

## FEATURE EXTRACTION FOR URBAN AND AGRICULTURAL DOMAINS USING ECOGNITION DEVELOPER

Mamta Rana, Subham Kharel\*

Regional Remote Sensing Centre-South, NRSC, Indian Space Research Organisation, Bengaluru- mrana.geo@gmail.com,  
\*[kharelsubham1@gmail.com](mailto:kharelsubham1@gmail.com)

Commission III, WG III/10

**KEYWORDS:** Classification, Rule Based, Support Vector machine, Feature Space Optimization. KNN

### ABSTRACT:

Feature extraction has always been a challenging task in Geo-Spatial studies both in urban areas as well as in agricultural areas. After the evolution of eCognition Developer, different segmentation techniques and classification algorithms which help in automating feature extraction have been developed in recent years which have been a boon for scientists and people conducting research in the field of geomatics. This research reflects a study depicting the potential of eCognition Developer in extracting features in Agricultural as well as urban areas using various classification techniques. Rule Based and SVM Classification techniques were used for feature extraction in urban areas whereas Feature Space Optimization and K-Nearest Neighbor were used for classifying agricultural features. Results reflect that rule based classification yields more accurate results for urban areas whereas Feature Space Optimization along with object-based classification gave more accuracy in case of agricultural areas.

### 1. INTRODUCTION

Remote Sensing is the acquisition of information about an object or phenomenon without making physical contact with the object. On a broader sense Remote Sensing is explained to be the art of obtaining information about an object, area, or phenomenon by a device that is not in contact with the area, object, or phenomenon under the study. This concept was totally limited to pixel based classification. Due to advancements in scientific techniques over the years, the trend has shifted to Object Based Image Analysis (OBIA) as pixel based techniques cannot interpret the spatial-photo elements such as texture, context, and shape due to high variability in high resolution satellite images that confuses general pixel-based classifiers and lead to low classification accuracy.

This weakness has been overcome using OBIA which groups a set of pixels into objects based on similar spectral signatures or different variables such as soils, buildings, plots, etc. Over the years, different types of classification techniques have been developed under OBIA in order to simplify and refine classifications for an image. Today, researchers are conducting research on the use of OBIA for classifying vegetative crops, urban areas and also geological features and have obtained higher accuracies as compared to classical pixel based image segmentation/classification.

#### 1.1 Areas of Study

The areas taken for study were Bengaluru and Tiptur both located in the state of Karnataka as displayed in Figure 1. Bengaluru, popularly known as the Silicon Valley of India (lying in between 77°25'00"E to 77°44'00"E longitude and 13°04'00"N to 12°49'00"N latitudes) is having the fifth largest urban agglomeration in India was used for the study involved with classification of urban features.

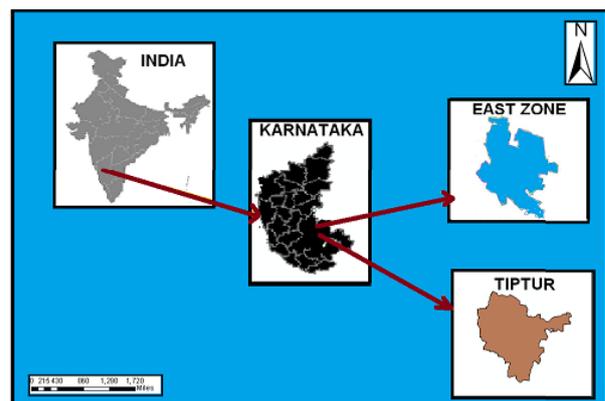


Figure 1: Study Area

Tiptur lying in between 76° 28' 41" E longitudes and 13° 15' 30" N latitudes, is a taluk and sub-divisional head of Tumkur Smart city which is very famous for its coconut plantations, was taken for study involved with the classification of agricultural areas.

#### 1.2 Research Questions

The following research questions were set-up for the study:

- Which object based method is best suited for classification of urban features in Indian cities like Bengaluru?
- Which object based method is best suited for classification of agricultural crops for an area of dense vegetation like Tiptur?
- What is the efficiency and ease of working with OBIA techniques when compared to methods of classification used during older days for classifying objects?

