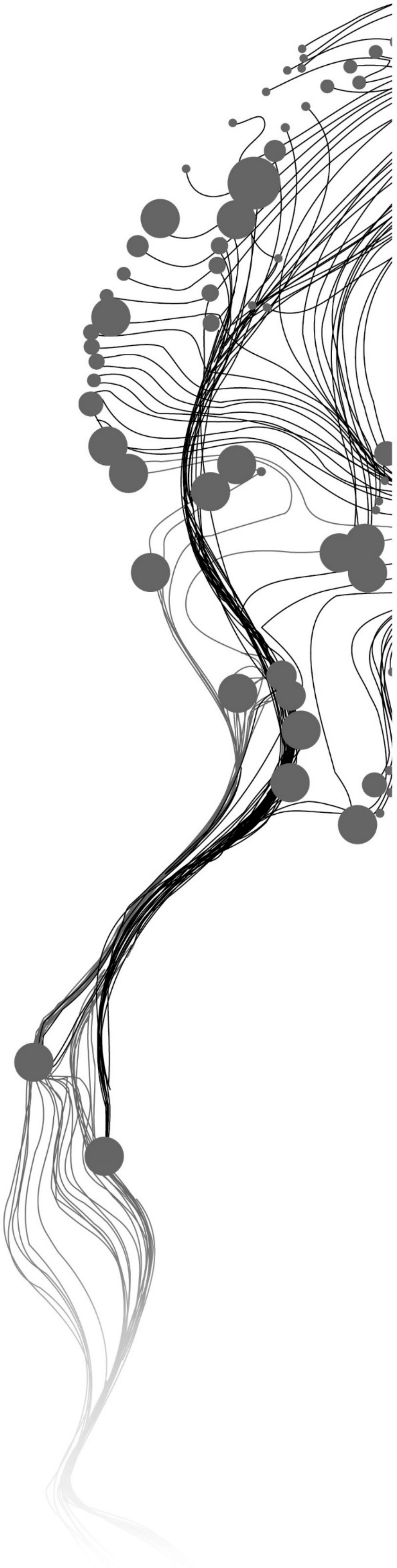
Robustness of Rule Sets Using VHR Imagery to Detect Informal Settlements – A Case of Mumbai, India

VICKYSON NAOREM
February, 2016

SUPERVISORS:
Ms. Monika Kuffer, MSc (1st Supervisor)
Drs. J.J. (Jeroen) Verplanke (2nd Supervisor)



Robustness of Rule Sets Using VHR Imagery to Detect Informal Settlements – A Case of Mumbai, India

VICKYSON NAOREM
Enschede, The Netherlands, February, 2016

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Urban Planning and Management

SUPERVISORS:
Ms. Monika Kuffer, MSc (1st Supervisor)
Drs. J.J. (Jeroen) Verplanke (2nd Supervisor)

THESIS ASSESSMENT BOARD:
Prof.dr. P.Y. Georgiadou (Chair)
Dr. M. Netzband (External Examiner, Ruhr-Universität Bochum)

ABSTRACT

Robust monitoring approaches of informal settlements using very high resolution (VHR) satellite imagery can deliver essential information for supporting policies aiming to reduce the proliferation of such settlements and assist in the formulation of pro poor policies, and moreover, can complement census methods or participatory approaches. With the increasing availability of VHR satellite imagery, detection of the informal settlements benefits from the conceptualization of location specific knowledge in the form of a local slum ontology. In this study, we developed the local slum ontology for Mumbai, India, by incorporating local knowledge. Then, we translated the local ontology into a rule set using Object Based Image Analysis (OBIA) to identify informal settlements by using spatial, geometry, spectral indices and texture. The method was applied to three subsets of a Worldview-2 imagery to analyse the robustness of the initial rule set with the help of membership functions. After the classification, we refined the rules to extract object features based on spectral information, contextual feature and shape. The results showed that the parameters of NDNB of spectral index and GLCM feature of texture found to be effective in all three subsets in the identification of the informal settlements. The results suggest that the rule sets used in this study can be applied to other study areas of Worldview-2 imagery for identification of the informal settlements with minimum adaptation.

Keywords: Robustness, rule sets, object based image analysis, informal settlements, urban

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to Mrs. Monika Kuffer and drs. Jeroen Verplanke, my thesis supervisors, for his patient guidance, enthusiastic encouragement and useful critiques of this thesis. I would also like to thank Mr. Ronald Soubam and Mr. Kakchingtabam Lenin Sharma for their support in the fieldwork.
I would also like to thank Dr. Divyani Kohli for the valuable advice in learning the software of eCognition and also for the guidance of questionnaire.

My grateful thanks are also extended to the sponsor who granted me the EMMA-West 2013 scholarship for the financial support throughout the course.

I extend my sincere thanks to our Department of Urban and Regional Planning and Geo-Information Management (PGM) for their support throughout my thesis.

I would also like to thank my friends and my classmates especially Mr. Jati Pratomo for their heartfelt support and enthusiasm.

Finally, I wish to thank my family for their support and encouragement throughout my thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

DEM	Digital Elevation Model
EO	Earth Observation
FCM	Fuzzy Certainty Measure
GEOBIA	Geographical Object Based Image Analysis
GIS	Geographical Information System
GIScience	Geographical Information Science
GLCM	Grey Level Co-occurrence Matrix
GSO	Generic Slum Ontology
km	Kilometre
km^2	Square kilometre
m	Metre
NIR	Near-Infrared Red
NDVI	Normal Difference Vegetation Index
OBIA	Object Based Image Analysis
OOA	Object Oriented Image Analysis
r	Degree of robustness (measurement)
R_o	Initial rule set
R_f	Adapted rule set
SPSS	Statistical Package for Social Sciences
Subset 1	Referenced imagery
Subsets 2 & 3	Test imagery
UN	United Nations
VHR	Very High Resolution
WV2	WorldView-2

1. INTRODUCTION

This chapter provides the justification of the study and the research problem. It also includes the main objective, research objectives and their research questions, significance of the study and structure of the thesis.

1.1. Background and Justification

The availability of more job opportunities and better services close to the urban areas as compared to that of rural areas make the rural people migrate into urban areas. The urban slum population in the Global South is clearly proliferating due to urbanization (UN-Habitat, 2012). Due to dynamic and rapid growth of the population in the Global South, the economic instability and inadequate urban planning law enforcement drive rich people into gated communities thereby creating peculiar regions for their own security (UN-Habitat, 2012). Apparently, the gap causes farther exclusion and marginalization through restrictions on jobs, compounding gender differences, limiting social interactions and reducing social capital, increasing the crime and violence, with deteriorating quality of life as a result (UN-Habitat, 2012). Additionally, it makes the durable housing in the city unaffordable for the underprivileged people who have to settle in informal settlements. The growth of informal settlements is also due to the lack of governance which is inclusive approach in the public sector (Ezeah, Fazakerley, & Roberts, 2013; UN-Habitat, 2010). The rapid growth of informal settlements in cities of the Global South are a cause of concern to the respective government as it becomes proliferated, menaced and challenging task for the overall development of the countries. On the other hand, the cities need a high demand of workers in constructing the infrastructures, retail and many other sectors. At the same time, the need of jobs and livelihood also compel the workers to shift from rural areas to cities. However, the high cost of living in formal settlements in cities and other social factors make the workers to settle in the informal areas characterized by overcrowded, dilapidated housing and inadequate access to basic services (UN-Habitat, 2010). Moreover, location of the informal settlements are often at polluted and hazardous areas, areas close to the river and the places of unhygienic environment (Kit, Lüdeke, & Reckien, 2012; Kohli, Sliuzas, Kerle, & Stein, 2012; Kuffer, Pfeffer, Baud, & Sliuzas, 2013; Sirueri, 2015; UN-Habitat, 2010). Eventually, all these factors affect severely health conditions of the inhabitants of informal settlements. In addition, the social exclusion from the privileged communities can cause mental problems in the inhabitants (Subbaraman et al., 2014). In order to tackle these mentioned problems of informal settlements in a decisive and effective manner, there is a need to detect the location and boundaries of informal settlements. These measures will enable to provide such information to planning authorities and to monitor development dynamics which can assist in pro-poor policy formulation, as stepping stone towards inclusive development.

Several methods like census methods, participatory methods and advanced remote sensing image analysis have been introduced so far to identify informal settlements. However, the utility of the census method is limited by the long temporal gap between two census surveys due to the political influences (Ebert, Kerle, & Stein, 2009). The degree of the data aggregation can also be a problem due to variety in size of the households (Kohli et al., 2012). The participatory methods are helpful to incorporate spatial data and non-spatial data but time-consuming to collect data especially at large areas (Kohli et al., 2012). The above mentioned methods are sluggish for extracting information of large areas. In such circumstances, there is actual need for remote sensing technique without which it is difficult to extract information at large scale.

This shows the importance of the technique for identifying and detecting informal settlements at large scale so that the issues can be mitigated for their development.

Recently, image analysis of remote sensing techniques have been developed to such an extent that these instruments yield considerably comparatively faster in acquiring results to the mapping and measurements which can explain the arrangement and development of large-scale urban lands (Taubenböck & Kraff, 2014). With the increasing availability of very high resolution (VHR) satellite imagery, detection of the informal settlements can be further enhanced to provide the spatial information (Kohli, Warwadekar, Kerle, Sliuzas, & Stein, 2013). In addition, using high resolution remote sensing data can help to outline the informal settlements even in very complex regions and also to control fast growing cities (Netzband & Rahman, 2009). This makes it obvious that advanced remote sensing image analysis can contribute spatially the information of informal settlements faster as compared to the conventional census methods and participatory methods.

Though the crucial factor for housing of informal settlements is its location, the structure of informal settlements is highly flexible so as to circumvent the needs of the inhabitants (Hofmann, 2014). So, the housings required to be dynamic in nature so that it can move quickly or extend their housings if necessary. Apparently, this made a great influence on the location and characteristics of the informal settlements especially the overall configuration. However, informal settlements at the periphery changes faster as compared to the core areas (Kuffer, 2003). In some cases, even vertical growth can be noticed (Sliuzas & Kuffer, 2008, as cited in Hofmann, 2014). Ultimately, the conventional methods of detecting informal settlements stood just for the sake of existence but not for the welfare of informal settlements. So, consistent methods are required and can provide detail information in while collecting it at large areas.

Remote sensing methods should be integrating with contextual information to detect informal settlements at large areas in less time (Blaschke et al., 2014). In 2009, Ebert et al. (as cited in Kohli et al., 2013) stated that OBIA methods have the potential to capture the heterogeneity on ground by following a hierarchical procedure for object-classification and by including contextual information for objects and non-physical features for capturing the dynamic nature. In urban studies, this makes it possible to identify objects like buildings, roads and other anthropogenic features more accurately (Lang, 2008). Veljanovski, Kanjir, Pehani, Otir, & Kovai (2012) also suggested that the population dynamics can be monitored through object based image analysis (OBIA) using VHR imagery. Thus, it is evident that the OBIA is the optimal technique to detect and identify the informal settlements in a faster, robust and efficient way using the remote sensing data without accompanying the ground based mapping (Kuffer, 2003; Sliuzas & Kuffer, 2008).

The demand in mapping and monitoring informal settlements is very high (Kit et al., 2012; Sliuzas & Kuffer, 2008; Taubenböck & Kraff, 2014). Many individual approaches have been employed so far for detecting informal settlements from VHR imageries but there is no consensus on the best method to delineate or to analyse them from remote sensing data (Sliuzas & Kuffer, 2008). Hence, the individual characteristics of informal settlements were explained only by their respective approaches. This may result to different versions of identifying informal settlements which sooner or later lead to improper understanding the real phenomena of informal settlements. In a study, Kohli et al. (2012) introduced the general slum ontology (GSO) which defined the general characteristics of informal settlements and can be reapplied, adapted and extended for operational tasks. Without such a conceptualization (e.g. the GSO) description of informal settlements was relatively fuzzy (Hofmann, Blaschke, & Strobl, 2011), the local influence factors of the informal settlements were hardly addressed (Hofmann, 2014). The utility of such an conceptualisation (ontologies) that it can be used as input for the creation of an OBIA rule set Kohli et al. (2013).And rule sets are hardly robust in recent studies. In recent methods for automated detecting informal settlements from remote sensing data, there is still highly manual adaptation to achieve the local conditions (Hofmann, 2014). Mathenge (2011) recommended that there is a need to do research on the robustness of the rule set developed to another area. Thus there is still a need to analyse the robustness of the rule set of object based approach in the context of informal settlements using VHR imagery.

1.2. Research Problem

OBIA has many advantages in the field of analysing VHR Imagery as it functions on the image objects as a collection of pixels rather than on single pixels. So, speckles are avoided and huge feature space can be used for advanced object image analysis (Blaschke & Strobl, 2001). Furthermore, OBIA is an iterative process which functions with a random initial image segmentation and incorporated knowledge-based for improvement of image segments as per their own analysis (Baatz, Hoffmann, & Willhauck, 2008). As a result, the subsequent image objects do not focus on the accuracy but on the replica of the real objects that are to be identified. But the suitable scale factor remains unpredictable since the quality of the segmentation depends on the image object at local context. In the case of detecting informal settlements, the typical structural elements of informal settlements, such as small buildings, shadows of tall trees as well as small roads, need to be outlined well enough to indicate on a higher scale. However, Hofmann et al. (2011) mentioned that lesser the manual adaptions and interactions of the underlying rule sets, the more robustness will be the given rule sets. Consequently, more robust rule set will improve the automated detection of informal settlements.

Hofmann et al. (2011) developed one rule set with present knowledge and reapplied it to another city. On the other hand, we shall start the rule set with another city which was evaluated two years ago. Recently, additional indicators were successfully used to extract informal settlements in VHR imagery. For example, Hamedianfar & Shafri (2015) stressed the utility of spectral indices for differentiating roofs and roads. The rule set which we shall be taking as a base were adapted from the study of Kohli et al. (2013). Using spatial features, e.g. line feature, OBIA approach becomes more robust which helps in classifying between formal settlements and informal settlements due to the regularity of the pattern of buildings and roads.

OBIA was applied in many applications for extraction of information based on their own local frameworks of analysis. For instances, by Rizvi & Krishna Mohan (2012)stressed the difficulties in extracting information in a spectral and spatial heterogeneous urban landscapes through VHR imagery (case Mumbai, India) . Shekhar (2012) also applied OBIA to detect slums in Pune, India after separating all other classes from the rule sets. Recently, Kohli et al. (2012) addressed the GSO to develop the generic ontological framework and also provided a comprehensive description of spatial characteristics and their relationship from the image. However, not many studies have implemented the framework of GSO for the detection of informal settlements in various locations. However, only few researchers, for instance, Kohli et al. (2013) had analysed the transferability of OBIA using three different subsets of the same VHR imagery in Ahmedabad, India. Hofmann et al., (2011) also evaluated the robustness of rule sets in Cape Town, South Africa and identified the parameters to be changed for the extraction of informal settlements. Moreover, understanding the clear conceptualization of the informal settlements as an object of interest is a big challenge as informal settlements have diverse definitions and different appearances depending on their geographical location and local context (Kohli et al., 2013). This led to the research gap to challenge and practice more in different locations to make more reliable and transparent for supporting the automated detection of informal settlements. Hence this study is going to contribute to the understanding of robustness of rule set in OBIA for detecting informal settlements. In order to achieve this goal, the study will adapt the GSO to local context of informal settlements in Mumbai, India for extracting them using OBIA with VHR imagery. Moreover, it is vital to test its robustness rule set across different subsets.

1.3. Research Objective of the Study

The main objective and its research objectives are given with their respective questions of the study.

1.3.1. Main Objective

The main objective is to measure the robustness of the OBIA rule set for detecting informal settlements across different subsets using the case study of Mumbai (India).

1.3.2. Research Objectives and their respective Research Questions

- To study the local adaption of the GSO to Mumbai, an Indian mega-city by adding local knowledge.
  1. What are the spatial and contextual features of the informal settlements in the local context?
  2. What are the indicators that can be distinguished from VHR imagery?
- To develop the adapted rule set for detecting informal settlements based on the indicators identified from the local slum ontology.
  1. How is GSO adapted to local slum ontology to generate the indicators in local context?
  2. How are the indicators translated into image based parameters?
  3. What is the difference in accuracy between the results of classification quality of subset 1, subset 2 and subset 3?
- To explore the robustness of the set of parameters used in the OBIA rule set by applying it to other two subsets of VHR imagery.
  1. Which parameters of the rule set can be used to acquire similar results with no or little adaptation for detecting informal settlements?
  2. How are the stable parameters estimated in different subsets when the similar results are achieved?
  3. What are the result of the robustness of the developed rule set?

1.4. Conceptual Framework

This section discusses about the conceptual framework of the study (Figure 1) in relation to the research problem and it also highlight how the concepts and the links flow to accomplish the research under study. In this study, we focus on the measurement of robustness of OBIA rule set. How the concept of robustness arrived at a particular point is demonstrated with the support of three stages, namely, feature description of informal settlements, translation and development and image analysis. Each stage comprises of few concepts and link to one another to lead the adjacent stage. At the first stage of feature description of informal settlements, the author characterises the informal settlements at local context by adding the local knowledge to the concept of general ontology for slum. The second is translation of the local ontology for slum to develop a rule set in this stage. The result of the adapted rule set will be utilised for the OBIA approach for detecting the informal settlements in the third stage. The segmentation is the first process steps in OBIA. Then, the last but not least stage will deal with the measurement of robustness of the adapted rule set for OBIA approach using VHR imagery. However, in this study, the developed rule set will be tuning until the classification quality (q_i) in the test imagery (I_f) is at least similar as the classification quality (q_o) of the referenced imagery (I_o). Additionally, the measurement is mainly focussed on the rule set deviation of the relative change of the membership function of each classes.

In fact, informal settlements are highly dynamic and therefore, need to be measured regularly. Accordingly, the concept of the robustness of rule set has emerged to face the challenge of detection of informal settlements regularly by reapplying the same rule set in OBIA to a different area. To summarise the conceptual framework in short, local knowledge and the Generic Slum Ontology (GSO) will produce a local ontology that will lead to indicators and therefore rules for OBIA approach. Finally, the rule set will be evaluated through three steps of Hofmann et al., (2011) by comparing qi and q_o and relative deviation (d) of the membership function will be measured to contribute in the evaluation of the robustness measurement of the rule set if qi $\neq$ q_o.

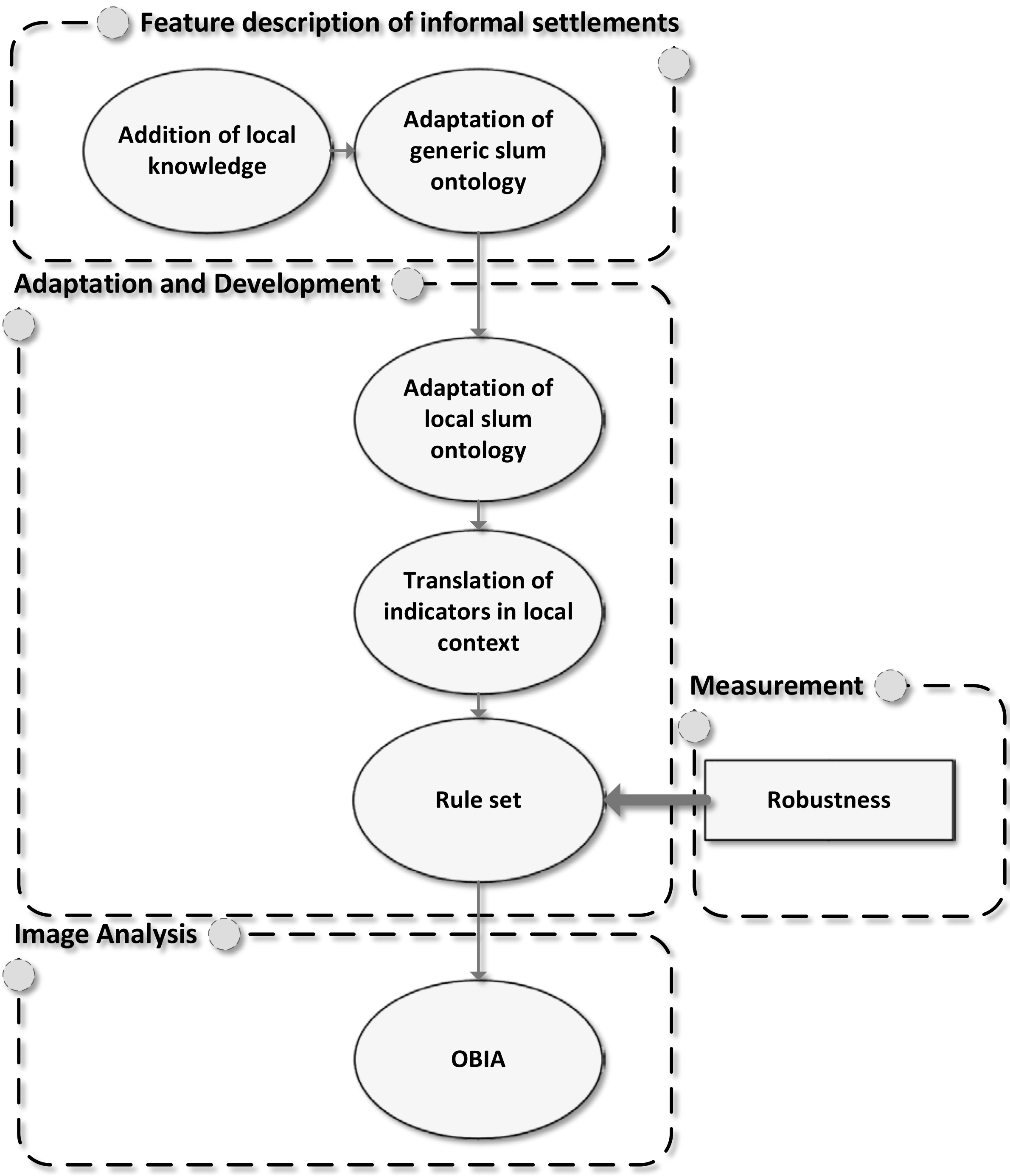


Figure 1: Conceptual framework for the study

1.5. Structure of Thesis

The structure of the thesis is elaborated below

Chapter 1 presented the research background and justification and talk about the research problem. It will also include the conceptualisation of the research from where the research objectives and the research questions are quantified. In this chapter, the main portion of the research approach would be discussed.

Chapter 2 reviewed the literature on the main concepts involved in the study. It mainly focussed on the concepts of the informal settlement in general and local contexts, ontology, OBIA and robustness.

Chapter 3 described the study area, data requirement, the scope and limitations and the methodological framework and its methods would be elaborated in details.

Chapter 4 displayed the results of segmentation, image analysis and adaptation of developed rule set and its deviation. It elaborated the processes with respect to their research questions. In this chapter, the main part of the study included the identification of the general slum ontology based on the morphological features, translation of image based parameters, and the result of the image analysis and robustness measurement.

Chapter 5 discussed the combined effects of the results from the chapter 5 by research questions.

Chapter 6 concluded the achievements based on the research objectives of feature characterisation of informal settlements, development of rule sets and robustness measurement.

2. LITERATURE REVIEW

In this chapter, the conceptual dimensions of the study required for the research and the structure of the concepts are discussed.

2.1. Concept of Informal Settlements

Until now, there is no unique definition of informal settlements. And the concept of informal settlements differs from place to place. Iinformal settlements were defined as illegal settlements which are situated either at the public or private land in a haphazard manner without planning regulations and adequate basic amenities (Ishtiyaq & Kumar, 2011). However, in the beginning of 21st century, UN-Habitat (2003) developed a definition of the term 'slum', as urban area which lacks one or more of the following conditions: access to safe water, access to acceptable sanitation, durable houses, sufficient living space and security of tenure.

