanced techniques and high dimensional data (hyperspectral data) is used. This made the method expensive and difficult to work with. When using high resolution imagery, other methods have been shown to be a better approach when compared to spectral based approaches in identifying urban land-cover classes, one such approach is known as object-based image analysis (OBIA). These methods identify individual objects, e.g., urban objects such as buildings and other materials. However, this is also the method's limitation when trying to identify regions or groups of objects. For example, when identifying settlements instead of individual buildings, the diverse materials on a single roof result in continuous roof lines, improper pixelation of building outlines, and incorrect dwelling delineation. In addition, OBIA methods can be computationally expensive when dealing with finer resolution data for a relatively large area.

Texture-based methods were also reviewed for urban and suburban analysis. These methods proved to be the most appropriate for classification of urban material, particularly at a region scale. Numerous texture measures have been proposed to address the recognition of different spatial textures found in urban areas. Examples include: the detection and monitoring of urban deprivation hotspots, quality of life index assessment, urban growth analysis, and many other applications. For this reason, texture-based methods were considered for the case study, which is to improve classification of human settlements. A study comparing texture-based methods for settlement classification using QuickBird imagery showed the GLCM and LBP algorithms to be the better performing methods.

Although the GLCM and LBP methods showed to be effective at classifying settlements, they performed poorly when presented with imagery acquired on different dates. The poor generalization ability was mainly caused by two dominant factors: viewing- and illumination geometry. In an attempt to improve generalization performance, a feature extraction method that is robust to these effects is required. The proposed solution is a method that is invariant to rotation which is expected to minimise the viewing geometry effects. To minimise illumination differences, input images are normalised using histogram equalisation or histogram matching techniques. However, this is not a general solution, as global histogram equalisation cannot correct intra-image (local) gray-scale variations. Since the LBP is, by definition, invariant to monotonic changes in gray scale, it was extended by supplementing it with an orthogonal measure of local contrast. This method, denoted as the extended LBP, is thus invariant with respect to position, scale and rotation while at the same time sensitive to local contrast features.

Since feature extraction is the key step in classification tasks, an image feature with the above-mentioned properties can be expected to yield good generalization performance in classifying settlements. To learn these features a good classification algorithm or classifier is important. The SVM has been showed to possibly provide the best classification performance when compared to other known classifiers. The SVM was the chosen classifier as it has the ability to find the right balance between accuracy attained on a given finite amount of training patterns, and effectively generalize to unseen data.

Chapter 3

Methods

3.1 Introduction

This chapter presents a detailed description and derivation of feature extraction and classification methods that are going to be used for the task of classifying human settlements. The extended LBP (LBP/VAR) texture feature extraction method is presented first, followed by the Support Vector Machine (SVM) classifier, used for the classification task.

3.2 Extended Local Binary Patterns

The extended LBP is a multiresolution LBP algorithm with a contrast extension presented as a joint distribution of gray-scale and rotational invariant LBP with the rotational invariant Variance measure (LBP/VAR). In this section, the derivation of the extended LBP algorithm proposed by Ojala *et al.* [90] is presented.

3.2.1 Gray-Scale Rotation Invariant Local Binary Patterns

Firstly, let us define texture T in a local neighbourhood of a gray-scale image as the joint distribution of the gray-levels of $P + 1$ $(P > 0)$ image pixels:

$$T = t(g_c, g_p), \tag{3.1}$$

where g_c is the gray-value of the centre pixel and $g_p(p = 0, \ldots, P - 1)$ corresponds to the gray-values of P pixels equally spaced on a circle of radius $R(R > 0)$ that form a circularly symmetric set of neighbours, see figure (3.1 pg. 28). The coordinates of the neighbours g_p are given by $(y_c - R\sin(2\pi p/P))$,$(x_c + R\cos(2\pi p/P))$, taking the centre as our origin $g_c(0, 0)$, the coordinates of the gray-values g_P are then $(-R\sin(2\pi p/P))$,$(R\cos(2\pi p/P))$. Gray-values that do not fall exactly in the centre pixel are estimated by interpolation.

Without losing information, we can subtract the centre pixel g_c from the values of the neighbours g_p:

$$T = t(g_c, g_0 - g_c, \ldots, g_{P-1} - g_c). \tag{3.2}$$

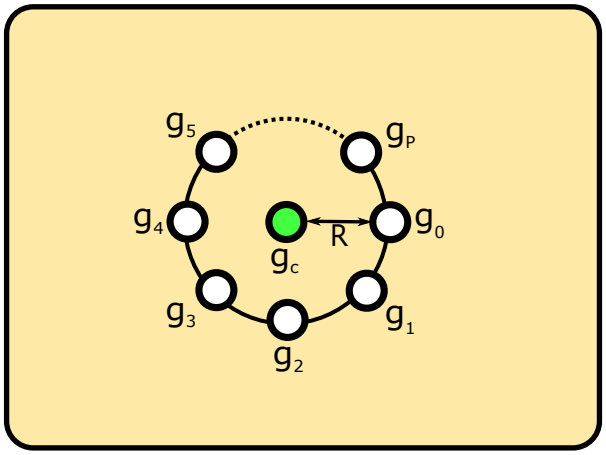


Figure 3.1: Local circularly symmetric neighbourhood set (P samples) of radius R.

Assuming g_c is independent of $(g_p - g_c)$, the distribution can factorised to:

$$T \approx t(g_c)t(g_0 - g_c, \ldots, g_{P-1} - g_c). \tag{3.3}$$

We can ignore $t(g_c)$ as it describes the overall luminance of an image, which is unrelated to local image texture:

$$T \approx t(g_0 - g_c, \ldots, g_{P-1} - g_c). \tag{3.4}$$

The above expression is invariant with respect to gray-scale shifts, but is affected by scaling. To achieve invariance with respect to any monotonic transformation of the gray-scale, only the signs of the differences are considered:

$$T \approx t(s(g_0 - g_c), \ldots, s(g_{P-1} - g_c)), \tag{3.5}$$

where

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0. \end{cases} \tag{3.6}$$

By assigning binomial coefficient 2^p to each sign $s(g_p - g_c)$, the above is transformed into a unique P-bit pattern code, that characterises the spatial structure of the local image texture:

$$\mathrm{LBP}_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p. \tag{3.7}$$

That is, the threshold of local neighbourhood is at the gray-value of the centre pixel into a binary pattern.

Rotation Invariance

To remove the effect of rotation, each LBP code must be rotated back to a reference position, effectively making all rotated versions of a binary code the same. This transformation can be defined as

$$LBP^{ri}_{P,R} = \min\{ROR(LBP_{P,R}, i)|i = 0, 1, \ldots, P-1\}, \tag{3.8}$$

where ri stands for "rotation invariant". The function $ROR(x, i)$ performs a circular bit-wise right shift on the P-bit number x i times.

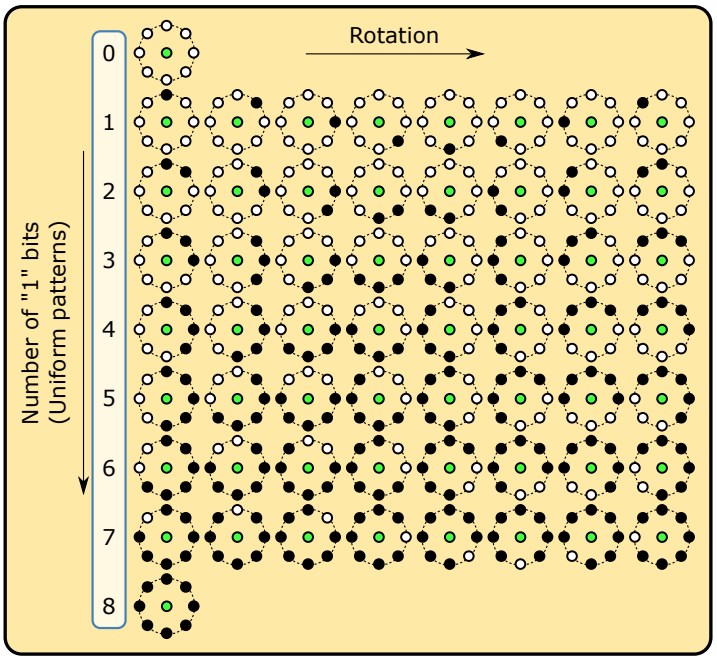


Figure 3.2: The Look Up Table (LUT): to achieve rotation invariance the LUT stores all the possible uniform patterns i.e., for P = 8, nine "uniform" patterns with the numbers (0 – 8) corresponding to their unique $\mathrm{LBP}^{\mathrm{riu2}}_{8,\mathrm{R}}$ codes.

In an attempt to improve the rotation invariant LBP^{ri} features, Ojala *et al.* introduced "uniform patterns" (uniform circular structures, as illustrated in figure 3.2). To formally define the "uniform" patterns, a uniformity measure U ("pattern"), which corresponds to the number of spatial transitions (bitwise 0/1 changes) in the "pattern" is needed. A pattern that has a value of U that is at most 2 is labelled as "uniform" resulting in the following operator for gray-scale and rotation invariant texture description:

$$\mathrm{LBP}^{\mathrm{riu2}}_{P,R} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \leq 2 \\ P+1 & \text{otherwise,} \end{cases} \tag{3.9}$$

where

$$U(LBP_{P,R}) = \begin{array}{l} |s(g_{P-1} - g_c) - s(g_0 - g_c)| \\ + \sum_{p=0}^{P-1} |s(g_P - g_c) - s(g_{P-1} - g_c)|. \end{array} \tag{3.10}$$

Superscript $(\mathrm{riu2})$ is $r$otation $i$nvariant "$u$niform" binary patterns that have U value of at most 2. By definition, exactly $P + 1$ "uniform" binary patterns can occur in a circularly symmetric neighbour set of P pixels. Equation (3.9) assigns a unique label to each of them corresponding to the number of "1" bits in the pattern $(0 \rightarrow P)$, while the "nonuniform" patterns are grouped under the "miscellaneous" label $(P + 1)$, see figure (3.3).

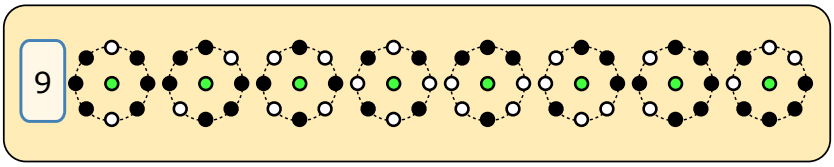


Figure 3.3: Examples of nonuniform patterns that are labelled as 9.

In practice, the mapping from $LBP_{P,R}$ to $LBP^{riu2}_{P,R}$, which has $P+2$ distinct output values, is best implemented with a lookup table of 2^P elements. The final texture feature employed in texture analysis is the histogram of the operator outputs (i.e., pattern labels) accumulated over a texture sample.

3.2.2 Rotational Invariant Variance Measures (Contrast Extension)

To incorporate the contrast of local image texture, we measure it with a rotation invariant measure of local variance. Like the LBP, the rotation invariant local variance can be measured in a circularly symmetric neighbour set as

$$\mathrm{VAR}_{P,R} = \frac{1}{P}\sum_{p=0}^{P-1}(g_p - \mu)^2, \text{where } \mu = \frac{1}{P}\sum_{p=0}^{P-1} g_p \tag{3.11}$$

is the sample mean. $\mathrm{VAR}_{P,R}$ is, by definition, invariant against shifts in gray-scale and rotation along the circular neighbourhood but it is not invariant to global contrast changes. A rotation invariant description of texture in terms of texture patterns is obtained with the joint distribution of LBP and local variance, denoted by $LBP^{riu2}_{P,R}/VAR_{P,R}$, as shown in figure (3.4).

3.2.3 Multiresolution

A general rotation-invariant operator for characterizing the spatial pattern and contrast of local image texture using circular symmetric neighbour set of P pixels with radius R has been

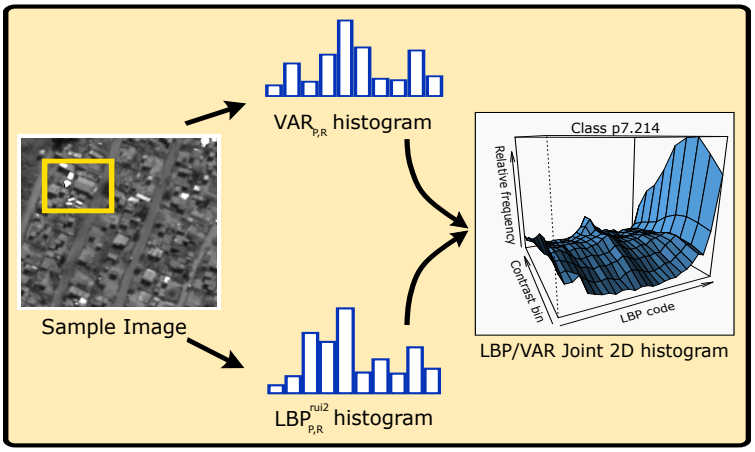


Figure 3.4: Example of extracted LBP and local variance (VAR) features (1D histograms) for settlement type p7.214 used to calculate a joint 2D histogram feature , denoted as $LBP^{riu2}_{P,R}/VAR_{P,R}$.

presented. Quantization of the angular space and spatial resolution can be realized by altering parameters P and R. A simple method of enlarging the spatial support area is to combine the information provided by multiple operators of varying (P, R), see figure (3.5). The resulting multiresolution LBP and LBP/VAR feature vector is simply the concatenation of the component LBP and LBP/VAR feature vectors corresponding to the choices of (P, R).

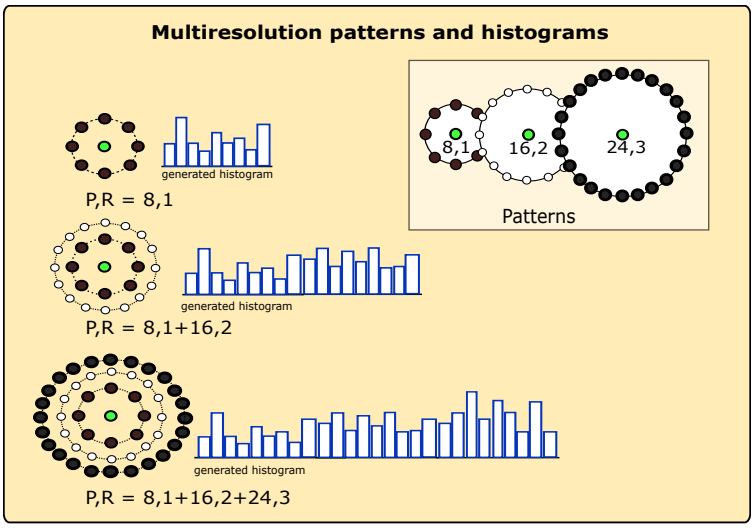


Figure 3.5: Multiresolution features obtained by concatenating LBP features extracted at multiple (P, R) parameters.

3.3 Support Vector Machines

A Support vector machine algorithm is a supervised non-parametric machine learning algorithm. That is, the method is presented with a set of labelled data instances where there are

no assumptions made on the underlying data distribution. Given the training data, the SVM finds a separating hyperplane in some feature space induced by the kernel function while all the computations are done in the original space [126]. An optimal separating hyperplane is the one that separates the data with maximum margin, minimising misclassification obtained in the training step. The description and specification of the separating hyperplane does not use all the available training examples, selected training samples or *support vectors* near the decision boundary are the only points that define the hyperplane of maximum margin. This can provide good generalization performance in pattern classification problems [17].

SVMs are particularly appealing in the remote sensing field due to their ability to find the right balance between accuracy attained on a given finite amount of training patterns, and the ability to generalize to unseen data [83, 73]. Comparisons of the SVM with conventional classifiers confirmed the good performance of the SVMs [33, 50, 94, 78]. In an assessment study by Huang *et al.* [50], the SVM algorithm was found to be competitive when its performance was compared with the best available classification methods, ML classifier, neural network classifiers (NNC) and decision tree classifiers in land cover classification. The study also revealed that the performance of the SVM classifier is influenced by the configurations of SVM kernel type and kernel parameter. Pal *et al.* [94] investigated SVMs for classification in remote sensing and suggested that the SVM can achieve higher accuracies than either of the ML or ANN classifiers. The underlying principle that benefits SVMs is that, they minimise classification error on unseen data without prior assumptions made on the probability distribution of the data (structural risk minimisation) while statistical techniques such as maximum likelihood estimation usually assume that data distribution is known a priori [83]. SVMs have demonstrated impressive performance in classifying hyperspectral data acquired from the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) [39, 38], in classifying urban environments [136] and in spectral and spatial classification of hyperspectral data [31]. As described above, SVMs have shown remarkable abilities to deal with remote sensing data and have proved better suited to cope with the extremely high dimensionality of the data (e.g. hyperspectral data), and with the limited availability of training samples in remote sensing applications [104].

3.3.1 Theoretical development of SVMs

In this section we only give a very brief introduction to SVM's, a well-organised SVM tutorial can be found in [12, 25]. The underlying principle behind SVM is the learning process that follows what is known as structural risk minimisation (SRM). Consider training data

$\mathbf{S} = \{(\mathbf{x}_i, y_i) \in P(\mathbf{x}, y)\},\ i = 1, \ldots, N$, where for each sample, $\mathbf{x}_i \in \mathbb{R}^d$, belongs to a class labelled by $y_i \in \{+1, -1\}$ and $P(\mathbf{x}, y)$ is some unknown probability distribution from which data **S** is drawn, i.e., the data are assumed "iid" (independently drawn and identically distributed). We need to find a deterministic mapping function, $f(\mathbf{x}, \alpha) : \mathbf{x}_i \mapsto y_i$ based on a sample data **S**, where functions $f(\mathbf{x}, \alpha)$ themselves are labelled by the adjustable parameters α. A particular choice of α generates what is called a "training machine". As a result, the expected risk of a learning machine (R) is

$$R(\alpha) = \int \frac{1}{2}|y - f(\mathbf{x}, \alpha)|\mathrm{d}P(\mathbf{x}, y). \tag{3.12}$$

The quantity $R(\alpha)$ (known as the risk error) is the true mean error, but it is not useful unless we can have an estimate of $P(\mathbf{x}, y)$. The empirical risk $R_{\text{emp}}(\alpha)$ is a fixed number defined as

$$R_{\text{emp}}(\alpha) = \frac{1}{2l}\sum_{i=l}^{l}|y_i - f(\mathbf{x}_i, \alpha)|, \tag{3.13}$$

where the quantity $\frac{1}{2}|y_i - f(\mathbf{x}_i, \alpha)|$ is the loss.

