Urban Sprawl Analysis in Kutupalong Refugee Camp, Bangladesh

Filip Loncar

Dissertation submitted in partial fulfilment of the requirements for the Degree of Mestre em Ciência e Sistemas de Informação Geográfica (Master in Geographical Information Systems and Science)

NOVA Information Management School

Urban Sprawl Analysis in Kutupalong Refugee Camp, Bangladesh

Dissertation supervised by

Professor Doutor Pedro Cabral

October 2021

DECLARATION OF ORIGINALITY

I declare that the work described in this document is my own and not from someone else. All the assistance I have received from other people is duly acknowledged and all the sources (published or not published) are referenced.

This work has not been previously evaluated or submitted to NOVA Information Management School or elsewhere.

Lisbon, 02 October 2021

Filip Loncar

Digitally signed by cp-LONCAR
Date: 2021.10.08 18:02:54 +02'00'

[Digital signature]

Acknowledgments

First and foremost, I would like to express my gratitude to my family for their unconditional love and support. You are my safe harbour, thank you for believing in me. I hope I make you proud.

To Nadja, you are not only my sister but a best friend, I would not be who I am today without you in my life.

I want to thank my colleagues and supervisors at UNHCR for allowing me to study and telecommute abroad in Lisbon.

I would like to thank my supervisor Pedro Cabral for his support throughout this work, encouragement, and expertise that helped me complete the research until the end. I am honoured to have studied at NOVA Information Management School, and I would like to express appreciation to my professors for the knowledge they gave me.

To my colleagues and friends that I made in Lisbon, to people that I met and created beautiful memories, you hold a very special place in my heart, and you helped me grown so much into the person I am today. To my beauties, my best friends - Eftychia, Ian and Pablo who I love so much.

To Matheus for his love and support. Thank you for being an amazing partner.

Urban Sprawl Analysis in Kutupalong Refugee Camp, Bangladesh

ABSTRACT

Urban sprawling is a common phenomenon associated with geographical and political challenges such as refugee settlements and environmental extremes. Urban sprawl related to refugee or habitation settlement has been an area of active interest because of humanitarian and environmental problems. For example, higher rates of urban sprawling are positively correlated with higher rates of deforestation. The present study explored the viability and reproducibility of different classification techniques in assessing urban sprawl among Rohingya refugees in the Kutupalong refugee camp in South-Eastern Bangladesh. Two classification techniques were used to assess the urban sprawl among the study population. These classifications include the Support Vector Machine and Maximum Likelihood Classifier. The sprawl was measured based on the classification of urban ad non-urban classes, according to the topography of the camps. The study showed that urban class exhibited exponential growth from 2.01 km^2 to 5.37 km^2 within nine months based on Support Vector Machine Classifier, while Maximum Likelihood Classification detected 3.2 km^2 to 7.8 km^2 of urbanization. On the contrary, the non-urban class shrunk from 12.58 km^2 to 9.95 km^2 during the same period with Support Vector Machine and 11.3 km^2 to 6.7 km^2 with Maximum Likelihood Classification. The Support Vector Machine yielded better overall accuracy performance compared to Maximum Likelihood Classification.

KEYWORDS

Urban Sprawl

Refugee Camp

Unmanned Aerial Vehicle

Support Vector Machine

Maximum Likelihood Classification

ACRONYMS

FN – False Negative

FP – False Positive

MLC - Maximum Likelihood Classification

OA – Overall Accuracy

PA – Producer Accuracy

SVM - Support Vector Machine

TN – True Negative

TP – True Positive

UA – User's Accuracy

UAV - Unmanned Aerial Vehicle

Table of Contents

Index of Figures

Index of Tables

1. Introduction

Approximately one million refugees of the Rohingya minority population in Myanmar crossed the border to Bangladesh seeking shelter from systemic operation and prosecution (Faye 2021). This caused significant expansions of the Kutupalong refugee camp within two months and also a reduction in the vegetation in surrounding forests. Different humanitarian and Human Rights Organizations demanded frameworks camp monitoring and environmental impact analysis (Sahana, Jahangir, and Anisujjaman 2019). The refugee camp is situated in Ukhia, CoxBazar, Bangladesh. Remote sensing has become popular in the field of humanitarian action because it is an independent and reliable source of information that allows both a quick response to emergencies and monitoring of gradual changes that are associated with human settlements, including rehabilitation, sprawling, migration, and refuge (Stefan Lang et al. 2020) (A. Braun, Lang, and Hochschild 2016) (Blaschke et al., 2014; Lang et al., 2015). Remote sensing is extremely important when observations in the field are not possible manually due to limited budget, legal barriers, and security aspects (Chen n.d.) . The observation of specific places from space is not only crucial for decision making involving responses to natural disasters and emergencies concerning the human race but also helps to develop a general understanding of an area and the way trends and temporal dynamics have shaped the special and spatial patterns (Chang et al., 2011; Bello & Aina, 2014). The oppressions and extortions of the Rohingya refugees caused the Kutupalong refugee camp to expand (Honest et al. n.d.). The refugee camps nowadays are more permanent than simple transitory settlements, therefore are considered as urban areas. The variables such as size, population density, layout, infrastructure concentration, socio-occupational profile, and trading activities are supporting factors to consider refugee camps as urban areas (Montclos & Kagwanja, 2000). The present work explores the sprawling dynamics of the Rohingya refugees at the Kutupalong refugee camp through remote sensing techniques.