In the Indian context, under the section 3 in the Slum Areas (Improvement and Clearance) Act, was setup in 1956, slum areas are defined as areas where "the buildings are unfit for human habitation or are by cause of dilapidation, overcrowding, faulty arrangement, and design of buildings, narrowness or faulty arrangement of streets, lack of ventilation, light or sanitation facilities, or any combination of these factors, are harmful to safety, health or morals" (Census of India, 2013, p. 2). Then, the Census of India had reformulated the definition of slums in 2001 based on the same idea without considering the population size (Census of India, 2013, p. 4). Subsequently, a topology of slums merged in the Census, specifically, Notified, Recognised and Identified (Census of India, 2013, p. 4). A notified slum is a slum acquainted by Sate, Union Territories Administration or Local Government, Housing and Slum Boards but not formally notified as slum under any act (Census of India, 2013, p. 5). A recognised slum is the one notified by any level Government authorities under any Act whereas a recognised slum was not under any Act (Census of India, 2013, p. 5). The last but least, an identified slum is comprised of an area where at least 300 people or 60-70 households in overcrowded dwellings settle in a area deprived from hygienic environment, infrastructure, proper sanitary and water facilities and the areas is recorded in the charge register by an officer from Directorate of Census Operations (Census of India, 2013, p. 5).

Though informal settlements and slums are not synonymous to each other, they share many characteristics. However, a slum is informally originated either from an informal settlement or a formal settlement and came to exist as a slum illegally after the long period of time if not recorded by the concern authorities. Thus, all slums would not be defined as informal settlements but all informal settlements can be of slums because acoording to , all informal. Since informal settlement and slum are often used interchangeably, the author uses the word, "informal settlements" for this study.

2.1.1. Dynamic of Informal Settlements

Within the concept of informal settlements, one of the most emerging effects is the dynamic of the informal settlements in order to adapt the condition of the situation. For instance, due to the overcrowding, the households need to increase more rooms in order to adjust their capacity of living space. The changes mostly occurred at the periphery rather than the core of the informal settlements (Hofmann et al., 2011). A good understanding of dynamics of informal settlements is necessary to detect then in VHR imagery (Hofmann, Strobl, Blaschke, & Kux, 2008). Furthermore, transferring the methods is one of the challenging tasks in Earth Observation (EO). However, the EO methods should be transferable to meet the need of the observation of automated methods in an effective and efficient manner (Sliuzas, Mboup, & de Sherbinin, 2008).

2.2. Concept of OBIA

Object Based Image Analysis (OBIA) is an iterative process in which segmentation is the initial process then followed by classification using the rule set generated through local knowledge in terms of spatial, spectral and contextual characteristics (Baatz et al., 2008). Furthermore, OBIA goes for the aggregate of pixels rather than being pixel based. Moreover, OBIA is useful for extracting features with similar spectral reflectance and is capable of developing semantic information particularly in a complex urban context. In addition, OBIA is capable to export the result in vector format. Lang & Blaschke (2006), added another advantages of OBIA, it plays a vital role in bridging between the remote sensing and Geographical Information System (GIS).

Many acronyms of OBIA have been emerged in last decade. The term, "Object Oriented Image Analysis (OOA)" are used e.g. by Mathenge (2011) for the research of the application of object oriented image analysis in slum identification; Kohli et al. (2012) for slum ontology for image-based classification; and Martha, Kerle, van Westen, Jetten, & Kumar (2011) for segment optimization and data-driven thresholding for knowledge-based landslide detection. Some of the researchers used another term, Geographic Object-based Image Analysis (GEOBIA) by Ardila, Bijker, Tolpekin, & Stein (2012) in the extraction of tree crown objects in urban areas using VHR satellite images; and Blaschke et al. (2014) for the emergence of geographic object-based image analysis - towards a new paradigm. Consequently, GEOBIA seems to be represented as a sub-discipline of GIScience whereas OBIA is almost too comprehensive (Blaschke, 2010).

2.2.1. Segmentation

Segmentation is the first step of the OBIA where the partition of an image takes place in order to generate segments or regions representing ground objects(Trimble, 2015a). Multi-resolution segmentation is a method where the approach begins with bottom-up of segmentation on the basis of a pairwise region merging technique (Benz, Hofmann, Willhauck, Lingenfelder, & Heynen, 2004). The translated local ontology would be used for segmentation of the imagery to obtain the image object. Though the segmentation influences a lot the quality of the result due to heterogeneity of the image pixels, the focus of the study is not on segmentation. Thus, the segmentation will be done by using the estimated scale parameter 2 (ESP2) tool generated by Drӑguţ, Tiede, & Levick (2010) for the estimation of the scale parameter of the segmentation. It will help in saving time and produce the appropriate one for using OBIA approaches. The tool measures the changing aspects of local variance (LV) from an object level to another so that the suitable scale parameter can be genarated. According to Drӑguţ, Tiede, & Levick (2010), the equation of rate of change (ROC) is given below.

$$ROC = \left[\frac{L - (L - 1)}{L - 1}\right] * 100$$

where L = LV at target level and $L - 1$ = LV at next lower level.

The final result of the scale parameter is produced by the ESP tool using the equation. The segmented objects build the imput for the classification in a next step.

2.2.2. Classification

Besides using only spectral features like pixel based classicisation, the interesting fact of OBIA is that it is totally based on the image objects and classifies them by the features of shape, size, texture, context using information of the local knowledge.

According to Trimble, (2015b), the image object generated from the segmentation are classified and refined by the knowledge of the intrinsic features, topology and context. The intrinsic features can be determined by the physical characteristics of the image objects, for instance, colour, texture and form of the objects. The topological features can be determined by the geometric relationships or surroundings of the image objects. For example, distance to a certain area or being in a certain area with the image. The context features can be determined by the symbolic relationships of the image objects among each other. For example, informal settlements are mostly located by the low-lying areas which can be used as the indicator for classifying informal settlements.

2.2.2.1. Fuzzy systems

Fuzzy classification is the process which plays a key role in classifying the fuzziness among classes. It involves three steps, namely, fuzzification, fuzzy rule base classification and defuzzification (Trimble, 2015b). The unique feature about the fuzzy classification is that the fuzzification assigns and allows image objects to have a membership of more than one class but varying the degree of the membership function. A fuzzy rule can have one single condition or several conditions for assigning objects to a specific class to maintain the hierarchal classification (Trimble, 2015b). This expresses the fuzziness of the uncertainties of the class description. Likewise, the local knowledge from local experts could be summarized by incorporating different feedbacks in the same description.
The first part, fuzzification, defines the transition from a crispy value range to a fuzzy membership value. Each membership value ranges from 0 to 1. The membership function defines the membership value of each parameter used in the classification. Since the informal settlements are dynamic in nature, fuzzy certainty measure of the fuzzy-logic approach quantify the robustness of the rule set as fuzzy sets because the fuzzy logic approach evaluated the understanding that was estimated rather than specific (Hofmann et al., 2011). Defuzzification is the technique in which the results of multi-valued logic variables of each parameter are quantified to get common crisp result for measurement of robustness of the adapted rule set (Zadeh, 1965). So, the defuzzification can be utilised to generate the overall results for each subset.

2.2.2.2. Utilization of spectral indices

One of the emerging parameters which was very useful for classifying the roofs of the buildings from the roads is the utilisation of spectral indices using coastal band and NIR1 band of WorldView-2 satellite imagery practised by Hamedianfar & Shafri (2015). The parameters is effective for increasing the quality of the classification in OBIA. Moreover, due to variation of local definition of informal settlements from place to place, it would be helpful to use these spectral indices as an additional indicators for classifying the particular features.

2.3. Concept of Ontology

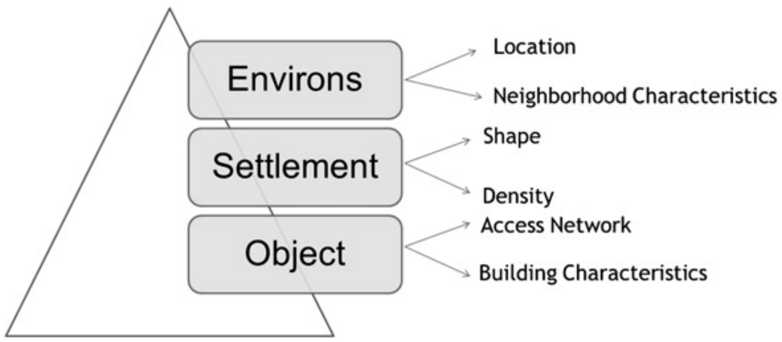


Figure 2: Representation of concepts with a hierarchy of respective general indicators at three spatial levels (figure adopted from Kohli et al., 2012)

An ontology, derived from the field of philosophy (Deved¿ic, Djuric, & Ga¿evic, 2009), is a knowledge representation about a concept which deals with the study of integration and their relationships into

coherent process to detect complex concepts (Lüscher, Weibel, & Mackaness, 2008). Hofmann et al. (2008) quoted that there are two domains in the image analysis, namely, image domain and real world domain. The image domain defines the general characteristics whereas the real world domain defines the properties that can be detected from the image. Since there is no specific term that defines informal settlements, the parameter settings are always required to be tuned (Sliuzas et al., 2008). Thus, Kohli et al. (2012) had developed the generic slum ontology for the conceptualisation at local context as shown in the figure 1.
In the figure 1, the hierarchy of the general concept at three spatial levels are shown with their respective indicators. According to Kohli et al. (2012), the characterisation of informal settlements originated from the object level to the environs level. Though the generic slum ontology need to be applied to the local context for extracting informal settlements, it plays an important role of knowledge integration into image based parameters in the OBIA using VHR imagery. It gives a formulized characterisation of informal settlements which can be translated to the local context. Thus, local knowledge is needed in the local ontology to be translated and reapplied in a rule set (Blaschke et al., 2014).

2.3.1. Generic Slum Ontology

The methontology is a method to develop the ontology for a specific local context (Lopez, Gomez-Perez, Sierra, & Sierra, 1999). It comprised of three phases, namely, specification, conceptualization and implementation. The first phase, specification, deals with the conversion of the general characteristics of the informal settlements into the specific features of the informal settlements at local context by incorporating the local knowledge from local experts. For example, maps of informal settlements drawn by local experts can be used to specify the informal settlements. The second phase, conceptualization, arises to develop the feature description of the informal settlements at the local context to define the specific informal settlement. The last phase, implementation, is the local adaptation of generic slum ontology to deal with interpreting the local indicators to physical factors that can be translated into image based parameters to develop the rule set and to proceed the classification in OBIA.

2.3.2. Local Slum Ontology (LSO)

The local ontology is the concept and their relationships of physical features of informal settlements after adaptation of GSO at the local context in order to translate the image based parameters of OBIA (Kohli et al., 2012). A LSO always needed to be developed since the generic slum ontology cannot be used for translating the image based parameters for classification. Hence, the indicators always need to address in the local context, more specifically the local characteristics of the informal settlements.

2.4. Concept of Rule Set

According to Trimble (2015), the rule set is defined as a set of parameters where the information of different characteristics are written to the particular object in the image based classification. In fuzzy systems, fuzzy membership function helps in assigning the rule in each class feature to classify the image. Moreover, the purpose of the rule set is to set the rules and to classify the desired features in the image analysis as accurate as possible. The rule to classify each feature might have one or more class descriptions. For instance, Hamedianfar & Shafri (2015) used the class descriptions of Normal Difference Vegetation Index (NDVI) in the rule set to classify grass and trees. Incorporation of thematic layers in the rule set improve the quality of segmentation and classification of OBIA. (Sori, 2012).

2.5. Concept of Robustness

In the context of OBIA, rule sets are robustness if they generate similar results with similar quality on similar images with minimum adaptation effort (Hofmann et al., 2011). A ruleset is easily transferable if it requires minimal manual adaptations to provide comparable results for different imaging conditions

(Kohli et al., 2013). According to Hofmann (2014), the term robustness can be quantified by two determining factors such as classification quality and deviation of the rule set. Four deviations of the rule set are as follows:

a) Addition or removal of classes

This factor considers to evualate when class features in the developed rule set are not equal to the number of class features in the adapted rule set. For example, removal of vegetation class in the classification when reapplied with adapted rule sets.

b) Changing logical operators

This factor measures the degree of change of the logic operator connected with one or more different parameters in a class feature. For example, "OR" logic operator changes to "AND" operator to consider the classification allows to use parameter when the class feature is avaibale and vice versa.

c) Changing relational operators

This factor quanitfies the number of change of relational operater when there is one or more terms in the class description of the class feature. For example, changing the operator of "+" to "-"

d) Changing the thresholds

This factor measures the deviation of relative change of the membership function for each parameter. For example, shifting the central values or expansion of the value range in the membership function.

2.5.1. Comparison of data conditions from different researchers

The comparison, shown in the table, brought a lot of contribution for the selection of the category of the image based parameters for classification. And it helps in strengthening the local knowledge by supporting the spectral features but not only morphological characteristics of the informal settlements.

Considerations	Divyani Kohli's research work	Peter Hoffman's research work	Alireza Hamedianfar's research work	Inferences
Site area	Ahmedabad, India	Cape Town, South Africa and Rio de Janeiro, Brazil	Kuala Lumpur, Malaysia	Contrast of formal and informal were observed
Site characteristics	Slums located mainly along the river and mainly close to factories	Quite contrast, no mix up of land cover of built-up areas	Metal roofs mostly found in industrial buildings and concrete tile and asbestos roofs in residential buildings.	Rule set developed on the basis of all subsets characteristics.
Sensors	GeoEye -1 imagery	IKONOS and Quickbird imageries	WorldView-2 imagery	Similar imagery recommended spectral index features
Methods	Qualitative approach	Quantitative approach	Qualitative approach, images are used from same imagery	Quantitative approach was recommended for quality research
Segmentation type	Multiresolution	Multiresolution	Multiresolution	Multiresolution
Imagery sources applied for rule sets	Same source	Different sources	Same source	Same source provide less reliability.
Purpose	Identifying slums	Identifying informal settlements	Detailing types and coverage of intra-urban features	Parameters can be adapted for rule set development
Main OBIA features used in rule sets	Spectral range, geometry, texture and association	Geometry, association, spectral range	Spatial, spectral, and textural features and even several spectral indices	More focus on features of same sensor
Quality	47% to 68% in pixel accuracy	0.14 in area based	88%, 88% and 86% in pixel accuracy	Relatively low as a whole

Table 1: Comparison of data conditions from different researchers

2.6. Summary

In this chapter, the concepts used in the study were discussed in a manner that they all are linked to the methods in the next chapter. Moreover, there were more image based parameters that had been found from the literature review also discussed in this chapter.

3. DATA AND METHODOLOGY

3.1. Study Area

The study was conducted in Mumbai city which is located at the western part of India with the urban population of 18,394,912 as per Indian Census (2011). The city area covers the airport area with large areas of the informal settlements as shown in the figure 4. According to Kuffer et al. (2013), the physical features of informal settlements in Mumbai are characterised as small building size, clustering of small buildings, organic pattern, different roof materials and close to transport infrastructures. More or less vegetation is are widely spread throughout the image shown in the figure 4. Three subset images are chosen to reflect different type of areas having different morphological and patterns, for instance, variability of shape and configuration.

- Subset 1 is the west-southern part of the image from the airport area and covers vegetation, river, buildings and informal settlements loacted in Dharavi. The area is of 1.90 square kilometres (subset area 1532.5 metres by 1237 metres).
- Subset 2 was of area of 1.15 square kilometres, in which the subset has informal settlements,and majority of the area are of formal settlements and the location consists of Chunabhatti (North), Swadehi Mill, Kurla (South), Qureshi Nagar, Samrat Ashok Marg (Near Eastern Express Highway/Agra-Mumbai Road (National Highway 3)) (1152 metres by 994 metres).
- Subset 3 is part of the outer city and contains a mixture of vegetation, buildings and slums with slums located mainly close to old factories. The location consists of Khar Railway Colony, Nirmal Nagar, Golibar slums, Government Colony Bandra (Northeastern part), Western Express Highway with area of 0.80 square kilometre (1000 metres by 800 metres).

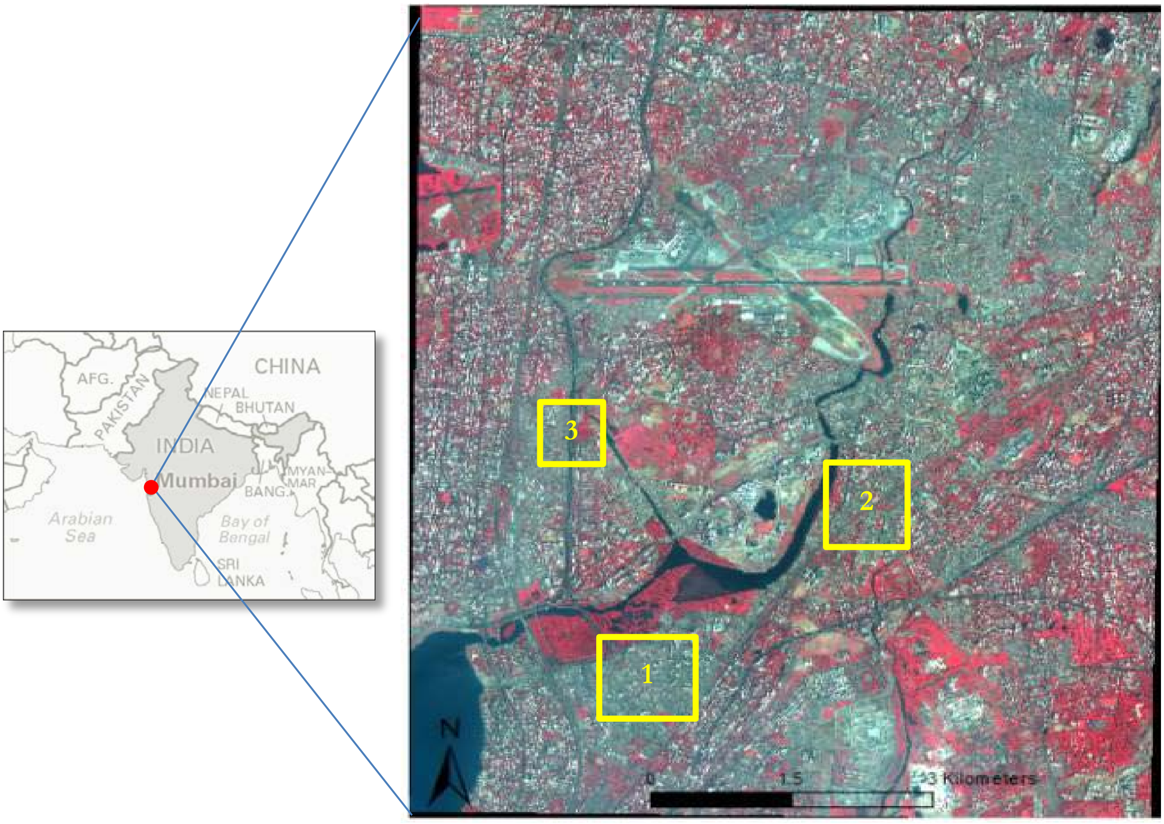


Figure 3: The central part of Mumbai provided by DigitalGlobe, 2009. Three subsets of the study area are shown inside the yellow box of the right image.

Small subsets were taken to analyse the process of OBIA in faster way and also chosen the different scene. Moreover, the area of the informal settlements are mostly clustered which makes the image analysis difficult to extract small roads within settlements.

3.2. Data and Software Requirement and Acquisition

In this study, the VHR imagery of Worldview-2 which was acquired on 2009 with a resolution of 0.5 m panchromatic and 2 m multispectral were used for OBIA approach to test the robustness of the rule sets. The pre-processing correction of the imagery used for the analysis was done by the image supplier. The merge of the resolution was performed to produce a pan-sharpened image through the principal component analysis with a 0.5 m resolution so as to implement the data fusion without losing the spectral characteristics of the imagery (Lu & Weng, 2005). There were few ground truth data which was limited for the verification of the image analysis. For the data collection, the paper map was used for collecting the ground points with images. All the processes will be used in the software, eCognition Developer. Charts and tables were generated through Microsoft Excel.

In image analysis, misclassification of the target objects leads to the initialisation of fieldwork. However, in Mumbai, access to the informal settlements are not an easy task due to the complexity of narrow roads, high density of settlements and close to hazardous areas. In the fieldwork, a number of 150 ground control points was collected within the subsets of VHR imagery as shown in the table 3 where the sample are selected by quota sampling. However, more the number of sample points, the quality of analysis would be better. Since the data was relevant for the accuracy assessment of the quality, the fieldwork were taken place in second two weeks of October, 2015 for 10 days. In short, the descriptions of the data requirements in the study were given below in the table 2.

Data	Year/Configuration	Source	Remarks/Status
WorldView-2 imagery (WV-2 satellite imagery)	2009 (2 metres in RGB* and 0.5 metre in panchromatic band)	DigitalGlobe	VHR Imagery
GIS layers	2015	Mapzen	Roads, buildings and water.
DEM	2015	USGS	
150 ground points	Second two weeks of October, 2015 for 10 days	Primary data	Were collected in fieldwork

Table 2: Descriptions of data acquisitions

*RGB = Red band, green band and blue band

And we added the remaining 175 points manually in order to assess the classification of all classes. Moreover, we assumed in the study that no discrepancy would not influence the quality of assessment due to the variability of time of data collection with that of the data analysis.

3.2.1. Beneficial Features of WorldView-2 Imagery

According to DigitalGlobe (2010), the satellite WorldView-2 used in this study was launched in 8 October, 2009. It provides 8-band. The sun-synchronous satellite operates at an altitude of 770 km with an average revisit time of 1.1 days. It also provided very high resolution of 0.46 m panchromatic resolution and 1.85 m multispectral resolution respectively and was also capable of collecting up to 1 million km^2 in area per day.

Thus WorldView-2 imagery is a very high resolution sensor having 4 new bands namely, coastal, yellow, red edge and NIR2 (DigitalGlobe, 2010). Such new color bands might be useful for generating and adapting the spectral indexes for better identifying commonly confused classes. According to DigitalGlobe

(2010), the dyanamic range is 11–bits per pixel and the specification of the sensor bands were shown in the table 2.

Sensor bands	Wavelength (in nanometers)
Panchromatic	450 – 800
Coastal	400 – 450
Blue	450 – 510
Green	510 – 580
Yellow	580 – 625
Red	625 – 690
Red Edge	705 – 745
NIR1	770 – 895
NIR2	860 – 1040

Table 3: Descriptions of sensor bands of WorldView-2 imagery

Since there were few more colour bands mentioned in the table 4 of the sensor bands descriptions, these bands could be used to generate new spectral indexes for distinguishing two confusing classes, for instance, roads and roofs. This index suggested by Hamedianfar & Shafri (2015) to distinguish roads and roofs by uses the spectral indexes of NIR2 and coastal bands.

3.2.2. DEM & GIS layers

DEM (as illustrated in the annexure 12, annexure 13 and annexure 14) was downloaded from the Earth Explorer[1] for the study area of Mumbai to analyse whether low-lying areas played a key role for detecting informal settlements. Thus low-lying areas can be used as an indicator in the rule sets as discussed in the section 2.4.

3.2.3. Software limitations

eCognition (formally known as Definiens) mainly used in this study for object based classification. Furthermore, ENVI software was used for extracting the feature of GLCM, and ArcGIS software was for converting roads and water to raster. While generating the GLCM, the window size was selected by visualisation as shown in the annexure 20. The first and foremost step to start the classification was to segment the object and then develop the rules to classify image objects (Trimble, 2015b). The main drawback of the software was that is took a lot of time while running the process of segmentation and fuzzy classification.

3.3. Methodology of the Study

The first step was to pre-process the WorldView-2 imagery for radiometric calibration and atmospheric correction before using them. Moreover, in order to detect the informal settlement, we needed to determine the definition of informal settlements at local context and develop a strategy for extracting local expert knowledge of the informal settlements at local context.

3.3.1. Feature description of informal settlements

As discussed in the section 2.1.1 that every informal settlement has its own dynamic nature and different characteristics from others. In the study, local ontology described the local features of informal settlements mainly based on the morphological features through the questionnaire given in the annexure 6. Since there is fieldwork only for the ground point collection, the local knowledge was acquired through the questionnaire conducted. Furthermore, the visual interpretation of the imagery given the questionnaire

[1] Earth Explorer, "n19_e072_1arc_v3", "SRTM v3" [http://earthexplorer.usgs.gov/download/options/8360/SRTM1N19E072V3/], 19 July 2015

through local experts also provide more information to the local knowledge by drawing informal settlements as examples for all subsets as well as by mentioning the physical factors to capture informal settlements in all subsets.

For the feature characterisation of informal settlements, the first phase of the methontology was elaborated below to show how to meet the research objective 1.