According to the SRM principal, the risk of a learning machine $R(\alpha)$ is bound by the sum of the empirical risk estimated from training samples (R_{emp}) and a confidence interval (Ψ): $R(\alpha) \leq R_{\text{emp}} + \Psi$ [126]. The strategy of SRM is to keep the empirical risk (R_{emp}) fixed and to minimise the confidence interval (Ψ), or to maximise the margin between a separating hyperplane and closest data points. A separating hyperplane refers to a plane in a multidimensional space that separates the data samples of two classes, see illustration in figure (3.6). Under this scheme, SVMs minimise classification error on unseen data without prior assumptions made on the probability distribution of the data.

3.3.2 Linear SVM: Optimal Separating Hyperplane (OSH).

In its simplest form, SVM is a linear binary classifier that assigns a given test sample a class from one of the two possible labels.

3.3.2.1 Separable Case

Suppose we have some hyperplane which separates the positive from the negative examples (i.e., a separating hyperplane), see figure 3.6-A. The points **x** which lie on the hyperplane satisfy

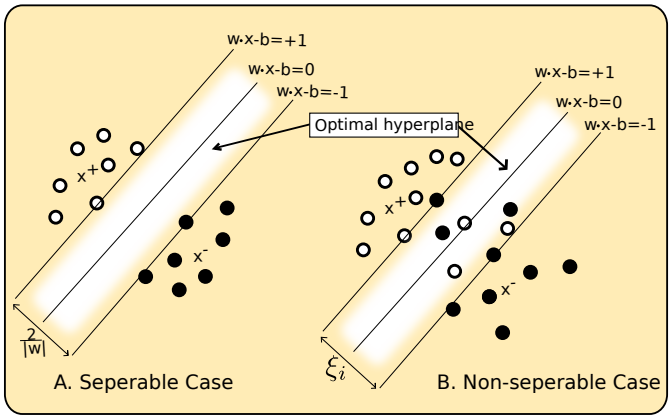


Figure 3.6: The optimal separating hyperplane between (A) separable samples and (B) non-separable data samples.

$\mathbf{w} \cdot \mathbf{x} + b = 0$, where $\mathbf{w}$ is normal to the hyperplane, $|b|/||\mathbf{w}||$ is the perpendicular distance from the hyperplane to the origin, and $||\mathbf{w}||$ is the Euclidean norm of $\mathbf{w}$ [12]. For the linearly separable case, all the training data $\mathbf{S}$ satisfy the following constraints:

$$\mathbf{w} \cdot \mathbf{x}_i + b \geq +1 \quad \text{for} \quad y_i = +1, \tag{3.14}$$

$$\mathbf{w} \cdot \mathbf{x}_i + b \leq -1 \quad \text{for} \quad y_i = -1. \tag{3.15}$$

Linear discriminant and perceptron learning algorithms essentially do the same thing, they find a linear separator by adjusting weights on misclassified examples. The above inequalities can written in the form:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \quad \forall i. \tag{3.16}$$

For linearly separable classes x^+ and x^- (see figure 3.6-A) it is easy to show that the maximum margin is simply $2/||\mathbf{w}||$ (i.e., the margin is inversely proportional to the norm of $\mathbf{w}$). Thus, the pair of hyperplanes which gives the optimal separating hyperplane can be found by minimizing the square of the norm $\mathbf{w}$ (i.e., minimize $||\mathbf{w}||^2$), subject to constraints (3.16). This can be solved by constructing a Lagrangian function

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2}||\mathbf{w}||^2 + \sum_{i=1}^{N} \alpha_i - \sum_{i=1}^{N} \alpha_i y_i (\mathbf{w} \cdot \mathbf{x}_i + b), \tag{3.17}$$

where the coefficient α_i is a Lagrange multiplier, and by transforming it into the corresponding dual Lagrangian by imposing the optimal conditions,

$$\frac{\partial L}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i = 0, \tag{3.18}$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{N} \alpha_i y_i = 0. \tag{3.19}$$

The result is a quadratic programming problem with linear constraints

$$W(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_i), \tag{3.20}$$

$$\alpha_i \geq 0, \tag{3.21}$$

$$\sum_{i=1}^{N} \alpha_i y_i = 0, \tag{3.22}$$

that presents just a global maximum and can always be exactly solved efficiently. The resulting solution has the property that

$$\mathbf{w} = \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i \tag{3.23}$$

and in fact, often most of the coefficients α_i are equal to zero. The only positive coefficients correspond to the points that lie closest to the hyperplane are the support vector points and the solution is represented as a linear combination of only these points. The final decision function can be written as

$$f(x) = \mathbf{w} \cdot \mathbf{x} + b, \tag{3.24}$$

$$f(x) = \text{sign}\left(\sum_{i=1}^{s} \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{x}) + b \right), \tag{3.25}$$

where the index i runs only on the support vectors s. This property is known as sparseness [26].

3.3.2.2 Non-separable Case: Soft Margin Hyperplane

It is not always easy to classify patterns accurately using only linear separable decision boundaries, especially if the data points of different classes overlap one another. In this case the data

is said to be inseparable, see figure 3.6-B. One way to separate the training data with a minimal number of errors is to introduce some non-negative variables (known as slack variables) $\xi_i \geq 0, \quad i = 1, \ldots, N$. We can now minimize the function

$$||\mathbf{w}||^2 + C \sum_{i=1}^{N} \xi_i, \tag{3.26}$$

subjected to the constraints (known as Soft Margin constraints):

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i, \quad i = 1, \ldots, N, \tag{3.27}$$

$$\xi_i \geq 0, \quad i = 1, \ldots, N. \tag{3.28}$$

The constant C controls the number of allowed training errors. The Lagrangian now becomes:

$$L(\mathbf{w}, b, \alpha, \xi, \mu) = \frac{1}{2}||\mathbf{w}||^2 + \sum_{i=1}^{N} \alpha_i[1 - \xi_i - y_i(\mathbf{w} \cdot \mathbf{x}_i + b)] - \sum_{i=1}^{N} \mu_i \xi_i + C \sum_{i=1}^{N} \xi_i, \tag{3.29}$$

where μ_i are Lagrange multipliers introduced to enforce positivity of the slack variables ξ_i. The dual problem:

$$\text{maximize:} \quad L_D = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j), \tag{3.30}$$

$$\text{subjected to:} \quad 0 \leq \alpha_i \leq C, \tag{3.31}$$

$$\sum_{i=1}^{N} \alpha_i y_i = 0. \tag{3.32}$$

The solution is again given by

$$\mathbf{w} = \sum_{j=1}^{s} \alpha_j y_j \mathbf{x}_j, \tag{3.33}$$

where s is the number of support vectors. Ultimately, the only difference from the optimal hyperplane case is that the values α_i now have an upper bound of C.

3.3.3 Nonlinear SVM: Kernel Functions

To generalize the above method to nonlinear decision functions, the support vector machine maps the input data onto a higher dimensional (Euclidean or Hilbert) space and defines a separating hyperplane there. The mapping of input data makes it possible to transform nonlinear relations within the data into linear ones [26]. The input vector **x** is mapped into a higher feature space H through a mapping function Φ,

$$\Phi : \mathbb{R}^d \mapsto H. \tag{3.34}$$

Notice in the training problem in equation (3.32), the data is in the form of dot products, $\mathbf{x}_i \cdot \mathbf{x}_j$. The training algorithm in the high dimensional space H would then only depend on data through dot products in H, i.e. on functions of the form $\Phi(x_i)\Phi(x_j)$. $\Phi(x)$ represents the input vector **x** in the high-dimension space H. If there is a kernel function K such that $K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(x_i)\Phi(x_j)$, we would only need to use K in the training algorithm without knowing the explicit form of Φ. Replacing the inner product in (3.32) by the kernel function K:

$$\text{maximize:} \quad L_D = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j), \tag{3.35}$$

$$\text{subjected to:} \quad 0 \leq \alpha_i \leq C, \tag{3.36}$$

$$\sum_{i=1}^{N} \alpha_i y_i = 0. \tag{3.37}$$

The SVM is now a non-linear classifier, and the final decision function can be written as

$$f(x) = \text{sign}\left(\sum_{i=1}^{s} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \right). \tag{3.38}$$

A kernel that can be used to construct a SVM must meet Mercer's condition. The following two types of kernels meet this condition [50]:

- The polynomial kernels,

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^p, \tag{3.39}$$

- and the radial basis function (RBF) kernel,

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{||\mathbf{x}_i - \mathbf{x}_j||^2}{2\sigma^2}\right). \tag{3.40}$$

3.3.4 Multi-class SVM

In the above theoretical development, the SVM was developed as a binary classifier where the class labels can only take two class labels, e.g., ± 1. This is a problem for multiple class problems, such as remote sensing problems. Various approaches have been proposed to address this problem. The simple SVM binary classifier is adjusted to operate as a multi-class classifier using methods such as one-against-all and pairwise method.
Consider N-class problems:

Pairwise method, constructs a machine for each pair of classes, resulting in $N(N-1)/2$ machines. When applied to a test pixel each machine gives one vote to the winning class, and the pixel is labelled with the class having most votes.

One-against-all method, breaks the N-class case into N two-class cases, in each of which a machine is trained to classify one class against all others. When applied to a test pixel, a value measuring the confidence that the pixel belongs to a class can be calculated from equation (3.38) and the pixel is labelled with the class with the highest confidence value.

3.3.4.1 Performance evaluation

The main task of a classifier is to learn how to classify unseen data from a set of given examples, it is common to evaluate classifiers performance as a number of correctly predicted instances divided the total number of tested samples [63], that is:

$$\text{Accuracy} = \frac{\text{Number of correctly classified samples}}{\text{Total number of samples}} \times 100\%. \tag{3.41}$$

The performance of a classifier depends on the number of available training samples and the feature vectors. Depending on the available data, classification accuracy can be measured in one of the following ways:

- either by splitting the data into two-thirds to be used for training and a third to be used for testing,

- by dividing the training set into mutually exclusive and equal-sized subsets and for each subset the classifier is trained on the union of all the other subsets (which is known as *cross-validation*)

If data is scarce, it is sensible to use cross validation in order not to waste any data, which could be useful to enhance classifier performance; all data is used both for training the classifier and for testing its performance. Classification accuracy may yield unsatisfactory performance due to high dimensionality, not large enough training set, inappropriate features or algorithms. If the evaluation of classifier shows poor generalization ability, various factors must be examined: perhaps the number of features is too large relative to the number of training samples, relevant features for the problem are not being used or the number of parameters associated with the classifier is large, the dimensionality of the problem is too high, the classifier is too intensively optimised on the training set (over-training), the selected algorithm is inappropriate. To fix these, a larger training set may be needed or parameter tuning for algorithm is needed.

3.4 Summary

A detailed description and theoretical development of the extended LBP and SVM classifier for the task of classifying settlements was presented. The extended LBP method was based on a joint distribution of LBP and local variance (To incorporate the contrast of local image texture) proposed by Ojala *et al.* [90]. The multiresolution gray-scale and rotational invariance of the LBP, the rotational invariance of local variance and "uniform" patterns were the main properties of the method. The theoretical development of the SVM was presented based on [12, 25]. The main properties of the SVM theoretical development presented were the linear SVM optimal separating hyperplane (for separable and non-separable cases), kernel functions for nonlinear SVM and multi-class SVM. The design of the experiment and implementation of these methods is presented in chapter 4.

Chapter 4

Experimental Design

4.1 Introduction

An investigation of whether the addition of contrast features to the normal LBP algorithm can aid in improving the classification of settlement type of aerial images was conducted. To evaluate the classification accuracy improvements and the significance of the contrast features in classifying settlement type, the extended LBP (LBP with local contrast features) was compared with the original LBP method (which omits the local contrast components).

The design and implementation of this investigation is described in this chapter. The chapter starts with a section presenting a detailed review of the data set, these include the study area, data properties (e.g., sensor type) and the selection of training and testing set. The chapter ends with several implementations, such as, the process of extracting image features from the input data using LBP and LBP/VAR algorithms and the evaluation process (training and testing) of the SVM classifier are presented. Figure 4.1 presents an illustration of the experimental design and implementation.

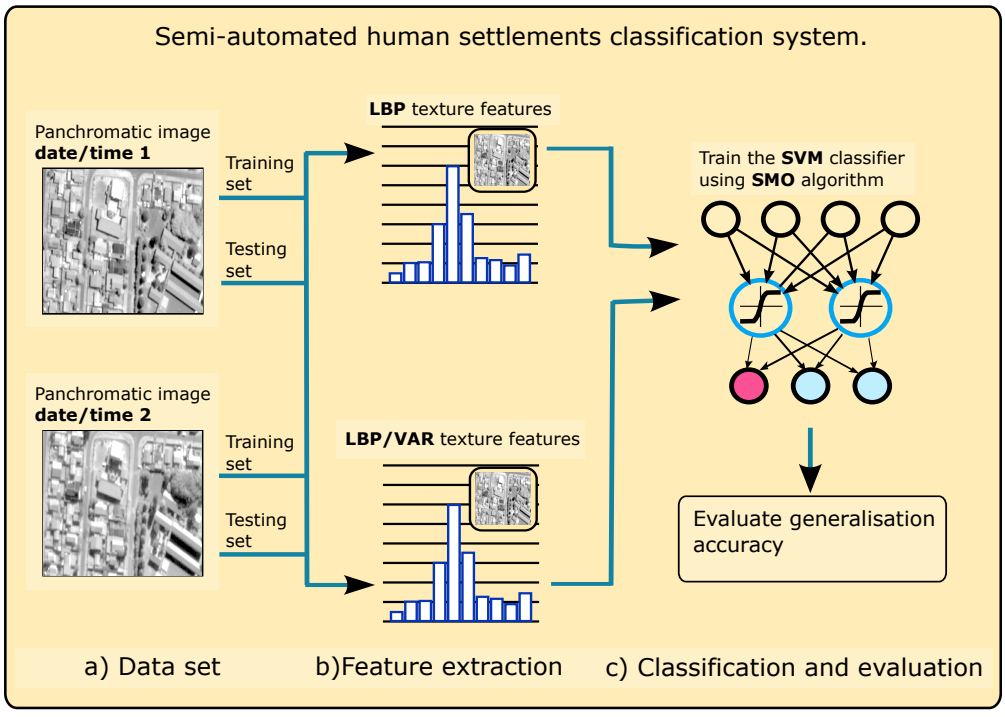


Figure 4.1: A summary of the experiment layout.

4.2 Data Preparation

High spatial resolution (panchromatic) aerial images were acquired over the same area at different dates. The difference in acquisition dates allowed the investigation of the robustness of the LBP/VAR algorithm to large differences in viewing- and illumination geometries between the images, and the generalization evaluation of the classifier. From the acquired urban scenes, training and testing samples were labelled according to settlement types available in the study area. This section presents the acquisition and selection of settlement classes in more detail.

4.2.1 Data Acquisition

4.2.1.1 RS sensor

The images were acquired using a Digital Mapping Camera (DMC), panchromatic airborne sensor. The system consists of eight independent camera modules (units) that capture a central perspective view. The combined image of the seperate lenses produce higher optical performance than can be achieved in a single, larger-diameter lens. For the simultaneous collection of colour and false colour infrared images, four multi-spectral (MS) camera heads (red, green, blue and NIR) are incorporated in the camera base unit, see illustration in figure (4.2a). The DMC captures a square pixel footprint where all camera heads are exposed simultaneously, the image is thus frozen in one shot, see illustration in figure (4.2b). This minimizes unfavourable influences due to airspeed fluctuation, sudden aircraft movement, or objects moving within the frame [24].

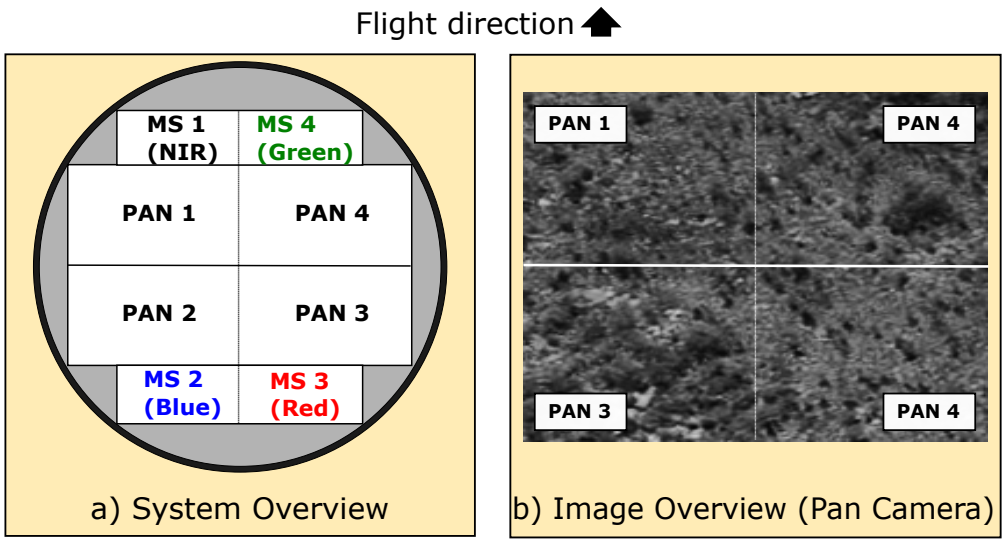


Figure 4.2: System and image overview of a Digital Mapping Camera.