\* Corresponding author

## 2. RELATED WORK

1. Jixian Zhang *et al.* (2013)[1] during their study used OBPA to extract information about LiDAR point clouds. An OBIA is designed to classify airborne LiDAR point clouds in urban areas. Surface growing algorithm is used to make clusters of point clouds. A SVM is used for classifying segments and connecting component to optimize the original results. Three datasets having varying point densities & complexities are used for obtaining results. Experiments show that the method is capable of classifying urban point clouds with accuracy more than 92.34% and Kappa coefficient greater than 0.8638. However, the method fails to extract power lines due to similarities with vegetation. This method has limitations like misclassification of vegetation and buildings and time-consuming feature calculations.[1]

2. B. Mougel *et al.* (2008)[2] reported that VHSR allows monitoring of vegetation at sub-metric and metric scales making individual trees detectable. Hence, disclosing new methods for precise agriculture for orchards & other crops. The authors present methodologies for extraction and classification information about agriculture from satellite images. Plot mapping, tree crop detection, cropping-system characterization, species identification are the objectives set by authors. The method allowed discrimination between tree crop structures and their ages with good accuracies in almost all cases. The author suggests that deeper analysis of vegetation, texture, and other factors can certainly improve accuracy.[2]

3. J. Torres-Sánchez *et al.* (2015)[3] generated an automated threshold algorithm developed in OBIA framework was tested in UAV images taken on different herb crops. The goal was to distinguish vegetation and bare soil in which accuracy of around 90% was achieved. Two cameras namely visible, and visible + infrared was tested. The research aimed at developing an algorithm using OBIA based on Otsu's technique, and explains how the results are affected by segmentation parameters. It was applied for vegetation extraction in satellite images captured with the two sensors and taken over fields of three crops viz., maize, sunflower & wheat. The authors concluded that increasing object size will diminish the object error until an optimum is reached. After this value, increase in object size produces big errors.[3]

4. Rudolph Joshua Candare *et al.* (2016)[4] in the present work describes various methods to map different high value crops in Philippines using LiDAR technology. Because of high resolution data, the author uses OBIA, LiDAR data & Orthophoto. As objects can't be separated by simple thresholds SVM was used. But, SVMs suffers from issues that can significantly affect the results. After performing segmentation in eCognition, the optimization and extraction equations of hyper-planes was carried out in Matlab. The authors reports overall accuracy to be more than 90% in many areas.[4]

5. Neha Gupta *et al.* (2014)[5] demonstrated OBIA feature extraction by the use of eCognition which classifies remotely-sensed data taking object features, like spatial, contextual and spectral information. The author noticed that classification accuracy improved by object-oriented approach. The procedure proves efficient due to the use of object's multi-feature. The OBIA guarantees higher accuracy using high resolution images. Manual adjustment of various parameters makes objects adaptive for required conditions and data. The author proves that the OBIA will become a future trend for the classification of high resolution remotely sensed data.[5]

6. R. T. Alberto *et al.* (2016)[6] quoted that LiDAR and Aerial imagery have great possibilities for Agricultural LULC mapping. But, these images result in pixels which are shadowy. The author tried to Manage shadowed regions for classification without enhancing images. Segmentation was carried out at three scales and tested in order to segment ground features as only ground features are disturbed due to the shadows. Segmentation at level 25 was optimal fit for non-shadowed and shadowed areas. SVM was performed and Radial Basis Function kernel was applied for extracting classes. Training points were selected separately for shadowed areas to create additional classes for reducing misclassifications. The results indicated that accuracy of the LULC done at Level 25 segmentation scale for shadow detection & classification was comparatively higher with respect to higher scales of segmentation.[6]