According to Kuffer et al. (2013), informal settlements in Mumbai can be divided into five categories which were given in the annexure 5. The physical characteristics of the informal settlements had been used as the benchmark to identify as the informal settlements during the fieldwork for collecting the ground truth. And the remaining built-up areas are considered as the formal settlements.

To define local characteristics of informal settlements in Mumbai, we developed a questionnaire for local experts on informal settlements in Mumbai. The questionnaire was sent to six local experts via email (see annex 6). The questionnaire purposely focussed on extracting local knowledge of the informal settlements irrespective of the nationality of the expert. However, the experts should had experience before in remote sensing and urban planning projects in Mumbai. With the help of a literature review and the questionnaire provided by Kohli et al. (2012), the indicators were collected and to extract the local ontology at three levels from local experts. It had been made in a semi-structured way so that the local experts can give their opinions on questions. Besides, those indicators which were not expressed during the questionnaire were to be considered as the limitation. So, the outcome of the questionnaire could be used to some extent by discussing with other characteristics of the WV-2 imagery because those inputs were resulted from individual experts which might cause biased. Moreover, the information from the questionnaire was used for the local adaptation of GSO for translating into image based parameters.

As discussed in the section 2.3, the strategy for eliciting local knowledge was based on the three processes of methontology suggested by Lopez et al. (1999). Those steps were specification, conceptualisation and implementation. The result provided answers to first part of the research objective 2 i.e. adaptation of ontology.

The step focussed on the determination of the set of terms, its scope and level. In this phase, the feature characterisation of informal settlements was the focus and was as the basis for developing the ontology. Hence, the information of specific characteristics of the informal settlements was collected from local experts on the basis of the GSO to have a general overview of the informal settlements at local context.

3.3.1.1. Informal settlements mapped by local experts

After gathering from the local knowledge from local experts, physical factors were extract from the questionnaires to conceptualise the main indicators. The main indicators contributed to the implement for the local adaptation of GSO into image based parameters.

The general slum ontology includes the information of informal settlements from different contexts. In order to it break down into the local context, it is necessary to conceptualise by determining the indicators which are important for the local context. The local characteristics of informal settlements from the experts were organized according to the six general indicators given in GSO that are grouped in three spatial levels (i.e. object level, settlement and environs level). Inevitably, the indicators need to be parameterised into image based parameters.

For the translation of the ontology and the development of the rule set, the second phase and third phase of the methontology was involved as elaborated below to achieve the research objective 2.

The environs (neighbour) relationship would be given less priority because was not suited to the feature characteristic of informal settlements of subset 1 while adapting the local knowledge. From the scientific article of Kohli et al. (2013), the rule set which was utilised in Ahmedabad was used for the development of the first indicators since the study areas was also an Indian city of Ahmedabad. Additionally, some new indicators were added and some poorly-performing indicators were removed to enhance the quality of classification result.

In the last phase of the methontology, the process of implementation would guide the general ontology to incorporate the local context. The local indicators of informal settlements compiled in the process of conceptualisation were utilised to initiate the development of local ontology. The local indicators were arranged in an order such that each indicator would be assigned to suit the particular image based parameter. A developed rule set that only included successful parameters was used for the image classification. Those successful parameters were be supported by specific image characteristics given by the local experts.

3.3.2. Image Analysis

For the image analysis, the softwares of eCognition, ArcGIS and ENVI were used. The outcome of the image based parameters were used for the image analysis. These segmented objects were used to assign the particular class features according to the rule sets. Here, the classification was used by inserting membership functions with the help of class description in each class feature. The rational was that this membership function allowed to integrate the local ontology based on local knowledge. The result of the image classification quality was used to judge the changes between developed rule set and adapted rule sets. The degree of resistant change without changing the initial parameter could provide the measurement of the robustness of the rule sets. As discussed in chapter 2.5, the classification quality should be similar to conclude the robustness of rule sets. If not, deviation was considered to quantify the measurement.

For the translation of ontology and the development of rule set, image analysis was done in two steps, segmentation and OBIA approach. The pre-processing was done followed by two steps which are elaborated as follows:

Data preparation had been done assisted by roads from OpenStreet map (OSM) and then subsets were selected from WV2 imagery. Atmospheric correction was conducted to ensure that the analysis was not much influenced by difference in atmospheric conditions. And pan-sharpening had also applied to increase the spatial resolution of the imagery. Steps were mentioned below: Radiometric calibration of the raw data (VHR imagery) was conducted for the top of atmosphere (TOA). And then FLAASH of ENVI software was applied for the surface reflectance of the atmospheric correction. The procedure for pre-processing was followed through the user guide of the ENVI software, (2009). While preparing road data, steps taken up for data preparation in ArcGIS Desktop 10 were as follows: minimum bounding box of the clipped road as per each subset and then adding a new filed in the attribute table to identify if the pixel is road or not. Finally, union was used to combine all the features together to form the road features for each subset which would be used for developing the rule sets as well as classifying the particular features.

Before segmentation, the estimated scale parameter tool for multiresolution segmentation was used to generate meaningful segments. The tool of automation segmentation v.0 tool for multiresolution was acquired from Chandi Witharana[2]. The shape and compactness of the scale parameter are inserted by trial and error. The multiresolution segmentation was used for the image segmentation for all subsets of WV-2 imagery. The rational is that the multiresolution segmentation is a bottom-up segmentation which can be execute on an existing image object level or a pixel level for developing a new image object on a new image object level. This fitted the purpose of the study. The obtained image objects were used for the image analysis and further assigned a class description based on the spectral, colour, indices, spatial and contextual features. Thus a rule set was developed.

In order to measure the segmentation quality of the study, we used the location discrepancy of a segmented object to a referenced object. The measure used was the locational based metric (D_{sr}) (Montaghi, Larsen, & Greve 2013). The rational why the location based approach was used but not area based approach was that there was a high chance to mislead to the cause of superimposing more

[2] Chandi Witharana, "Segmentation Automation ver 1" in "eCognition Community""
[http://community.ecognition.com/home/SegmentationAutomation_ver_1.dcp/] 17 December, 2015

segments by taking bigger reference polygons or smaller scales parameter developed in the segmentation. Forty image objects were manually delineated for the segmentation quality assessment from each subset. Image objects were selected from each subset in such a way that formal buildings, vegetation and informal settlements taken in total of 40 polygons. Additionally, more segments were chosen from the formal building footprints because they had a clearer boundary after segmentation.

$$D_{sr} = \frac{1}{n}\sum_{i=0}^{n}\sqrt{\left(X_{s(i)} - X_r\right)^2 + \left(Y_{s(i)} - Y_r\right)^2} \quad (1)$$

where D_{sr} is the average of the distance in the Euclidian plane between the centroid coordinates of the i^{th} segmented object ($X_{s(i)}$ and $Y_{s(i)}$) and the centroid coordinates of the reference object (X_r and Y_r)

When D_{sr} tends to be zero, the quality of the segmentation increases whereas increase of the D_{sr} values represented the under- and over-segmentation of the object generated.

The reference polygons were either delineated manually and few were taken from the open street maps[3]. These references objects were used to check the quality of segmentation. This was done by superimposing one above another by means of intersection in ArcGIS. Then, we generated the near table for the centroid distances between referenced objects and segmented objects to calculate the quality of the segmentation with the help of the equation 1. So, the more the distance between them, the lesser will be the quality. Additionally, there are no discrete rules or criteria so far for the quality threshold to justify how good the quality is. Thus, the result of the image segmentation would be taken for further processes of the image analysis of OBIA approach.

[3]Mapzen, "Metro Extracts: City-sized portions of OpenStreetMap, served weekly" [https://mapzen.com/data/metro-extracts], 9 October, 2015

The image based parameters were used for the development of the initial rule set to perform the classification of the referenced imagery (subset 1). Each parameter was assigned with threshold values in the membership function of the class description. The class description was used for the classification of the image object after the process of segmentation. Using the rule sets, we executed OBIA by using the initial rule set for subset 1 as well as adapted rule sets for subset 2 and subset 3. Rule sets for subset 2 and subset 3 were adapted until an achievable classification quality of referenced imagery was achieved.

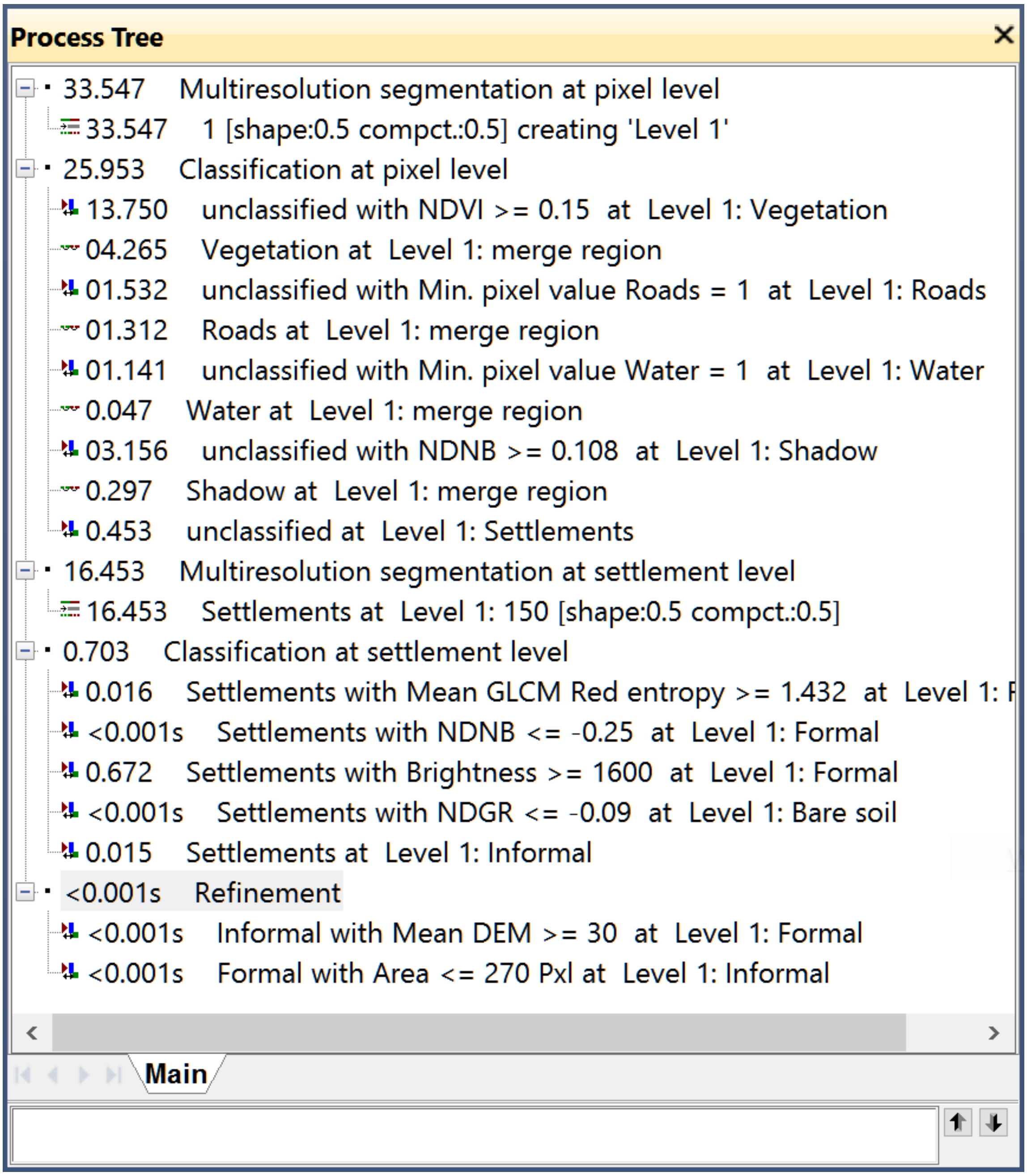


Figure 4: Process tree of the developed rule set

Thus three subsets were used, in subset 1 and subset 2, seven classes i.e. vegetation, bare soil, roads, formal settlements, water, shadow and informal settlements were classified whereas six classes were classified in subset 3 because the subset 3 did not have the class feature of water. As input, the segmented objects at pixel level were used for the classification. Eventually, class features of vegetation, roads and water were classified first and then merged to the unclassified objects for further processes. Again, the

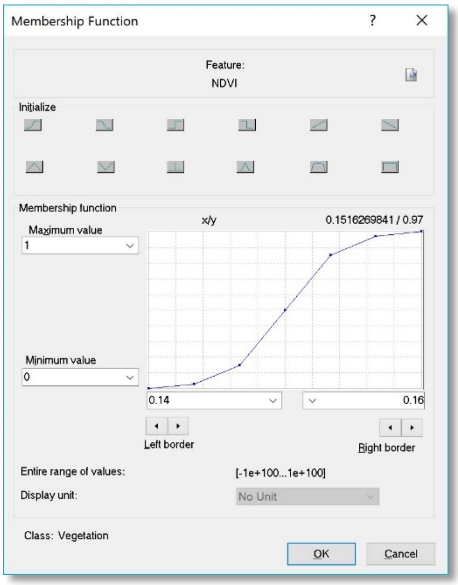


Figure 5: Diagram illustrating how the membership values are assigned for NDVI

unclassified settlements were merged and convert into settlement level to classify the formal settlements first. Thereafter, bare soil was classified. Once all other classes were classified, the remaining objects were assigned to the class informal settlements.

The rational of the processes that had been used in the study was to contribute to the quantitative results in defining the measurement of the robustness. Furthermore, the fuzzy rule sets are used to measure the reliability of the classification and to give the membership values to evaluate each object of a class, no class or several classes of each parameter. To quantify the value of the relative deviation of feature parameters, a curve so called membership function as shown in the figure 5 had been assigned to each parameter with the membership value through the class description. The type of membership function were provided in the annex 9, 10 ad 11.

In the figure 5, image object would be assigned to the desired features with the help of class description where each feature consists of at least one membership function defining the degree of membership (μ) in the range $0\geq \mu\geq 1$. As illustrated in the figure 5, the membership function for each class were assigned in the class description. To assign the membership value for the parameter, different types of membership function were used to define different value distributions and class transitions. Hence, the system provided support for the classification as well as measurement of the robustness. Specific types of membership functions used in the study were depicted in the table 4.

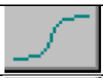
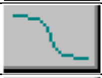
Type of membership functions	Function form
	Larger than
	Smaller than
	Full range

Table 4: Type of membership function used in the study

Feature classes of the classification

Features	Parameters	Formulas or Category
Bare soil	NDNB	$\frac{Green - Red\ edge}{Green + Red\ edge}$
Informal area	Area	Geometry-
Formal area	$GLCM_{red\ entropy}$, NDNB, Brightness	Spatial, Texture, Spectral index
Vegetation	NDVI	$\frac{Near\ InfraRed - Red}{Near\ InfraRed + Red}$
Water	Pixel based	Spatial
Shadow	NDNB	$\frac{Blue - Near\ InfraRed}{Blue + Near\ InfraRed}$
Roads	Pixel based	Spatial

Table 5: Feature descriptions of the tools used in the software (in the world of image analysis)

Parameters	Functions
NDNB	Classifying the bare soil from the formal area as well as classifying the shadow effectively
Brightness	Used in the distinguishing the formal area from the settlements
$GLCM_{red\ entropy}$	Aiding in classifying the formal area from the informal area by density of the reflectance
NDVI	Used in classifying the vegetation
Pixel based	Few classes were used as parameters through GIS layers like roads and water.

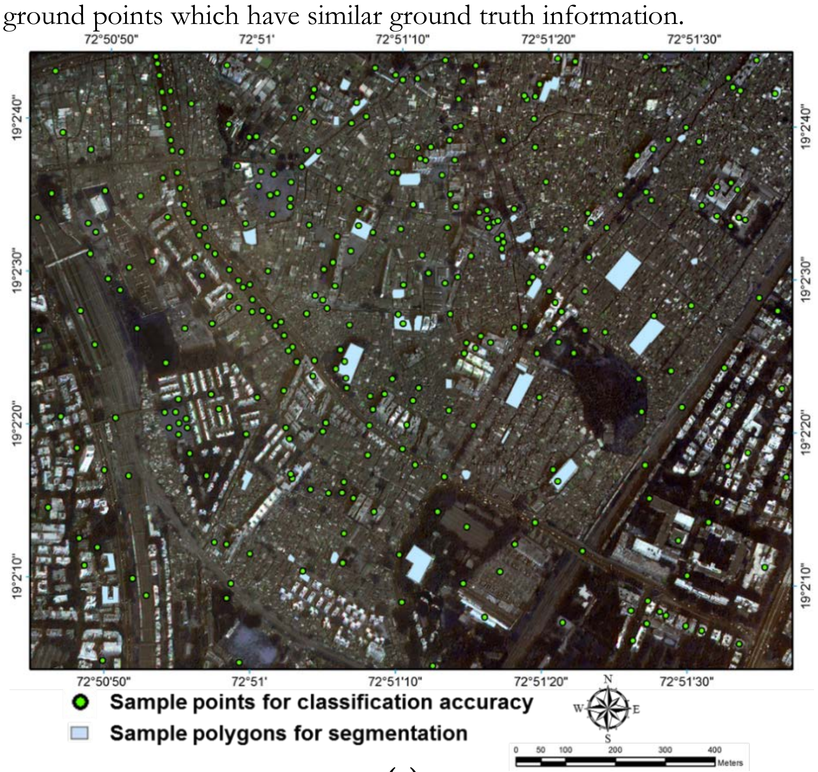
Table 6: Function of image based parameters

Assembling other features related to the textures and locational relativity for classification

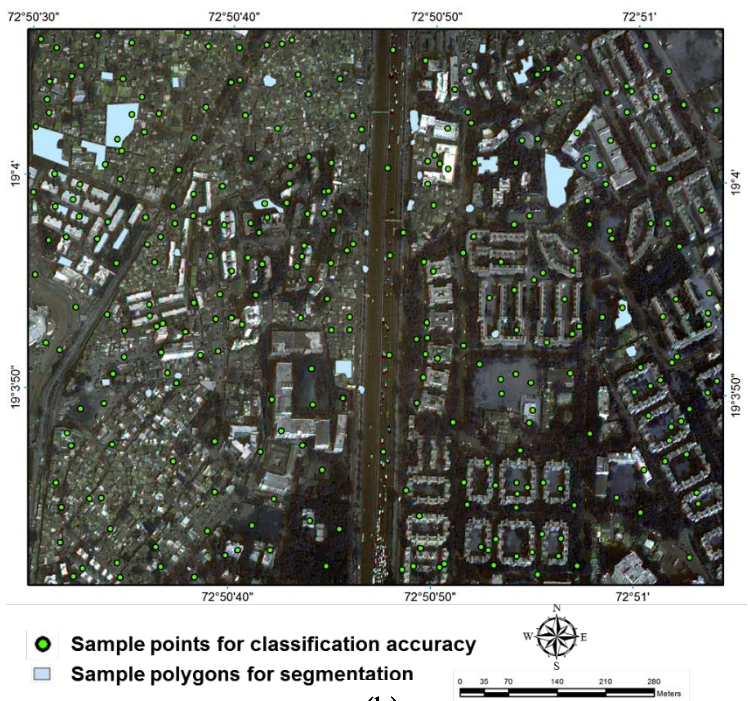
The rational that we generated the GLCM red entropy in ENVI software was that Haralick feature inbuilt in eCognition software was possible to generate texture of $GLCM_{red\ entropy}$ only in 8 and 11 bits whereas WV-2 satellite imagery was of 16-bits. And conventional Haralick feature took lot of time to work with the imagery. The texture $GLCM_{red\ entropy}$ were generated by using the software of ENVI with the parameters of with the window size of 3X3 and the offsets of 2 by 2 because we did not consider the neighbourhood level for detecting the classes. So, it is not necessary to observe the larger context and its relationship with adjacent features. Roads and water were generated by using the thematic layer from Mapzen (2015) in the ArcGIS software and then rasterised to use as one of the layer in the classification. Furthermore, the topographic layers of elevation were also used as one of the parameters for the classification on the basis of the topological features in order to define the geometric relationships between informal settlements and formal settlements.

3.3.2.1. Accuracy Assessment

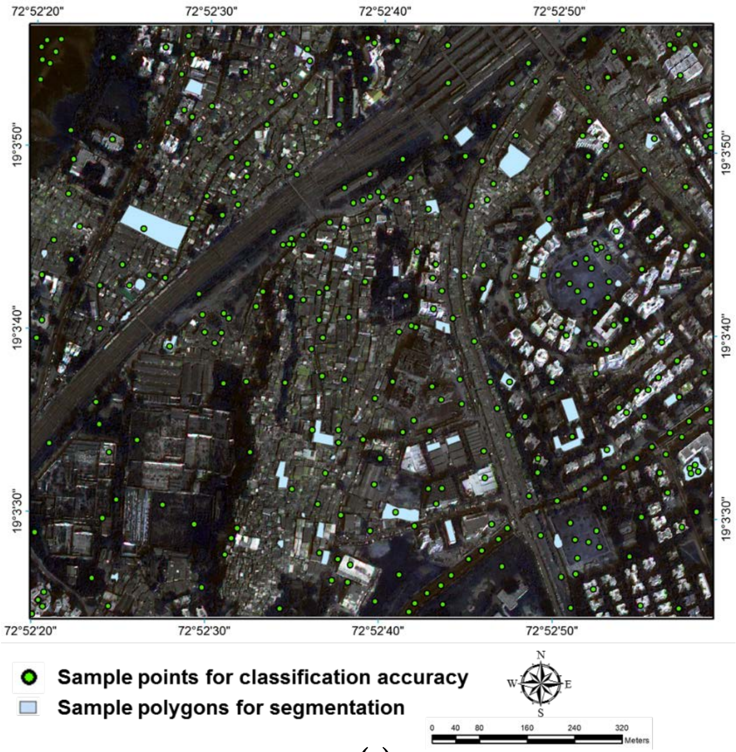
The classification was mandatory to assess the classification quality to verify the overall results of the classification. The confusion matrix was used to measure the accuracy (Lillesand, Kiefer, & Chipman, 2008). For this purpose ground truth information as well as manually generated points through visual interpretation were used. So, as shown in the figure 6, 50 points for each class and 25 for water class features were used because the sample of 35 points for informal settlements and 15 points for formal settlements were collected during the fieldwork. Henceforth, to limit this problem the accuracy assessment in the study was used through ground truth information and manually inserted points through visual observation. The manual points were cross-checked with google earth imagery and also referring to the ground points which have similar ground truth information.



(a)



(b)



(c)

Figure 6: Maps showing ground truth information of subset 1 in (a), subset in (b) and subset 3 in (c)

As per Lee, Shan, & Bethel (2003), the mathematical equation of the quality percentage for the quality control was given as follows:

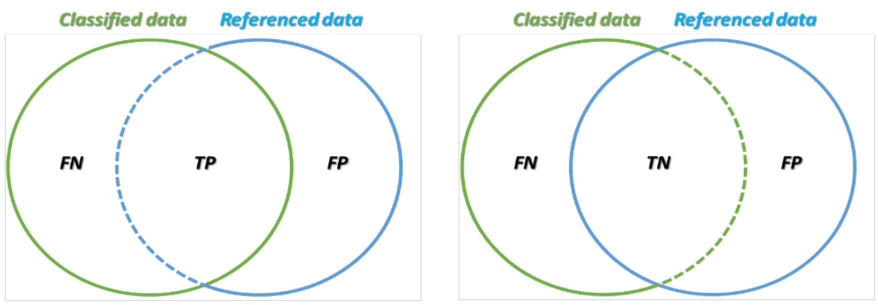
$$Completeness = \frac{TP}{(TP+FN)} \tag{2}$$

$$Correctness = \frac{TP}{(TP+FP)} \tag{3}$$

$$Quality\ Percentage = 100 \times \frac{TP}{(TP+FP+FN)} \tag{4}$$

where *TP; FP* and *FN* were False Negative, False Positive and True Positive

FP was derived from the union of the sets of referenced data and classified data subtracted by the set of classified data whereas *FN* was derived from the union of the sets of referenced data and classified data subtracted by the set of referenced data (ground truth). Therefore, *FN* and *FP* should be focussed to reduce in order to increase the quality (see figure 7 (a)).