1.1 Objectives and Research Question

The research question focuses on how urbanization has changed following the outbreak of the Rohingya emergency by analyzing four (4) drone images from 2017 and 2018.
The research objective is to answer the following questions:

1. Which machine learning classifier technique yields better performance in urban sprawl classification in Refugee camp context?
2. How much km^2 has urban class increase over the period of one (1) year in Kutupalong Refugee Camp?

2. Literature review

2.1 Image classification in Refugee Camps

Braun in (A. B. Braun 2020) conducted a study that suggested a workflow based on spaceborne radar imagery for measuring the expansion of Rohingya settlements showed a decrease in forest cover. The authors used 11 image pairs of sentinel-1 and ALOS 2 and a DEM (digital elevation model) for supervising land cover classification. The image pairs were trained on automatically derived reference areas retrieved from different multispectral images. Such approaches reduce the need for required user input as well as increase their transferability (Lambert et al. 2013). There was a decrease in vegetation cover of around 1500 hectares, out of which 20% values were attributed to the expansion of the camp, while 80% was due to deforestation. The data matched the findings from the previous studies. The time-series analysis showed that the reduced impact of the seasonal variations on the results and the accuracies reached was between 88% and 95%. The major variables include vegetation indices that were based on synthetic aperture radar backscatter intensity, but the topographic parameters were important too. Labib et al in (Labib and Hossain 2018)evaluated the environmental costs of Rohingya settlements. The focus was to estimate vegetation loss at the Kutupalong-Balukhali expansion area. Change vector analysis was used to estimate the forest cover loss using Landsat 7 images. Then the authors used the carbon sequestration capacity of the respective forest land to estimate the total loss. The study showed that 572 hectares were deforested to set up camp that accounted for an approximate loss of 365,288 Great Britain Pounds per year for the Bangladesh government. There were significant differences in the findings of Labab et al. in (Labib and Hossain 2018) that of Braun (A. B. Braun 2020) because the latter reported around 1200 hectares of land were deforested while the former showed that 575 hectares were deforested. Such differences in the findings of the two authors are not surprising considering the fact that Labab (Labib and Hossain 2018)measured the area of the land that was initially set up what the development of the camp. On the other hand, Braun (A. B. Braun 2020) estimation was based on time-series data

after one year of Labab (Labib and Hossain 2018) as well as their data depicted the urban sprawl features of the Rohingya population within a span of one year. Finally, it would also be possible that the Remote Sensing imaging techniques might have contributed to the differences in the assessment of the changes in land area as estimated by the two authors during two different periods.

2.2 Drone Imagery in Refugee Emergencies

The term drone refers to different types of aerial vehicles ranging from small enough to be handheld to those that could be used for dropping bombs. The use of drones in the history of humanitarian aids could be traced back to 2014, when they were used to track and identify disaster-affected areas, search and rescue operations, and procurement and the delivery of aid materials (Yaacoub et al. 2020). The cutting-edge technology used by drones has made it a robust choice for mapping, surveying, and delivering rescue services (Burgess 2018). On the other hand, the maneuvering of drones and monitoring of marginalized populations such as refugees draw ethical dilemmas (Horsman 2016). It is contended that the marginalized voices of the disaster victims must be protected, while the expectations of individuals suffering from the apprehensions of security due to flying drones and satellite imagery should be optimized (Wang, So, and Smith 2015). Nevertheless, the use of drones has been acknowledged as a major facilitator for reducing conflicts as well as an approach for satellite imagery (Rotte 2016). For example, the satellite sentinel project was launched to monitor threats to security across individuals habituating the Sudan-South Sudan border and the clashes in Mali (Charbonneau, 2017 ; Heisbourg, 2013). To a large extent, the use of satellite imagery and drone images prevented conflicts (Villa et al., 2016 ; Rohi et al., 2020) . The satellite sentinel project has been in practice since 2010 and is used to build maps and software. They are used to map forest cover, too (Hansen et al. 2008) Click or tap here to enter text.

2.3. Remote sensing in monitoring the urban expansion of refugee camps

The use of high spatial resolution satellite images for efficient support of refugee settlement and planning humanitarian aid is growing (van Westen 2013; Boccardo 2013 ; Quinn et al. 2018). A study explored the role of an integrated approach for dwelling classification from VHR satellite images by applying convolutional neural network models as input data for object-based image analysis knowledge-based semantic method (Corbane et al. 2021). Unlike the standard pixel-based classification methods that are used ad applied for the CNN model, the integrated approach aggregate CNN. On the other hand, the object-based accuracy of spatial images is widely used in the mapping (Denis et al. 2016 ; Tiede et al. 2017). Accuracies were as high as 90% for each applied parameter of precision, recall, and F1 score (Corbane et al. 2021). Therefore, it is concluded that synchronizing the CNN models with the OBIA capabilities is necessary for carrying out dwelling extraction and classification through algorithm refinements (Friedl et al. 2010).