7. Lei Ma *et al.* (2014)[7] report that accurate and efficient cultivated land information extraction is required for the development of precise agriculture. In recent years UAVs are being used for monitoring natural resources due to their great availability, miniaturized sensors, and the ability to use UAVs at low altitudes. The author proposes an information extraction method for cultivated lands based on TCLE method. More spatial information of a region was used for classification. which comprises of three steps namely triangulation construction, image segmentation, and triangulation clustering by the use of AUTOCLUST. Three UAV images of Deyang, China were experimented using eCognition and TCLE to extract information for cultivated lands. Results of the experimentation depict that there is no need for providing training samples in TCLE method. TCLE has high level of automation and is an equivalently accurate when compared to ECLE. When compared to ECLE, TCLE can extract coherent cultivated land with very less noise.[7]

8. Stéphane Dupuy *et al.* (2012)[8] report that urban sprawl dynamics are strong with growth in population and land crisis. The increase in city area generally called as urban sprawl leads to reduction in agricultural spaces. The irreversible part is the conversion of agricultural land into urban areas. Therefore, spatial-temporal methods for locating and quantifying reduction in agricultural land at local and national scales was need to be done. This study focused on designing a method for monitoring LULC enhanced due to urban sprawl. OBIA was used for extracting artificial areas images and an artificial patch was built periphery artificial areas. Diachronic analysis conducted on artificial patch maps enabled characterization of urban sprawl. 1996–1997 (Indian Remote Sensing) and 2009 (RapidEye) data were used. An increase in artificial areas was noted ranging from 113,000 ha (1997) to 133,000 ha (2009), i.e., 18% in 12 years. The method can be used for homogenous spatial-temporal analysis for assessing LULC in future for agro monitoring.[8]

9. Yafit Cohen *et al.* (2000)[9] show the development of a crop recognition system which is knowledge based, and integrates remote sensing and GIS in a hierarchical fashion. Spot pan and Landsat TM images were merged in order to decrease the heterogeneity after enhancing the field boundaries. NDVI maps that are multi-temporal were generated from the two images was classified in eight different crop types by the use of unsupervised classification. The relations between crop types, natural vegetation, phenological precipitation, spectral and soil types were derived which were used as binary rules in a knowledge-based crop recognition system experimentally. The paper concludes that the using knowledge-based rules gives

improves 9% accuracy when solely compared to unsupervised classification.[9]

10. Valentine Lebourgeois *et.al* (2000)[10] describe that Sentinel-2 images are used as they tend to help in improving monitoring of crops globally in harsh situations where satellite data is disturbed by presence of cloud cover. For reducing these problems, performance of Random Forest (RF) classifier or object-based approach was optimized and analyzed and applied to multisource satellite images for producing LULC maps of agriculture in Madagascar area. Initially, less number of variables was used as inputs. Research was done for analysing the importance of each data source (HSR and VHSR) for classification and quantified the contributions they offered for accuracy assessment. Results indicate that RF classifier optimization reduced the number of variables by 1.5 to 6 fold and also reduced image processing time. Overall accuracy of 91.7% and 64.4% for cropland and subclass levels of crops were achieved using traditional rule based classification at all levels.[10]

### 3. SATELLITE DATA USED FOR THE STUDY

The spatial and non-spatial datasets used during the study are given below:

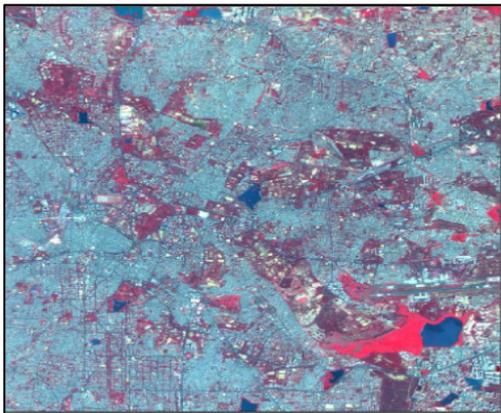


Figure 2: Cartosat 2S and LISS IV Merged Data

The following datasets were used for Bengaluru City:

1. CartoSat-2 data was used for studying features of the East Zone of Bengaluru City
  - a. PAN (Panchromatic with 0.6 m spatial resolution),
  - b. HRMX (Multi spectral camera with 1.6 m spatial resolution)
2. LISS IV data