(a) **(b)**

Figure 7: Diagram demonstrating the indicators of measuring quality control of the classification in (a) and (b)

The equations (2) represented the completeness which measured the ability of classification out of the classified data whereas the equations (3) represented the ability of truthiness measured from the ground data.

The above equation (4) implied that more the true positives the higher the quality whereas the more false positives and false negatives, the lower the quality. The reason of the false negatives were that those classified pixels are the products of the manual method which could not be counted as the accurate result of the classification(Lee et al., 2003). Hence, the quality control was conducted only for the class features of the informal settlements as the class features were the main focus of the study. Once the assessment of quality control was accomplished, the class descriptions of each rule set would be taken to explore the robustness measurement of rule sets. In the class description, the membership function of each parameter would be analysed to produce the central value as well as the value range for relative change of rule set deviation.

3.3.3. Robustness Measurement of the developed rule set

As discussed in the section 2.5, two determining factors i.e. classification quality of informal settlements and its deviations of rule set was measured to evaluate the robustness of rule set. The measurement of robustness of rule set was adopted from research of Hofmann et al. (2011) to quantify the rule set. There were three steps to check how the robustness of a rule set could be quantified. First one is given in such a way that if the classification quality of subset 2 or subset 3 was achieved with the developed rule set with the classification quality of subset 1 with any adaptation, the rule set is robust. The second was given that if the quality of subset 2 or subset 3 was not achieved with developed rule set, then the rule set was adapted until the quality of subset 1 was achieved and the deviation of relative change of membership was considered during the measurement of robustness. The third was that if the quality of subset 2 or subset 3 was not at all achieved with the quality of subset 1, then the developed rule set was adapted and also deviation of member function was also considered during the measurement. All above three steps were measured by using the equation 8. Hence the classification qualities for subset 2 and subset 3 were first compared with the qualities of subset 1. If the classification quality of a particular subset was found equal to or greater than that of subset 1, the determining factors was not considered in the equation 8 where the mathematical equation was given to measure the robustness of a subset. The quality of subset was measured on the basis of the map delineated by local experts during the questionnaire. Mathematically, the calculation of the area was done by using the equations of 2, 3 and 4. Regarding the deviation of rule set mentioned by Hofmann et al. (2011), out of all deviations, the least deviation of relative change was the deviation of relative change of membership function. In addition, the definition of robustness was focussed on the one which was change with the minimum adaptation. Hence, deviation of relative change of membership function was considered while evaluating the deviation of membership function. Moreover, the deviation of membership function was demonstrated in the figure 8 showing that the relative change of central value and value range of a membership function. Then the equations (5) and (6) gave the condition of the measurement of central value as well as value range were measured respectively.

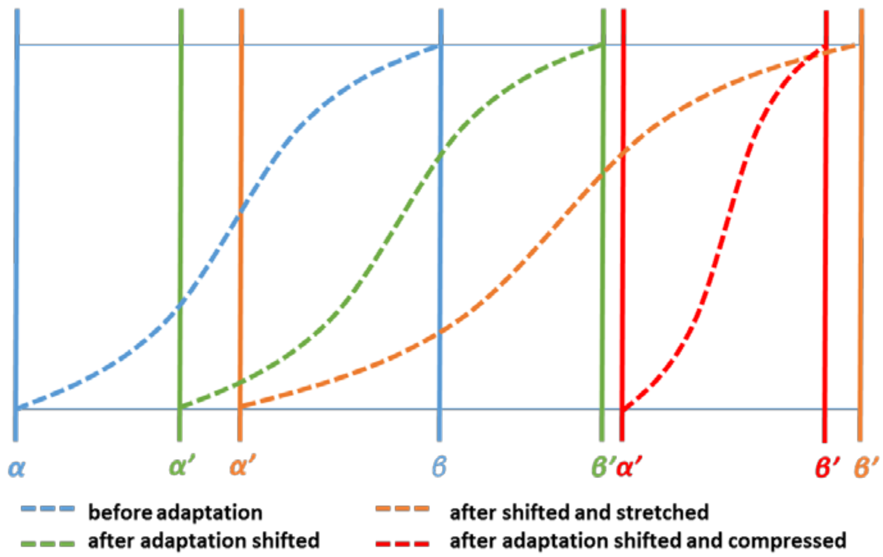


Figure 8: α' and β' depicts the changed lower and upper bound of the membership function where different colours of the given membership function demonstrating the shifting, stretching and compressing of the membership function respectively after adaptation (Adapted from Hofmann et al., 2011)

Mathematically, as per Hofmann et al. (2011), the equations of the relative change δa and δv are given below:

$$\delta a = \begin{cases} \left(1 - \frac{a_f}{a_o}\right) & \text{for } a_f > a_o (\text{positive shift}) \\ 0 & \text{for } a_f = a_o (\text{no shift}) \\ \left(1 - \frac{a_o}{a_f}\right) & \text{for } a_f < a_o (\text{negative shift}) \end{cases} \tag{5}$$

and

$$\delta v = \begin{cases} \left(1 - \frac{v_f}{v_o}\right) & \text{for } v_f > v_o (\text{stretch}) \\ 0 & \text{for } v_f = v_o (\text{linear shift}) \\ \left(1 - \frac{v_o}{v_f}\right) & \text{for } v_f < v_o (\text{compression}) \end{cases} \tag{6}$$

where

δa is the shifting-change of the membership function and δv is the stretching-change of the membership function

a_f and a_o are the parameters that deals with change of the shifting and stretching in the adapted rule set (r_f) and the initial rule set (r_o) respectively

v_f and v_o are the parameters that deals with change of the stretching in the adapted rule set (r_f) and the initial rule set (r_o) respectively

From the equations 7 and 8, we can summarise the measurement of the deviations in the rule set to quantise the robustness of the particular rule set. Furthermore, it depicted whether the parameters are stable or not throughout the subsets. The main equations 7 and 8 can be taken up to quantify the robustness measurements since the research focussed mainly on the degree of the membership function rather than changing classes, changing logical operators, changing relational connectors, changing rules.

We measured the robustness of the rule set based on the method of Hofmann et al. (2011), using the following equations. The equation (7) gives the relative change of the adapted rule set and using the equation (8), all steps given by Hofmann et al., (2011) was measured to provide the robustness of a particular rule set:

$$\delta F_i = \delta a_i + \delta v_i \tag{7}$$

where δF_i is the relative change of a fuzzy membership function in the i^{th} subset
δa is the relative change for shifts in the i^{th} subset
δv is the relative change for stretch or compression in the i^{th} subset

$$r_i = \frac{q_i/q_o}{d+1} \tag{8}$$

where r_i is the robustness of the rule set in the i^{th} subset
d is the rule set deviation ($= \sum_{o=1}^{f} \delta F_i$)
q_i is the quality of classification of test imagery (I_f) using the rule set (R_f)
q_o is the quality of classification of referenced imagery (I_o) using the rule set (R_o)

$$r = \frac{1}{n} \sum_{i=0}^{n} r_i \tag{9}$$

where r is the average of the robustness obtained in subset 2 and subset 3

Using the equation (7), the relative change of the rule set were measured. Thereafter, the summation can give the overall deviation (d) in the rule set. Furthermore, in this study, only this form were analysed due to the time constraint and this type of form influenced more based on the definition of the robustness.
Using the equation (8), the robustness of the rule set by deviation and quality was measured to get the final result of the study and the average of robustness in both cases were measured by using the equation (9).
According to Hofmann et al. (2011), quantifying the deviations of rule sets would be measured by the degree of the membership function of each rule developed for different classes. Each class might have one or more than condition. This degree of few changes of membership function would give the measurement of robustness of the particular rule set. The changes were quanitified by the relative change δa and δv where δa is the shifting-changes and δv is the stretching-changes respectively. In the figure 8, the change of shifting and stretching were demonstrated to visualise how the membership function of the paramters were considered to provide the measurement of the robustness in particular rule set.

3.4. Summary

In summary, we specified the feature description of informal settlements of Mumbai, then we had shown how the local knowledge from local experts had been conceptualised into three levels of hierarchal characteristics in order to develop the local ontology of informal settlements based on the morphological features given by Kohli et al. (2012) and to define the informal settlements in the local context. While defining the informal settlements, there were many more indicators regarding the sensor specification of the imagery which also brought a lot of influence, for instance, the advantage of new bands in assembling the spectral indices as indicators for developing the rule sets. Then, the classification quality was compared to that subset 1 and classification of all subsets was performed with developed rule set to check the deviation was required to consider in calculating the robustness of rule set. Finally, we quantified the robustness of the developed rule set using two criteria i.e. classification quality and its deviations.
As a whole, the methodological framework are visually interpreted in the annexure 4. And the legend was given in the annexure 2. We would discuss how the research questions of each research objective

could be answered and the methods were processed to link each other in the methodological framework as illustrated in the annexure 1.

4. RESULTS

4.1. Feature description of informal settlements in Mumbai by local experts

4.1.1. Specification

Based on the questionnaires, we developed to local adaptation of GSO in order to translate into image based paramters. Out of six, only two were responded during the questionnaire. The initial part of the questionnaire was conducted to determine the terminology and the local definition of the informal settlements in Mumbai. The feature characteristics of informal settlments of are summarised in the table 3. This information was used for conceptualising the local knowledge for informal settlements and completing the set of indicators for the image based paramters. Feature descriptions of informal settlements listed in the table 7 also provided the answer to the first research objective of the study.

General indicators	**Observed local charateristics of informal settlements**
Characteristics of building structure	Building type, number of floors, floor area, building height, roof materials
Road network (type and nature of roads)	Type of roads, percentage paved, width, layout
Density and compactness (amount of built-up in a settlement)	Percentage of roof coverage, open spaces and vegetation
Shape of the settlement	Shape of informal settlements
Site condition (location of the settlement with respect to hazards)	Site condition of informal settlements (Steep slopes, low-lying areas)
Neighbourhood location (relation to surrounding areas)	Relationship with surrounding areas of slum

Table 7: Local characteristics of general indicators from local experts

In the table 3, the local characteristics associated with the general indicators were observed that had been asked in the questionnaire for local experts. Moreover, it was refined and filtered to the local context of Mumbai. Since the image characteristics of the subsets taken for the study varied to large extent, the neighbourhood level (environ level) into considered during the analysis. Visual interpretation from the imagery was found hazy to delineate informal settlements because informal settlements had found low contrast compared to formal settlements. Hence, the final extraction of the informal settlements was given at the settlement level.

4.1.1.1. Informal settlements mapped by local experts

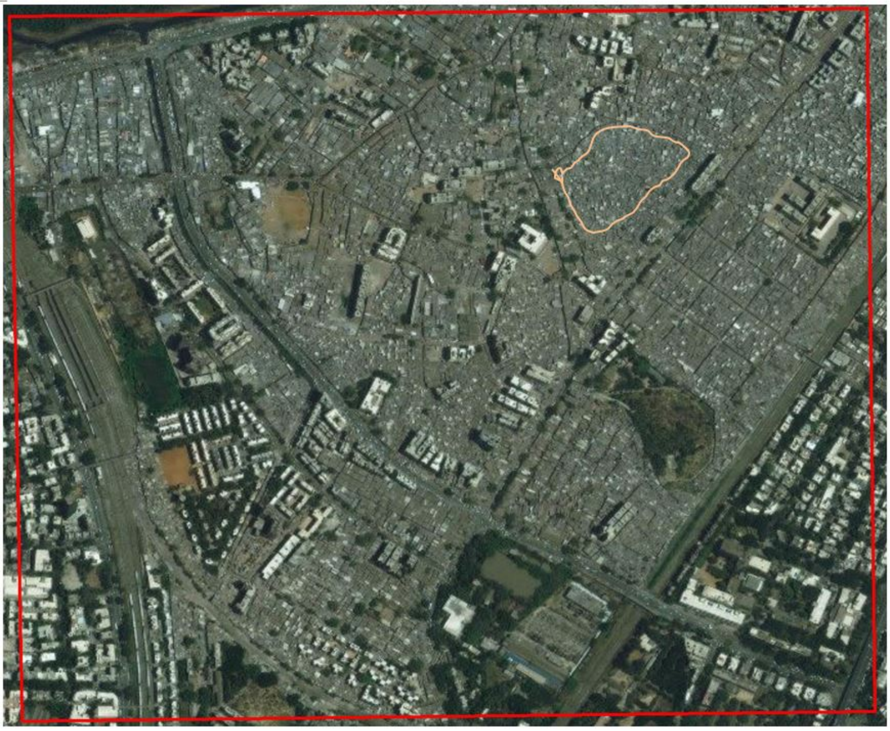


Figure 9: Subset 1 an example of informal settlement, Mumbai (Dharavi)

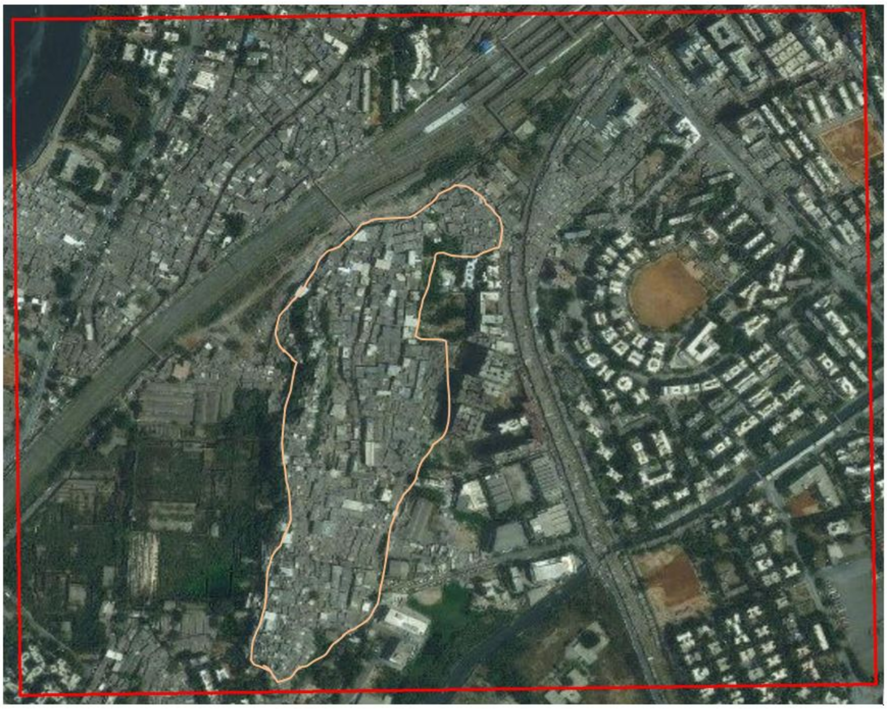


Figure 10: Subset 2 an example of informal settlement, Mumbai (Kurla)

Figure 11: Subset 3 an example of informal settlement, Mumbai (Khar Road East)

In the above figures 9, 10 and 11, the informal settlements mapped by local experts are examples to demonstrate the physical factors that they used while delineating them in the satellite imagery. Moreover, the manual delineation by local experts would be used as referenced data for assessing the classification quality for measurement of robustness of rule set. The physical factors were also noted down in the questionnaire. These factors were summarised after organising and structuring the concept of informal settlements at three spatial levels. From the view of local experts especially, texture, pattern and size were mainly emphasized while mapping in the given satellite imagery.

4.2. Translation into local ontology

4.2.1. Conceptualization

In the figure 12, physical characteristics that were shown were used for selecting indicators from the building characteristics. Only footprint, shape, orientation and type can be easily extracted via remote sensing. One more important characteristrics is that the variability roof material in informal settlements of Mumbai.

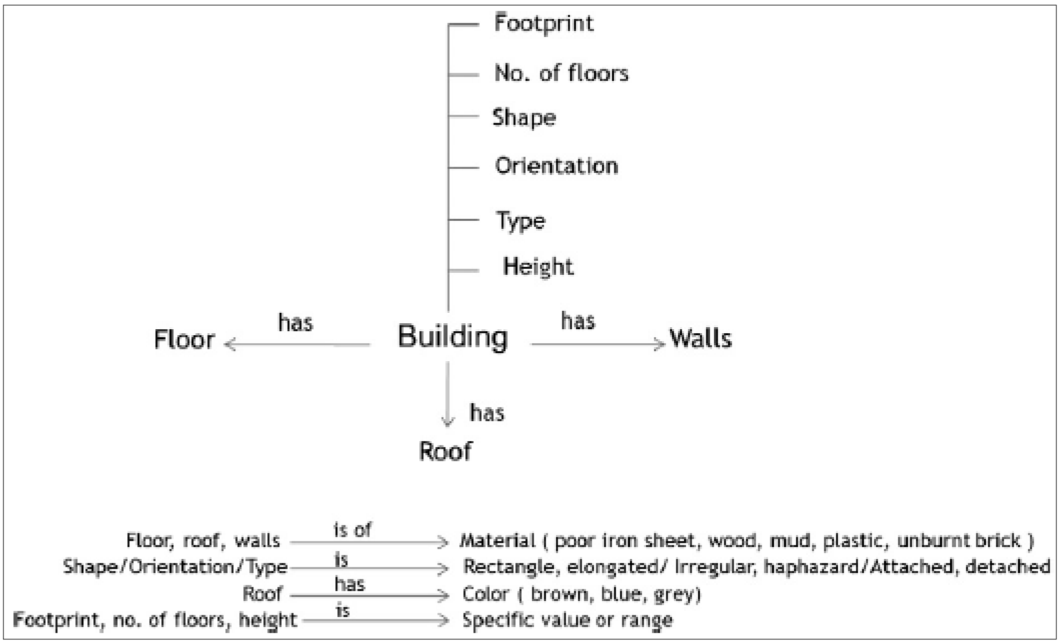


Figure 12: Local characteristics of building structure from local experts (image modified from Kohli et al., 2012)

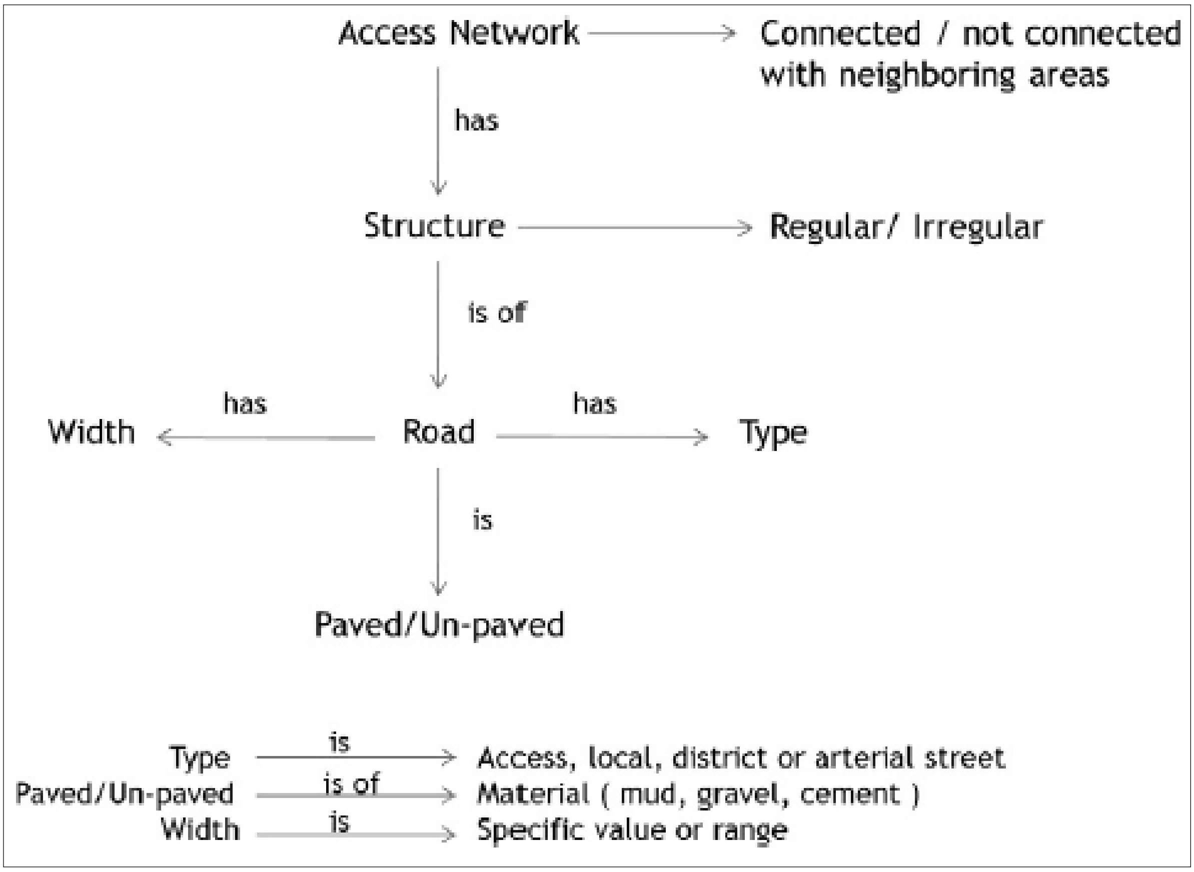


Figure 13: Local characteristics of road network from local experts (image modified from Kohli et al., 2012)

In the figure 13, type of the road network found in Mumbai's informal settlements were: paved/un-paved, structure and connectivity to neighbouring areas whereas no with data available due to lack of visibility in VHR imagery. Moreover, ancillary data were not found.

In the figure 14, the informal settlements had situated at the site which was constructed along the railway lines, major roads and highways. The features of the road network were type, paved/un-paved, structure and connectivity to neighbouring areas.

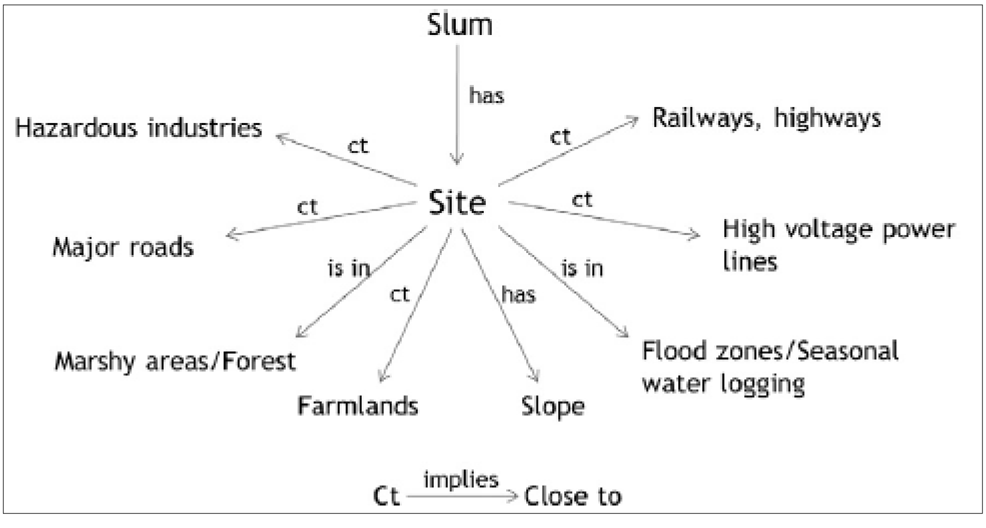


Figure 14: Local characteristics of site conditions from local experts (image modified from Kohli et al., 2012)

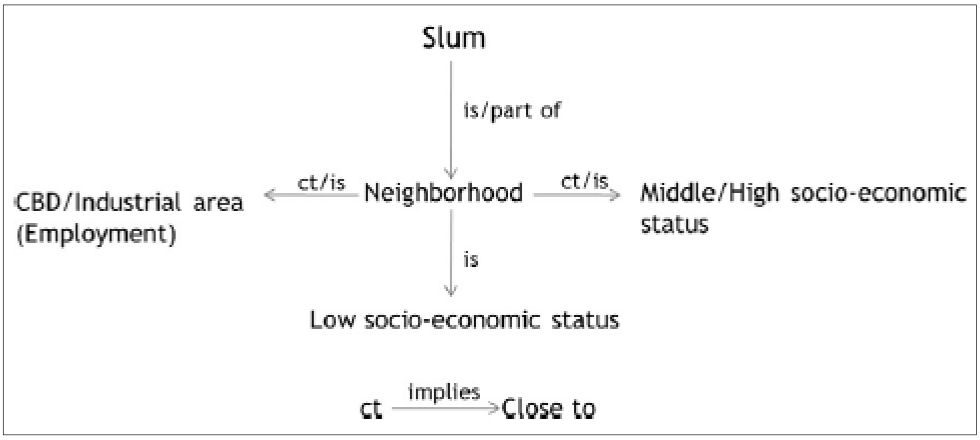


Figure 15:: Local characteristics of neighbourhood location from local experts (image modified from Kohli et al., 2012)

In the figure 15, the informal settlements had neighbourhood close to the middle/high income status of planned areasand close to industrial areas for employment. These characteristics were good when it was dealt with the level environs and not able to visualise from the image observation through remote sensing. It would help to find out easily to some extent where the informal settlements could be high chance to figure it out.