2.4 Support Vector Machine

Support vector machine classifier is among the many classifiers that are unanimously acknowledged in the field of remote sensing (Hassan et al. 2020). It is a binary classifier and is dependent on the concept that training samples, which are nearer concurrences to the boundaries of the class. The SVM classifier focuses on finding the optimal hyper-plane that separates the samples of the training input into many classes, and the samples of training data are close to the boundaries of the class and at a lesser distance to hyper-plane are taken as the support vectors that are to be used for actual training (Hassan et al. 2020) . Considering the dynamic and convoluted model of urban expansion and urban sprawl rate, it is necessary to handle continuous and categorical variables as well as non-normalized data and non-linear relationships (Karimi et al. 2019 ; A. Braun and Hochschild 2017a). These bottlenecks are

addressed through the binary support vector machine that is configured by regulating a penalty parameter by selecting the best kernel function's parameter.

2.5 Maximum Likelihood Classification and Urbanization

Maximum Likelihood Classification (MLC) is a statistical estimate that each class in each band is normally distributed, and the probability of a given pixel belonging to a specific class is calculated (Sisodia, Tiwari, and Kumar 2014). Mapping and monitoring of urban changes through remote sensing are difficult because of the complex land-use patterns. Although the different image processing techniques have been administered, they are confounded by atmospheric conditions, illumination, and surface moisture. MLC is one of the most popular methods of remote sensing, where the likelihood of a pixel is defined as the posterior probability. The method has an advantage over the probability theory, but care should be taken so that sufficient ground truth data is taken into account for allowing the estimation of the mean vector and the variance-covariance matrix of the population that is investigated. Moreover, the inverse-matrix of the variance-covariance matrix could become unstable when there is a high correlation between two bands or if the ground truth data is homogenous. Under such circumstances, the number of bands should be verified and reduced through the principal component analysis and regression (Hogland, Billor, and Anderson 2013). If the distribution of the population is not aligned with the normal distribution, the MLC method is not applicable as an imaging technique. On the other hand, artificial neural networks using deep learning techniques are used to address the limitations related to the assumptions required for MLC. Monitoring land cover and urbanization have been the major hallmarks of remote sensing. The changes in land cover over time and space helps to assess urbanization and urban sprawl. In this regard, the Landsat imagery has continuously evolved from its inception. The evolutions were primarily based on statistical assumptions and classification algorithms mediated through pixel resolution. The reliability of land cover and land usage patterns from remotely sensed data is dependent on the accuracy and reproducibility of different classification parameters (Rimal et al. 2019). Rimal in (Rimal et al. 2019) compared two land usage imageries (SVM and MLC) to map urbanization in the Kathmandu

Valley, Nepal. The study was conducted with the land usage imageries in the referred geographic location from 1988 to 2016. The authors showed that SVM was a more accurate classifier of urbanization and land usage patterns compared to MLC. The advantage of SVM on the accuracy of land usage data is attributed to the training of the image samples and not on the laws of probability for explaining a pixel. These findings suggest that land usage patterns or urbanization dynamics should be estimated through different Landsat imagery approaches (A. Braun and Hochschild 2017b). Such assumptions would enhance the reliability of the findings.

3. Data and methods

This chapter describes the methodology used to develop the research and introduces the study area and data used. The diagram in Figure 1 explains the workflow of the research. The study is based on four different images from 4 different dates to analyze urbanization within the area of sprawl. Moreover, this research explores different remote sensing supervised classification methods (Support Vector Machine and Maximum Likelihood Classification) to see which performs better analyzing urbanization in refugee settlements.