The following datasets were used for Tiptur:

1. CartoSat-1 & LISS IV Merged data was used for the Tiptur District
  - a. PAN (Panchromatic Camera with 0.6 m spatial resolution),
  - b. HRMX (Multi spectral camera with 1.6 m spatial resolution)
2. LISS IV data



Figure 3: LISS IV image of Tiptur District

## 4. RESEARCH METHODOLOGY

### 4.1 OBIA of East Zone of Bengaluru

i. **Segmentation:** The segmentation was carried out for a portion of Bengaluru City, East Zone at scale 25 in which different band/ mean layer values used were 3, 1, and 1 to enhance the effect of NIR band for Urban Indices generation, shape factor was 0.3 and compactness was 0.7 (considering objects of interest) as shown in Figure 4.

ii.

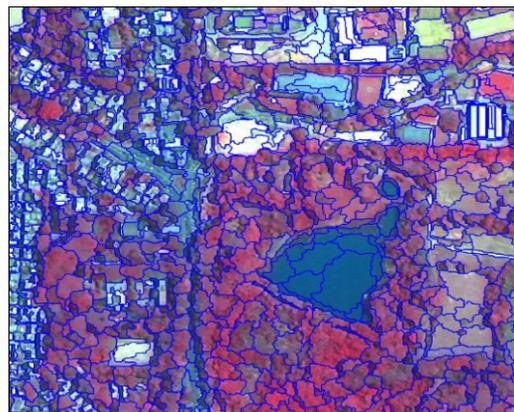


Figure 4: Segmentation (Level 30)

iii. **Classical Rule Based Method:**

**1. Arithmetic Indices:** As a merged image was used, the homogeneity of the layer values were modified. Hence, Using trial and error method two indices were generated for the image:

$$a. \text{INDEX 1} = \frac{\text{LAYER A} - \text{LAYER B}}{\text{LAYER A} + \text{LAYER B}}$$

$$b. \text{INDEX 2} = \frac{\text{LAYER C} - \text{LAYER A}}{\text{LAYER C} + \text{LAYER A}}$$

**2. Classification:** Classification was done considering parameters such as INDEX 1, INDEX 2, Brightness, Mean Layer 3 and Area

**a. Vegetation Mask:** INDEX 1 > 0.23 was the threshold condition used which helped in generating vegetation mask in a broader sense.

**b. Water Mask:** The following three boundary conditions were used for generating water mask:

1. INDEX 1 > 0.39, INDEX 2 < 0.85

2. Brightness  $\geq 20$ , Brightness  $\leq 30-2$
  3. Mean Layer 3  $\geq 30$ , Mean Layer 3  $\leq 42$
- c. Turbid water:** Water mask obtained was not very proper. Hence, an additional class for turbid water was extracted using following conditions:
1. INDEX 2  $< 0.6$
  2. Brightness  $> 25$
- d. Building Layer with Open Spaces:** The following conditions were used to extract urban features which included building layer and open spaces.
1. INDEX 1  $< 0.2$
  2. Brightness  $> 30$ , Brightness  $< 95$
  3. Area  $< 1400$  pixels
- e. Open Space Layer:** Finally threshold conditions were used for separation of open spaces and vacant plots from the building layer, thresholds of which are given below:
1. INDEX B  $> -0.16$
  2. Brightness  $> 50$
- f. Validation:** Validation of the above-generated database was done using earlier building layer from previous projects carried out at the centre.
- iv. Support Vector Machine Classification**  
Three to four training samples were taken for vegetation, water mask, building layer and open spaces & vacant plots and SVM classification was run.



Figure 5: Segmentation (Level 15)

## 4.2 OBIA OF TIPTUR FOR COCONUT EXTRACTION

### 4.2.1 Segmentation:

Multi-Resolution Segmentation at level 15 was for classifying Coconut Area using fused Cartosat-1 and LISS IV data of Tiptur District, Karnataka displayed in Figure 5. The shape factor was 0.4 (weighted as high as necessary) and compactness 0.6 (depending on properties of objects of interest in the satellite image).