Each camera offers a pixel size of 12×12 μm and a 12-bit per pixel radiometric resolution

for each of the panchromatic and colour channels with a 0.12 m focal length at a maximum aperture of f/4. Detailed camera parameters of the data used for the study are reported in table 4.1.

Table 4.1: Camera Parameters for Virtual Image (High and Colour Resolutions).

Resolution	Focal length (m)	Sensor size (Pixel)	Pixel size (μm)	Principal point (mm)
PAN	0.12	13824×7680	12	X = 0.0 Y = 0.0
MS (before PPS)	0.12/4.75	3075×2048	12	X = -0.646 Y = 0.646
MS (after PPS[1])	0.030	3456×1920	12	X = 0.0 Y = 0.0

[1]The results of the Platform calibration were generated with DMC Postprocessing SW (PPS), Version 5.2, from Intergraph Z/I Imaging photogrammetric product suite.

4.2.1.2 Study area

The Mamelodi area, part of the City of Tshwane Metropolitan Municipality, north-east of Pretoria (Gauteng, South Africa) shown in figure (4.3 pg 43) was chosen as the study area. The area was suitable for our research as it has a high percentage of formal and informal areas within a single square kilometre image scene and includes new estate developments and upgrades (i.e., informal to formal housing). Settlements in this area comprise several informal and formal settlement characteristics (subclasses) ranging from housing being either dense or sparse, with-/without backyard structures, ordered or unordered and having dirt or tarred roads. The area was also chosen because of the lack of remote sensing and GIS-based research on informal settlements in this particular area.

4.2.1.3 Cross-dates

To effectively evaluate the generalization ability of the SVM classifier, high spatial resolution data was acquired over the same area (Mamelodi) on different dates (i.e., year 2010 and 2012, denoted by d_1 and d_2 respectively). Each date of acquisition contained multiple "images", as shown in figure (4.4 pg 43). To reduce differences in image properties, the image pairs were converted to 8-bit colour depth and histogram matched (to reduce brightness differences). With the relevant data acquired, the next step is to select and extract sample classes, from the Mamelodi scene, to be used as training and testing data sets.

4.2.2 Sample Selection

The study area contained a large variety of settlements, ranging from formal suburbs to very informal settlements consisting of small shacks. A variety of key characteristics (tone, shape,

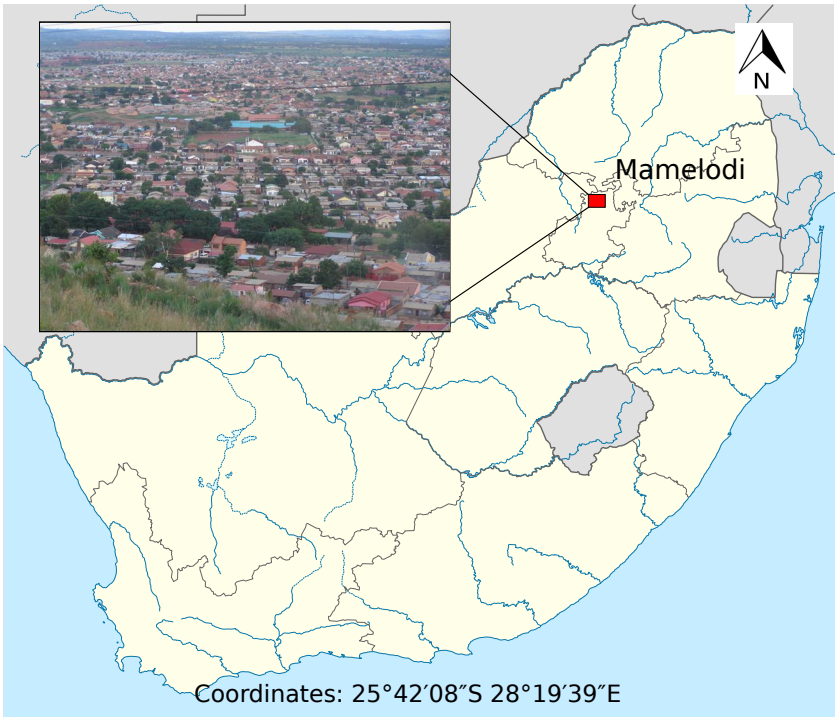


Figure 4.3: Map showing Mamelodi area found in Gauteng, South Africa.

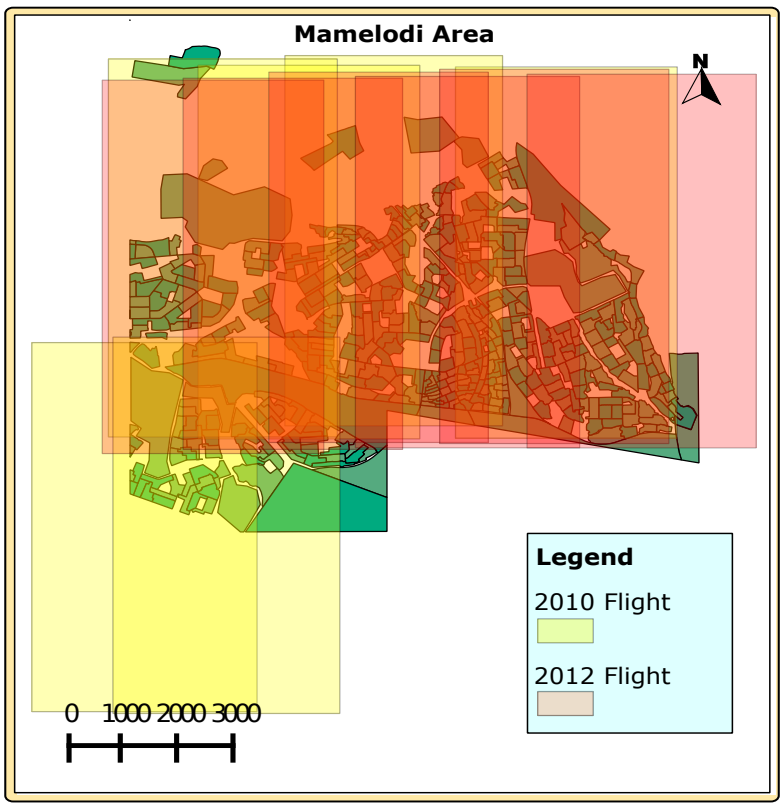


Figure 4.4: Data acquisition scenes for 2010 (date 1) and 2012 (date 2).

size, pattern, texture, site and context differences) were used to *visually* delineate and select settlement types (classes/samples) into polygons to generate subclasses, see . The latter was done manually, using visual interpretation to label the polygons into codes/categories, e.g. formal and informal settlements were denoted by codes $7.1xx$ and $7.2xx$ respectively (xx denotes the informal/formal respective subclasses/extensions) and non-built areas were denoted as code 20.100, see illustration in figure (4.5) and class descriptions in table (4.2 pg. 46).

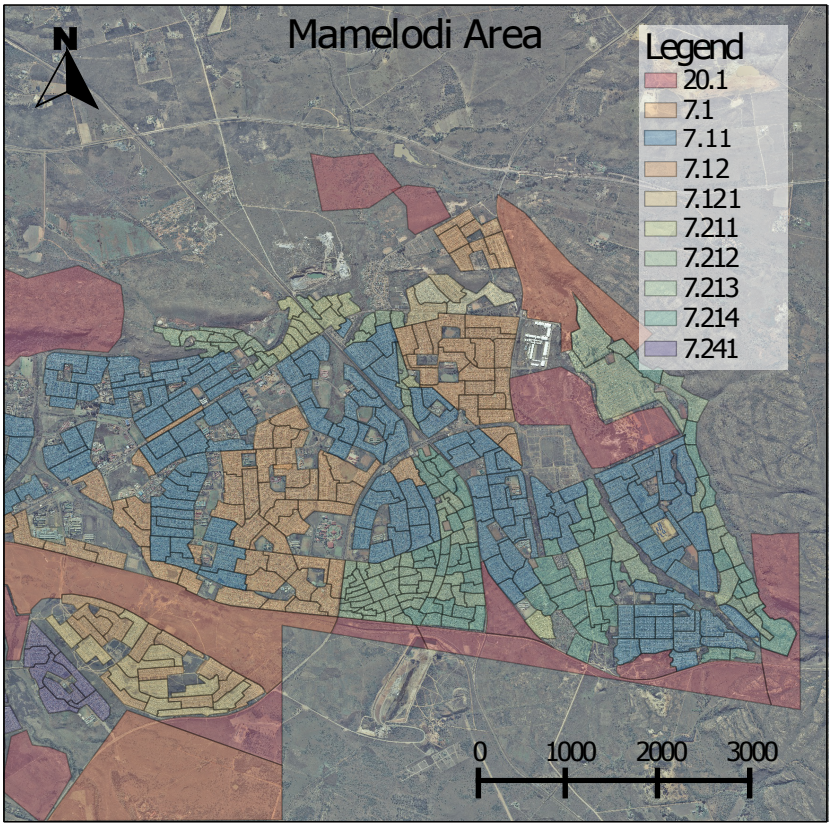
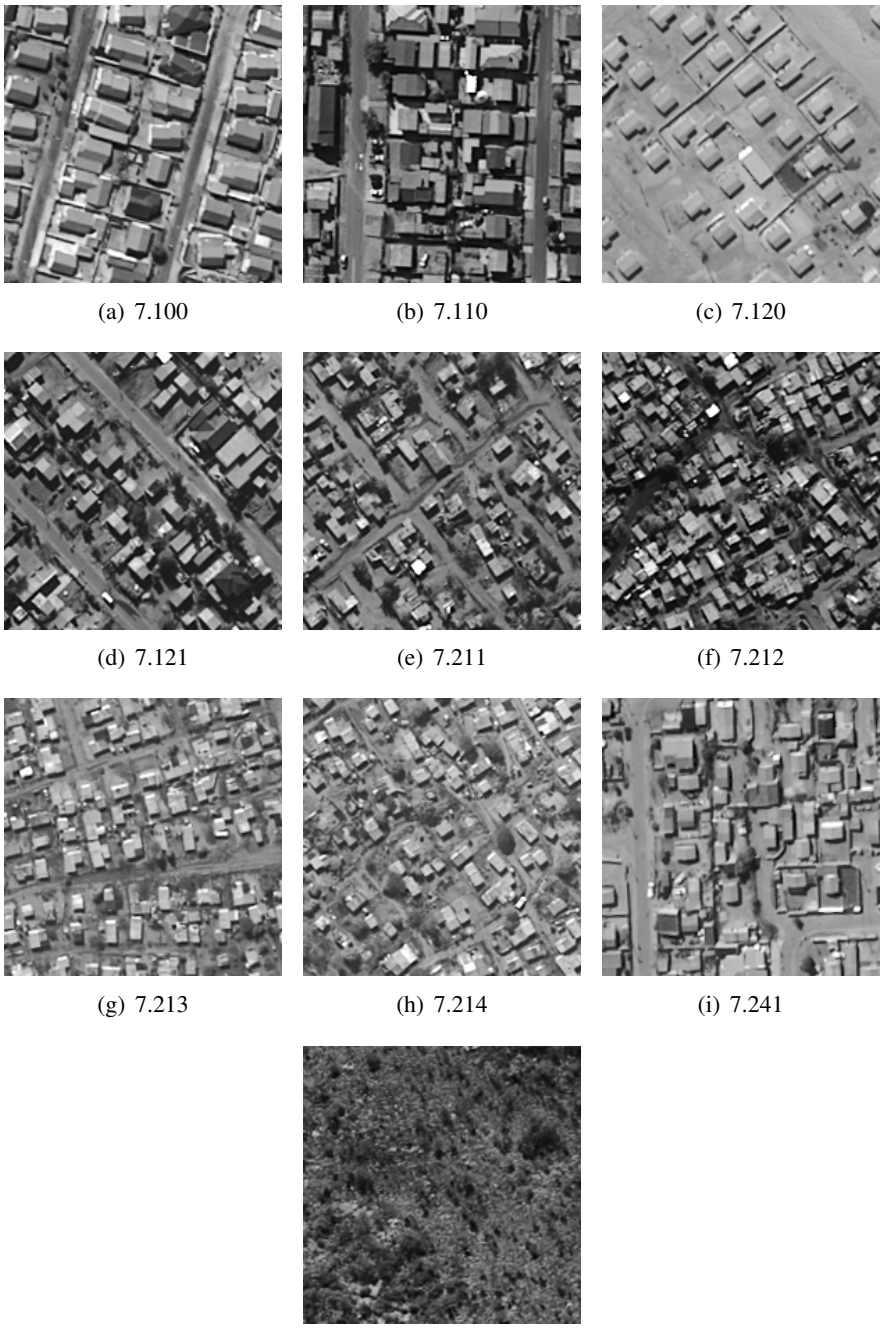


Figure 4.5: Settlement type categories over the Mamelodi area.

4.2.2.1 Training and Testing datasets

On both data sets (date 1 and 2) the polygons were divided into non-overlapping subsets A and B where these are used interchangeably for training and testing the classifier. This distinction is important, because the samples within the training set may overlap, but the training and testing sets do not overlap. From each polygon keeping the label (A or B) intact, square tiles (of size 120 m $\times$ 120 m) were extracted from random locations entirely within the demarcated polygons. Examples of the extracted square tiles for each settlement classes are presented in figure (4.6 pg. 45).



(a) 7.100 (b) 7.110 (c) 7.120

(d) 7.121 (e) 7.211 (f) 7.212

(g) 7.213 (h) 7.214 (i) 7.241

(j) 20.100

Figure 4.6: Examples of the settlements classes found in the Mamelodi area.

Table 4.2: Description of the extended land use codes to differentiate settlements found in Mamelodi Area.

Class	Description
7.100	Formal residential
7.110	Formal housing with backyard shacks
7.120	Low-income housing (RDP[a])
7.121	Low-income housing (RDP) with backyard structures
7.211	Informal housing (shacks with loose layout and dirt roads)
7.212	Unordered informal housing – typically dense
7.213	Ordered informal housing – typically sparse
7.214	Unordered informal housing – typically sparse with some ad hoc dirt roads
7.241	Dense low-income housing, tarred roads, gridded layout, typically no pitched roofs
20.100	Non-built-up area representing vegetation and bare areas.

[a]RDP housing is a South African low cost subsidy housing project

4.3 Extended LBP design

In order to successfully characterize the effect of contrast (variance measures) on imagery acquired at different dates from the chosen study area, a comparison was made between the normal LBP (without variance measures) and the LBP/VAR (with variance measures) algorithms, see illustration in figure (4.1 pg. 40). This section presents a detailed review of the image feature extraction process, from the input data, using the LBP and LBP/VAR algorithm.

4.3.1 Extracting LBP features

Extracting LBP features involves two main processes:

1. **Measuring the LBP code**:

   First construct regular circular neighbourhoods with P $(\mathrm{P} > 1)$ image pixels and radius $R(\mathrm{R} > 0)$, with the coordinates of the gray values g_P being $(-R\sin(2\pi p/P))$,$(R\cos(2\pi p/P))$ at $g_c(0,0)$ (gray values that do not fall exactly in the pixel centre are estimated by interpolation). The LBP code is measured from the pattern (circular neighbourhood) using equation (3.7 pg. 28). Three patterns (P,R) were considered: for samples P = 8, 16, 24 having radius R = 1,2,3 respectively, see illustration in figure (4.7).

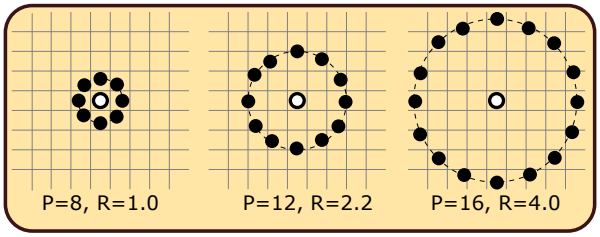


Figure 4.7: Local circularly symmetric neighbourhood sets, three cases.

2. **Labelling LBP code**:

   A look-up table containing all the uniformity measures (LBP code labels) was constructed for a given pattern (P,R), see illustration in figure (3.2 pg. 29). Using the look-up table, uniformity measures U with a value of at most 2 are stored as uniform patterns with bin labels $[0, P-1]$ while the non-uniform patterns are stored as bin label $(\mathrm{P}+1)$, where bins $[0, P-1]$ correspond to a texture feature from equation 3.10, see figure (3.3 pg. 30). Examples of texture primitives captured by the uniform patterns of LBP are shown in figure (4.8).

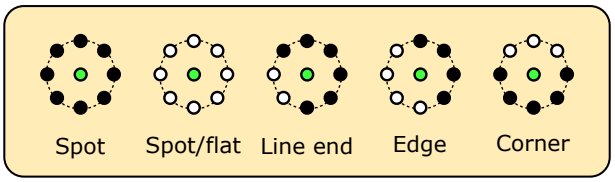


Figure 4.8: Different texture primitives detected by the uniform patterns of LBP.

An illustration showing the measurement of the LBP codes and labels (primitives detected) for a pattern size P = 8 is presented in figure (4.9).

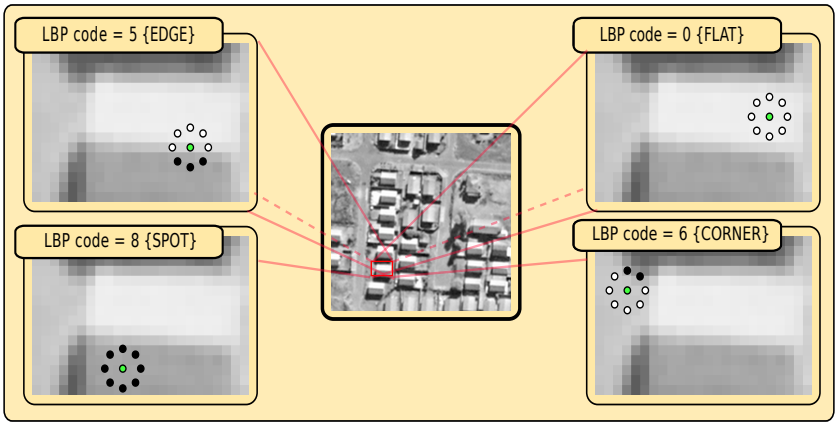


Figure 4.9: Extraction of LBP features for $\mathrm{P} = 8$.