The attributes from different aspects of the feature characteristics would be clapped together with the mapping by local experts to summarise the finalised indicators. Furthermore, it would be needed to be ready to proceed for the next section for translating the image based parameters.

4.2.1.1. Local adaptation of GSO

In the table 8, the image based parameters were translated through the local knowledge of local experts and visual observation. This table 8 summarises by incorporating the section of mapping by local experts and local characteristics of the indicators at three spatial levels. The final indicators from the questionnaire are used in the OBIA. This is followed by OBIA parameterisation to develop the rule set and to run in the OBIA process.

Visual observation in image domain	**Inputs in real world domain from local experts**	**Quantitative Indicators**
Irregular pattern in buildings	Irregular building arrangement	Shape/Geometry
Very difficult to point out the footprints	Size of buildings	Geometry
Less access compared to planned areas	Lack of visibility of road network	Texture
Along the railway lines	Location of large number of buildings near railway lines	Geometry
Close to sites of employment opportunities and planned areas	City fabrics, i.e. often located in between well-developed areas	Association
Variable colours	Same building with varying colours of roofs	Spectral
Variable shapes and sizes	Settlements of different shapes and sizes	Geometry
Variable size of buildings	Varying degrees of deprivations residents face in terms of physical aspects.	Spectral
Variable densities	Settlements range from more organized layouts with medium density till very dense irregular layouts to some extent also caused by the topography of Mumbai	Texture

Table 8: Visual observation and local knowledge from local experts

From the table 8, the overall feature types (quantitative indicators) for classification were found to be spectral, geometry, texture, association. But association would not be considered since the neighbourhood level would not be disussed as mentioned in the specification. However, the indicators elicited from the local knowledge were cross-checked with the characteristics of WV-2 satellte imagery to perform the classification better. So, majority of the indicators were mostly inculcated from the sensor properties rather than the local knowledge because the image characteristics of subset 1 was widely spread Dharavi. So, the neighbourhood level was neglected while developing the rule set parameters.

4.2.2. Implementation

In the table 9, the image based parameters had been developed by using the local adaptation of GSO at three spatial levels. These parameters would be used for developing the rule set and finding the values for the membership function of the parameters. Then, the measurement of robustness would be done by quantising the differences among the respective parameters.

4.2.2.1. Local indicators at three levels of the informal settlements in Mumbai

Level	**General indicators**	**Elements at image domain**	**Local knowledge at real world domain**	**Defining the image based parameters to develop rule set**
Environs	Site location (location of the settlement with respect to hazards)	Slope, pattern, secondary data	Waterlogging, along railway lines, proximity to high voltage power lines (pylon)	Association – distance to features
	Neighbourhood characteristics (relation to surrounding areas)	Pattern, secondary data	Often located in between well-developed areas, close to sites of opportunities for employment	Association – distance to planned areas and employment opportunities
Settlement level	Density and compactness (amount of built-up in a settlement)	Pattern	Highly compact compared to planned areas	Association – distance to less compact areas
	Shape of the settlement	Texture	Planned areas with medium density and very dense irregular layouts	Texture – entropy, contrast, variance, mean
			Lack of visibility of road network and low vegetation	Geometry – area of vegetation, service area of road network
Object level	Characteristics of building structure	Shape	Rectangular	Geometry – rectangular
		Size	10 – 30 m²	Geometry – area
		Material	Iron, abestos, plastics, concrete	Spectral – layer mean values
	Road network (type and nature of roads)	Shape	Irregular	Geometry – rectangular
		Type	Unpaved, concrete, tarmac (metalled), Paved	Spectral – length/width ratio, layer mean value
		Widths	1 – 2 m²	Geometry – area

Table 9: Local adaptation of GSO to develop the image based parameters for Mumbai

4.3. Image Analysis

4.3.1. Segmentation

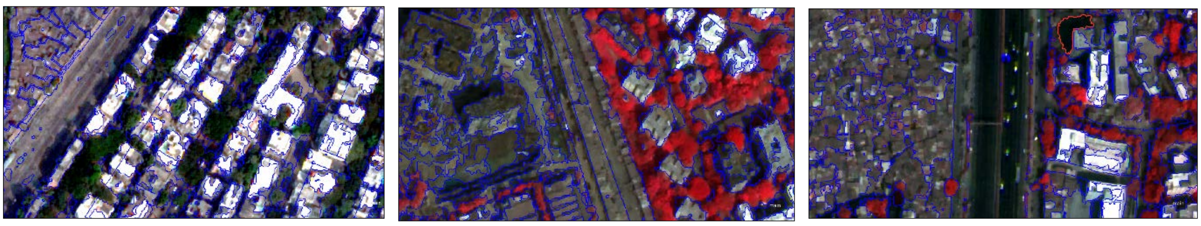
After pre-processing phase, we segment the image for each subset by using the Multiresolution segmentation in the eCognition 9.0 software. Furthermore, we also incorporated the thematic layers of water and roads which enhanced the segmentation quality and also for the particular feature while classification. So, the scale parameter is implemented in each subset with the value of shape and compactness as 0.5 and 0.5 respectively. We evaluate the segmentation from 40 polygons as the referenced objects at settlement level. After the scale parameter applied in each subset, the average the centroid distance between the referenced objects and the segmented objects are found in the range of the 3.75 metres to 5 metres. According to the equation 1, the locational discrepancy is estimated as follows:

a. for subset 1 is 4.97 metres
b. for subset 2 is 3.81 metres
c. for subset 3 is 3.99 metres

Results are drawn from the average of centroid distance of segmented polygons with the reference objects as shown the statistical mean in the annexure 7.

The average centroid distance D_{sr} was found to be larger at lower scale of the subset 1 whereas smaller at larger scales of other two subsets. The over-segmentation of the subset 1 at lower scale produces larger value because the referenced objects superimpose with many segmented objects. Although the subset 2 is smaller in scales and applied all the subsets with equal scale parameters, the centroid distance of subset 3 is found to be larger due to image variations of the subsets. It is clear that the segmentation cannot be estimated and implied by the same scale parameters even though they are from the same sensor. So, the result is still correct in the study and take into consideration for the image analysis because the optimal segmentation parameter is conducted by trial and error. And we assumed it to be representative for the image analysis.

The visual observation from the results for each subset is illustrated in the figure 15. We observe that informal settlements have groups of buildings in many of segmented objects because buildings of informal settlements have lower contrast. In formal settlements, discrete buildings and trees are observed. Integration of thematic layer of roads and GLCM $_{\text{Red entropy}}$ help in segmentation as observed in the figure 15.



(a) **(b)** **(c)**

Figure 16: Results of segmentation for each subset (a) for subset 1(b) for subset 2 (c) for subset 3

4.3.2. OBIA approach

After finishing the task of segmentation, the developed rule set after integrating the local knowledge was applied in the in the subset 1. The classification was performed with the membership function in the class descriptions. Out of the classification, seven class features were classified in the subset 1 during the image analysis. The accuracy assessment was conducted in the Erdas Imagine software to produce the overall accuracy of the particular classification. The results assisted in comparing the classification accuracy of subsets with respect to that of subset 1. It was cleared that the comparison of accuracy played an important role to justify whether the relative function of membership function was to be measured or not in measuring the robustness of the rule set. For subset 1, the class features of vegetation was classified with the same accuracy of 100% (see annexure 15) on both user's accuracy and producer's accuracy. This shows that the NDVI parameter of the spectral indices used in the rule set for the classification of vegetation was very effective. Furthermore, class features of roads and water were also classified as 100% (see annexure 15) in accuracy because the thematic layers were used to classify them which in turn increased not only segmentation quality but also classification accuracy with 100% (see annexure 15) on both user's accuracy as well as producer's accuracy. The NDNB parameter of the spectral indices was also found to be potential in classifying the shadow with the accuracy of 96% on both user's accuracy as well as producer's accuracy. The formal settlements were classified by using the combination of three parameters namely GLCM, NDNB and brightness. The result of the formal settlements was found to be less in accuracy of 76% (see annexure 15) on both as compared to other class features because there were confusion by mixing with the class features of bare soil and informal settlements which led to the increase of false-positives. In addition, different types of roof materials in the subsets gave inconvenient for the parameters to predict the appropriate numerical values for classifying formal settlements. However, working with the fuzzy membership functions which improved the performance of classification for formal settlements. The bare soil was found to be classified with accuracy of 72% (see annexure 15) because the spectral values were similar with old roof materials. Lastly, the remaining unclassified features had been assigned as informal settlements and found with the accuracy of 72% (see annexure 15). The overall classified result showed that the class feature of informal settlements was found to be classified throughout the subset 1 because of the hierarchy of the classification rules that all remaining features had been assigned as informal settlements. And also, existence of the faded and old roof materials gave more challenging to provide the threshold for classifying informal settlements. All image based parameters was used for the classification of all class features except Area parameters of the geometry category for informal settlements. Moreover, we found that parameters of Area and DEM refined the classification because these parameters reduced the false-positives of formal settlements on informal settlements and false-positives of informal settlements on formal settlements respectively during the classification of subset 1. The overall accuracy of the classification for subset 1 was found to be 87% (see annexure 15) as shown in the figure 17 (a). Moreover, the accuracy result in the subset 1 was used for the threshold for comparison of accuracy. The comparison provided whether further analysis were to be measured for deviation of relative change of membership function or not so that the evaluation of the robustness measurement of rule set could be completed.

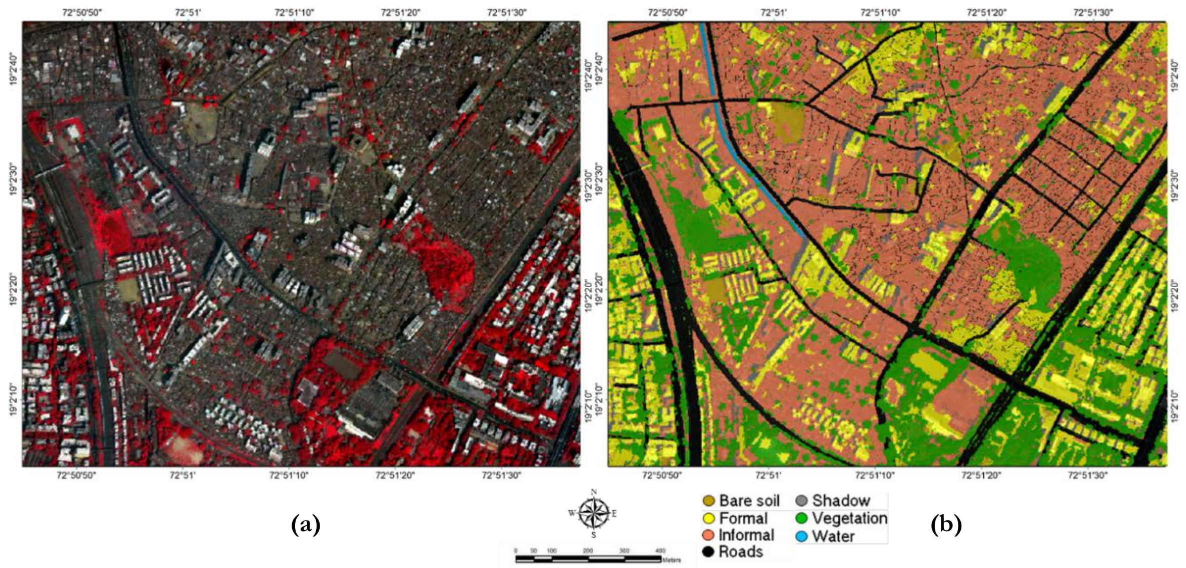


Figure 17: False Colour Composite (FCC) Image on the left (a) and classified image on the right (b) for subset 1

After classification of the subset 1, the same developed rule set was reapplied on both subset 2 and subset 3 to measure their accuracies. The classification results of subset 2 and subset 3 were illustrated in the figure 19 (a) for subset 2 and figure 19 (b) for subset 3. And the accuracies with developed rule set were found to be 79.08% (see annexure 16) and 82% (see annexure 18) for subset 2 and subset 3 respectively which were found lesser accuracies compared to that of subset 1. This implies that we tuned up manually the parameters to adapt the developed rule set until the classification quality of subset 1 was achieved. After adaptation of the developed rule set to their respective subsets, overall classification accuracy for subset 2 and subset 3 after adaptation were found to be 85.54% (see annexure 17) and 86.33% (see annexure 19) respectively. Henceforth, we considered not only qualities but also the deviation of relative change of membership functions in evaluating the robustness of rule set in both case.

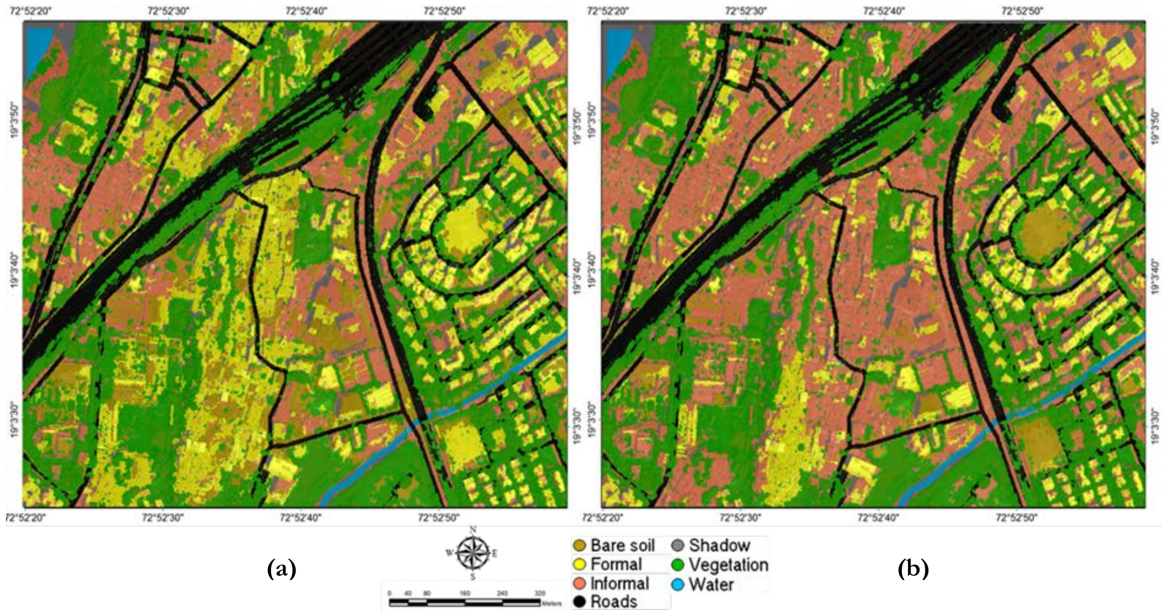


Figure 18: Image classification for subset 2 before adaptation and after adaptation

Figure 19: Image classification for subset 3 before adaptation and after adaptation

4.4. Robustness Measurement of Rule Set

The main focus of the evaluation was on the informal settlements to evaluate the rule set. In case of subset 1, 19036.69 square metres of informal settlements was correctly classified whereas 5283.27 square metres were omitted and 12024.28 square metres were wrongly classified. In case of subset 2 before adaptation, we found that the informal settlements was correctly classified by the area of 2000.69 square metres whereas 1446.31 square metres were omitted and 32010 square metres were wrongly classified. In case of subset 3 before adaptation, 14558.87 square metres of informal settlements were classified correctly whereas 9298.88 squares were omitted and 1303.79 square metres were wrongly classified. The condition after adaption of rule set, the classification of subset 2 was obtained with the correctly classified area of 24458.17 square metres whereas the informal area of 12676.39 square metres was omitted and 10039.86 square metres were wrongly classified. In case of subset 3 after adaptation, 12799.34 square metres of informal settlements were classified correctly whereas 7321.04 squares were omitted and 3063.32 square metres were wrongly classified. Using the equations 2, 3 and 4, the completeness, the correctness and the quality percentage for informal settlements of all subsets were drawn in the table 11 quantitatively and in the figures 20, 21 and 22 qualitatively.

Condition	Subsets	True Positives (TP) (in metres)	False Negatives (FN) (in metres)	False Positives (FP) (in metres)	Completeness (C_1)	Correctness (C_2)	Quality percentage (QP)
Original	Subset 1	19036.69	12024.28	5283.27	61.29%	78.28%	52.38%
Before Adaptation	Subset 2	2000.69	1446.31	32010.00	58.04%	5.88%	5.64%
	Subset 3	14558.87	9298.88	1303.79	61.02%	91.78%	57.86%
After Adaptation	Subset 2	24458.17	12676.39	10039.86	65.86%	70.90%	51.85%
	Subset 3	12799.34	7321.04	3063.32	63.61%	80.69%	55.21%

Table 10: Classification qualities of all subsets and also qualities of subset 2 and subset 3 after adaptation

Moreover, the figure 20 and the figure 21 were illustrating the classification of subset 2 and subset 3 with the condition of rule set before and after adaptation. In the figure 20 and we still observed that there was over-segmented in the classification of subset 1 and subset 3 because the value of false-negatives were comparatively high in all subsets. In the table 11, we found that the classification quality of subset 1 was 52.4%. The classification quality of subset 2 was 56% which was terribly low compared to that of subset 1 whereas the classification quality of subset 3 was found more than that of subset 1 with the value of 58. Therefore, deviation of relative change of membership function was considered in case of subset 3 whereas in case of subset 2, deviation would not be considered in measuring the robustness o rule set because the quality was more than the quality with developed rule set.

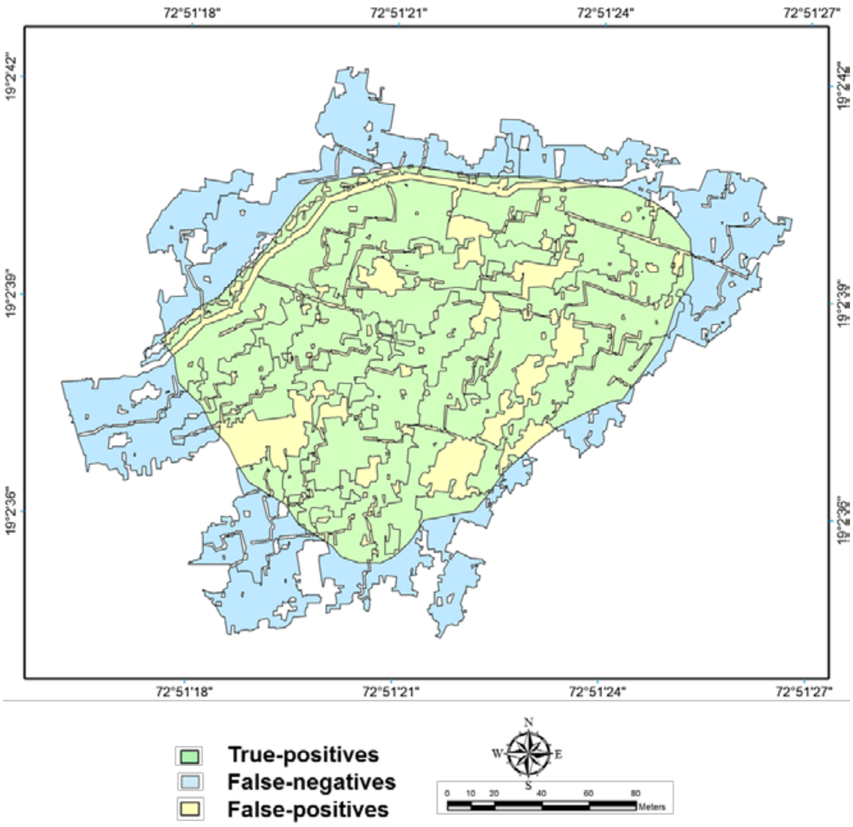


Figure 21: Visualisation of classification quality for subset 1

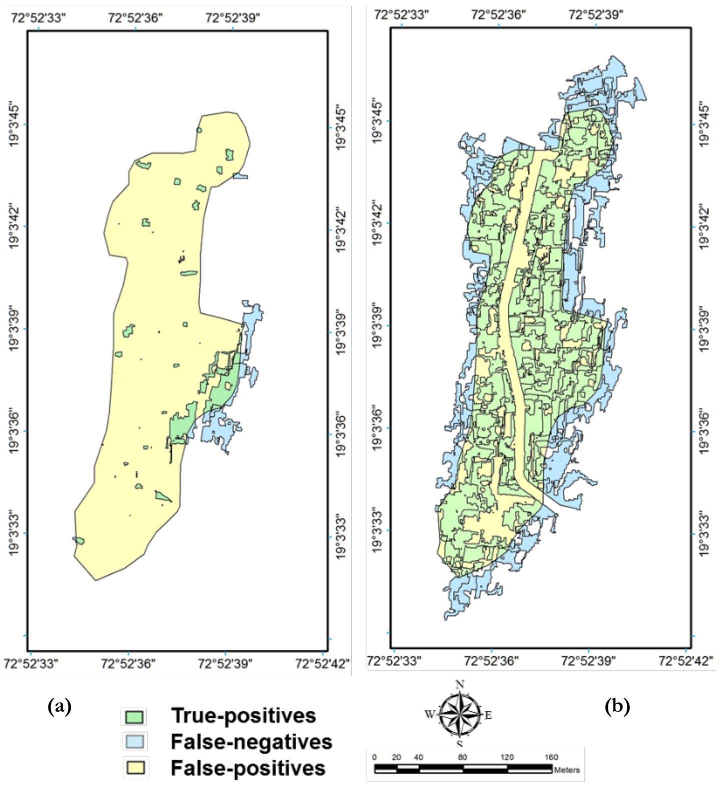


Figure 20: Visualisation of classification quality for subset 2 with (a) before adaptation and (b) after adaptation

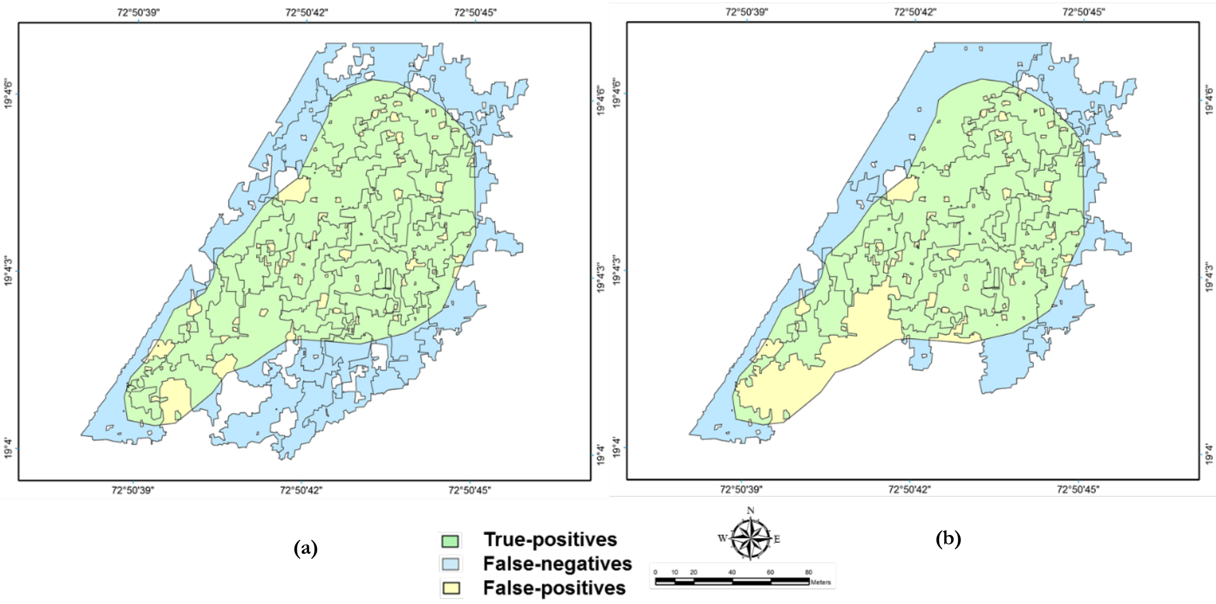


Figure 22: Visualisation of classification quality for subset 3 with (a) before adaptation and (b) after adaptation

After measuring the quality criterion, the deviation of relative change of the membership function was illustrated in the table 12. Adaptation of developed rule set were also demonstrated how the local ontology were translated into image based parameters. Moreover, the assignment of membership function and result of the determining factors for deviation of relation change of member functions for all subsets were illustrated in the annexure 9 for subset 1, annexure 10 for subset 2 and annexure 11 for subset 3. Using equations 5 and 6, the determining factor of deviation such as value range and central value was found out whereas the equation 7 provided the results of deviation for each parameters used in the class features as the values were resulted in the table 12. In the table 12, we found that the parameters of thematic attribute of roads, GLCM red entropy and NDGR were found to be stable throughout the classification by getting the deviations below 0.5 in all subsets. However, absence of water class in subset 3 obtained the individual score with the value of 2 which in turn reduced the stability of the parameter. The overall deviation (d_1) for subset 2 is 6.04 where deviation (d_2) for subset 3 after adaptation is 6.89.