### 4.2.2 Feature Space Optimization(FSO):

Feature Space Optimization (FSO) tool applied for feature selection in order to attain ideal results by providing a set of training samples and feature names, the FSO tool determines which features provide the best class separability. NBI (Normalized Blue Index) was generated as an arithmetic index. Normalized Blue Index was then added while classifying since it clearly segregates the water and weed and doesn't interfere with the Dark Coconut Class. Figure 6 displays the methodology followed during the study.

### 4.2.3 KNN Classification(KNN):

KNN classification was carried out for Tiptur by traing different classes of samples for different objects based on parameters such as NDVI and NBI at level 15 segmentation.

### 4.2.4 Accuracy Assessment

The results generated for both urban and agricultural classification were subjected to accuracy assessments in which twenty random points at different spatial locations within the classified output were selected for each generated class as displayed in Figure 7. Accuracies were calculated using the following equations:

$$\text{a. Percentage Accuracy} = \frac{\text{Number of Objects Correctly Classified}}{\text{Total Number of Objects in Sample Space}} * 100 \text{---} \quad \text{---(1)}$$

$$\text{b. Overall Accuracy} = \frac{\sum \text{Percentage Accuracy of Each Class}}{\text{Total Number of Classes}} \text{----} \quad \text{----(2)}$$