4.3.2 Extracting LBP/VAR features

Circular neighbourhoods used in extracting LBP features, are used to calculate Variance measures as illustrated in equation (3.11 pg. 30). Variance measures yield continuous values, thus need to be quantized. To achieve this, the Variance measure distribution is calculated over the entire dataset. Using ***R*** *Project for Statistical Computing*[2], percentiles (bin breaks) are measured from the calculated Variance distribution, such that they are distributed according to the Variance measures for different number-of-bins choices (which are $3 \rightarrow 20$ bins in this case). A summary of the process is presented below:

1. For a given circular neighbourhood pattern size (P, R), we calculated LBP features and $(3 \rightarrow 20)$ bin breaks from the quantized Variance distribution.

2. A 2D joint distribution histogram for a given pattern size P, R is then constructed by incrementing the count in the bin corresponding to local Variance and the LBP code, as shown in figure (4.10). For an illustration showing a 2D joint histogram distribution using the LBP/VAR algorithm, please refer to figure (3.4 pg. 31).

[2]R is a free software environment for statistical computing and graphics (source: http://www.r-project.org).

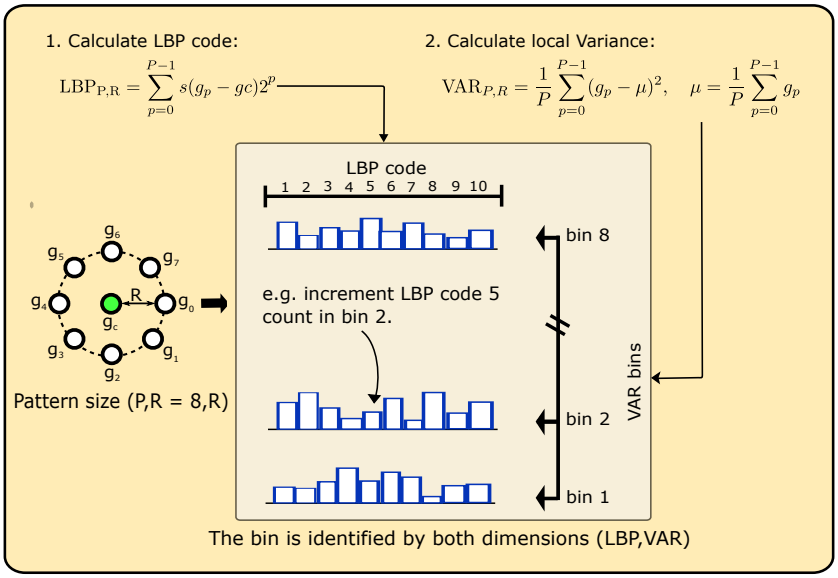


Figure 4.10: Calculation of the LBP/VAR features for a given pattern size P, R.

3. For a given pattern size (P, R), the process yields $(P + 2)$ LBP features and $(3 \rightarrow 20)$ local Variance features. For example, consider a 2D joint distribution histogram for pattern size (P,R = 8,1) and number-of-bins = 3. This yields a 2D histogram of (8+2 LBP uniform features $\times$ number-of-bins = 3) features/elements. . These elements are converted to a 1D list (where feature/element (f0) = LBP/VAR(0,0), f1 = LBP/VAR(0,1) to f29 = LBP/VAR(8,3) in the 2D feature space) which form input to the classification step, see illustration in table (4.3).

Table 4.3: An example of the feature sets for LBP/VAR ($P, R = 8, 1$ for 3 bins) which form input to the classifier using three sample images, where attributes $f0$ to $f29$ are the calculated features/elements.

f0	f1	f2	f3	f4	f5	$\cdots$	f29	f29	class
0.0349	0.0514	0.0353	0.0503	0.0645	0.0562	$\cdots$	0.0499	0.0487	20.100
0.0347	0.0538	0.0360	0.0533	0.0738	0.0600	$\cdots$	0.0520	0.0511	20.100
0.0329	0.0501	0.0355	0.0525	0.0696	0.0617	$\cdots$	0.0517	0.0541	20.100
0.0283	0.0365	0.0296	0.0459	0.0766	0.0490	$\cdots$	0.0367	0.0442	7.100
0.0275	0.0335	0.0258	0.0404	0.0574	0.0419	$\cdots$	0.0320	0.0365	7.100
0.0296	0.0357	0.0286	0.0467	0.0828	0.0508	$\cdots$	0.0358	0.0421	7.100
0.0328	0.0513	0.0512	0.0949	0.1146	0.0987	$\cdots$	0.0483	0.0469	7.110
0.0325	0.0531	0.0497	0.1039	0.1335	0.1058	$\cdots$	0.0521	0.0542	7.110
0.0326	0.0525	0.0502	0.1038	0.1316	0.1053	$\cdots$	0.0532	0.0547	7.110

4.4 Implementation

This section presents the evaluation process involving the training and testing of the SVM classifier, including statistical measures. The data of a specific date is divided into subsets A and B, to provide spatially separate training and testing data. Generalization classification accuracy for the different texture features algorithms was determined by evaluating the performance of the true positive values of the Support Vector Machine classifier.

4.4.1 Training and testing SVM

The SVM was trained and tested using Weka's[3] implementation of John Platt's [103] *Sequential Minimal Optimization* (SMO) algorithm with a polynomial kernel (i.e, $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^p$). Training (subset A) and testing (subset B) images acquired on the same and different dates (denote as same- and across-date) were used interchabley to evaluate the classifier over six combinations, see figure (4.11).

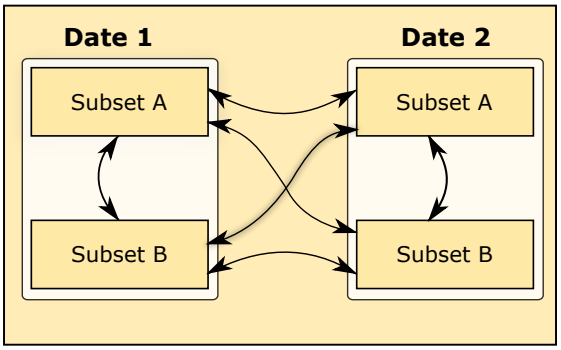


Figure 4.11: Training and test between same and cross date classes.

For simplicity, date-1 subset-A and date-2 subset-B are denoted as A_{d1} and B_{d2} respectively.

4.4.2 Statistics

To obtain the standard deviations on various classification results, the following procedure was used to evaluate a given configuration using data sets X and Y ($X = A_{d_1}$ and $Y = B_{d_1}$ respectively):

1. Train the SVM using the whole of set X.

2. Partition set Y in 10 folds using stratified sampling to preserve relative class frequency.

[3]Weka is a collection of machine learning algorithms for data mining tasks (source: http://www.cs.waikato.ac.nz/ml/weka).

3. Evaluate the SVM (trained on X) on each of the 10 folds of Y, obtaining one accuracy figure for each fold.

4. Exchange X and Y, and repeat 1–3.

The process described above, denoted by $X \rightleftharpoons Y$, produces 20 individual values for each accuracy metric, which are then used to calculate a mean and standard deviation for each metric. We distinguish between two classes of test, namely same-date (when both training and test sets are derived from the same-date aerial image) and cross-date (when two aerial images acquired at different times/dates are used). The difference in performance between these to classes highlights the degree to which a particular classifier is invariant to viewing- and illumination geometry and generalisation ability.

4.5 Summary

This chapter presented the design of the experiment, discussing the process of image feature extraction from aerial images using the extended LBP algorithm for the training of the SVM classifier for evaluation. The aerial images were acquired in the same area under different conditions (date and season). The extraction of features from the images was achieved using LBP and extended LBP methods. These features were used to train and evaluate the SVM classifier using John Platt's [103] SMO algorithm (using a polynomial kernel) over the six possible date combinations. Results obtained from the experiment are presented in chapter 5.

Chapter 5

Results and Discussion

5.1 Introduction

This chapter presents the results obtained with the system described in Chapter 4. To recap, the chapter starts with an overview of the experiment, including the parameters used. Results are then presented in three parts; the first presents results of the SVM classifier overall accuracies including same- and cross-dates evaluations and the effects of the number-of-bins parameter. The second part presents the results of the SVM classifier per-class (i.e., per settlement type) True Positive (TP) evaluations. Lastly, to investigate parameters that yield maximum performance, the evaluation of the SVMs SMO algorithm parameters are presented. A discussion of the results is presented in three parts according to the results sections mentioned above.

5.2 Experiment overview

In remote sensing, image textures can occur at arbitrary spatial resolutions and rotations, and may be subjected to varying illumination conditions. Due to the variety of texture, an ideal texture description is the one that can efficiently discriminate different types of textures (such as distinguishing human settlements) and is robust to viewing- and illumination-geometry effects. The LBP algorithm with Variance measures (LBP/VAR) has the potential of being an ideal texture description as it incorporates invariance with respect to spatial scale, orientation, while considering gray scale variations. To investigate the LBP/VAR robustness to viewing- and illumination-geometry effects and ability to efficiently discriminate different types of textures the following experiment was conducted:

1. High resolution data was acquired at different date from an area containing a variety of settlements.

2. The data set was prepared (i.e., it was labelled and class samples were collected).

3. The extended LBP (LBP/VAR) features were extracted over spatially non-overlapping areas.

4. The SVM classifier was trained and evaluated for generalization performance of LBP and LBP/VAR over both same-date and cross-date subsets.

5. To enhance classifier performance, the classifier was evaluated at various parameter settings.

5.2.1 Experiment parameters

5.2.1.1 Input data

For detailed acquisition, processing and extraction of sample classes, please refer to section (4.2). This section serves as a summary of the data set used for the experiment. Aerial imagery was used for the experiment with imagery captured at different dates (2010 and 2012) in Mamelodi area (South Africa). From a total of ten labelled polygon classes, square tiles (input data) of size (120 m $\times$ 120 m) were extracted from random locations within the demarcated polygons. Each class was divided into two subsets (subset A and B) for each date; for the number of extracted square tile for each class and subset see table (5.1). The number of input data (square tiles) are not equal for both dates (see table (5.1)) since more scenes were available from the 2012 acquisition campaign.

Table 5.1: The number of patterns in each class, for each subset.

	2010 Data		2012 Data	
Class	Subset A	Subset B	Subset A	Subset B
7.100	1926	1775	2158	2192
7.110	3183	3644	2834	3275
7.120	410	312	452	374
7.121	148	142	365	303
7.211	142	187	155	190
7.212	573	500	518	516
7.213	884	236	593	135
7.214	750	682	670	610
7.241	180	311	545	619
20.100	4241	2518	4607	3193
Total	12437	10307	12897	11407

5.2.1.2 LBP, VAR and LBP/VAR parameters

A detailed LBP, VAR and LBP/VAR feature extraction process is presented in section (4.3 pg. 47). To recap, the LBP, VAR and LBP/VAR requires two input variables that define the circular neighbourhood: pattern size P, and radius R. For the task of classifying human settlements, the LBP and extended LBP (local contrast extension) features were measured for a range of pattern sizes (P,R): 8,1; 16,2; 24,3. The range of pattern sizes (P,R) parameter is important for the extraction of both small and large settlement features. Each pattern generates P+2 features ($[0, P]$ uniform patterns + 1 miscellaneous bin for non-uniform patterns), see illustration for P=8, R=1 example in figure (3.2 pg. 29).

Local contrast measures (VAR) have continuous outputs and need to be quantised first. This is done by calculating the global contrast distribution and constructing bin breaks for a set number-of-bins, see section (4.3.2 pg. 48). To determine the least-complex set of features that yields optimal performance, the number of bins parameter is investigated from 3–20 bins. These experiment parameters (P,R), including the number of bins for contrast measures (VAR) and the total number of generated features for LBP, VAR and LBP/VAR are reported in table 5.2.

Table 5.2: Number of extracted features from each image.

Sample (P,R)	LBP[a]	LBP/VAR[b]
8,1	10	10 × (3,...,20)
16,2	18	18 × (3,...,20)
24,3	26	26 × (3,...,20)
8,1+16,2	10 + 18	28 × (3,...,20)
8,1+16,2+24,3	10 + 18 + 26	54 × (3,...,20)

[a] LBP features = [0,P] (uniform patterns) + 1 miscellaneous bin.
[b] Contrast bins (VAR) = 3,...,20 bins

5.2.1.3 SVMs SMO parameters

The effectiveness of SVM depends on the selection of kernel and the kernel's parameters. Typically, each combination of parameter choices is checked using cross validation, and the parameters with best cross-validation accuracy are picked [49]. These are then trained on the whole training set using the selected parameters to be tested in classifying new data. For a detailed implementation of training and testing the SVM please refer to section (4.4.1 pg. 50). The chosen SVM SMO algorithm used a polynomial kernel $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^p$, which requires two

parameters, denoted as SMO(C, p), where C is the complexity value and p is the polynomial kernel exponent value. The best combination of C and p was investigated by a grid search with exponentially growing sequences of $C \in \{2^0, 2^1, \ldots, 2^7\}; p \in \{1, 2, 3\}$ for each C value.

5.3 Experiment Results

This section presents the LBP and LBP/VAR results obtained from the experiment outlined in chapter 4. The results are split in three subsections where the first section presents the classification accuracies. Secondly, results investigating classification accuracies per class are presented and lastly, results for various SMO parameters are presented.

5.3.1 SVM classifier overall accuracies

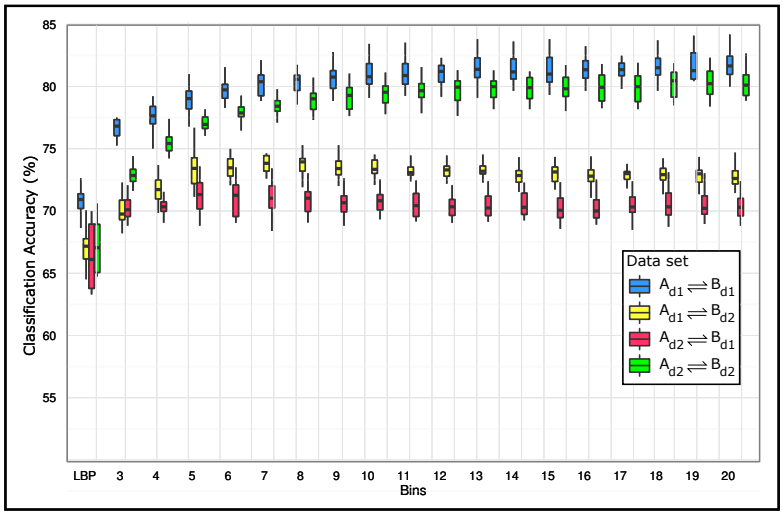
The SVM classifier was trained and evaluated for generalization performance of LBP and LBP/-VAR for settlement type classification over images acquired on the same-date and at different dates (cross-date). The classifier's overall accuracy is defined as the number of correctly classified samples divided by the total number of samples multiplied by 100. The overall classification accuracy values for both LBP and LBP/VAR features are presented on figure (5.1 pg. 56) and table (5.3 pg.57). The highest overall classification accuracy value for each date configuration is shown bold.

5.3.2 Per-Class evaluation

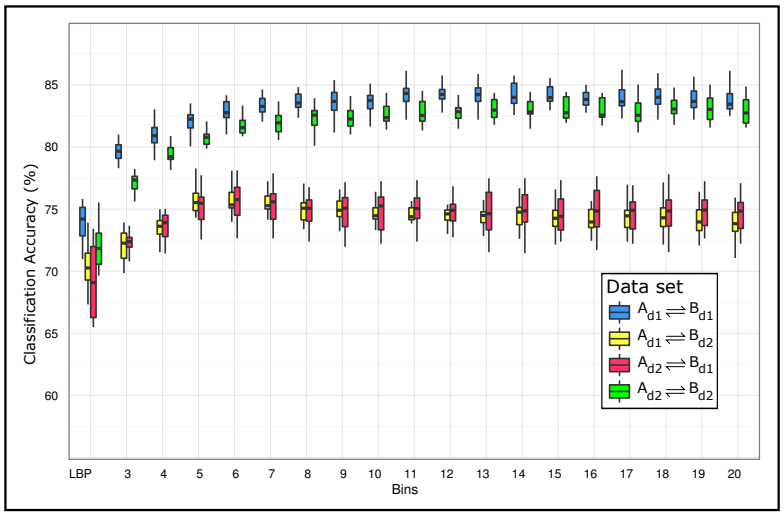
True positive rates for each class (settlement type) for a given numbers of bins (the lowest number-of-bins that yields optimal performance: $\approx$ 6 bins) were evaluated and are presented in table (5.4). Settlement class true positive rates of 80% or greater are shown in bold.