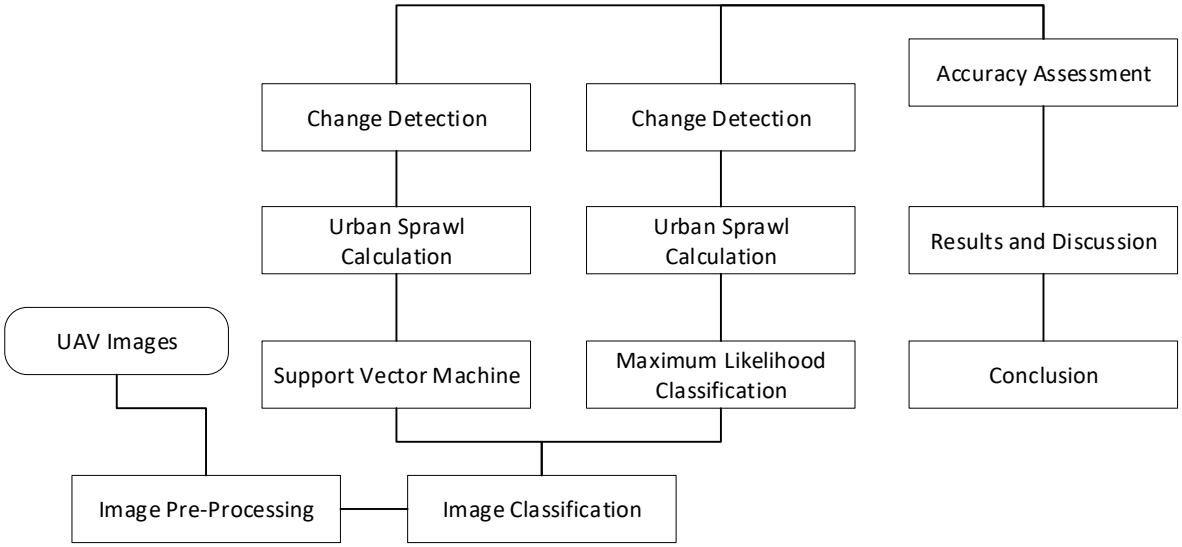

Figure 1 Data and methodology used for image classification and result validation

3.1 Data

3.1.1 Remote Sensing Data

The methodology for the present study included acquisition of the images from IOM Bangladesh – Needs and population Monitoring (NPM) Cox's Bazar Rohingya Refugees

Settlements UAV Imagery which were stored at data.humdata.org. Images were chronologically taken from December 2017 to September 2018 following the outbreak of violence in August 2017 in Rakhine State, Myanmar. The Imagery type is UAV with the resolution of 10 cm, projection type WGS84 _Zone 46 N. While there were images from the entire Cox's Bazar Refugee and different camps stored in repository, this research only focuses on the Kutupalong camp, which is the world's biggest of the camp (according to United Nations 2020 Review) that located in Cox's Bazar region.

3.1.2 Refugee camp extent

The spatial database in shapefile format with outline of camps, settlements, and sites where Rohingya refugees are staying in Cox's Basar has been acquired from data.humdata.org provided by the Inter Section coordination Group. The database contains the camp-block boundary (admin level-2 or camp sub-division) of Rohingya refugees in Cox's Bazar, Bangladesh. Since the research is excluding the surrounding camps, the shapefile was cut leaving out only the Kutupalong camp and its extensions.

3.2. Study Area

Kutupalong refugee camp is located in southeastern Bangladesh along the border with Myanmar. The camp administrative area is defined by following coordinates 21.2126°N 92.1634°E , and with the total area of 14.5 km2 it hosts population of 860,356 registered refugee individuals and 187,423 families (as of June 30, 2020, UNHCR). When Rohingya minority left Myanmar's adjacent Rakhine state as a consequence of religious and ethnic persecution, which culminated in brutal crackdowns and systematic executions beginning on August 25, 2017, Kutupalong became the world's biggest refugee camp. The study area in this research focuses on Kutupalong camp and its extensions located in Ukhia Upzalla in district of Cox's Bazar region and its extensions. The Figure 2 below represents the study area