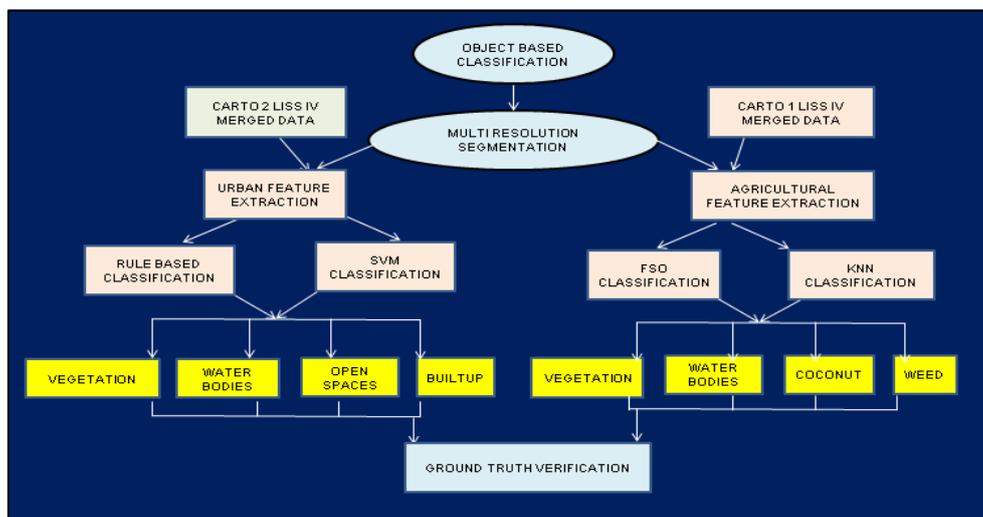


Figure 6: Methodology Flowchart

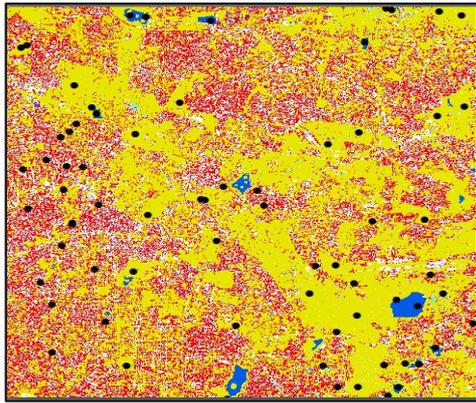


Figure 7 : Rule Based Classification

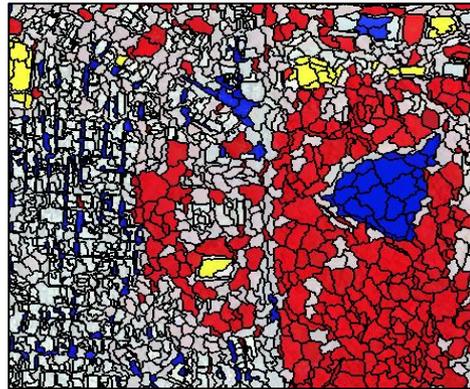


Figure 9: SVM Classification

## 5 RESULTS AND DISCUSSIONS

**5.1 Urban Feature Extraction:** Multi resolution segmentation was attempted for extraction of various urban classes for the study area which was carried out at a scale of Level 30. These parameters were varied in order to get image objects having homogeneous pixels with defined edges of visible objects. The image objects obtained were used for fuzzy rule based classification and extraction. By taking a scale parameter of 30, shape factor of 0.3 and compactness of 0.7, the image objects obtained were of homogenous pattern. It was observed that there was a little problem in extraction of water bodies but better results were obtained immediately after specifying one more class for a water body layer and classifying it using rule based classification. All layers were successfully extracted thereafter. The classified output based on rule based classification for urban features is displayed in Figure 8.

However, according to the accuracy assessment done using the satellite image, results depicted that rule based classification yielded far more accurate results than Support Vector Machine for the East Zone of Bengaluru.

Later SVM classification was run on the same segmented image in which roads and water bodies were creating much problem as both the objects were of darker shades.

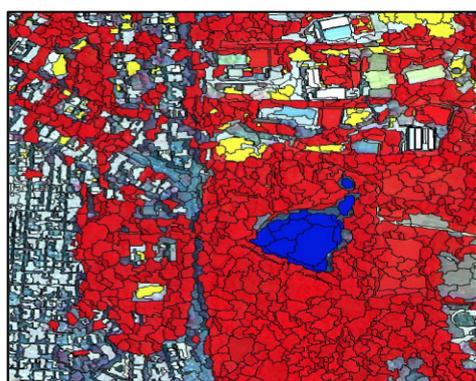


Figure 8: Rule Based Classification

Also, misclassifications between buildings and vegetation were observed. Both the results were checked with reference to a building layer of the city from previous projects. Figure 9 shows the classified outputs of SVM classification for urban extraction.

**5.2 Agriculture Feature Extraction:** After FSO classification 9 Features Selected for Classification. The nine selected features were: Standard Deviation Layer 4, Maximum Differences, Asymmetry, Standard Deviation Layer 3, Border Index, NDVI : Arithmetic Function Generated using mean layer . Standard Deviation Layer 1, Compactness, Mean Layer 4 and NBI as per earlier observations, Weed Classes were not separated clearly in FSO Classification and mixed with water class only. Figure 10 displays the plot for the separation matrix obtained during Bayes Classification and Figure 11 displays the classified output for the same.

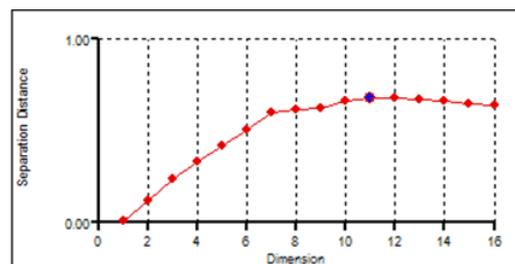


Figure 10: FSO Chart

After attempting KNN classification, it was observed that Barren Patch and other classes is coming more and even Water weed are classified as same class in to FSO/ Bayes and SVM classification. But in KNN and SVM classification, the interested class coconut is separated from other fields (red and smooth texture) better than Bayes Classification.