5.3.3 SVM's SMO parameter test

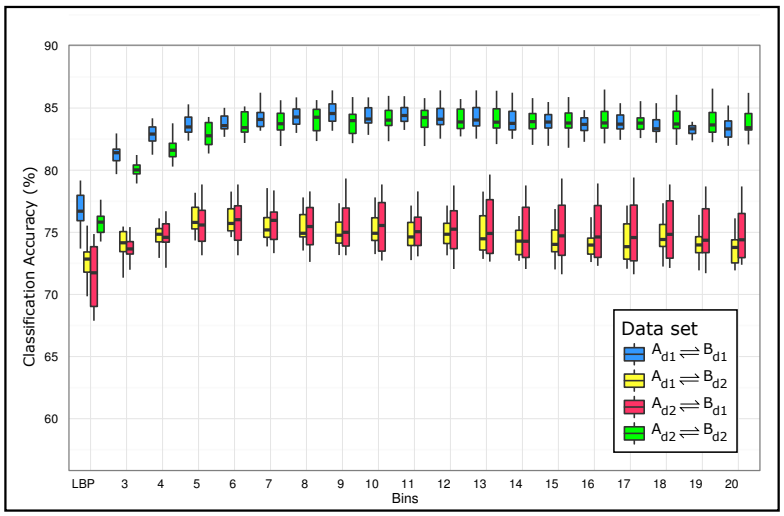
Lastly, a test on the SMO algorithm for the optimal parameters for training the SVM classifier over the six possible combinations was conducted and presented in table (5.5 pg. 59). The test was done using the a pattern size $(P, R = 8, 1 + 16, 2 + 24, 3)$ for number-of-bins = 6. Two parameters were considered, SMO(C, p), where C is the complexity value and p is the polynomial kernel exponent value (i.e, $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^p$). The experiment tested the SMO algorithm for C values $[C = 2^i, i = 0, 1, \ldots, 7]$ with exponents $[p = 1, 2, 3]$ for each C value.



(a) **P,R = 8,1**



(b) **P,R = 8,1+16,2**



(c) **P,R = 8,1+16,2+24,3**

Figure 5.1: Multiresolution boxplots showing the overall classification accuracies (%) for $\mathrm{LBP}^{\mathrm{riu2}}_{\mathrm{P,R}}$ (LBP) and $\mathrm{LBP}^{\mathrm{riu2}}_{\mathrm{P,R}}/\mathrm{VAR}_{\mathrm{P,R}}$ (3-20 Bins) of same- and cross-date imagery.

Table 5.3: LBP and LBP/VAR overall classification accuracies (%) with standard deviation.

Data set	$\mathbf{LBP}^{\mathbf{riu2}}_{\mathbf{P,R}}$	$\mathbf{LBP}^{\mathbf{riu2}}_{\mathbf{P,R}}/\mathbf{VAR}_{\mathbf{P,R}}$							
		3	6	8	10	12	14	16	19 (bins)
P,R=8,1									
$A_{d1} \rightleftharpoons B_{d1}$	70.82 (1.08)	76.54 (0.99)	79.65 (0.85)	80.44 (1.08)	80.56 (0.98)	81.09 (1.11)	81.37 (1.11)	81.41 (1.00)	**81.65 (1.12)**
$A_{d2} \rightleftharpoons B_{d2}$	67.10 (2.07)	72.92 (0.74)	77.931 (0.69)	78.82 (0.8)	79.22 (0.99)	79.75 (1.05)	79.79 (0.98)	79.97 (1.11)	**80.30 (1.16)**
$A_{d1} \rightleftharpoons A_{d2}$	66.59 (2.10)	70.72 (0.98)	74.241 (1.44)	74.43 (1.04)	74.49 (1.08)	74.55 (0.89)	74.44 (0.96)	**74.84 (1.10)**	74.40 (0.99)
$A_{d1} \rightleftharpoons B_{d2}$	67.10 (1.43)	70.01 (1.14)	73.565 (1.10)	**73.79 (0.86)**	73.45 (0.83)	73.19 (0.70)	72.76 (0.89)	72.77 (0.87)	72.82 (0.92)
$B_{d1} \rightleftharpoons B_{d2}$	69.26 (1.02)	73.12 (1.03)	75.868 (1.88)	76.02 (2.09)	**76.09 (2.01)**	76.06 (1.80)	75.60 (1.71)	75.29 (1.65)	75.62 (1.64)
$A_{d2} \rightleftharpoons B_{d1}$	66.43 (2.70)	70.19 (0.86)	**71.06 (1.45)**	70.93 (1.18)	70.70 (1.06)	70.415 (0.93)	70.53 (0.93)	70.28 (1.03)	70.56 (1.05)
Average	67.88 (1.73)	72.25 (0.96)	75.39 (1.23)	75.74 (1.18)	75.75 (1.16)	**75.84 (1.08)**	75.75 (1.10)	75.76 (1.13)	75.89 (1.15)
P,R=8,1+16,2									
$A_{d1} \rightleftharpoons B_{d1}$	73.95 (1.51)	79.57 (0.93)	82.81 (0.89)	83.54 (0.88)	83.77 (0.97)	84.19 (0.88)	**84.22 (0.97)**	83.98 (0.86)	83.81 (1.06)
$A_{d2} \rightleftharpoons B_{d2}$	71.85 (1.56)	77.20 (0.69)	81.72 (0.711)	82.37 (0.96)	82.65 (0.83)	82.83 (0.78)	83.01 (0.90)	83.00 (0.92)	**83.11 (0.99)**
$A_{d1} \rightleftharpoons A_{d2}$	70.45 (2.44)	74.04 (1.30)	**79.65 (2.21)**	79.20 (2.18)	78.95 (2.33)	78.94 (2.39)	79.27 (2.64)	79.46 (2.81)	79.06 (2.68)
$A_{d1} \rightleftharpoons B_{d2}$	70.30 (1.63)	72.04 (1.27)	**75.76 (1.18)**	74.92 (1.02)	74.77 (0.99)	74.66 (1.03)	74.62 (1.31)	74.27 (1.14)	74.10 (1.21)
$B_{d1} \rightleftharpoons B_{d2}$	73.91 (1.96)	76.85 (1.25)	**79.55 (2.35)**	79.18 (2.33)	78.96 (2.57)	79.08 (2.40)	79.16 (2.47)	78.96 (2.17)	79.04 (2.02)
$A_{d2} \rightleftharpoons B_{d1}$	69.31 (3.07)	72.30 (0.78)	**75.51 (1.67)**	74.78 (1.30)	74.70 (1.59)	74.65 (1.47)	74.83 (1.71)	74.83 (1.97)	74.75 (1.45)
Average	71.63 (2.03)	75.33 (1.03)	79.17 (1.50)	79.00 (1.44)	78.97 (1.55)	79.06 (1.49)	**79.18 (1.67)**	79.08 (1.65)	78.98 (1.57)
P,R=8,1+16,2+24,3									
$A_{d1} \rightleftharpoons B_{d1}$	76.56 (1.57)	81.402 (0.86)	83.79 (0.63)	**84.32 (0.88)**	84.25 (0.80)	84.32 (0.99)	83.95 (0.98)	83.82 (0.91)	83.37 (0.65)
$A_{d2} \rightleftharpoons B_{d2}$	75.87 (1.20)	80.057 (0.56)	83.67 (0.95)	84.05 (1.03)	**84.16 (0.98)**	84.07 (0.95)	83.88 (1.00)	84.00 (1.18)	83.89 (1.18)
$A_{d1} \rightleftharpoons A_{d2}$	73.84 (1.84)	76.207 (1.16)	79.97 (2.51)	**80.08 (2.23)**	79.73 (2.91)	80.01 (2.94)	79.54 (2.97)	79.61 (3.10)	79.52 (3.05)
$A_{d1} \rightleftharpoons B_{d2}$	72.65 (1.35)	74.044 (1.19)	**76.01 (1.17)**	75.45 (1.17)	75.14 (1.28)	74.88 (1.10)	74.26 (1.19)	74.09 (1.12)	74.12 (1.37)
$B_{d1} \rightleftharpoons B_{d2}$	75.99 (1.59)	78.212 (1.09)	**80.46 (2.26)**	80.266 (2.40)	79.98 (1.89)	80.03 (2.12)	79.55 (2.41)	79.58 (2.22)	79.54 (2.17)
$A_{d2} \rightleftharpoons B_{d1}$	71.66 (2.50)	73.719 (0.82)	**75.88 (1.70)**	75.50 (1.88)	75.47 (2.06)	75.30 (2.06)	75.04 (2.34)	75.09 (2.39)	75.00 (2.31)
Average	74.43 (1.68)	77.27 (0.95)	**79.96 (1.54)**	79.94 (1.60)	79.79 (1.65)	79.77 (1.69)	79.37 (1.82)	79.37 (1.82)	79.24 (1.79)

Table 5.4: LBP and LBP/VAR per-class True Positive (TP) rate values (%) with standard deviation for number-of-bins = 6.

Data set	**TP values** (%)										
	Average	7.100	7.120	7.121	7.211	7.212	7.213	7.214	7.110	7.241	20.100
P,R = 8,1											
$A_{d1}B_{d1}$	79.65 (0.85)	**87.89 (3.44)**	**82.05 (5.00)**	**88.54 (6.26)**	9.76 (12.00)	36.53 (8.73)	61.99 (6.04)	35.08 (19.76)	60.74 (7.20)	71.24 (7.41)	**96.63 (1.25)**
$A_{d2}B_{d2}$	77.93 (0.69)	**83.15 (1.950)**	**80.55 (4.13)**	**80.92 (7.38)**	7.52 (3.66)	48.49 (11.36)	58.99 (8.20)	15.62 (17.05)	65.23 (8.31)	61.99 (5.92)	**97.44 (0.83)**
$A_{d1}A_{d2}$	74.24 (1.44)	**80.13 (4.88)**	78.71 (11.48)	58.94 (28.98)	2.48 (4.77)	32.30 (11.57)	54.00 (6.65)	13.95 (13.24)	61.23 (16.95)	42.24 (7.98)	**97.96 (0.73)**
$A_{d1}B_{d2}$	73.57 (1.10)	**80.31 (6.46)**	75.39 (15.04)	58.95 (29.09)	0.69 (2.13)	22.88 (15.18)	49.60 (5.43)	40.73 (17.28)	55.51 (12.00)	44.02 (9.45)	**97.50 (1.22)**
$B_{d1}B_{d2}$	75.87 (1.88)	78.85 (4.07)	76.53 (10.82)	70.78 (15.18)	0.36 (1.59)	21.28 (16.82)	49.00 (8.28)	48.32 (7.09)	53.93 (20.60)	40.18 (20.85)	**97.20 (1.03)**
$A_{d2}B_{d1}$	71.06 (1.45)	76.24 (2.84)	74.20 (11.61)	66.51 (21.63)	0.67 (2.06)	31.17 (9.65)	51.18 (7.00)	11.27 (12.15)	57.47 (24.94)	32.66 (6.63)	**95.61 (0.85)**
P,R = 8,1 16,2											
$A_{d1}B_{d1}$	**82.81 (0.89)**	**90.11 (2.94)**	**86.39 (4.04)**	**91.08 (6.96)**	44.08 (10.21)	37.48 (10.24)	64.55 (6.48)	48.03 (19.68)	68.04 (6.11)	**81.61 (6.09)**	**96.04 (0.76)**
$A_{d2}B_{d2}$	**81.72 (0.71)**	**85.78 (2.40)**	**84.38 (4.50)**	**89.33 (5.38)**	47.49 (8.87)	51.33 (13.96)	59.66 (7.02)	38.52 (25.95)	69.08 (7.00)	72.34 (4.38)	**97.01 (0.69)**
$A_{d1}A_{d2}$	79.65 (2.21)	**84.91 (3.76)**	**85.50 (5.29)**	63.77 (25.98)	25.60 (23.52)	39.51 (11.81)	59.41 (11.27)	34.25 (11.91)	72.36 (13.79)	48.41 (16.30)	**97.96 (0.80)**
$A_{d1}B_{d2}$	75.76 (1.18)	**82.46 (7.04)**	**80.53 (10.60)**	69.61 (26.42)	21.16 (18.84)	22.49 (13.85)	53.18 (8.39)	47.34 (23.56)	57.98 (17.71)	42.23 (20.99)	**96.89 (1.19)**
$B_{d1}B_{d2}$	79.55 (2.35)	**83.82 (4.03)**	**82.99 (5.73)**	75.57 (16.02)	28.44 (23.20)	32.10 (25.75)	53.58 (12.47)	58.10 (20.46)	51.71 (23.91)	40.63 (18.60)	**96.90 (1.53)**
$A_{d2}B_{d1}$	75.51 (1.67)	**82.46 (2.40)**	**81.35 (4.81)**	68.07 (21.96)	22.94 (23.10)	38.81 (10.69)	53.61 (8.53)	22.44 (13.74)	65.18 (19.16)	38.98 (8.92)	**94.90 (0.85)**
P,R = 8,1 16,2 24,3											
$A_{d1}B_{d1}$	**83.79 (0.63)**	**91.04 (2.46)**	**86.54 (3.07)**	**92.12 (7.38)**	50.36 (10.06)	39.76 (11.76)	67.35 (7.51)	52.66 (21.23)	70.03 (6.12)	**87.89 (5.66)**	**95.56 (1.12)**
$A_{d2}B_{d2}$	**83.67 (0.95)**	**87.94 (1.68)**	**86.35 (3.32)**	**92.89 (2.55)**	56.75 (11.22)	56.40 (15.87)	63.91 (5.83)	45.09 (27.97)	71.86 (5.33)	75.82 (4.96)	**96.78 (0.70)**
$A_{d1}A_{d2}$	79.97 (2.51)	**84.11 (6.02)**	**85.50 (7.63)**	60.39 (28.92)	37.46 (30.71)	35.27 (10.77)	63.63 (11.18)	38.49 (10.78)	76.96 (10.56)	48.15 (25.93)	**97.81 (0.78)**
$A_{d1}B_{d2}$	76.01 (1.17)	**82.88 (8.21)**	**82.19 (9.40)**	67.36 (28.41)	37.21 (27.19)	24.65 (15.07)	49.86 (5.69)	45.33 (31.81)	61.35 (15.32)	35.49 (23.31)	**96.94 (1.45)**
$B_{d1}B_{d2}$	**80.46 (2.26)**	**85.07 (7.36)**	**83.15 (6.76)**	79.43 (14.35)	44.78 (33.45)	43.69 (27.74)	56.08 (6.89)	57.71 (21.88)	55.32 (21.72)	37.40 (18.88)	**96.55 (2.44)**
$A_{d2}B_{d1}$	75.88 (1.70)	**82.58 (6.73)**	**80.02 (8.04)**	71.98 (19.89)	30.73 (29.06)	39.35 (17.59)	58.07 (8.52)	29.30 (9.08)	65.90 (14.20)	38.98 (15.44)	**94.19 (1.09)**

The default parameters, SMO$(1, 1)$, which are parameters used on the experiments, are compared with the classification accuracy averages obtained by the SMO with adapted parameters and are presented in table (5.5 pg. 59).

Table 5.5: John Platt's sequential minimal optimization algorithm optimal complexity parameter (C) and exponent value (p) test for LBP/VAR pattern size ($P, R = 8, 1 + 16, 2 + 24, 3$) with number-of-bins = 6. Evaluating overall classification accuracies (%) with standard deviation.

Data set	$C = 1$			$C = 2$		
	$p = 1$	2	3	1	2	3
$A_{d1} \rightleftharpoons B_{d1}$	83.785 (0.628)	90.950 (9.117)	90.451 (9.613)	82.159 (2.759)	77.198 (4.490)	74.953 (4.525)
$A_{d2} \rightleftharpoons B_{d2}$	80.903 (5.265)	86.871 (10.338)	79.560 (6.547)	82.696 (3.044)	82.171 (11.398)	91.732 (8.509)
$A_{d1} \rightleftharpoons A_{d2}$	79.973 (2.511)	77.331 (1.350)	80.504 (7.001)	80.014 (2.656)	79.864 (11.151)	90.268 (11.646)
$A_{d1} \rightleftharpoons B_{d2}$	74.869 (3.460)	77.061 (4.729)	73.287 (5.102)	73.728 (1.709)	76.169 (5.196)	82.867 (12.027)
$B_{d1} \rightleftharpoons B_{d2}$	80.428 (2.289)	79.847 (1.999)	79.548 (1.720)	84.619 (9.179)	79.767 (1.931)	87.733 (12.616)
$A_{d2} \rightleftharpoons B_{d1}$	87.172 (10.171)	76.514 (4.157)	82.628 (5.475)	76.720 (5.750)	88.018 (11.766)	78.518 (8.622)
Average	**81.188 (4.054)**	**81.429 (5.282)**	**80.996 (5.910)**	**79.989 (4.183)**	**80.531 (7.655)**	**84.345 (9.658)**
	$C = 4$			$C = 8$		
$A_{d1} \rightleftharpoons B_{d1}$	79.367 (3.303)	79.159 (8.663)	78.626 (3.895)	83.150 (0.952)	83.161 (4.180)	93.041 (10.605)
$A_{d2} \rightleftharpoons B_{d2}$	88.030 (5.257)	80.093 (8.614)	79.970 (3.342)	79.353 (5.917)	78.253 (4.290)	74.808 (8.946)
$A_{d1} \rightleftharpoons A_{d2}$	82.571 (6.660)	78.567 (3.029)	79.202 (6.700)	87.426 (10.112)	75.519 (2.121)	77.869 (1.703)
$A_{d1} \rightleftharpoons B_{d2}$	74.406 (1.277)	72.272 (1.276)	78.476 (7.797)	73.744 (1.272)	76.423 (8.398)	76.709 (3.829)
$B_{d1} \rightleftharpoons B_{d2}$	80.056 (2.383)	75.191 (4.180)	86.208 (8.064)	79.497 (2.182)	84.641 (7.091)	74.638 (4.915)
$A_{d2} \rightleftharpoons B_{d1}$	74.362 (1.605)	78.595 (12.718)	88.057 (11.983)	73.312 (1.463)	88.419 (14.522)	75.760 (4.869)
Average	**79.799 (3.414)**	**77.313 (6.413)**	**81.757 (6.964)**	**79.414 (3.650)**	**81.069 (6.767)**	**78.804 (5.811)**
	$C = 16$			$C = 32$		
$A_{d1} \rightleftharpoons B_{d1}$	80.109 (3.851)	81.497 (0.836)	77.566 (4.860)	81.893 (0.898)	80.301 (2.963)	85.890 (14.332)
$A_{d2} \rightleftharpoons B_{d2}$	82.847 (5.148)	79.532 (7.074)	82.959 (13.165)	79.425 (5.453)	95.509 (10.042)	89.138 (10.257)
$A_{d1} \rightleftharpoons A_{d2}$	74.162 (2.178)	76.946 (1.141)	84.504 (12.917)	77.184 (1.987)	86.505 (14.506)	77.861 (1.713)
$A_{d1} \rightleftharpoons B_{d2}$	84.446 (11.835)	78.738 (12.370)	72.157 (1.078)	82.551 (13.224)	77.378 (10.131)	76.904 (6.778)
$B_{d1} \rightleftharpoons B_{d2}$	77.857 (3.755)	87.541 (13.947)	76.657 (4.463)	88.466 (11.300)	79.590 (1.738)	81.585 (6.415)
$A_{d2} \rightleftharpoons B_{d1}$	76.624 (4.904)	84.845 (13.630)	79.421 (4.688)	78.954 (8.530)	72.469 (1.503)	90.119 (10.476)
Average	**79.341 (5.279)**	**81.517 (8.166)**	**78.877 (6.862)**	**81.412 (6.899)**	**81.959 (6.814)**	**83.583 (8.329)**
	$C = 64$			$C = 128$		
$A_{d1} \rightleftharpoons B_{d1}$	81.556 (1.050)	76.216 (5.090)	74.392 (4.077)	79.598 (3.419)	89.803 (13.991)	85.982 (8.247)
$A_{d2} \rightleftharpoons B_{d2}$	82.317 (0.887)	82.981 (1.249)	86.855 (14.257)	82.047 (0.986)	86.184 (14.200)	82.972 (1.086)
$A_{d1} \rightleftharpoons A_{d2}$	76.433 (1.719)	80.160 (12.261)	87.896 (8.903)	75.813 (1.679)	79.095 (3.140)	77.869 (1.706)
$A_{d1} \rightleftharpoons B_{d2}$	74.452 (5.331)	86.184 (10.546)	85.704 (11.590)	74.240 (3.925)	96.011 (8.758)	72.154 (1.076)
$B_{d1} \rightleftharpoons B_{d2}$	78.771 (1.866)	75.170 (4.130)	79.657 (1.413)	83.643 (9.726)	78.919 (1.906)	79.673 (1.514)
$A_{d2} \rightleftharpoons B_{d1}$	79.766 (10.601)	77.241 (5.073)	87.790 (10.275)	72.259 (0.937)	71.172 (1.093)	87.066 (13.496)
Average	**78.883 (3.576)**	**79.659 (6.392)**	**83.716 (8.419)**	**77.933 (3.445)**	**83.531 (7.181)**	**80.953 (4.521)**