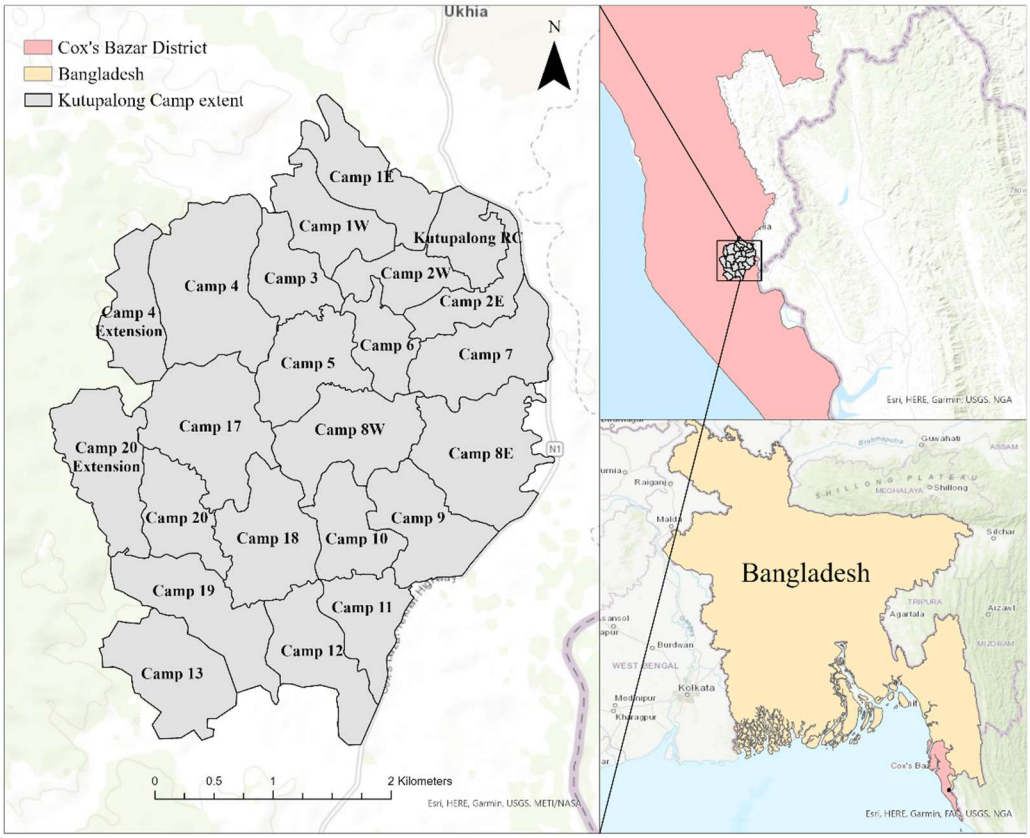

Figure 2 Study Area

The shapefile of study area is divided per camp into 23 sub-regions (polygons) with total area consisting of 14.5 km^2. The biggest camp being Camp 4 with area of 1.15 km^2 being on the northwest side of the study region to smallest being camp 6 in mid-eastern side with 0.36 km^2. Table 1 shows overview of camps and area in km^2

Camp name	**Area km^2**
Camp 4	1.155838
Camp 8E	0.957176
Camp 17	0.954713
Camp 8W	0.772624
Camp 19	0.770065
Camp 20 Extension	0.766568
Camp 13	0.754222
Camp 18	0.752133
Camp 7	0.714525

Camp 9	0.649482
Camp 1E	0.63397
Camp 12	0.631523
Camp 5	0.615671
Camp 1W	0.534719
Camp 4 Extension	0.497775
Camp 10	0.496448
Camp 20	0.489401
Camp 11	0.466304
Camp 3	0.453837
Camp 2W	0.392101
Camp 2E	0.391093
Kutupalong RC	0.387556
Camp 6	0.361309

Table 1 Camp names and total area (km^2)

Figure 3 Aerial Image of Kutupalong Refugee Camp (UNHCR 2020)

3.3. Image classification

This section explains the process and methods used to classify images the UAV imagery of our study area. The initial research and classification schema was focusing on land cover and land use changes in the study area and detect changes of physical attributes over the time. However, due to physical nature of the study area, infrastructure, barren land, different types

of waters and higher computing power required to produce results, the researcher decided to only focus to 2 classes instead of initial 6. The initial results were unsatisfactory, and it was difficult to differentiate how the camps developed. Based on previous studies done with classifying urbanization it was concluded that the supervised classification is superior to unsupervised and produces better accuracy. However, unsupervised classification was performed regardless to see the map output and results. Running the unsupervised classification with 3 classes however showed aggregation of barren land and urban areas, water bodies and vegetation, and high part of misclassified urban land as non-urban. The researcher decided to dismiss the map outputs of unsupervised classification and proceed with supervised.

3.3.1 Supervised Classification

The suggested technique consists of a two-different classification assessment. The performance of various classifiers with found distinct dataset combinations from different dates in the same study area, to determine which of the two will yield superior results for urbanization. As the aim of the research is to understand which supervised classification provides best results for mapping the urbanization in refugee settlements, MLC and SVM were tuned in.

Two detect and understanding urbanization, the researcher has decided to create two classes – Urban and Non-Urban. The table 2 explains the categories and definitions of the classes:

Class	Level 1 class	Description
Urban	Residential	Refugee housing units
	Commercial	Residential areas
	Industrial	Warehouses
Non-Urban	Agriculture	Farmlands
	Green Space	Grasslands, shrublands
	Waterbody	Natural or artificial waterbodies
	Undeveloped	Vacant land, bare land or land under construction

Table 2 Classification nomenclature

3.3.2. Image classification accuracy metrics

When evaluating the effectiveness of a classification model applied to remote sensing data, accuracy measures are used to determine how near the model's predictions are to reality. As a result, accuracy evaluation compares the predicted labels assigned to an item using an ML classifier to its actual label using ground-truth data (test dataset).

A confusion matrix (or error matrix) is commonly used to determine classification accuracy. The classification accuracy is confusion matrix table which shows correspondence between the classification result and reference image (images being ground truth data in this research). This enables for more in-depth examination than a simple fraction of right classifications (accuracy). If the data set is imbalanced, that is, when the number of observations in various classes varies considerably, accuracy will produce false results (Maxwell & Warner, 2020). Table 3 illustrates an example of binary confusion matrix for a two-class problem. The 'True Positives' are values which are classified by the model as 'Positive', and are really 'Positive', while 'False Positive' are values which are classified as 'Positive' but are actually 'Negative'. With the same reasoning, we can understand that terms 'True Negative' and 'False Negative' mean.