	Descriptions of OBIA feature			Subset 1		Subset 2		Deviations			Subset 3		Deviations		
Level	**Class Feature**	**Parameter**	**Feature Type**	**Value Range (v_r)**	**Central Value (a_r)**	**Value Range (v_1)**	**Central Value (a_1)**	**δv_1**	**δa_1**	**δF_1**	**Value Range (v_2)**	**Central Value (a_2)**	**δv_2**	**δa_2**	**δF_2**
Object level	**Vegetation**	NDVI	**Spectral index**	**0.50**	**0.20**	**0.86**	**0.22**	**0.70**	**0.11**	**0.80**	**0.53**	**0.23**	**0.06**	**0.14**	**0.20**
	Roads	**Thematic attributes**	**Spectral**	**1.00**	**0.50**	**1.00**	**0.50**	**0.00**	**0.00**	**0.00**	**1.00**	**0.50**	**0.00**	**0.00**	**0.00**
	Water	**Thematic attributes**	**Spectral**	**1.00**	**0.50**	**1.00**	**0.50**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**1.00**	**1.00**	**2.00**
	Shadow	NDNB	**Spectral index**	**0.50**	**0.06**	**0.80**	**0.13**	**0.60**	**0.56**	**0.04**	**0.56**	**0.10**	**0.13**	**0.72**	**0.84**
Settlement level	**Formal**	GLCM Red entropy	**Texture**	**1.94**	**1.07**	**2.11**	**1.07**	**0.09**	**0.00**	**0.09**	**1.87**	**0.93**	**0.04**	**0.13**	**0.09**
		NDNB	**Spectral index**	**2.20**	**1.10**	**0.57**	**0.05**	**0.74**	**1.05**	**1.78**	**0.43**	**0.04**	**0.81**	**1.04**	**1.85**
		Brightness	**Spectral**	**1423.26**	**1565.94**	**1987.76**	**1749.71**	**0.40**	**0.12**	**0.51**	**1461.72**	**1661.19**	**0.03**	**0.06**	**0.09**
	Bare soil	NDGR	**Spectral index**	**0.24**	**0.05**	**0.31**	**0.06**	**0.31**	**0.14**	**0.45**	**0.23**	**0.05**	**0.03**	**0.05**	**0.01**
	Formal	Mean (DEM)	**Geometry**	**19.00**	**10.50**	**37.56**	**16.78**	**0.98**	**0.60**	**1.57**	**11.30**	**6.29**	**0.41**	**0.40**	**0.81**
	Informal	Area	**Geometry**	**469817**	**234909.5**	**287070.00**	**143538.00**	**0.39**	**0.39**	**0.78**	**233077.00**	**116541.50**	**0.50**	**0.50**	**1.01**
									$\sum\delta F_1 = 6.04$					**$\sum\delta F_2 = 6.89$**	

Table 11: Adaptation of developed rule set and deviations between developed rule set and adapted rule sets

5. DISCUSSION

The absence of a generalised definition of informal settlements has triggered the elicitation of local knowledge from local experts. From local experts, the *spatial and contextual features* of informal settlements in the local context are acquired. This features relate to the morphology (physical characteristics) of informal settlements such as, spatial, geometry and texture. The findings show that the employed spectral indices (NDNB and NDGR) have an important role in extracting informal settlements, but the spectral indices were not mentioned by local experts as spatial and contextual features. However, we used them for the local adaptation of the GSO, coming not from local experts but added from literature.
The *indicators of informal settlements in VHR* based on local experts imagery are irregular patterns of buildings, varying building sizes, less access to the buildings located inside settlements and varying roof colours (roof materials) in informal settlements. Few indicators like varying roof colours in VHR imagery are related with the spectral properties. So, spectral indices plays a key role in easily translating the GSO to the local context and in the development of the rule set for detecting informal settlements. Earlier in a study, the spectral indices were mainly employed to distinguish between roof and non-roof surfaces in the research work of Hamedianfar & Shafri (2015), but in our study, the spectral indices are used to distinguish formal, bare soil and informal settlements. So, the spectral indices are added during the development of the rule set.
The indicators that are translated into the local ontology through the *adaptation of GSO* are spatial, geometry, texture and spectral indices. Since the GSO was developed mainly on the basis of the morphological features, there should be an awareness of other indicators during the development of rule set which can highly contribute to the local adaptation of the GSO. The elicited local knowledge are always necessary to filter all feasible indicators collated from local experts depending on the local context because, sometimes, the indicators do not influence much in the process of classification.
The indicators are acquired after elicitation of local knowledge from local experts whereas the spectral indices are acquired from literature. The GSO adaptation helps in translating *indicators into the image based parameters*. The method was strengthened by the inclusion of spectral features besides physical features. These features are translated into OBIA. The geometry feature (DEM) is translated into quantitative factors in order to read the thematic layers for classification. The texture feature is translated in order to enhance the contrast between the formal settlements and informal settlements. The spectral indices are translated to distinguish the class features of vegetation, shadow and bare soil which are not visible with the naked eye from the image domain, using the spectral indices with NIR spectral band which is beyond the visible range.
Comparing the *difference in accuracy* between the classifications of different subsets, we found a slightly lower accuracy compared to the classification accuracy of subset 1. One factor contributing to the lower accuracy to some extent is the segmentation scale, which is of concern as it requires to estimate an optimal segmentation scale to make the classification quality comparable. This may influence the classification quality which might lead to misinterpretation of further processes in measuring the robustness. The non-random type of sampling of ground truth data should be considered. This was done because of limit accessibility of informal settlements to capture the unbiased reference data. In our study, we classified seven class features. So, we used classification accuracy as the supplement of the classification quality for measuring the robustness. Moreover, the classification accuracy allows to check the accuracy among the classes. The robustness of the rule set was considerable with the deviation of relative change of the membership functions. The referenced polygons was taken majority from formal buildings because formal buildings has clear boundary and good for assessment of segmentation quality. After extracting areas of

formal buildings having the advantage that they are clearly visible in VHR imagery, the remaining areas are assigned as informal settlements after all other class features are classified during the classification.
The *adaptation of parameters* to detect informal settlements could have been improved and found more stable parameters by the combination of parameters. The performance of parameters dropped when the first parameter was optimized. For instance, in case of detecting formal settlements, three OBIA parameters are used at a time where first parameter is the GLCM red entropy, after optimizing this parameter the remaining parameters showed problems with the adaptation. In case of inter-class comparison, the parameter NDNB did not require much adaptation. For example while working to detect shadow the parameter NDNB showed good potential for detecting formal settlements.
Estimation of *stable parameters* in different subsets was done by selecting the image based parameters which have found the membership values below 0.50 in different subsets. Regarding the classification using membership function, samples for the classification must be taken for each rule in order to generate the fuzzy membership function automatically in the class description rather than manually inserting the membership function type used for the classification. Moreover, fuzzy membership function plays an important role in incorporating local knowledge while developing the rule set in the study. Eventually, the application of the elicitation of local knowledge can be valuable. The determination of the spectral indices of the subset size are mainly based on the pixel sizes. The results would have been better if we consider the same size of area for all subsets. This implies that the deviation of the rule set is highly influenced by the scales adopted. If all other parameters such as changing classes, changing operators, changing connectors, changing rules, etc. were included as additional conditions in measuring the deviations of the rule sets, then it might contribute to strengthen the identification of the stable parameters. For example, the deviation of relative change of membership function as obtained in the case of water class features in subset 3 would have been reduced from the value of 2 to a lesser value.
The *robustness of the developed rule* set is found to be 0.14 out of 1. This is very low like in many other research studies as discussed in section 12.5.1. The robustness of the rule set could be higher, increasing up to 0.62 out of 1, when considering those cases in which classification quality is as close to the quality of referenced imagery i.e. subset 1 and ignoring deviation of relative change of membership function.

6. CONCLUSION

This study addressed three sub-objectives, first the local adaption of the GSO to Mumbai by adding local knowledge, second, the development of an adapted rule set for detecting informal settlements and third, the exploration of robustness of the set of parameters used in the OBIA rule set by applying it to other two subsets of VHR imagery.

In this study, the local knowledge from local experts are categorized at two spatial levels for *local adaptation of GSO* to Mumbai using three specific subsets. These subsets were taken from Mumbai, a city of India to present the physical characteristics of the informal settlements. We have the indicators of spatial, geometry and texture from local experts which are translated to local context. We adopted spectral indices as an important feature for extracting informal settlements. These spectral indices are specific toWV-2 satellite imagery with its very unique 8 bands. In our study, we have only tested these indices three subsets. Therefore, we cannot generalise that the information from the local ontology is of lesser importance then the spectral indices. However, the use of GSO in our study is minimal because we used spectral based indicators in majority from literature rather than the physical factors from the elicitation of local knowledge from local experts.

Spatial, geometry, spectral indices and texture are set in translating image based parameters at two levels for *developing the adapted rule set*. In order to initiate the classification, the multiresolution segmentation was applied with a segmentation the quality of 5 m tolerance determined by locational discrepancy of locational based approach. The classification was carried out with the membership functions and we found the classification accuracy was less in subset 2 and 3 as compared to that of subset 1. Moreover, we also found the addition of thematic layers of roads help in improving the segmentation quality as well as the classification accuracy. Since there is no optimal scale parameter to segment image objects for the image analysis, the integration of a thematic layer of formal buildings can be used to improve both the segmentation as well as the classification quality.

To *evaluate the robustness* of initial rule set, we used two criteria i.e. the classification quality and the deviation of relative change of the membership functions from the reference map of subset 1. We found that the degree of robustness is 0.14 (out of 1) which is considerably low as is also found in many studies (therein the literature review section 2.5.1). For future research, one can incorporate criteria other than the deviation adopted in this study to get more insights into the deviation. For example, due to the absence of class feature of water in subset 3, the membership value of water was found to be 2 which could have been turned into 0 if "OR" logic operator was applied. In this study, parameters such as NDGR of spectral index and GLCM texture are found to be most stable in all subsets under considerations. Therefore, these parameters can be used to strengthen the robustness in future research.

REFERENCES

Ardila, J. P., Bijker, W., Tolpekin, V. A., & Stein, A. (2012). Context-sensitive extraction of tree crown objects in urban areas using VHR satellite images. *International Journal of Applied Earth Observation and Geoinformation*, *15*, 57–69. doi:10.1016/j.jag.2011.06.005

Baatz, M., Hoffmann, C., & Willhauck, G. (2008). Progressing from object-based to object-oriented image analysis. In *Object-Based Image Analysis* (pp. 29–42). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-77058-9_2

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *58*(3-4), 239–258. doi:10.1016/j.isprsjprs.2003.10.002

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. doi:10.1016/j.isprsjprs.2009.06.004

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., … Tiede, D. (2014). Geographic Object-Based Image Analysis - Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, *87*, 180–191. doi:10.1016/j.isprsjprs.2013.09.014

Blaschke, T., & Strobl, J. (2001). What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. *GeoBIT/GIS*, *14*(6), 12–17. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-0035382401&partnerID=tZOtx3y1

Census of India. (2013). Primary Census Abstract for Slum. Retrieved June 2, 2015, from http://www.censusindia.gov.in/2011-Documents/Slum-26-09-13.pdf

Deved¿ic, V., Djuric, D., & Ga¿evic, D. (2009). *Model Driven Engineering and Ontology Development. Engineering* (2nd Ed.). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-00282-3

DigitalGlobe. (2010). The Benefits of the Eight Spectral Bands of Worldview-2. Retrieved August 19, 2015, from c-agg.org/cm_vault/files/docs/DG-8SPECTRALBANDS-WP.pdf

Drăguţ, L., Tiede, D., & Levick, S. R. (2010). ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, *24*(6), 859–871. doi:10.1080/13658810903174803

Ebert, A., Kerle, N., & Stein, A. (2009). Urban social vulnerability assessment with physical proxies and spatial metrics derived from air- and spaceborne imagery and GIS data. *Natural Hazards*, *48*(2), 275–294. doi:10.1007/s11069-008-9264-0

ENVI. (2009). *ENVI Atmospheric Correction Module: QUAC and FLAASH user's guide*. ITT Visual Information Solutions. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:ENVI+Atmospheric+Correction+Module:+QUAC+and+FLAASH+user's+guide#0

Ezeah, C., Fazakerley, J. A., & Roberts, C. L. (2013). Emerging trends in informal sector recycling in developing and transition countries. *Waste Management (New York, N.Y.)*, *33*(11), 2509–19. doi:10.1016/j.wasman.2013.06.020

Hamedianfar, A., & Shafri, H. Z. M. (2015). Detailed intra-urban mapping through transferable OBIA rule sets using WorldView-2 very-high-resolution satellite images. *International Journal of Remote Sensing*, *36*(13), 3380–3396. doi:10.1080/01431161.2015.1060645

Hofmann, P. (2014). Defining Robustness Measures for OBIA Framework: A Case Study for Detecting Informal Settlements. In *Global Urban Monitoring and Assessment through Earth Observation* (pp. 303–324). CRC Press. doi:10.1201/b17012-21

Hofmann, P., Blaschke, T., & Strobl, J. (2011). Quantifying the robustness of fuzzy rule sets in object-based image analysis. *International Journal of Remote Sensing*, *32*(22), 7359–7381. doi:10.1080/01431161.2010.523727

Hofmann, P., Strobl, J., Blaschke, T., & Kux, H. (2008). Detecting informal settlements from Quickbird data in Rio de Janeiro using an object based approach. In *Object-Based Image Analysis* (pp. 531–553). doi:10.1007/978-3-540-77058-9_29

Indian Census. (2011). Mumbai (Greater Mumbai) City Population Census 2011 | Maharashtra. Retrieved August 19, 2015, from http://www.census2011.co.in/census/city/365-mumbai.html

Ishtiyaq, M., & Kumar, S. (2011). Typology of informal settlements and distribution of slums in the NCT of Delhi. *Journal of Contemporary India Studies: Space and Society, Hiroshima University*, *1*, 37–46. Retrieved from home.hiroshima-u.ac.jp/hindas/PDF/Ishtiyaq_and_Kumar(2011).pdf

Kit, O., Lüdeke, M., & Reckien, D. (2012). Texture-based identification of urban slums in Hyderabad, India using remote sensing data. *Applied Geography*, *32*(2), 660–667. doi:10.1016/j.apgeog.2011.07.016

Kohli, D., Sliuzas, R. V., Kerle, N., & Stein, A. (2012). An ontology of slums for image-based classification. *Computers, Environment and Urban Systems*, *36*(2), 154–163. doi:10.1016/j.compenvurbsys.2011.11.001

Kohli, D., Warwadekar, P., Kerle, N., Sliuzas, R., & Stein, A. (2013). Transferability of Object-Oriented Image Analysis Methods for Slum Identification. *Remote Sensing*, *5*(9), 4209–4228. doi:10.3390/rs5094209

Kuffer, M. (2003). Monitoring the dynamics of informal settlements in Dar Es Salaam by remote sensing: Exploring the use of SPOT, ERS and small format aerial and photography. In M. Schrenk (Ed.), *Proceedings of the CORP* (pp. 473–484). Vienna, Austria: Technical University, Vienna.

Kuffer, M., Pfeffer, K., Baud, I., & Sliuzas, R. V. (2013). Analysing sub-standard areas using high resolution remote (VHR) sensing imagery. In *Proceedings of the14th N-AERUS / GISDECO conference* (pp. 1–16). Enschede, Netherlands.

Lang, S. (2008). Object-based image analysis for remote sensing applications: Modeling reality – Dealing with complexity. In *Object Based Image Analysis* (pp. 1–25). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-77058-9_1

Lang, S., & Blaschke, T. (2006). Bridging remote sensing and GIS - What are the main supportive pillars? In *1st International Conference on Object-based Image Analysis (OBIA 2006)*. Retrieved from http://www.isprs.org/proceedings/XXXVI/4-C42/Papers/01_Opening Session/OBIA2006_Lang_Blaschke.pdf

Lee, D. S., Shan, J., & Bethel, J. S. (2003). Class-Guided Building Extraction from Ikonos Imagery. *Photogrammetric Engineering & Remote Sensing*, *69*(2), 143–150. doi:10.14358/PERS.69.2.143

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2008). *Remote sensing and image interpretation*. John Wiley & Sons. Retrieved from https://books.google.nl/books?id=XaIsAQAAMAAJ

Lopez, M. F., Gomez-Perez, A., Sierra, J. P., & Sierra, A. P. (1999). Building a chemical ontology using Methontology and the Ontology Design Environment. *IEEE Intelligent Systems*, *14*(1), 37–46. doi:10.1109/5254.747904

Lu, D., & Weng, Q. (2005). Urban Classification Using Full Spectral Information of Landsat ETM Imagery in Marion County , Indiana. *Photogrammetric Engineering & Remote Sensing*, *71*(11), 1275–1284.

Lüscher, P., Weibel, R., & Mackaness, W. A. (2008). Where is the Terraced House? On the Use of Ontologies for Recognition of Urban Concepts in Cartographic Databases. In *Headway in Spatial Data Handling* (pp. 449–466). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-68566-1_26

Mapzen. (2015). Metro Extracts: City-sized portions of OpenStreetMap, served weekly. Retrieved October 9, 2015, from https://mapzen.com/data/metro-extracts

Martha, T. R., Kerle, N., van Westen, C. J., Jetten, V., & Kumar, K. V. (2011). Segment Optimization and Data-Driven Thresholding for Knowledge-Based Landslide Detection by Object-Based Image Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(12), 4928–4943. doi:10.1109/TGRS.2011.2151866

Mathenge, C. W. (2011). *Application of Object Oriented Image Analysis in Slum Identification and Mapping - The Case of Kisumu, Kenya*. ITC - Faculty of Geo-Information Science and Earth Observation, University of Twente. Retrieved from http://www.itc.nl/library/papers_2011/msc/upm/wanjiku.pdf

Montaghi, A., Larsen, R., & Greve, M. H. (2013). Accuracy assessment measures for image segmentation goodness of the Land Parcel Identification System (LPIS) in Denmark. *Remote Sensing Letters*, *4*(10), 946–955. doi:10.1080/2150704X.2013.817709

Netzband, M., & Rahman, A. (2009). Physical characterisation of deprivation in cities: How can remote sensing help to profile poverty (slum dwellers) in the megacity of Delhi/India?). In *Urban Remote Sensing Event, 2009 Joint* (pp. 1–5). IEEE.

Rizvi, I., & Krishna Mohan, B. (2012). Object-based analysis of WorldView-2 imagery of urban areas. In *2012 IEEE International Geoscience and Remote Sensing Symposium* (pp. 431–434). IEEE. doi:10.1109/IGARSS.2012.6351546

Schiewe, J., & Gähler, M. (2008). Modelling uncertainty in high resolution remotely sensed scenes using a fuzzy logic approach. In *Object-Based Image Analysis* (pp. 755–768). Berlin, Heidelberg: Springer Berlin

Heidelberg. doi:10.1007/978-3-540-77058-9_41
Shekhar, S. (2012). Detecting Slums From Quick Bird Data in Pune Using an Object Oriented Approach. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XXXIX-B8*(September), 519–524. doi:10.5194/isprsarchives-XXXIX-B8-519-2012
Sirueri, F. O. (2015). *Comparing spatial patterns of informal settlements between Nairobi and Dar es Salaam*. University of Twente Faculty of Geo-Information and Earth Observation (ITC). Retrieved from http://www.itc.nl/library/papers_2015/msc/upm/sirueri.pdf
Sliuzas, R. V., & Kuffer, M. (2008). Analysing the Spatial Heterogeneity of Poverty Using Remote Sensing: Typology of Poverty Areas using Selected RS Based Indicators. In C. Jürgens (Ed.), *Proceedings of the EARSel Joint Workshop: Remote Sensing New Challenges of High Resolution* (pp. 158–167). Bochum, Germany.
Sliuzas, R. V., Mboup, G., & de Sherbinin, A. (2008). *Report of the Expert Group Meeting on Slum Identification and Mapping*. Retrieved from www.alnap.org/pool/files/egm-slum-mapping-report-final.pdf
Sori, N. D. (2012). *Identifying and Classifying Slum Development Stages from Spatial Data*. University of Twente Faculty of Geo-Information and Earth Observation (ITC), Enschede, Netherlands. Retrieved from http://www.itc.nl/library/papers_2012/msc/upm/dinsasori.pdf
Subbaraman, R., Nolan, L., Shitole, T., Sawant, K., Shitole, S., Sood, K., … Patil-Deshmukh, A. (2014). The psychological toll of slum living in Mumbai, India: a mixed methods study. *Social Science & Medicine (1982)*, *119*, 155–69. doi:10.1016/j.socscimed.2014.08.021
Taubenböck, H., & Kraff, N. J. (2014). The physical face of slums: a structural comparison of slums in Mumbai, India, based on remotely sensed data. *Journal of Housing and the Built Environment*, *29*(1), 15–38. doi:10.1007/s10901-013-9333-x
Trimble. (2015a). *eCognition® Developer 9.1 Reference Book*. Munich, Germany: Trimble Documentation. Retrieved from http://www.ecognition.com/
Trimble. (2015b). *eCognition® Developer 9.1 User Guide*. Munich, Germany: Trimble Documentation. Retrieved from http://www.ecognition.com/
UN-Habitat. (2003). *The Challenge of Slums*. Routledge. doi:10.4324/9781849772907
UN-Habitat. (2010). *State of the World's Cites 2010/2011: United Nations Human Settlements Programme*. UN-HABITAT. Retrieved from www.unhabitat.org
UN-Habitat. (2012). *State of the World's Cites 2012/2013: United Nations Human Settlements Programme*. Nairobi, Kenya: UN-HABITAT. Retrieved from www.unhabitat.org
Veljanovski, T., Kanjir, U., Pehani, P., Otir, K., & Kovai, P. (2012). Object-Based Image Analysis of VHR Satellite Imagery for Population Estimation in Informal Settlement Kibera-Nairobi, Kenya. In *Remote Sensing - Applications*. InTech. doi:10.5772/37869
Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353. doi:10.1016/S0019-9958(65)90241-X