5.4 Discussion

This section presents the discussion of the results presented on the previous section and is split into three subsections. Discussions of the LBP and LBP/VAR classification accuracies, the per-class true positive values and lastly the SVM's SMO parameter test analysis are presented.

5.4.1 SVM classifier overall accuracies

To recap, table (5.6) shows the LBP and LBP/VAR overall classification average values over all date combinations ($A_{d1/d2} \rightleftharpoons B_{d1/d2}$) calculated for different number-of-bins. For detailed results refer to table (5.3 pg. 57).

Table 5.6: Overall classification accuracy averages (%) with standard deviations.

No. of bins	P,R = 8,1	8,1+16,2	8,1+16,2+24,3
LBP[1]	67.88 (1.73)	71.63 (2.03)	74.43 (1.68)
3	72.25 (0.96)	75.33 (1.03)	77.27 (0.95)
6	75.39 (1.23)	79.17 (1.50)	79.96 (1.54)
8	75.74 (1.18)	79.00 (1.44)	79.94 (1.60)
10	75.75 (1.16)	78.97 (1.55)	79.79 (1.65)
12	75.84 (1.08)	79.06 (1.49)	79.77 (1.69)
14	75.75 (1.10)	79.18 (1.67)	79.37 (1.82)
16	75.76 (1.13)	79.08 (1.65)	79.37 (1.82)
19	75.89 (1.15)	78.98 (1.57)	79.24 (1.79)

[1]LBP features without Contrast measurements.

The overall classification accuracy averages showed that the LBP/VAR outperforms the normal LBP for any given pattern size (P, R), see table (5.6). The LBP/VAR that showed the lowest overall classification accuracy averages (i.e., number-of-bins = 3) outperformed the normal LBP by approximately 4%. The classification performance was dependent on the building sizes detected on the images, that is, the overall classification increased when increasing pattern size (P, R). When the LBP pattern size (P, R) was increased with a larger pattern size and a larger radius (from $P, R = 8, 1$ to $P, R = 8, 1 + 16, 2$), see figure (5.1(b) pg. 56), both the performance of the LBP and LBP/VAR increased, where the normal LBP ranged from 68% to 72% and the LBP/VAR for number-of-bins = 3, was improved from 72% to 75%. Lastly, the settlement classification accuracy for the largest LBP pattern size investigated (from $P, R = 8, 1$ to $P, R = 8, 1 + 16, 2 + 24, 3$) with 54 LBP features and $54 \times (3, \ldots, 20)$ LBP/VAR features, as

shown in table (5.2 pg. 54), were measured and presented in figure (5.1(c) pg 56). The performance of the LBP ranged from 68% to 74% while the LBP/VAR for number-of-bins = 3, ranged from 72% to 77%. The LBP/VAR showed better performance over the LBP in all the pattern sizes considered, however the LBP/VAR performance was also dependent the number-of-bins parameter and was not consistent across all date combinations.

5.4.1.1 LBP/VAR number-of-bins parameter

The LBP/VAR classification accuracy showed a strong dependence on the number-of-bins parameter (i.e., contrast). Notable classification accuracy improvements were observed for number-of-bins = (3 to 6 bins) throughout the considered pattern sizes (P,R). Subsequent number-of-bins (6 to 20 bins) showed minor improvements in classification accuracies, see illustration in figure (5.1 pg. 56). For simplicity, the optimal number-of-bins parameter was the lowest number-of-bins that yield acceptable classification accuracies. Therefore, the highest classification accuracies (the turning points) were not observed for the number-of-bins (3 to 20 bins) investigated in the experiment. The optimal number-of-bins parameter for the experiment was observed (≈ 6 bins), where this was approximated visually from the boxplots and tables, see figure (5.1 pg. 56) and overall classification accuracy averages in table (5.6). The results showed that the optimal number-of-bins to be approximately the same for all pattern sizes considered. These results showed that the LBP/VAR (for number-of-bins = 6) outperformed the LBP by $\approx 7\%$ on average, see table (5.6).

5.4.1.2 Date effect

An experiment comparing classification accuracies of data acquired over the same area at different dates was done and reported in table (5.3 pg. 57). Averages of the same- and across-date (images from different dates) were calculated and summarised in table (5.7).

Features that are calculated from the same-date scenery are expected to have greater similarity than across-date features, and thus yield better classification accuracy than that of the cross-date scenery. This difference in classification accuracy between the same-date images and the cross-date images (denoted as *cross-date effect*) is mainly due to differences in viewing- and illumination-geometry, see illustration in figure (5.2) where the brightness differences in the roof sections (large highlighted building), shadow length and orientation are apparent.

Minor cross-date effects were observed for normal LBP with same-date classification accuracies having better results. The same-date averages ranged from 69% to 76% while for

Table 5.7: The same/cross -date overall classification accuracy averages (%) with standard deviations for various LBP/VAR configurations.

No. of		Same-date			Cross-date	
Bins	P,R = 8,1	8,1+16,2	8,1+16,2+24,3	8,1	8,1+16,2	8,1+16,2+24,3
LBP	68.96 (1.58)	72.90 (1.54)	76.21 (1.39)	67.34 (1.81)	70.99 (2.28)	73.54 (1.82)
3	74.73 (0.86)	78.38 (0.81)	80.73 (0.71)	71.01 (1.00)	73.81 (1.15)	75.55 (1.07)
6	78.79 (0.77)	82.26 (0.80)	83.73 (0.79)	73.68 (1.46)	77.62 (1.85)	78.08 (1.91)
8	79.63 (0.97)	82.95 (0.92)	84.19 (0.96)	73.79 (1.29)	77.02 (1.71)	77.82 (1.92)
10	79.89 (0.99)	83.21 (0.90)	84.21 (0.89)	73.68 (1.25)	76.84 (1.87)	77.58 (2.03)
12	80.42 (1.08)	83.51 (0.83)	84.19 (0.97)	73.56 (1.08)	76.83 (1.83)	77.56 (2.06)
14	80.58 (1.05)	83.62 (0.94)	83.92 (0.99)	73.33 (1.12)	76.97 (2.03)	77.10 (2.23)
16	80.69 (1.06)	83.49 (0.89)	83.91 (1.04)	73.29 (1.16)	76.88 (2.02)	77.09 (2.21)
19	80.98 (1.14)	83.46 (1.03)	83.63 (0.92)	73.35 (1.15)	76.74 (1.84)	77.04 (2.23)

Figure 5.2: Differences in viewing- and illumination-geometry found in the data set.

cross-date evaluation they ranged from 67% to 74%, showing only a $\pm 2\%$ difference, see table (5.7). The extended LBP (LBP/VAR) results showed noticeable cross-date effects compared to the LBP, see figures (5.1(a), 5.1(b) and 5.1(c) pg. 56). The LBP/VAR same-date averages for number-of-bins = 19 ranged from 81% to 84% while for cross-date evaluation they ranged from 73% to 77% which is a $\pm 7\%$ difference, see table (5.7). From these results it is clear settlement classification accuracy benefits significantly from the addition of contrast measures. However, the classification accuracies were improved more on the same-date images while on the cross-date imagery only minor improvements were observed, thus increasing the gap between the two accuracy measures, as illustrated in table (5.7). This trend was also observed in other similar studies, where while they managed to improve classification accuracies for cross-date images, the counterpart (same-date images) was also improved, resulting in cross-date effects being noticeable [70, 125].

5.4.2 Per-Class Evaluation

In an attempt to identify the cause of the minor classification accuracy improvements for the cross-date scenery we measured the true positive (TP) rates for each settlement type considered. TP rates for each settlement type for the lowest number-of-bins that showed optimal performance, which was a number-of-bins = 6, were evaluated and presented in table (5.4 pg. 58) and summarised in table (5.8).

Table 5.8: Same/cross -date true positive rate value averages per-class (%) with standard deviations of $\mathrm{LBP}^{\mathrm{riu2}}_{\mathrm{P,R}}/\mathrm{VAR}_{\mathrm{P,R}}$ for **number-of-bins = 6**.

Class	**P,R=8,1**		**P,R=8,1+16,2**		**P,R=8,1+16,2+24,3**	
	Same-date	Cross-date	Same-date	Cross-date	Same-date	Cross-date
7.100	81.30 (4.570)	76.21 (12.24)	85.39 (4.270)	82.59 (6.610)	86.45 (3.200)	82.73 (7.958)
7.110	85.52 (2.690)	78.88 (4.560)	87.95 (2.670)	83.41 (4.310)	89.49 (2.070)	83.66 (7.080)
7.120	84.73 (6.820)	63.80 (23.72)	90.21 (6.170)	69.26 (22.60)	92.51 (4.970)	69.79 (22.90)
7.121	8.640 (7.830)	10.05 (2.640)	45.79 (9.540)	24.54 (22.17)	53.56 (10.64)	37.55 (30.11)
7.211	42.51 (10.05)	26.91 (13.31)	44.41 (12.10)	33.23 (15.53)	48.08 (13.82)	35.74 (17.79)
7.212	60.49 (7.120)	50.95 (6.840)	62.11 (6.750)	54.95 (10.17)	65.63 (6.670)	56.91 (8.070)
7.213	25.35 (18.41)	28.57 (12.44)	43.28 (22.82)	40.53 (17.42)	48.88 (24.60)	42.71 (18.39)
7.214	62.99 (7.760)	57.04 (18.63)	68.56 (6.560)	61.81 (18.64)	70.95 (5.730)	64.88 (15.45)
7.241	66.62 (6.670)	39.78 (11.23)	76.98 (5.240)	42.56 (16.20)	81.86 (5.310)	40.01 (20.89)
20.100	97.04 (1.040)	97.07 (0.960)	96.53 (0.730)	96.66 (1.090)	96.17 (0.910)	96.37 (1.440)
Average	**78.79 (0.77)**	**73.69 (1.47)**	**82.27 (0.80)**	**77.62 (1.850)**	**83.73 (0.790)**	**78.08 (1.910)**

We have established that settlement type classification is not an easy task as settlements

have many things in common, this becomes even more difficult task when the classes have few differences. In particularly, informal settlements are the hardest to distinguish as they have subtle differences between them, see illustration in figure (5.3) where the house materials are similar but differ in the layout or in how they are ordered. This is made more apparent in table (5.4), where the TP rates (%) differences between the formal (7.1xx excluding 7.121) and non-built (20.100) classes performed fairly well when compared to informal (7.2xx) class categories regardless of the LBP pattern size.

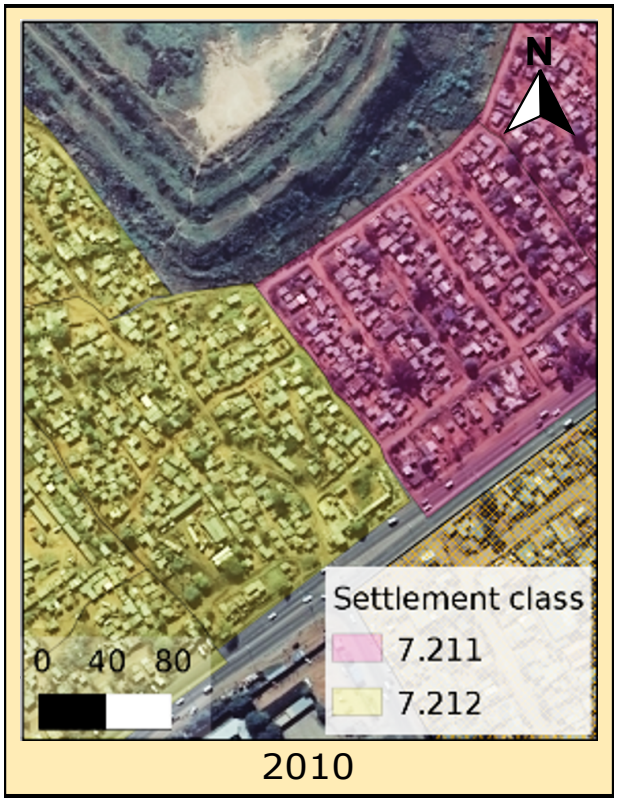


Figure 5.3: Subtle differences in settlement classes found in the data set.

In some cases some of the classes (e.g., $7.121, 7.211, 7.212, 7.213$ and 7.214) performed poorly even on same-date imagery, the classes may be too similar to discern or may be biased towards other classes, however, this may suggest a revision on the these particular classes. For example, combining some of the classes (e.g., $7.211, 7.212, 7.213$ and 7.214) may improve classification accuracy. Its was also noted that more structured classes (7.100 and 7.231) performed better than the unordered ones, also, these classes proved to be more robust to the cross-date effect. The decreased robustness of the classification of the unordered informal settlements is not surprising since these classes exhibit greater internal heterogeneity, see figure (5.4).

The cross-date effects were more pronounced as per-class cross-date evaluation results were lower when compared to the same-date TP values, see averages (shown in bold) in table (5.8).

Figure 5.4: Differences in internal heterogeneity found in the data set.

These results show that seasonal variations did not affect the cross-date accuracies in our experiment, that is, the difference in the amount or colour or the structure of vegetation had no effect on classification performance. The non-built class (20.100), representing vegetation and bare areas, performed better than all the classes under consideration, where the class showed TP values of over 95%. An example of the effects caused by vegetation/seasonal differences is illustrated in figure (5.5), where the amount of trees, colour and structure are more apparent in the 2012 scene than the 2010 scene.

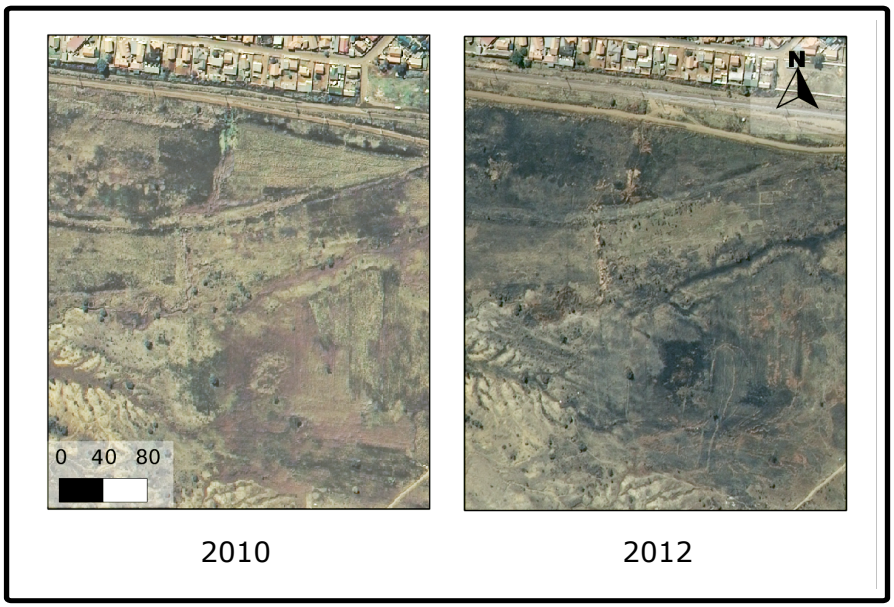


Figure 5.5: An illustration of seasonal differences found in the data set.

5.4.3 SVM's SMO parameter test

A test of the SVM's SMO parameters was done using the a pattern size ($P, R = 8, 1 + 16, 2 + 24, 3$) for number-of-bins = 6, see results in table (5.5 pg. 59). The results were summarised, taking only the overall classification accuracy average values, as shown in table 5.9.