Actual class	Predicted	
	Positive	Negative
Positive	True Positive	False Negative
Negative	False Positive	True Negative

Table 3 Binary confusion matrix(Luque et al. 2019)

From the confusion matrix, and following research development, we can compute several accuracy metrics, such as:

- Overall Accuracy : calculated by summing the number of correctly classified values and dividing by the total number of values. The correctly classified values

are located along the upper-left to lower-right diagonal of the confusion matrix. The overall accuracy value is represented in percentage (%) and is calculated as follows:

$$Overall\ Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

- User's accuracy: The probability is calculated by dividing the number of properly predicted values by the total number of values projected to belong to a class. User's accuracy is from the standpoint of the map user.

$$User\ Accuracy = \frac{TP}{(TP + FP)}$$

- Producer accuracy: The number of properly identified pixels in each category (on the major diagonal) divided by the number of reference pixels "known" to be of that category yields the following results (the column total).

$$Producer\ Accuracy = \frac{TP}{(TP + FN)}$$

- Kappa Coefficient: The kappa coefficient assesses the degree of agreement between categorization and truth values. A kappa value of one indicates complete agreement, whereas a value of zero indicates no agreement.

3.3.3. Urban Sprawl Analysis with Shannon's Entropy

The study aims to analyze the process of the bult-up camp sites over period of one (1) year in Cox's Bazaar, Bangladesh. While there is a wide variety of metrics that are used to measure the degree of an urban sprawl, this research has adopted the Shannon's Entropy Index. The Shannon Entropy is an index or indication which is capable of computing spatial concentration or dispersion in any spatial unit . The entropy values vary from 0 to 1, which 0 meaning that entropy values are maximally concentrated in one region, while 1 means that values are unevenly dispersed across space. The entropy value increases as built-up regions are dispersed from a city core or road network. This demonstrates whether the urban growth is more scattered or dense (Tewolde & Cabral, 2011). The model is calculated using the formula below:

$$Hn = -\sum_{i=1}^{n} pi \log(1/pi)/\log(n)$$

Where $pi = xi \,/\, \sum_i^n xi$ and xi is the density of land development, that is equals to the amount of built-up land divided by the total amount of land in the i^{th} of n total zones.

4. Results and Discussion

4.1 Support Vector Machine

Support Vector Machine (SVM) supervised classification has been performed over the 4 UAV Images from Kutupalong Refugee Camp region from 4 different dates to detect urban sprawl and camp expansion. Since Analyzing UAV images with high resolution requires stronger computing power, and after multiple failing attempts to perform SVM classification using original resolution, image size was reduced using resample function to X being 0.5 and Y being 0.5.

Image segmentation is key component in object-based classification workflow. The segmented images produced have grouped neighboring pixels together that are similar in color and shape. The segmented images produced for 4 different dates were acceptable. Following acceptance of segmented images, the training samples for 2 classes were collected for urban and non-urban class. SVM has been performed with default 500 maximum number of samples per class, with active chromaticity color and mean digital number activated.

Maps in *Figure 2* are outputs of Support Vector Machine supervised classification. The red polygons are representing urban areas, while the grey ones are the non-urban.

Upon completion of classification, the researcher will calculate the actual classified areas in square kilometers from all the 4 map outputs to compare the results. Firstly, the SVM classified maps will be converted to vector layer that is shape file, using reclassified and export raster to polygon functions. Upon conversion, using polygons that have same class were merged Calculate Geometry Attributes function was performed in order to n calculate the total square kilometer area of each class and compare between the dates.

Class/Date	24-12-17	12-02-18	31-07-18	24-09-18
Urban	2.01	2.91	4.63	5.37
Non-Urban	12.58	11.68	9.22	9.95

Table 4 SVM: Representing number of square kilometres (km2) for each of the classes in each date