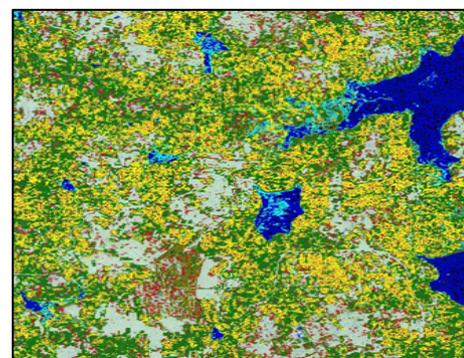


Figure 11: FSO Classification output

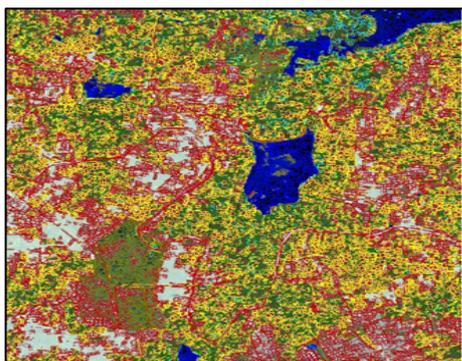


Figure 11: KNN Classification output

Later all these layers were extracted into shape file formats and verified in ArcGIS with respect to satellite data. The classified output for the agricultural areas of Tiptur after KNN classification is shown in Figure 11.

### 5.3 Accuracy Obtained

The results for accuracy assessments for urban and agricultural feature extraction are forecasted in the tables below.

Sl. No.	Class	Total points	Correctly Classified	Percentage Accuracy
1	Water	20	16	80
2	Vegetation	20	16	80
3	Buildings	20	12	60
4	Open Spaces	20	11	55

Table 1: Urban Features Accuracy using Rule Based Classification

Sl. No.	Class	Total points	Correctly Classified	Percentage Accuracy
1	Water	20	17	85
2	Vegetation	20	15	75
3	Buildings	20	14	70
4	Open Spaces	20	12	60

Table 2: Urban Features Accuracy using SVM Classification

Table 1 and Table 2 show the results for the accuracy assessment for urban feature extraction using Rule Based Classification and SVM Classification respectively. Figure 12 shows a plot of accuracy obtained after rule based classification and SVM classification against the separated classes after classification. Overall accuracies obtained after performing accuracy assessment for classified outputs for the East Zone of Bengaluru were 72.5 percent and 68.75 for rule based and SVM classifications respectively.

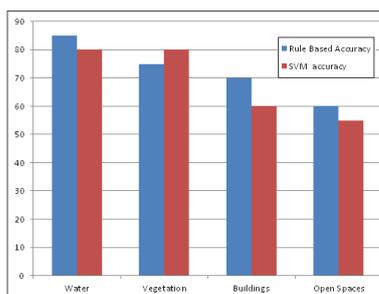


Figure 12: Urban Feature Extraction Accuracy Assessment Bar Plot

Sl. No.	Class	Total points	Correctly Classified	Percentage Accuracy
1	Water	20	17	85
2	Vegetation	20	16	80
3	Coconut	20	16	80
4	Weed	20	12	60

Table 3: Agricultural Features Accuracy using FSO Classifications

Sl. No.	Class	Total points	Correctly Classified	Percentage Accuracy
1	Water	20	17	85
2	Vegetation	20	16	80
3	Coconut	20	14	75
4	Weed	20	13	65

Table 4: Agricultural Feature Accuracy using KNN Classification

Table 3 and Table 4 show the results for the accuracy assessment for agricultural feature extraction using Feature Space Optimization and KNN Classification respectively.

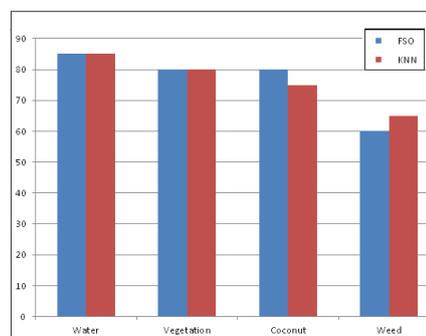


Figure 13: Agricultural Feature Extraction Accuracy Assessment Bar Plot

Figure 13 shows a plot of accuracy obtained after FSO classification and KNN classification against the separated classes after classification. Overall accuracies obtained after performing accuracy assessment for classified outputs for the Tiptur were 76.25 percent for both FSO and KNN classifications.