Table 5.9: John Platt's SMO algorithm (C, p) parameter test for overall classification averages (%) with standard deviations using LBP/VAR pattern size ($P, R = 8, 1 + 16, 2 + 24, 3$) with number-of-bins = 6.

C	$p = 1$	2	3
1	81.188 (4.054)	81.429 (5.282)	80.996 (5.910)
2	79.989 (4.183)	80.531 (7.655)	84.345 (9.658)
4	79.799 (3.414)	77.313 (6.413)	81.757 (6.964)
8	79.414 (3.650)	81.069 (6.767)	78.804 (5.811)
16	79.341 (5.279)	81.517 (8.166)	78.877 (6.862)
32	81.412 (6.899)	81.959 (6.814)	83.583 (8.329)
64	78.883 (3.576)	79.659 (6.392)	83.716 (8.419)
128	77.933 (3.445)	83.531 (7.181)	80.953 (4.521)

Classification averages for SMO parameters, SMO$(2, 3)$, SMO$(32, 3)$, SMO$(64, 3)$ and SMO$(128, 2)$, showed an improvement of approx 3% when compared to the default parameters SMO$(1, 1)$, see table (5.9 pg. 66). However, the default parameters SMO$(1, 1)$ also showed half the standard deviation, thus was regarded as the optimal parameters for the experiment. The SVM's SMO parameters were also evaluated for same- and cross-date effects, see table 5.10.

Table 5.10: John Platt's SMO algorithm optimal complexity parameter (C) and exponent value (p) test for LBP/VAR pattern size ($P, R = 8, 1 + 16, 2 + 24, 3$) with number-of-bins = 6. Same- and cross- date overall classification averages (%) with standard deviations.

C	Same-date			Cross-date		
	$p = 1$	2	3	1	2	3
1	82.34 (2.95)	88.91 (9.73)	85.01 (8.10)	80.61 (0.48)	77.69 (3.06)	78.99 (4.82)
2	82.43 (2.90)	79.69 (3.45)	83.34 (2.00)	78.77 (4.82)	80.95 (1.63)	84.85 (1.09)
4	83.70 (4.28)	79.63 (8.64)	79.30 (3.62)	77.85 (2.98)	76.16 (1.06)	82.99 (2.64)
8	81.25 (3.44)	80.71 (4.24)	83.92 (0.83)	78.49 (1.93)	81.25 (3.83)	76.24 (1.91)
16	81.48 (4.50)	80.51 (3.96)	80.26 (9.01)	78.27 (5.67)	82.02 (10.27)	78.18 (5.79)
32	80.66 (3.18)	87.91 (6.50)	87.51 (2.29)	81.79 (8.76)	78.99 (6.97)	81.62 (6.35)
64	81.94 (0.97)	79.60 (3.17)	80.62 (9.17)	77.36 (4.88)	79.69 (8.00)	85.26 (8.05)
128	80.82 (2.20)	87.99 (14.10)	84.48 (4.67)	76.49 (4.07)	81.30 (3.72)	79.19 (4.45)

The evaluation showed high sensitivity to cross-date evaluation, where a parameter that

showed high classification accuracies on same-date conversely yields lower classification accuracies on cross-date data. For example, parameters SMO$(1, 2)$ and SMO$(32, 2)$ showed classification accuracies of (88.91 ± 9.73 on same-date and SMO$(1, 2) = 77.69 \pm 3.06$ on cross-date) and (87.91 ± 6.50 on same-date and 78.99 ± 6.97 on cross-date) respectively. Irrespective of the evident cross-date effects, the results showed that the default parameters SMO$(1, 1)$ were sufficient for the experiment and were used to test and train the whole dataset to achieve optimal classification accuracies with respect to the classifier.

Chapter 6

Conclusions

6.1 Introduction

This chapter presents an overview of the dissertation including the main findings and conclusions. The chapter ends with recommendations for future work.

6.2 Thesis summary

In many parts of the developing world, census and socio-economic data is severely lacking, outdated, or not collected at neighbourhood scales. For government official or town managers to plan equitable solutions to improve living conditions and to be prepared in times of disaster, settlement studies key requirements are to keep track of informal settlements. Using remotely sensed data with pattern recognition techniques, these requirements can met in a timely, cost effective and repeatable manner. However, the diversity of land features, mixed-use settlements, terrain, and heterogeneity of building materials and neighbourhood structure in informal settlements worldwide will always limit the universal applicability of using a fixed set of indicators to identify these areas.

The main contribution of this work was to improve generalisation on settlement type classification of aerial imagery acquired at different dates. Such images (multi-temporal imagery) tend to exhibit high viewing- and illumination geometry effects, which result in a poor generalization performance in settlement type classification tasks. The study investigated the influence of contrast in settlement type classification tasks by measuring classification accuracies using LBP without contrast measures and LBP with contrast measures (LBP/VAR). This was achieved by recognizing fundamental properties of local image texture, i.e., a combination of structural and statistical approaches: the local binary pattern detects micro structures (e.g., edges, lines, spots, flat areas) while variance measures detects the underlying local contrast distribution. The extended LBP algorithm is based on a circularly symmetric neighbour set of pattern size P on a circle of radius R, denoted as $\mathrm{LBP}^{\mathrm{riu2}}_{\mathrm{P,R}}/\mathrm{VAR}_{\mathrm{P,R}}$. The parameter P controls the quantization of the angular space, whereas R determines the spatial resolution of the operator. In addition, the combination of multiple operators with different (P, R) allowed multiresolution evaluation.

6.2.1 Summary of the findings and sub-conclusions:

Classification Accuracy, The results showed that adding contrast features can improve classification accuracies for both same-date and cross-date analysis by $\pm 10\%$ and $\pm 5\%$, respectively. The classification accuracies showed a strong spatial dependence, where these were observed to increased with higher pattern sizes for both LBP and LBP/VAR.

Bin size parameter, The LBP/VAR classification accuracy showed a strong dependence on the bin size parameter where an increase of the number of bins showed an increase in classification accuracy. For simplicity, the small number-of-bins that showed good performance was chosen as the optimal number-of-bins parameter (i.e., number-of-bins = 6 for this study).

Cross-date effect, Even though the LBP/VAR was able to improve classification accuracies, the improvements were not the same for both same-date and cross-date classification accuracies (cross-date effect). The average classification accuracies for various pattern sizes (P, R) improved more on the same-date experiments while showing lower improvements on cross-date measures thus increasing the cross-date effect. The LBP was found to be more robust to cross-date effects compared to LBP/VAR, as it showed a difference in classification accuracy for both same- and cross-date classification. However, the significant increase in classification accuracy using LBP/VAR over LBP is large enough to overcome this deficiency.

Per-class analysis, True positive (TP) rates ($\%$) differences between the formal and non-built classes performed fairly well when compared to the informal class categories. More structured classes (formal settlements) perform better than the unordered ones (Informal settlements subclasses), in addition, these classes demonstrated to be more robust to the cross-date effect. Informal classes, especially with backyard structures, were observed to be more problematic as they exhibit higher internal heterogeneity in comparison to the formal counterpart. The informal classes may have been too similar to discern or the classifier was biased to other classes. Unfortunately, the study was limited to TP values only, thus we cannot conclude which class was confused which class. However, the results did show the LBP/VAR was robust to seasonal differences between the two dates (non-built class TP rates were 95%).

Classifier parameters, The optimal parameters for SMO algorithm used for training the SVM

classifier were investigated and the default $SMO(1,1)$ was found to be sufficient for the study. Thus the classification accuracies were obtained at an optimal configuration with respect to the classifier.

6.3 Conclusions and further research

The extended LBP offered strong spatial and temporal generalisation. Our findings suggest that adding the rotational invariant Variance measure to the rotational and gray-scale invariant Local Binary Pattern (LBP) played an important role in improving the classification of settlement type from aerial imagery. The experiments showed that optimal performance can be achieved without being to complex (i.e., using a small number-of-bins and default SVM parameters). Though cross-date effects were minimal on LBP with contrast features, while the added of contrast features were not as robust, the study showed major improvements in overall classification accuracies, especially under similar conditions (same-date images). There have been other studies that have shown the local spatial patterns and contrast features to be important in texture classification [77, 90, 120]. In conclusion, the study suggests that an contrast properties can be a useful feature in the implementation of an automated settlement monitoring system.

6.3.1 Suggestions for further research

1. Distinguishing formal from informal settlements using the extended LBP proved to be successful. However, classifying human settlement subclasses, especially informal settlement subclasses, was not as good and therefore needs further investigation. To improve the overall classification accuracies of the settlement subclasses, in particular informal settlements, a spatially explicit evaluation of informal settlement subclasses is required. A combination of approaches such as spectral, texture, geomorphology and road networks and dominant settlement materials (vegetation, soil, asphalt) has been shown to be a viable method in developing the correct combination of indicators for differentiating settlements [92]. The latter incorporated with the extended LBP method may significantly improve classification of settlement subclasses, however, this is not a trivial task and the added feature may quickly lead to large feature dimensions and result in a costly system.

2. Treating shadows prior to the extraction of features has been shown successful in distinguishing formal from informal settlements [71]. Used with the extended LBP, it is expected to produce a more generalised classification system. The correct combination

of indicators for the subclasses and prior shadow treatment, incorporated with the extended LBP generalisation classification accuracies can be expected to be more robust to viewing- and illumination-geometry effects.

BIBLIOGRAPHY

[1] AHUJA, N., AND ROSENFELD, A. Mosaic models for textures. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1 (1981), pp. 1–11.

[2] AMINIPOURI, M., SLIUZAS, R., AND KUFFER, M. Object-oriented analysis of very high resolution orthophotos for estimating the population of slum areas, case of Dar-Es-Salaam, Tanzania. In *Proceedings of the ISPRS conference: High-Resolution Earth Imaging for Geospatial Information* (2009), pp. 1–6.

[3] ARIVAZHAGAN, S., AND GANESAN, L. Texture classification using wavelet transform. *Pattern Recognition Letters 24*, 9 (2003), pp. 1513–1521.

[4] ARIVAZHAGAN, S., AND GANESAN, L. Texture segmentation using wavelet transform. *Pattern Recognition Letters 24*, 16 (2003), pp. 3197–3203.

[5] AZAR, D., GRAESSER, J., ENGSTROM, R., COMENETZ, J., LEDDY JR, R. M., SCHECHTMAN, N. G., AND ANDREWS, T. Spatial refinement of census population distribution using remotely sensed estimates of impervious surfaces in Haiti. *International Journal of Remote Sensing 31*, 21 (2010), pp. 5635–5655.

[6] BARROS FILHO, M., AND SOBREIRA, F. Urban textures: A multiscale analysis of socio-spatial patterns. In *10th International Conference on Computers in Urban Planning and Urban Management, Iguassu Falls, Brazil* (2007), pp. 1–12.

[7] BARROS FILHO, M., AND SOBREIRA, F. Accuracy of lacunarity algorithms in texture classification of high spatial resolution images from urban areas. In *XXI Congress of International Society of Photogrammetry and Remote Sensing* (2008), pp. 417–422.

[8] BENEDIKTSSON, J. A., PESARESI, M., AND AMASON, K. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *Geoscience and Remote Sensing, IEEE Transactions on 41*, 9 (2003), pp. 1940–1949.

[9] BLASCHKE, T., AND LANG, S. Object based image analysis for automated information extraction–a synthesis. In *Measuring the Earth II ASPRS Fall Conference* (2006), pp. 6–10.

[10] BOVOLO, F., BRUZZONE, L., AND MARCONCINI, M. A novel context-sensitive SVM for classification of remote sensing images. In *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on* (2006), IEEE, University of Trento, pp. 2498–2501.

[11] BRUZZONE, L., CHI, M., AND MARCONCINI, M. A novel transductive svm for semisupervised classification of remote-sensing images. *Geoscience and Remote Sensing, IEEE Transactions on 44*, 11 (2006), pp. 3363–3373.

[12] BURGES, C. J. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery 2*, 2 (1998), pp. 121–167.

[13] BURRUS, C. S., GOPINATH, R. A., GUO, H., ODEGARD, J. E., AND SELESNICK, I. W. *Introduction to wavelets and wavelet transforms: a primer*, vol. 23. Prentice hall Upper Saddle River, 1998.

[14] CAMPBELL, J. B., AND WYNNE, R. H. *Introduction to Remote Sensing*. The Guilford Press, 2011.

[15] CANADA, N. R. A Canada centre for remote sensing tutorial - fundamentals of remote sensing, 2001. Available from: `http://www.nrcan.gc.ca/sites/www.nrcan.gc.ca.earth-sciences/files/pdf/resource/tutor/fundam/pdf/fundamentals_e.pdf`.

[16] CENDRERO, A., LTTIG, G., AND WOLFF, F. C. *Planning the Use of the Earth's Surface*. Springer Verlag, 1992.

[17] CHEN, M., LIU, S., TIESZEN, L. L., AND HOLLINGER, D. Y. An improved state-parameter analysis of ecosystem models using data assimilation. *Ecological Modelling 219* (2008), pp. 317–326.

[18] CHEN, Y., AND DOUGHERTY, E. R. Gray-scale morphological granulometric texture classification. *Optical Engineering 33*, 8 (1994), pp. 2713–2722.

[19] COGGINS, J. M. A framework for texture analysis based on spatial filtering, 1983.

[20] COGGINS, J. M., AND JAIN, A. K. A spatial filtering approach to texture analysis. *Pattern recognition letters 3*, 3 (1985), pp. 195–203.

[21] COHEN, F. S., FAN, Z., AND PATEL, M. A. Classification of rotated and scaled textured images using Gaussian Markov random field models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 13*, 2 (1991), pp. 192–202.

[22] COLWELL, R. N., ESTES, J., AND THORLEY, G. *Manual of remote sensing.*, vol. 2. American Society of Photogrammetry, 1983.

[23] COLWELL, R. N., ULABY, F., AND SIMONETT, D. *Manual of remote sensing.*, vol. 1. American Society of Photogrammetry, 1983.

[24] CORPORATION, I. Digital mapping camera brochure, March 2014. Available from: `http://www.geospace.co.za/pdf/DMC\%20Brochure.pdf`.

[25] CORTES, C., AND VAPNIK, V. Support-vector networks. *Machine learning 20*, 3 (1995), 273–297.

[26] CRISTIANINI, N., AND SCHOLKOPF, B. Support vector machines and kernel methods: the new generation of learning machines. *Ai Magazine 23*, 3 (2002), pp. 31–42.

[27] DELL'ACQUA, F., AND GAMBA, P. Texture-based characterization of urban environments on satellite SAR images. *Geoscience and Remote Sensing, IEEE Transactions on 41*, 1 (2003), pp. 153–159.

[28] DONNAY, J.-P., BARNSLEY, M. J., AND LONGLEY, P. A. *Remote Sensing and Urban Analysis: GISDATA 9*, vol. 9. CRC Press, 2003.

[29] DUDA, R. O., HART, P. E., AND STORK, D. G. *Pattern classification*. John Wiley & Sons, 2012.

[30] ELLA, L., VAN DEN BERGH, F., VAN WYK, B. J., AND VAN WYK, M. A. A comparison of texture feature algorithms for urban settlement classification. In *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International* (2008), vol. 3, pp. III–1308 – III–1311.

[31] FAUVEL, M., BENEDIKTSSON, J. A., CHANUSSOT, J., AND SVEINSSON, J. R. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *Geoscience and Remote Sensing, IEEE Transactions on 46*, 11 (2008), pp. 3804–3814.

[32] FOGEL, I., AND SAGI, D. Gabor filters as texture discriminator. *Biological cybernetics 61*, 2 (1989), pp. 103–113.

[33] FOODY, G. M., AND MATHUR, A. A relative evaluation of multiclass image classification by support vector machines. *Geoscience and Remote Sensing, IEEE Transactions on 42*, 6 (2004), pp. 1335–1343.

[34] GAMBA, P., ALDRIGHI, M., AND STASOLLA, M. Robust extraction of urban area extents in HR and VHR SAR images. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 4*, 1 (2011), pp. 27–34.

[35] GAMBA, P., DELL 'ACQUA, F., AND TRIANNI, G. Satellite SAR and human settlement detection. In *Urban Remote Sensing Joint Event* (2007), pp. 1–4.

[36] GAMBA, P., PESARESI, M., MOLCH, K., GERHARDINGER, A., AND LISINI, G. Anisotropic rotation invariant built-up presence index: Applications to SAR data. In *Geoscience and Remote Sensing Symposium, IGARSS. IEEE International* (2008), vol. 5, pp. V–338–V–341.

[37] GRAESSER, J., CHERIYADAT, A., VATSAVAI, R. R., CHANDOLA, V., LONG, J., AND BRIGHT, E. Image based characterization of formal and informal neighbourhoods in an urban landscape. *Selected Topics in Earth Observation and Remote Sensing, IEEE Journal 5* (2012), pp. 1164–1176.

[38] GUALTIERI, J., AND CHETTRI, S. Support vector machines for classification of hyperspectral data. In *Geoscience and Remote Sensing Symposium. Proceedings, IGARSS. IEEE International* (2000), vol. 2, pp. 813–815.

[39] GUALTIERI, J. A., AND CROMP, R. F. Support vector machines for hyperspectral remote sensing classification. In *The 27th AIPR Workshop: Advances in Computer-Assisted Recognition* (1999), International Society for Optics and Photonics, pp. 221–232.

[40] HABITAT, U. The challenge of slums: global report on human settlements 2003. *London: Earthscan* (2003).

[41] HABITAT, U. State of the worlds cities. *The Millenium Development Goals and Urban Sustainability.London: Earthscan* (2006).

[42] HARALICK, R. M. Statistical and structural approaches to texture. *Proceedings of the IEEE 67*, 5 (1979), pp. 786–804.