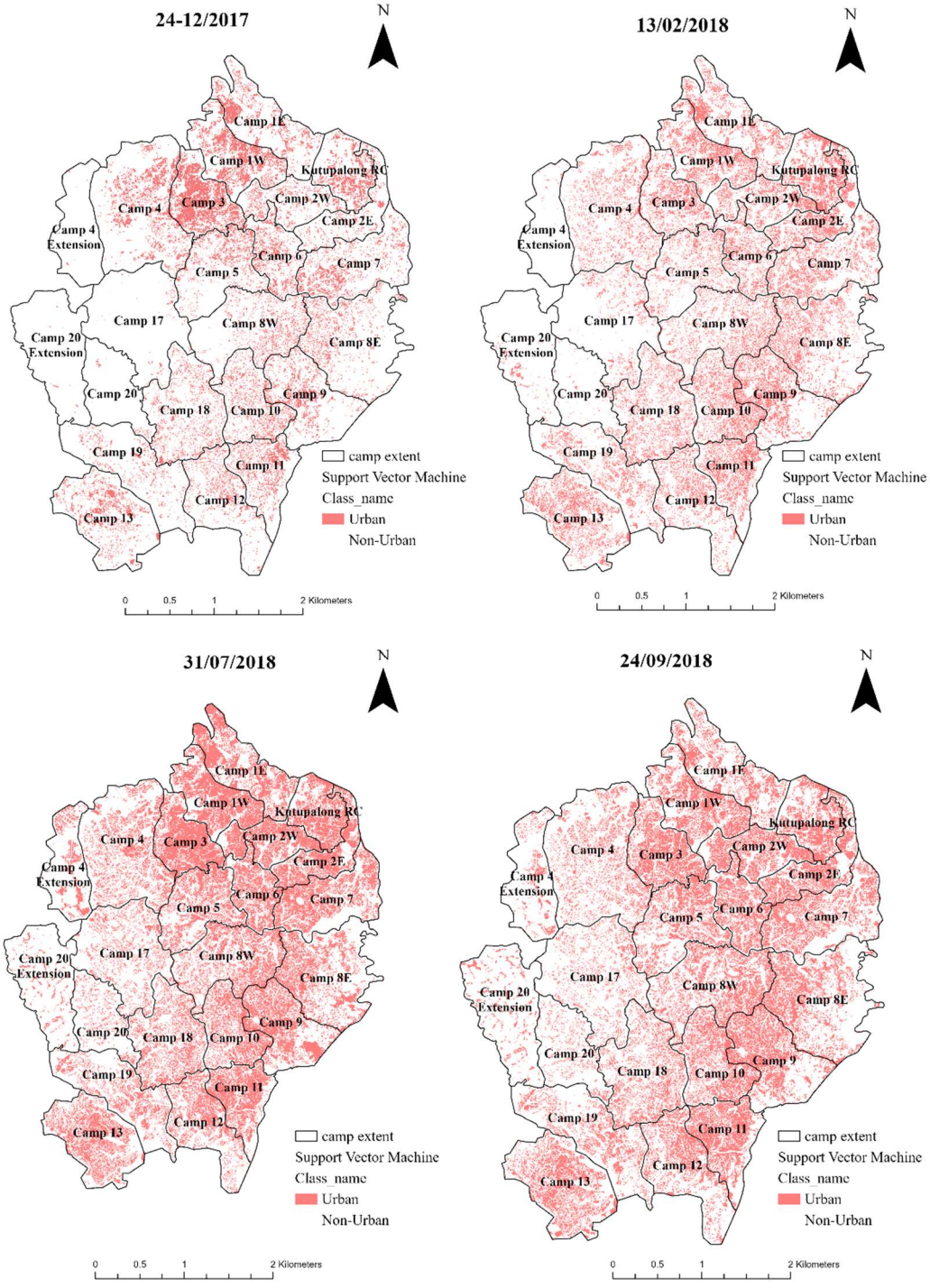

Figure 4 Support Vector Machine - Kutupalong Refugee Camp

With above table representing number of km^2 for four different dates for each of our classes, we can conclude that the urban class has an exponential growth for only around 9 months from 2.01 km^2 on the first image of 24/12/2017 to 5.37 km^2 on the last image of 24/09/2018, which reads 13.7% of total area on the first image to 35% total on the last image taken. The non-urban class however has reduced from 12.58 km^2 to **9.95** km^2 (from 86 % of the total area to 65%). The SVM classifier is showing higher are for non-urban class on 31-07-18 date compared to 24-09-18 date (9.22 km^2 to 9.95 km^2)

Even though the time spam of the 3rd and 4th image is not the longest (2 months compared to 2nd and 3rd image which is around 5 months) 2e can notice the highest growth of urban class between 31-07-18 and 24-09-18 , from 4.63 km^2 to 5.37 km^2 .

4.1.1 Change Detection – Support Vector Machine

In order to visualize the urban sprawl and detect changes between the dates, Change Detection Wizard of ArcGIS Pro has been utilized. The categorical change method of change detection has been configured over the 4 classified rasters images. Processing was set to full extent and class configuration is as follows:

- From Classes: Non-Urban
- To Classes: Urban and Non-Urban

This would give us the outputs where Non-Urban pixels have changed to Urban class, and where Non-Urban Class did not change. The smoothing Neighborhood was set to Non and output images are shown in Figure 3 and Figure 4 below. Figure 3 represents change analysis of subsequential images from first to last date, while the Figure 4 displays change analysis from the first date against the last.