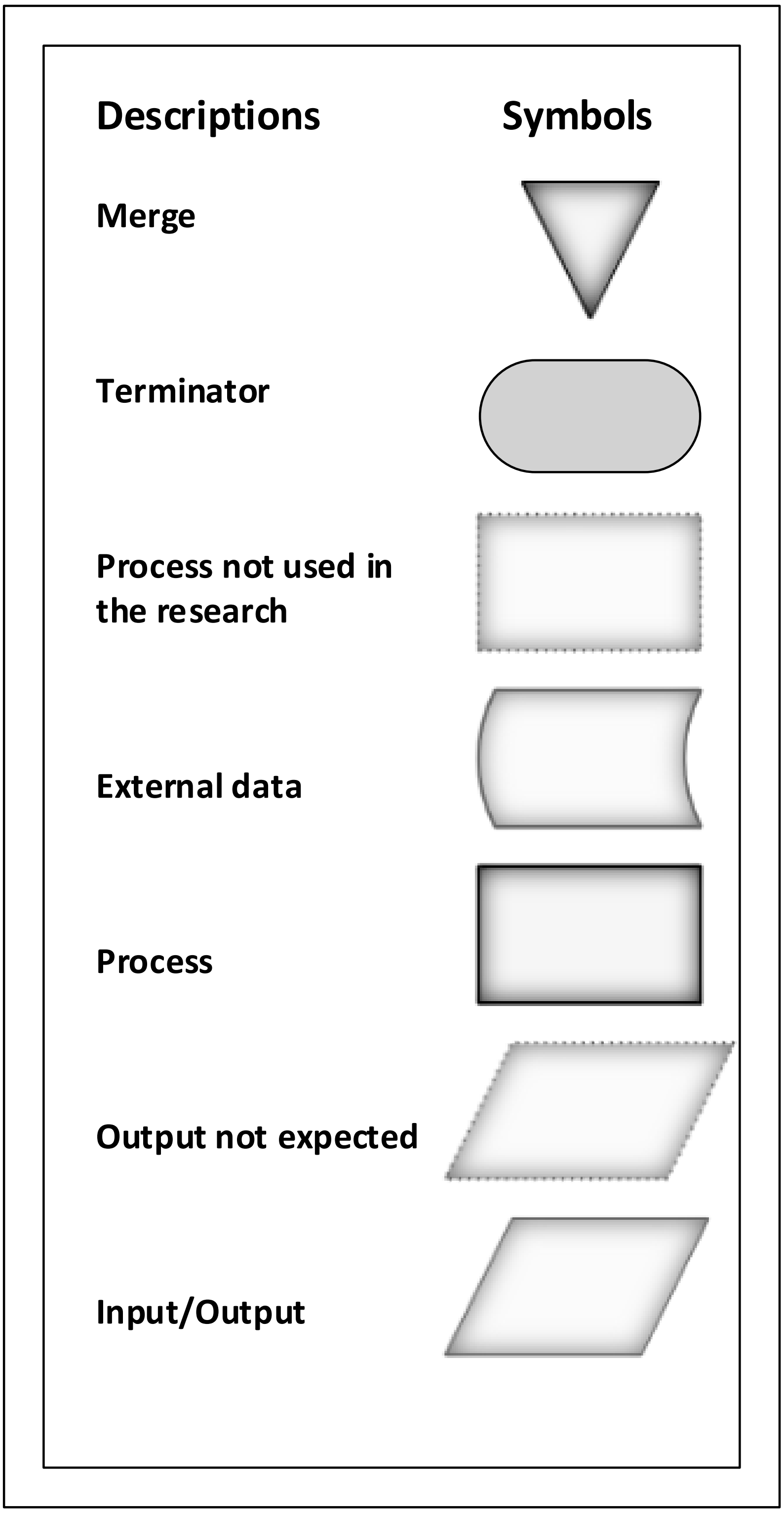
ANNEXES

Annexure 1: Table showing the link from the research objectives to the expected results

Research objectives	Research questions	Data and methods	Expected results
To study the local adaption of the GSO to Mumbai, an Indian mega-city by adding local knowledge	1. What are the spatial and contextual features of the informal settlements in the local context? 2. What are the indicators that can be distinguished from VHR imagery?	Literature review, spatial data, three subsets of Worldview-2 imagery, image interpretation	1. List of feature information in local framework 2. List of indicators
To develop the adapted rule set for detecting informal settlements based on the indicators identified from the local slum ontology	3. How is GSO adapted to local slum ontology to generate the indicators in local context? 4. How are the indicators translated into image based parameters? 5. What is the difference in accuracy between the results of classification quality of subset 1, subset 2 and subset 3?	Adaptation, additional indicators like spectral indices, OBIA, multiresolution segmentation	1. List of adapted rule set with image based parameters 2. List of parameters with variables 3. Accuracy of classified results with
To explore the robustness of the set of parameters used in the OBIA rule set by applying it to other two subsets of VHR imagery	6. Which parameters of the rule set can be used to acquire similar results with no or little adaptation for detecting informal settlements? 7. How are the stable parameters estimated in different subsets when the similar results are achieved? 8. What are the result of the robustness of the developed rule set?	Ground points, membership function, defuzzification, accuracy assessment	1. List of parameters 2. Degree of membership functions of each parameters for each subset 3. Result of robustness of two cases

Annexure 2: Legend of the methodology

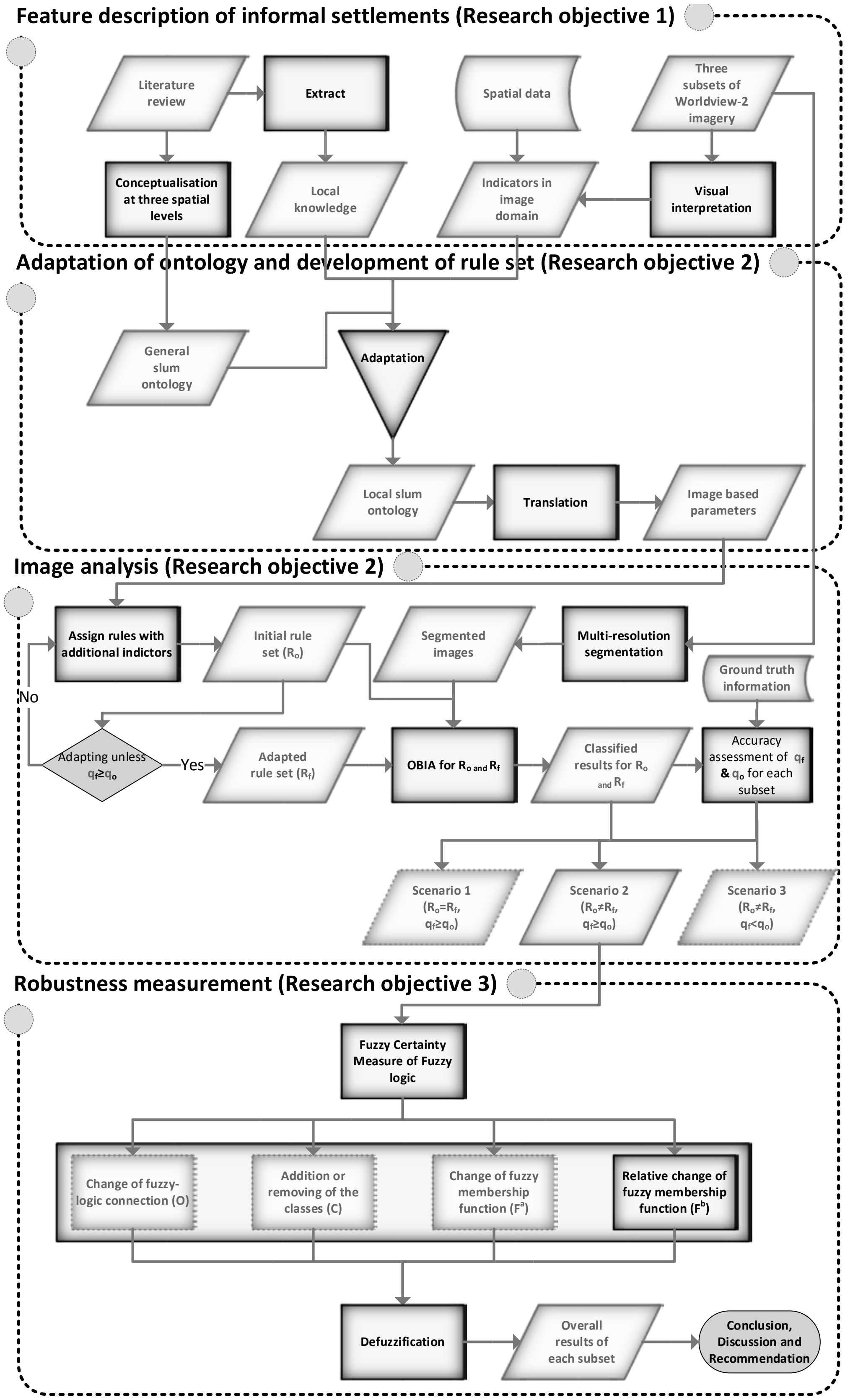
Legend



Annexure 3: Work timeline

Events	September				October				November				December				January				February			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Feature description of informal settlements																								
Extract information from the literature																								
Visual interpretation																								
Conceptualisation of slum ontology																								
Adaptation of ontology and development of rule set																								
Adaptation of ontology in local context																								
Translation into image based parameters																								
Image analysis																								
Finding the additional indicators																								
Assignment of rules for rule set																								
Segmentation																								
OBIA																								
Mid-term presentation																								
Robustness measurement																								
Generation of fuzzy membership function																								
Fuzzy logic approach																								
Accuracy assessment																								
Discussion and recommendation																								
Final report submission																								
Final presentation																								
Exceptional cases																								
Discussion with supervisors																								
Writing the updated work																								
Literature review																								

Annexure 4: Methodology



Annexure 5: Types of informal settlements in Mumbai (Source: Kuffer et al., 2013)

No.	Classification of deprived areas	Size of buildings footprints	Density of area	Pattern of area	Ground photo of sample area
1	Slum pocket	Small	High	Small areas with organic patterns often very temporary buildings (plastic etc.)	
2	Slum area, small buildings	Small	High	Poor area (poor condition) with small buildings and very organic layout	
3	Slum area, mix small/large buildings	Mix (small – medium sized buildings)	High	Diverse (some areas with and some without structure)	
4	Slum area, larger buildings and chawls	Medium	High	Larger buildings give somewhat more structure	
5	Basic formal	Medium	High	E.g. resettlement colonies area has some structure	

Questionnaire for Visual Interpretation of Informal Settlements in Mumbai, India

"Robustness of rule set using very high resolution (VHR) imagery to detect informal settlements – A Case of Mumbai, India"

Vickyson Naorem
ITC,
Faculty of Geo-Information Science and Earth Observation of the University of Twente
PO Box 335
7500 AE Enschede
The Netherlands
Phone:+31 (0)61 681 9911
Email: v.naorem@student.utwente.nl

"Robustness of rule set using very high resolution (VHR) imagery to detect informal settlements – A Case of Mumbai, India" is my research topic. Robust slum monitoring techniques help to large extent to identify the informal settlements, which is relevant for local authorities. With the increase use of VHR imagery in remote sensing, there is a need for the image analysis that provide an improved representation of complex environment. In order to understand the process in the local context, local expert knowledge is crucial. This will allow the construction of robusted rule ste for the imaage analysis of informal settlements. This research aims to contribute to improve the accuracy and consistency of spatial information on informal settlements.

This research emphasizes on developing local ontology and testing the rule set based on Object based Image Analysis (OBIA) to identify informal settlements in local context of Mumbai. The steps involved will be as follows:

- Formulation of feature descriptions of informal settlements in the local context of Mumbai.
- Development of indicators to be translated into image based parameters for identifying informal settlements based on the local context.

Note: This questionnaire is framed in the form of semi-structured way. So, there is always a way that the respondent can give their opinion other than the questions asked in the questionnaire. And it is generated on the basis of the one generated by Kohli, Sliuzas, Kerle, & Stein, 2012.

- ***Details of the Respondent***

Name……………………………………………….Country………………………… Date…………………

Department:…………………………………

Any experience in field of remote sensing in local context of Mumbai:………………………

- ***Definition of Informal settlements in Mumbai***

    - How would you define the definition of informal settlements in the local context (Mumbai)?

    - What are the types of informal settlements that you observe in Mumbai?

- ***Unique features in the local context of Mumbai***

What are the unique features that pop up in your mind when we mention 'informal settlements of Mumbai'? What are the typical characteristics that make these settlement different from other settlements?

- ***Morphological indicators for informal settlements of Mumbai***

According to Kohli et al., 2012, the indicators to identify informal settlements are divided into three level namely, object level, settlement level and environ (neighborhood) level. These indicators help in developing the understanding the structure and characteristics feature of the informal settlements in the local context to translate into image parameters. At object level, it helps to understand the internal characteristics (components) of the buildings. At settlement level, it helps in understanding the overall form/shape of the settlements. At last but least, the environ level helps in neighborhood characteristics of the settlement.

  - **Object level**

Indicators	Measure	Specification (please fill-up or tick your choose)	Other comments
Characteristics of building structure	Building type (e.g. attached, detached)		
	Number of floors		
	Floor area but not plot area (in square metres)	☐ <10 ☐ 10-20 ☐ 20-30 ☐ 30-40 ☐ >40	
	Building height (in metres)		
	Roof materials		
Road infrastructure (type and nature of roads)	Type of roads	☐ Primary ☐ Secondary ☐ Tertiary	
	Approximate percentage of pavement		
	Road width (in metres)	☐ <1 ☐ 1-2 ☐ 2-3 ☐ 3-4 ☐ >4	
	Road layout	☐ Regular ☐ Irregular	
	Variability of road types, widths	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	
	Connectivity with neighborhood	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	

- **Settlement level**

Indicators	Measure	Specification (please fill-up or tick your choose)	Other comments
Density and compactness (amount of built-up in a settlement)	Roof Coverage (%)		
	Open space (%)		
	Vegetation (%)		
	Number of buildings per square kilometre		
	Compactness	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	
Shape of the settlement	Common shape of the settlement (e.g. elongated, circular, etc)		
	Approximate pattern (e.g. irregular, uniform) and why?		

- **Environ level**

Indicators	Measure	Specification (please fill-up or tick your choose)	Other comments
Site condition (location of the settlement with respect to hazards)	Flood zone	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	
	Steep slopes	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	
	Proximity to railway lines, highways, major roads, airport	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	
	Proximity to high voltage power lines (Pylon)	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	
	Proximity to hazardous industries	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	
Neighborhood location (relation to surrounding areas)	Connected to infrastructure in neighboring areas	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	
	Close to/inside neighborhood of low socio-economic status	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	
	Close to	☐ Very low ☐ Low ☐ Moderate	

	employment areas (e.g. CBD, industries)	☐ Average ☐ High ☐ Very high	
	Close to neighborhood of middle/high socio-economic status	☐ Very low ☐ Low ☐ Moderate ☐ Average ☐ High ☐ Very high	

- **If possible, please suggest any other physical factors that you consider to be relevant for identifying the informal settlements in Mumbai?**

- ***Role of local knowledge in image analysis***

The image analysis cannot be fulfilled with the visual interpretation only. There is a need to include the local knowledge that can reduce the false positive errors during the validation. Moreover, the main focus of this part is to get local knowledge for developing the indicators into image based parameters that allow to identify the informal settlements.

Kindly delineate the boundary of informal settlements in the three images provided below as per your knowledge and answer the questions.

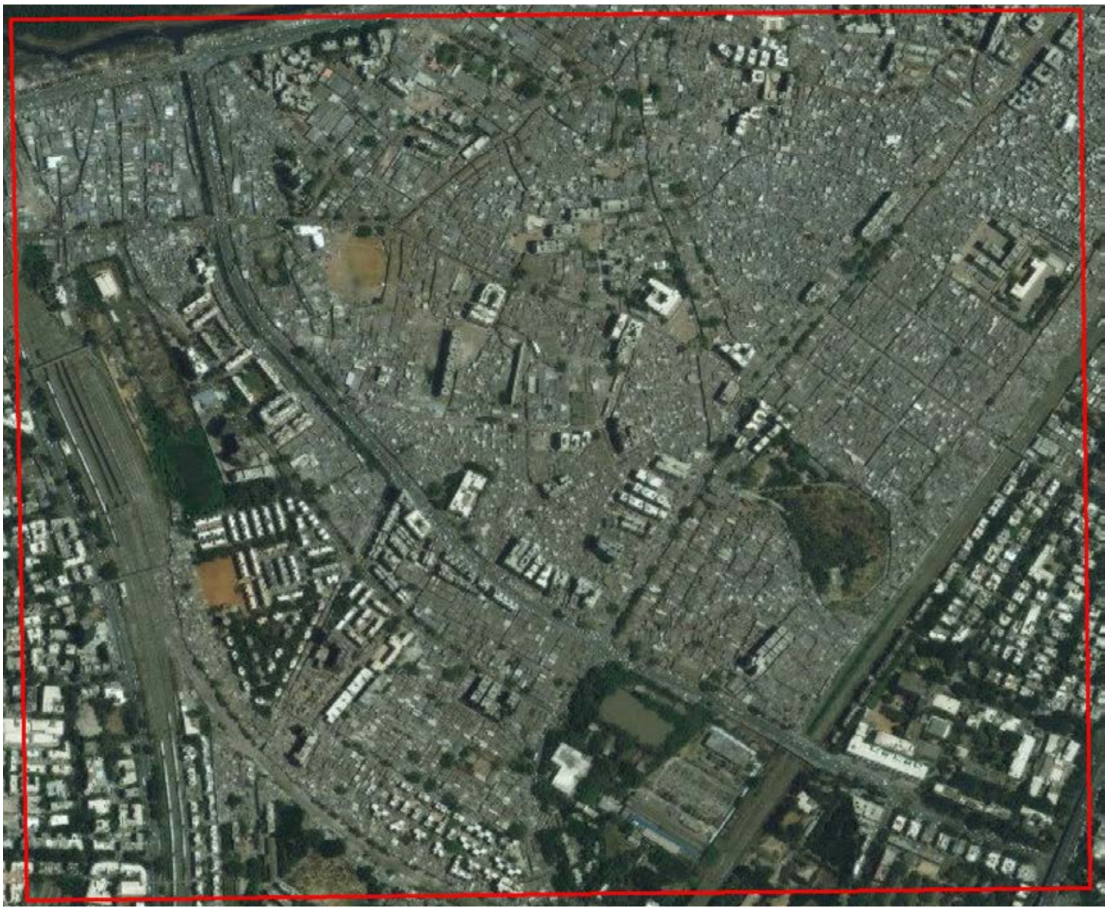


Figure 13: Subset 1 of informal settlement, Mumbai (Dharavi)

- What are the factors you used to draw the boundary of the first subset of informal settlements?

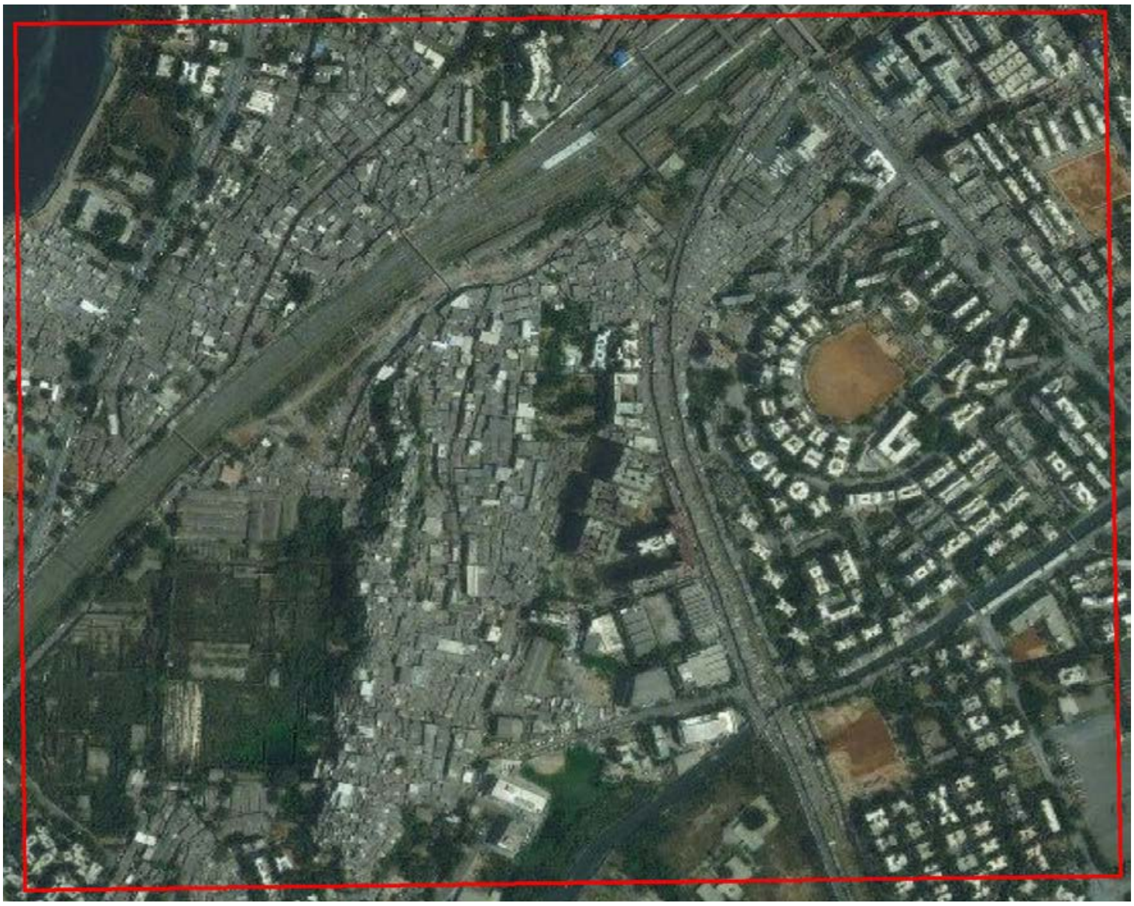


Figure 14: Subset 2 of *informal settlement, Mumbai (Kurla)*

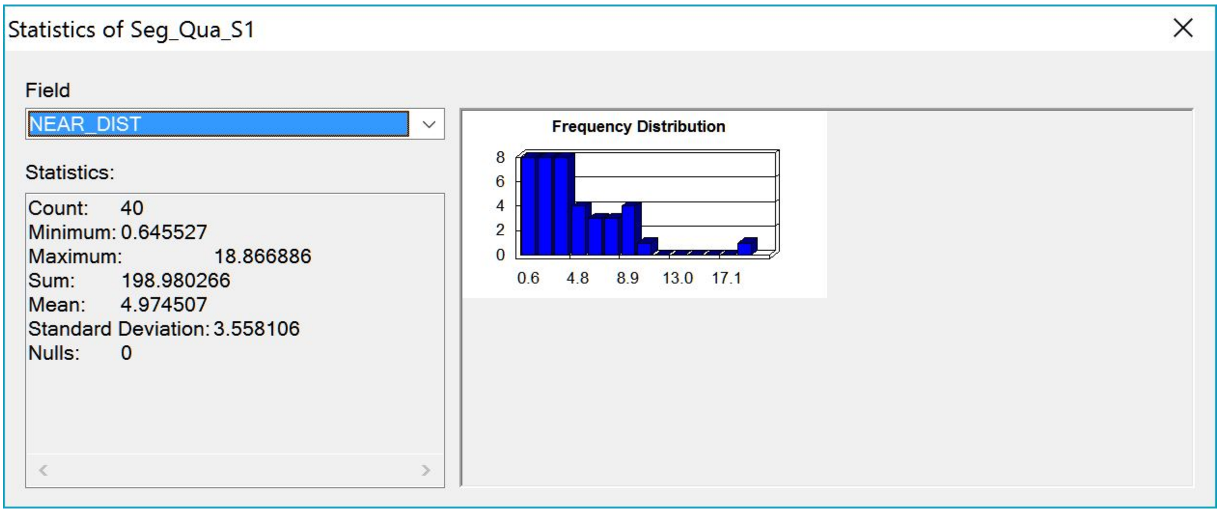
- **What are the factors you used to draw the boundary of the second subset of informal settlements?**