[43] HARALICK, R. M., SHANMUGAM, K., AND DINSTEIN, I. H. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, 6 (1973), pp. 610–621.

[44] HAY, G., AND CASTILLA, G. Object-based image analysis: Strengths, weaknesses, opportunities and threats (SWOT). In *Proc. 1st Int. Conf. OBIA* (2006), pp. 4–5.

[45] HEIDEN, U., ROESSNER, S., SEGL, K., AND KAUFMANN, H. Analysis of spectral signatures of urban surfaces for their identification using hyperspectral HyMap data. In *Remote Sensing and Data Fusion over Urban Areas, IEEE/ISPRS Joint Workshop 2001* (2001), IEEE, pp. 173–177.

[46] HEROLD, M., GARDNER, M. E., AND ROBERTS, D. A. Spectral resolution requirements for mapping urban areas. *Geoscience and Remote Sensing, IEEE Transactions on 41*, 9 (2003), pp. 1907–1919.

[47] HEROLD, M., LIU, X., AND CLARKE, K. C. Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering and Remote Sensing 69*, 9 (2003), pp. 991–1002.

[48] HOFMANN, P., STROBL, J., BLASCHKE, T., AND KUX, H. *Detecting informal settlements from Quickbird data in Rio de Janeiro using an object based approach*. Object-Based Image Analysis. Springer, 2008, pp. 531–553.

[49] HSU, C., CHANG, C., AND LIN, C. A practical guide to support vector classification, 2003.

[50] HUANG, C., DAVIS, L., AND TOWNSHEND, J. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing 23*, 4 (2002), pp. 725–749.

[51] HUCHZERMEYER, M., AND KARAM, A. *Informal settlements: a perpetual challenge?* Juta and Company Ltd, 2006.

[52] HUNG, W., CHEN, Y., AND CHENG, K.-S. Comparing landcover patterns in Tokyo, Kyoto, and Taipei using ALOS multispectral images. *Landscape and Urban Planning 97*, 2 (2010), pp. 132–145.

[53] INTERNATIONAL, P. A. P. A. History of aerial photography [web site], 2001. Available from: `http://www.papainternational.org/history.asp`.

[54] JAIN, A. K., DUIN, R. P. W., AND MAO, J. Statistical pattern recognition: A review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 22*, 1 (2000), pp. 4–37.

[55] JAIN, S. Use of IKONOS satellite data to identify informal settlements in Dehradun, India. *International Journal of Remote Sensing 28*, 15 (2007), pp. 3227–3233.

[56] JENEROWICZ, M., KEMPER, T., PESARESI, M., AND SOILLE, P. Post-event damage assessment using morphological methodology on 0.5 m resolution satellite data. *Italian Journal of Remote Sensing 42*, 3 (2010), pp. 37–47.

[57] JENSEN, J. R., AND COWEN, D. C. Remote sensing of urban/suburban infrastructure and socio-economic attributes. *Photogrammetric Engineering and Remote Sensing 65* (1999), pp. 611–622.

[58] KARISHMA, B., ANDRE, B., AND JARREL, W. Potential application of remote sensing in monitoring informal settlements in developing countries where complimentary data does not exist. In *Conference Report, Planning Africa* (2008).

[59] KHORRAM, S., NELSON, S. A. C., AND KOCH, F. H. *Remote sensing*. Siamak Khorram, 2012.

[60] KHUMALO, P., TAPAMO, J., AND VAN DER BERGH, F. Rotation invariant texture feature algorithms for urban settlement classification. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International* (2011), pp. 511–514.

[61] KIT, O., LÜDEKE, M., AND RECKIEN, D. Texture-based identification of urban slums in Hyderabad, india using remote sensing data. *Applied Geography 32*, 2 (2012), pp. 660–667.

[62] KOHLI, D., SLIUZAS, R., KERLE, N., AND STEIN, A. An ontology of slums for image-based classification. *Computers, Environment and Urban Systems 36*, 2 (2012), pp. 154–163.

[63] KOTSIANTIS, S. B., ZAHARAKIS, I., AND PINTELAS, P. Supervised machine learning: A review of classification techniques. *Informatica 31* (2007), pp. 249–268.

[64] LIANG, S. *Advances in Land Remote Sensing: System, Modelling, Inversion and Application*. Springer New York, NY, 2008.

[65] LILLESAND, T., AND KIEFER, R. *Remote sensing and image interpretation*. John Wiley and Sons Inc., New York, NY, 1987.

[66] LO, C. *Applied remote sensing.* Longman Scientific & Technical, Longman Group UK Limited, 1986.

[67] LOPEZ-ORNELAS, E., AND FLOUZAT, G. Implicit spatial information extraction from remote sensing images. In *Headway in Spatial Data Handling*. Springer, 2008, pp. 133–146.

[68] LU, D., AND WENG, Q. Use of impervious surface in urban land-use classification. *Remote Sensing of Environment 102*, 1 (2006), pp. 146–160.

[69] LU, D., AND WENG, Q. Extraction of urban impervious surfaces from an IKONOS image. *International Journal of Remote Sensing 30*, 5 (2009), pp. 1297–1311.

[70] LUUS, F. P. S., VAN DEN BERGH, F., AND MAHARAJ, B. T. J. The effects of shadow removal on across-date settlement type classification of QuickBird images. In *Geoscience and Remote Sensing Symposium (IGARSS), IEEE International* (2012), pp. 6196–6199.

[71] LUUS, F. P. S., VAN DEN BERGH, F., AND MAHARAJ, B. T. J. The effects of segmentation-based shadow removal on across-date settlement type classification of panchromatic QuickBird images. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 6*, 3 (2013), pp. 1274–1285.

[72] MALIK, J., AND PERONA, P. Preattentive texture discrimination with early vision mechanisms. *JOSA A 7*, 5 (1990), pp. 923–932.

[73] MANTERO, P., MOSER, G., AND SERPICO, S. B. Partially supervised classification of remote sensing images through SVM-based probability density estimation. *Geoscience and Remote Sensing, IEEE Transactions on 43*, 3 (2005), pp. 559–570.

[74] MANTHALKAR, R., BISWAS, P. K., AND CHATTERJI, B. N. Rotation and scale invariant texture features using discrete wavelet packet transform. *Pattern Recognition Letters 24*, 14 (2003), pp. 2455–2462.

[75] MATERKA, A., AND STRZELECKI, M. Texture analysis methods - a review. *Technical University of lodz, Institute of Electronics, COST B11 report, Brussels* (1998), pp. 9–11.

[76] MAYUNGA, S., COLEMAN, D., AND ZHANG, Y. A semi-automated approach for extracting buildings from QuickBird imagery applied to informal settlement mapping. *International Journal of Remote Sensing 28*, 10 (2007), pp. 2343–2357.

[77] MDAKANE, L., AND VAN DEN BERGH, F. Extended local binary pattern features for improving settlement type classification of QuickBird images. In *Proceedings of the Twenty-Third Annual Symposium of the Pattern Recognition Association of South Africa, PRASA* (2012), pp. 68–74.

[78] MELGANI, F., AND BRUZZONE, L. Classification of hyperspectral remote sensing images with support vector machines. *Geoscience and Remote Sensing, IEEE Transactions on 42*, 8 (2004), pp. 1778–1790.

[79] MESEV, V., GORTE, B., AND LONGLEY, P. A. Modified maximum-likelihood classification algorithms and their application to urban remote sensing. *Remote Sensing and Urban Analysis: GISDATA 9 9* (2003), pp. 62–84.

[80] MIRMEHDI, M., XIE, X., AND SURI, J. *Handbook of texture analysis*. Imperial College Press, 2009.

[81] MORGAN, J. L., GERGEL, S. E., AND COOPS, N. C. Aerial photography: a rapidly evolving tool for ecological management. *Bioscience 60*, 1 (2010), pp. 47–59.

[82] MOSHA, A. An evaluation of botswana's stratgies to regularize informal settlements. *Review of Urban & Regional Development Studies 8*, 1 (1996), pp. 46–65.

[83] MOUNTRAKIS, G., IM, J., AND OGOLE, C. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing 66*, 3 (2011), pp. 247–259.

[84] MURIUKI, G., MCALPINE, C., SEABROOK, L., AND BAXTER, G. The role of squatters in retention of native vegetation: A case study of the Chyulu Hills, Kenya. *Applied Geography 31*, 2 (2011), pp. 577–589.

[85] MYINT, S. W., GOBER, P., BRAZEL, A., GROSSMAN-CLARKE, S., AND WENG, Q. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment 115*, 5 (2011), pp. 1145–1161.

[86] NASA. Landsat program history, March 2013. Available from: `http://landsat.gsfc.nasa.gov/about/history.html`.

[87] NIEBERGALL, S., LOEW, A., AND MAUSER, W. Object-oriented analysis of very high-resolution Quickbird data for mega city research in Delhi/India. In *Urban Remote Sensing Joint Event, 2007* (2007), IEEE, pp. 1–8.

[88] NIEBERGALL, S., LOEW, A., AND MAUSER, W. Integrative assessment of informal settlements using VHR remote sensing data - the Delhi case study. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 1*, 3 (2008), pp. 193–205.

[89] OJALA, T., PIETIKÄINEN, M., AND HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition 29*, 1 (1996), pp. 51–59.

[90] OJALA, T., PIETIKAINEN, M., AND MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 24*, 7 (2002), pp. 971–987.

[91] OWEN, K. K., AND WONG, D. W. Exploring structural differences between rural and urban informal settlements from imagery: the basureros of Cobn. *Geocarto International* (2012), pp. 1–20.

[92] OWEN, K. K., AND WONG, D. W. An approach to differentiate informal settlements using spectral, texture, geomorphology and road accessibility metrics. *Applied Geography 38* (2013), pp. 107–118.

[93] PAGET, R. Texture modelling and synthesis. *Handbook of texture analysis. Imperial College Press, London* (2008), pp. 33–60.

[94] PAL, M., AND MATHER, P. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing 26*, 5 (2005), pp. 1007–1011.

[95] PAO, Y.-H. *Adaptive pattern recognition and neural networks*. Addison-Wesley Longman Publishing Co., Inc., 1989.

[96] PATINO, J. E., AND DUQUE, J. C. A review of regional science applications of satellite remote sensing in urban settings. *Computers, Environment and Urban Systems* (2013), 1–17.

[97] PESARESI, M. Texture analysis for urban pattern recognition using fine-resolution panchromatic satellite imagery. *Geographical and Environmental Modelling 4*, 1 (2000), pp. 43–63.

[98] PESARESI, M., AND BIANCHIN, A. *Recognizing settlement structure using mathematical morphology and image texture*. Taylor & Francis London, 2000.

[99] PESARESI, M., AND GERHARDINGER, A. Improved textural built-up presence index for automatic recognition of human settlements in arid regions with scattered vegetation. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 4*, 1 (2011), pp. 16–26.

[100] PESARESI, M., GERHARDINGER, A., AND KAYITAKIRE, F. Monitoring settlement dynamics by anisotropic textural analysis of panchromatic vhr data. In *Urban Remote Sensing Joint Event, 2007* (2007), IEEE, pp. 1–11.

[101] PESARESI, M., GERHARDINGER, A., AND KAYITAKIRE, F. A robust built-up area presence index by anisotropic rotation-invariant textural measure. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 1*, 3 (2008), pp. 180–192.

[102] PETROU, M., AND PETROU, C. *Image processing: the fundamentals*. Wiley. com, 2010.

[103] PLATT, J. C. 12 fast training of support vector machines using sequential minimal optimization, 1999.

[104] PLAZA, A., BENEDIKTSSON, J. A., BOARDMAN, J. W., BRAZILE, J., BRUZZONE, L., CAMPS-VALLS, G., CHANUSSOT, J., FAUVEL, M., GAMBA, P., AND GUALTIERI, A. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment 113* (2009), pp. S110–S122.

[105] PLOTNICK, R. E., GARDNER, R. H., AND O'NEILL, R. V. Lacunarity indices as measures of landscape texture. *Landscape Ecology 8*, 3 (1993), pp. 201–211.

[106] QUINLAN, J. R. *C4. 5: Programs for machine learning*, vol. 1. Morgan Kaufmann, 1993.

[107] RELLIER, G., DESCOMBES, X., FALZON, F., AND ZERUBIA, J. Texture feature analysis using a Gauss-Markov model in hyperspectral image classification. *Geoscience and Remote Sensing, IEEE Transactions on 42*, 7 (2004), pp. 1543–1551.

[108] RHINANE, H., HILALI, A., BERRADA, A., AND HAKDAOUI, M. Detecting slums from SPOT data in Casablanca Morocco using an object based approach. *J. Geographic Information System 3*, 3 (2011), pp. 217–224.

[109] RICHARDS, J. A. *Remote sensing digital image analysis: an introduction*. Springer, 2012.

[110] RICHASON JR, B. F. Introduction to remote sensing of the environment. *Laboratory manual for introduction to remote sensing of the environment.* (1978).

[111] SCHOWENGERDT, R. A. *Remote sensing: models and methods for image processing*. Academic press, 2006.

[112] SHENSA, M. J. The discrete wavelet transform: wedding the a trous and Mallat algorithms. *Signal Processing, IEEE Transactions on 40*, 10 (1992), pp. 2464–2482.

[113] SIVAKUMAR, K., AND GOUTSIAS, J. Morphologically constrained GRFs: applications to texture synthesis and analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 21*, 2 (1999), pp. 99–113.

[114] SOBREIRA, F. Squatter settlements consolidation: Spatial analysis in an agent-based environment. In *Proceedings, 4th International Space Syntax Symposium, London* (2003), pp. 16.1–16.12.

[115] SOHN, G., AND DOWMAN, I. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing 62*, 1 (2007), pp. 43–63.

[116] SOLOMON, C., AND BRECKON, T. *Fundamentals of digital image processing: A practical approach with examples in Matlab*. John Wiley & Sons, 2011.

[117] STASOLLA, M., AND GAMBA, P. Exploiting spatial patterns for informal settlement detection in arid environments using optical spaceborne data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36*, 3 (2007), pp. 31–36.

[118] STOW, D., LOPEZ, A., LIPPITT, C., HINTON, S., AND WEEKS, J. Object-based classification of residential land use within accra, Ghana based on Quickbird satellite data. *International Journal of Remote Sensing 28*, 22 (2007), pp. 5167–5173.

[119] SU, W., LI, J., CHEN, Y., LIU, Z., ZHANG, J., LOW, T. M., SUPPIAH, I., AND HASHIM, S. A. M. Textural and local spatial statistics for the object-oriented classification of urban areas using high resolution imagery. *International Journal of Remote Sensing 29*, 11 (2008), pp. 3105–3117.

[120] TAMURA, H., MORI, S., AND YAMAWAKI, T. Textural features corresponding to visual perception. *Systems, Man and Cybernetics, IEEE Transactions on 8*, 6 (1978), pp. 460–473.

[121] TECHNOLOGY, R. . S. Rocket & space milestones, 2006. Available from: `http://www.braeunig.us/space/`.

[122] TUCERYAN, M., AND JAIN, A. K. Texture analysis. *Handbook of Pattern Recognition and Computer Vision 276* (1993), pp. 277–310.

[123] UNSER, M. Texture classification and segmentation using wavelet frames. *Image Processing, IEEE Transactions on 4*, 11 (1995), pp. 1549–1560.

[124] VAN DE VOORDE, T., JACQUET, W., AND CANTERS, F. Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data. *Landscape and Urban Planning 102*, 3 (2011), pp. 143–155.

[125] VAN DEN BERGH, F. The effects of viewing- and illumination geometry on settlement type classification of quickbird images. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International* (2011), pp. 1425–1428.

[126] VAPNIK, V. N. *Statistical learning theory (adaptive and learning systems for signal processing, communications and control series)*. John Wiley & Sons, New York. A Wiley-Interscience Publication, 1998.

[127] VIJAYARAJ, V., CHERIYADAT, A. M., SALLEE, P., COLDER, B., VATSAVAI, R. R., BRIGHT, E. A., AND BHADURI, B. L. Overhead image statistics. In *Applied Imagery Pattern Recognition Workshop, 2008. AIPR'08. 37th IEEE* (2008), IEEE, pp. 1–8.

[128] WASKE, B., AND BENEDIKTSSON, J. A. Fusion of support vector machines for classification of multisensor data. *Geoscience and Remote Sensing, IEEE Transactions on 45*, 12 (2007), pp. 3858–3866.

[129] WEBER, C., AND PUISSANT, A. Urbanization pressure and modeling of urban growth: example of the Tunis Metropolitan Area. *Remote Sensing of Environment 86*, 3 (2003), pp. 341–352.

[130] WEEKS, J. R., HILL, A., STOW, D., GETIS, A., AND FUGATE, D. Can we spot a neighborhood from the air? Defining neighborhood structure in Accra, Ghana. *GeoJournal 69*, 1-2 (2007), pp. 9–22.

[131] WELCH, R. Spatial resolution requirements for urban studies. *International Journal of Remote Sensing 3*, 2 (1982), pp. 139–146.

[132] WENG, Q. *Remote sensing and GIS integration: Theories, methods, and applications*. McGraw-Hill Professional, 2009.

[133] ZHANG, D., ISLAM, M. M., AND LU, G. A review on automatic image annotation techniques. *Pattern Recognition 45*, 1 (2012), pp. 346–362.

[134] ZHANG, J., AND TAN, T. Brief review of invariant texture analysis methods. *Pattern Recognition 35*, 3 (2002), pp. 735–747.

[135] ZHOU, G., CHEN, W., KELMELIS, J. A., AND ZHANG, D. A comprehensive study on urban true orthorectification. *Geoscience and Remote Sensing, IEEE Transactions on 43*, 9 (2005), pp. 2138–2147.

[136] ZHU, G., AND BLUMBERG, D. G. Classification using ASTER data and SVM algorithms;: The case study of Beer Sheva, Israel. *Remote Sensing of Environment 80*, 2 (2002), pp. 233–240.