## 6 CONCLUSIONS

Digitization of different objects of a satellite imagery correctly is a very tedious and time consuming job. Recently developed segmentation techniques have automated this procedure which has made this complex procedure simple. The above analysis clearly describes the use of multi-resolution segmentation for classification of objects in the image using different techniques such as rule based, Support Vector Machine, Knowledge based Nearest Neighbour and Feature Space Optimisation. It was observed from the work that rule based classification was the best technique for urban feature extraction as there is a huge spectral variance between different pixels on the image for Bangalore and making use of SVM, KNN and FSO would be an impractical work. However, for classification of agricultural features, the best suited technique was KNN classification as the dominant class, coconut was extracted more accurately when compared to other classes.

## 7 FUTURE SCOPE OF STUDY

There is no doubt that research has been carried out for classifying objects in both urban and agricultural scenarios but still more work is needed to be done in order to develop newer methods and technologies for classifying even more complex objects such as vegetable plantation extraction, traffic density extraction and many more. Segmentation techniques do have the ability to conduct such rigorous tasks in just fractions of a second and therefore scientists and independent researchers should be encouraged more and more in this field as future trends in classification techniques will completely shift towards object based classification techniques if more accurate results could be obtained.

## 8 ACKNOWLEDGEMENT

Authors are thankful to Director, NRSC and CGM RRSCs and GM RRSC-S for their encouragement. We thank BBMP and other government organizations of Karnataka for providing necessary data for carrying out the research. We extend our sincere thanks to the technical as well as non-technical staff of RRSC-S for providing required support during the course of the work.

## REFERENCES

B. Mougel *et.al* (2008) "Classification and Information Extraction in Very High Resolution Satellite Images for Tree Crops Monitoring", CIRAD/UMR TETIS – Montpellier-FRANCE J.-M. Nicolas Telecom Paris – FRANCE

Jixian Zhang *et.al* (2013) "SVM-Based Classification of Segmented Airborne LiDAR Point Clouds in Urban Areas", *Remote Sens.* 2013, 5, 3749-3775

J. Torres-Sánchez *et.al*(2015), "An automatic object-based method for optimal thresholding in UAV images", *Journal of Computers and Electronics in Agriculture*, Volume 114 Issue C, June 2015, Pages 43-52

Lei Ma *et.al*(2014) "Object Based Agricultural Land Cover Classification Map of Shadowed areas from aerial image and Lidar data using Support Vector Machine ", *Journal of Applied Remote Sensing*, 083673-21, Vol. 8

Lei Ma *et.al*(2012) "An Object-Based Image Analysis Method for Monitoring Land Conversion by Artificial Sprawl Use of RapidEye and IRS Data ", *Remote Sensing*, MDPI, 2012, 4 (2), pp.404-423.

Neha Gupta *et.al* (2014), "Object based Information Extraction from High Resolution Satellite Imagery using eCognition", *IJCSI International Journal of Computer Science Issues*, Vol. 11, Issue 3, No 2, May 2014 ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784

R. T. Alberto *et.al*(2016) "Object Based Agricultural Land Cover Classification Map of Shadowed areas from aerial image and Lidar data using Support Vector Machine ", *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume III-7, 2016 XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic.

Rudolph Joshua Candare *et.al*(2016) "Mapping of High Value Crops through an Object-Based SVM Model Using Lidar Data And Orthophoto In Agusan Del Norte Philippines", *ISPRS*

*Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume III-7, 2016 XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic

Valentine Lebourgeois *et.al* (2000) "A Combined Random Forest and OBIA Classification Scheme for Mapping Smallholder Agriculture at Different Nomenclature Levels Using Multisource Data (Simulated Sentinel-2 Time Series, VHRS and DEM)", *Remote Sens.* 2017, 9, 259

Yafit Cohen *et.al* (2014) "Integration of remote sensing, GIS and expert knowledge in national knowledge-based crop recognition in Mediterranean environment", *International Archives of Photogrammetry and Remote Sensing*. Vol. XXXIII, Part B7. Amsterdam 2000.