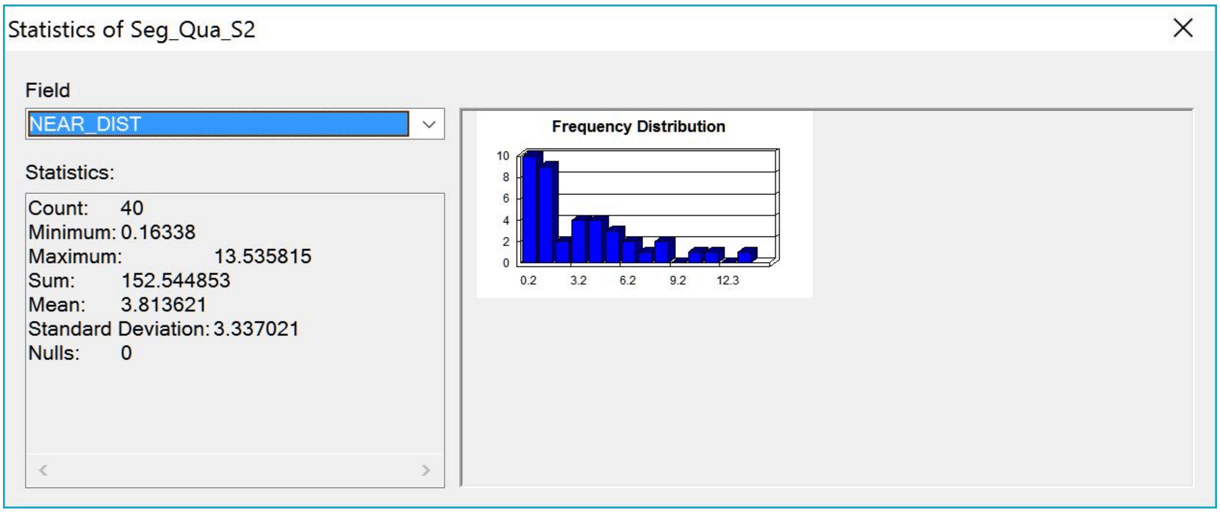
Figure 15: Subset 3 of informal settlement, Mumbai (Bairam Naupada)

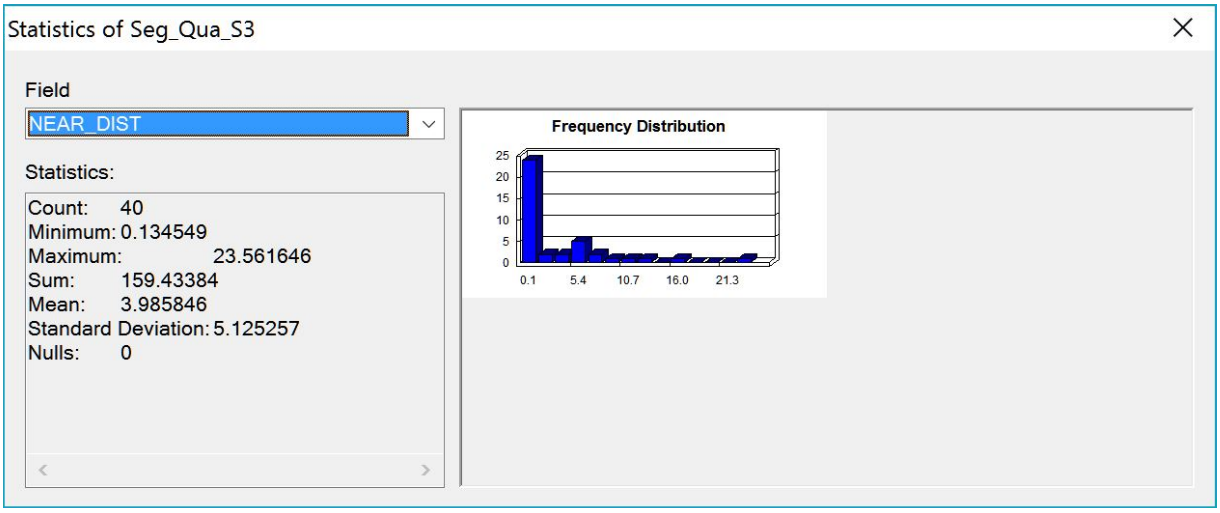
- **What are the factors you used to draw the boundary of the first subset of informal settlements?**

- **Any other comments**

Annexure 7: Segmentation quality of each subset







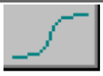
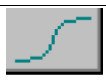
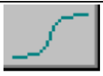
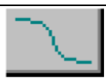
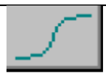
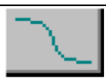
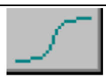
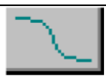
Annexure 8: Results of the segmentation automation

Image Object Information

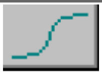
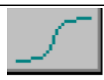
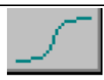
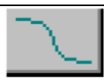
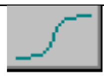
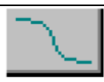
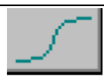
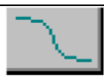
Feature	Value
Scene Related Features	
Scene features	**Scene variables**
Blue	1
Cnt	1
Comp_Incr	0.2000
CountLoops	101
Current scale_Level 1	150
Current scale_Level 2	400
Current scale_Level 3	700
Current scale_LV_graph	101
End	5
Green	1
Initial_Comp	0.1000
Initial_Scale	25
Initial_Shape	0.1000
Layer_Incr	3
Max_Comp	0.9000
Max_Scale	20
Max_Shape	0.9000
NIR	1
PAN	1
Red	1
Scale_Incr	5
Shape_Incr	0.2000
Start	1
Temp_Comp	1.100
Temp_PAN	0
Temp_Shape	1.100
Scene features	**Scene-related**
Number of image layers	8

Features | Classification | Class Evaluation

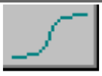
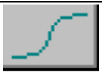
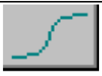
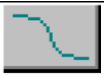
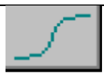
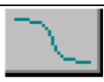
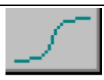
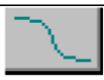
Annexure 9: Descriptions of parameters of rule set for subset 1

Level	Class Feature	OBIA Feature	Membership Function	Feature Type	Lower Border (μ_l)	Upper Border (μ_u)	Value Range (v_r)	Central Value (a)
Object level	Vegetation	NDVI		Spectral index	-0.05	0.45	**0.50**	**0.20**
	Roads	Minimum Pixel Value		Thematic attributes	0.00	1.00	**1.00**	**0.50**
	Water (if existed)	Minimum Pixel Value		Thematic attributes	0.00	1.00	**1.00**	**0.50**
	Shadow	NDNB		Spectral index	-0.31	0.19	**0.50**	**-0.06**
Settlement level	Formal	GLCM Red entropy		Texture	0.10	2.04	**1.94**	**1.07**
	Built-up area	NDNB		Spectral index	0.00	2.20	**2.20**	**1.10**
	Formal	Brightness		Spectral	854.31	2277.57	**1423.26**	**1565.94**
	Bare soil	NDGR		Spectral index	-0.17	0.07	**0.24**	**-0.05**
	Formal	Mean (DEM)		Geometry	1.00	20.00	**19.00**	**10.50**
	Informal	Area		Geometry	1.00	469818.00	**469817.00**	**234909.50**

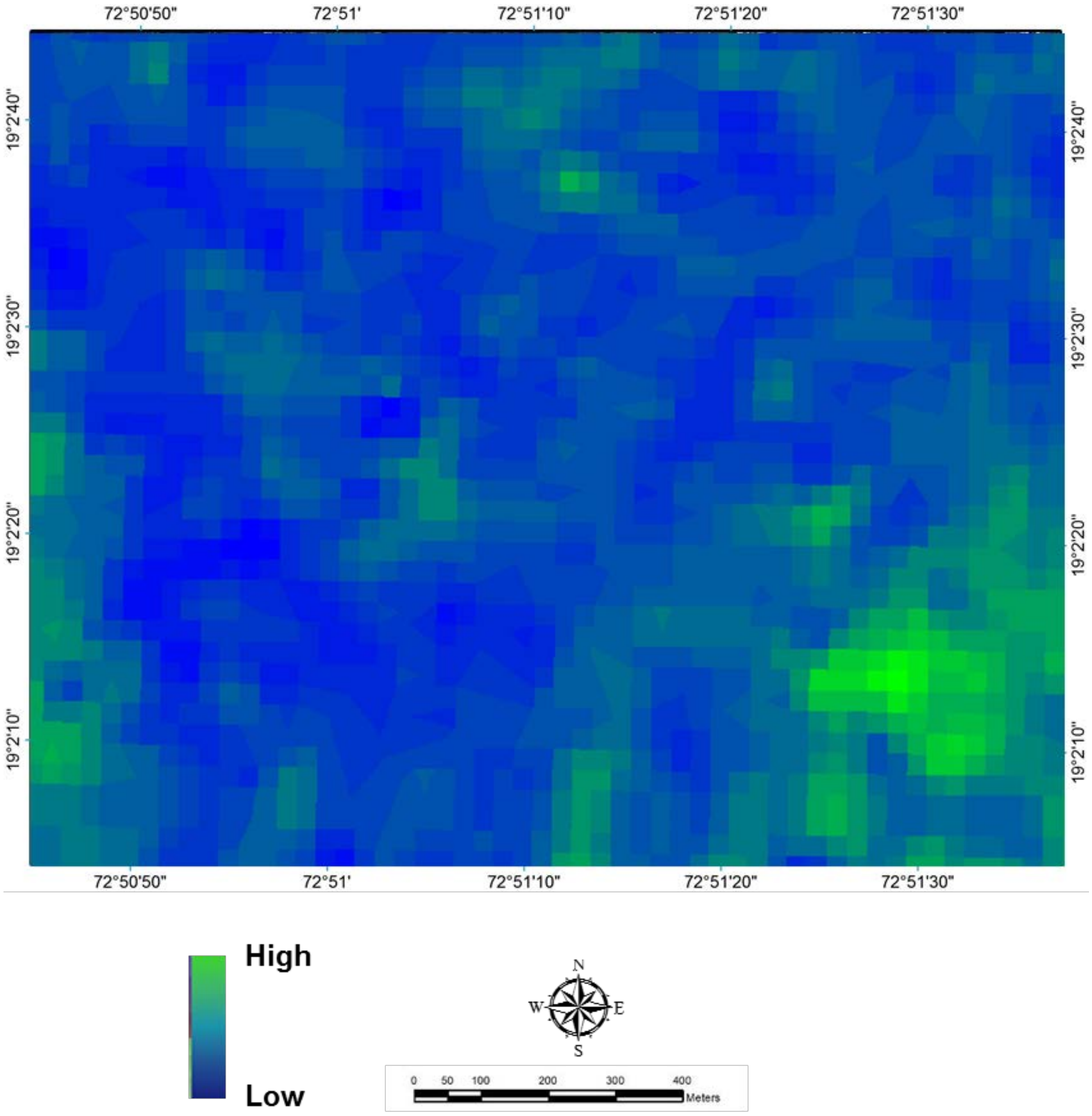
Annexure 10: Descriptions of parameters of rule set for subset 2

Level	Class Feature	OBIA Feature	Membership Function	Feature Type	Lower Border (μ_l)	Upper Border (μ_u)	Value Range (v_r)	Central Value (a)
Object level	Vegetation	NDVI		Spectral index	-0.21	0.65	**0.86**	**0.22**
	Roads	Minimum Pixel Value		Thematic attributes	0.00	1.00	**1.00**	**0.50**
	Water (if existed)	Minimum Pixel Value		Thematic attributes	0.00	1.00	**1.00**	**0.50**
	Shadow	NDNB		Spectral index	-0.53	0.27	**0.80**	**-0.13**
Settlement level	Formal	GLCM Red entropy		Texture	0.02	2.12	**2.11**	**1.07**
	Built-up area	NDNB		Spectral index	-0.34	0.24	**0.57**	**-0.05**
	Formal	Brightness		Spectral	755.83	2743.59	**1987.76**	**1749.71**
	Bare soil	NDGR		Spectral index	-0.22	0.10	**0.31**	**-0.06**
	Formal	Mean (DEM)		Geometry	-2.00	35.56	**37.56**	**16.78**
	Informal	Area		Geometry	3.00	287073.00	**287070.00**	**143538.00**

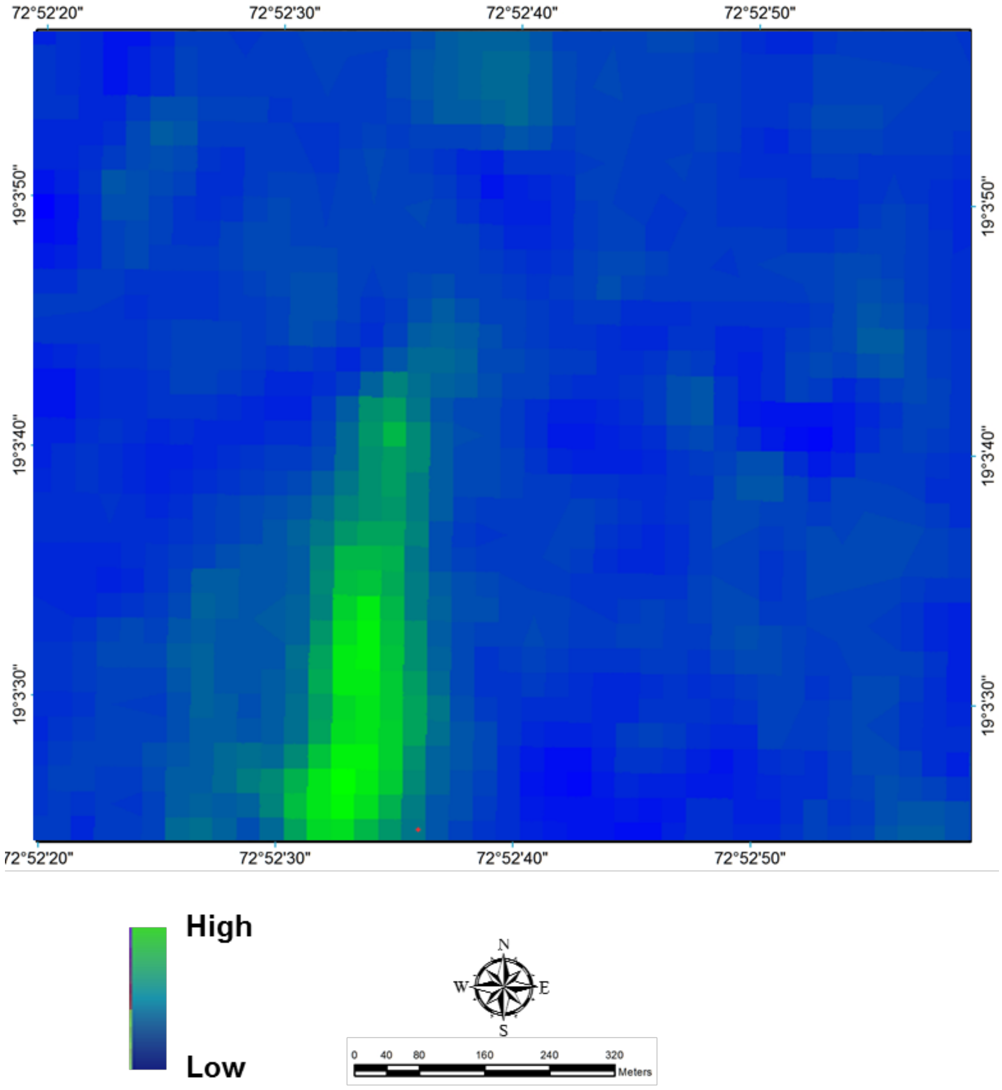
Annexure 11: Descriptions of parameters of rule set for subset 3

Level	Class Feature	OBIA Feature	Membership Function	Feature Type	Lower Border (μ_l)	Upper Border (μ_u)	Value Range (v_r)	Central Value (a)
Object level	Vegetation	NDVI		Spectral index	-0.04	0.50	**0.53**	**0.23**
	Roads	Minimum Pixel Value		Thematic attributes	0.00	1.00	**1.00**	**0.50**
	Water (if existed)	Minimum Pixel Value		Thematic attributes	0.00	0.00	**0.00**	**0.00**
	Shadow	NDNB		Spectral index	-0.38	0.19	**0.56**	**-0.10**
Settlement level	Formal	GLCM Red entropy		Texture	0.00	1.87	**1.87**	**0.93**
	Built-up area	NDNB		Spectral index	-0.26	0.17	**0.43**	**-0.04**
	Formal	Brightness		Spectral	930.33	2392.05	**1461.72**	**1661.19**
	Bare soil	NDGR		Spectral index	-0.17	0.06	**0.23**	**-0.05**
	Formal	Mean (DEM)		Geometry	0.64	11.94	**11.30**	**6.29**
	Informal	Area		Geometry	3.00	233080.00	**233077.00**	**116541.50**

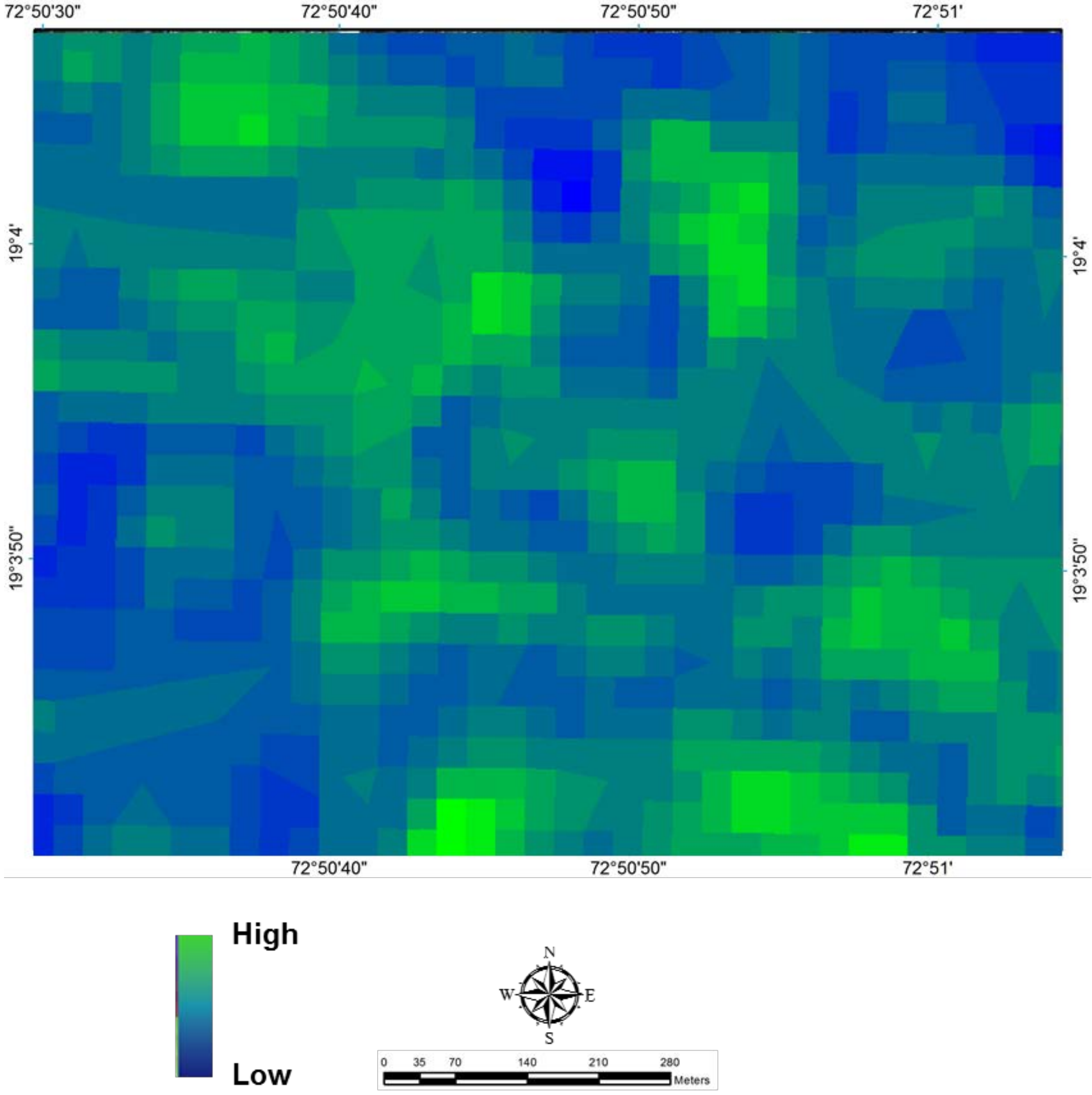
Annexure 12: (DEM) Digital Elevation Model of subset 1



Annexure 13: (DEM) Digital Elevation Model of subset 2



Annexure 14: (DEM) Digital Elevation Model of subset 3



Annexure 15: Classification Accuracy with developed rule set for Subset 1

		Referenced data								
		Vegetation	Shadow	Roads	Bare soil	Formal	Informal	Water	Total	User's accuracy
Classified data	Vegetation	50							50	100%
	Shadow		48				2		50	96%
	Roads			50					50	100%
	Bare soil				36	5	9		50	72%
	Formal				9	38	3		50	76%
	Informal		2		5	7	36		50	72%
	Water							25	25	100%
	Total	50	50	50	50	50	50	25	325	
	Producer's accuracy	100%	96%	100%	72%	76%	72%	100%		
Overall Accuracy	87.08%									
Kappa	0.85									

Annexure 16: Classification Accuracy for Subset 2 before adaptation

		Referenced data								
		Vegetation	Shadow	Roads	Bare soil	Formal	Informal	Water	Total	User's accuracy
Classified data	Vegetation	49			1				50	98%
	Shadow	1	42		2	5			50	84%
	Roads			50					50	100%
	Bare soil		1		20	6	14		41	49%
	Formal		2		24	39	4		69	57%
	Informal		5		3		32		40	80%
	Water							25	25	100%
	Total	50	50	50	50	50	50	25	325	
	Producer's accuracy	98%	84%	100%	40%	78%	64%	100%		
Overall Accuracy	79.08%									
Kappa	0.75									

Annexure 17: Classification Accuracy for Subset 2 after adaptation

	Referenced data									
		Vegetation	Shadow	Roads	Bare soil	Formal	Informal	Water	Total	User's accuracy
Classified data	Vegetation	49			1				50	98%
	Shadow	1	42		2	5			50	84%
	Roads			50					50	100%
	Bare soil				35	9	7		51	69%
	Formal		5		6	36	2		49	73%
	Informal		3		6		41		50	82%
	Water							25	25	100%
	Total	50	50	50	50	50	50	25	325	
	Producer's accuracy	98%	84%	100%	70%	72%	82%	100%		
Overall Accuracy	84.54%									
Kappa	0.83									

Annexure 18: Classification Accuracy for Subset 3 before adaptation

	Referenced data								
		Vegetation	Shadow	Roads	Bare soil	Formal	Informal	Total	User's accuracy
Classified data	Vegetation	48	2					50	96%
	Shadow		43			2	5	50	86%
	Roads		1	48				49	98%
	Bare soil		3		39	8		50	78%
	Formal				8	23		31	74%
	Informal	2	1	2	3	17	45	70	64%
	Total	50	50	50	50	50	50	300	
	Producer's accuracy	96%	86%	96%	78%	46%	90%		
Overall Accuracy	82.00%								
Kappa	0.78								

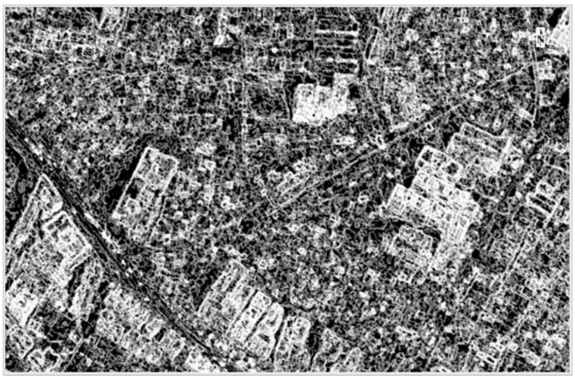
Annexure 19: Classification Accuracy for Subset 3 after adaptation

	Referenced data								
		Vegetation	Shadow	Roads	Bare soil	Formal	Informal	Total	User's accuracy
Classified data	Vegetation	48	2					50	96%
	Shadow		43			2	5	50	86%
	Roads		1	48				49	98%
	Bare soil		3		39	8		50	78%
	Formal				10	38	2	50	76%
	Informal	2	1	2	1	2	43	51	84%
	Total	50	50	50	50	50	50	300	
	Producer's accuracy	96%	86%	96%	78%	76%	86%		
Overall Accuracy	86.33%								
Kappa	0.84								

Annexure 20: GLCM feature different window sizes with the offset of 1 metre

(a) 7X7 window size

(b) 5X5 window size



(c) 3X3 window size