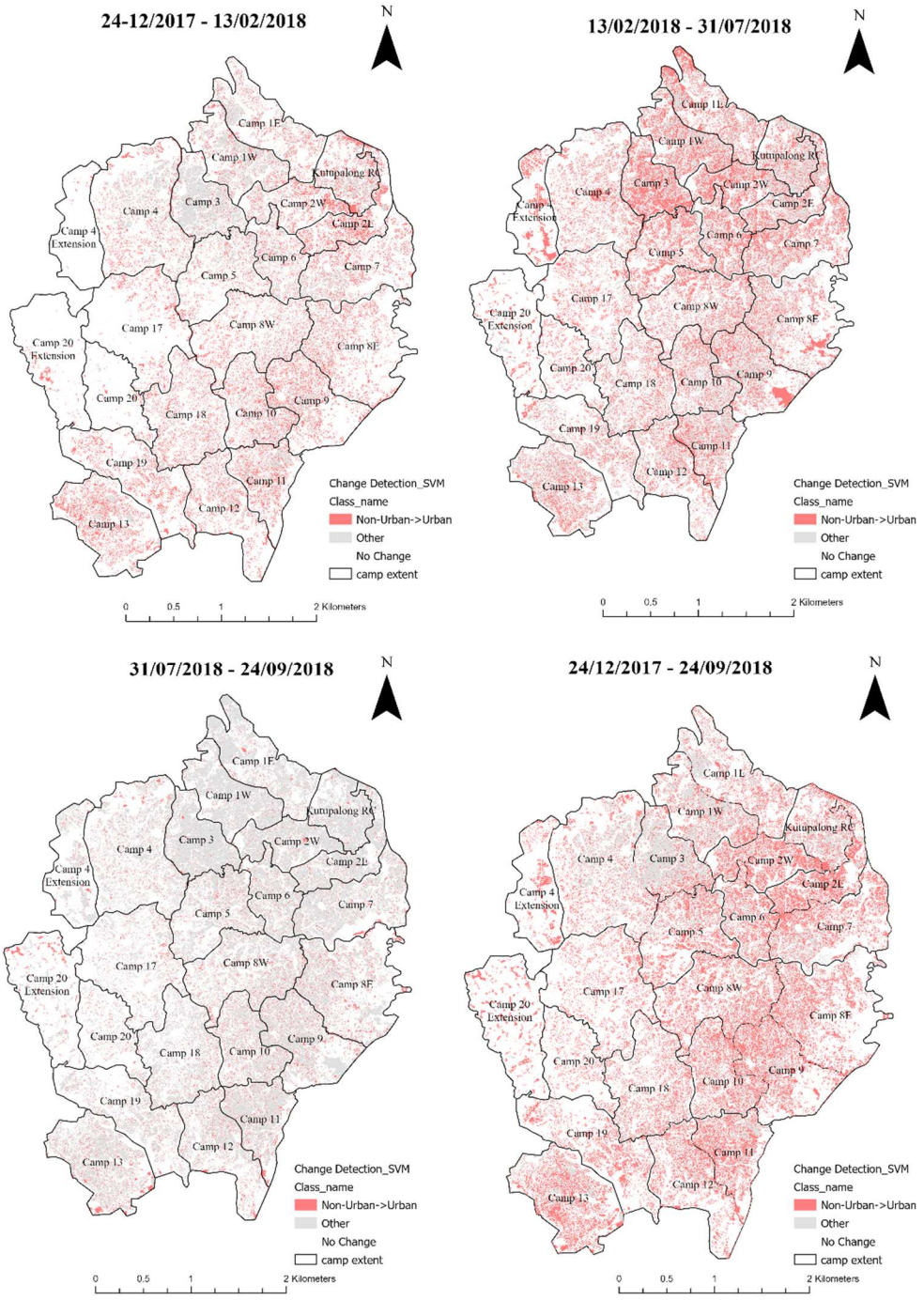

Figure 5 Change Detection Maps - Support Vector Machine

With overlaying the study area extent which has information on sub-camps and the classified raster image, we can compute the camp-vise statistics to understand the direction urban sprawl is taking. The below *table 5* represents camp wise area statistics and percentage of the urbanized units of the SVM supervised classification for four (4) different dates. The table for each class and each date shows area (in square kilometers) and percentage of the class withing the camp.

In terms of urban class, the Camp 3, Camp 1W, Kutupalong RC and Camp 6 had the highest value of urban class on the first image (24/12/2017) with 53%, 32%, 25% and 23% respectively, while in case of the last image (24/09/2018) Camp 10, Camp 11 , Camp 12 and Camp 13 showed highest values of urban class with 70%, 59%, 58% and 57% respectively. However, the highest increase of urban class withing camps comparing between the first and the last image can be noticed in Camp 10, Camp 11, Camp 12, and Camp 13 with increase of 45%, 43% 33% and 32% respectively. These camps that show the highest change in terms of urbanization are all located in the south of the study area.

Table 5 Camp wise area statistics of the urbanized units. The SVM supervised classification was used to analyse the urbanization.

Camp name	24-12-17		13-02-18		31-07-18		24-09-18	
	Urban (km^2)	Non-Urban (km^2)	Urban (km^2)	Non-Urban (km^2)	Urban (km^2)	Non-Urban (km^2)	Urban (km^2)	Non-Urban (km^2)
Camp 10	0.09	0.4	0.16	0.34	0.23	0.27	0.23	0.27
Camp 11	0.1	0.37	0.15	0.31	0.24	0.22	0.25	0.22
Camp 12	0.07	0.56	0.12	0.51	0.2	0.42	0.21	0.43
Camp 13	0.11	0.64	0.19	0.56	0.28	0.46	0.3	0.47
Camp 17	0.01	0.94	0.06	0.89	0.13	0.79	0.17	0.83
Camp 18	0.1	0.65	0.16	0.59	0.21	0.5	0.25	0.54
Camp 19	0.05	0.72	0.12	0.65	0.16	0.59	0.18	0.61
Camp 1E	0.11	0.52	0.14	0.49	0.19	0.33	0.31	0.45
Camp 1W	0.17	0.36	0.17	0.36	0.24	0.22	0.31	0.3
Camp 20	0	0.48	0.01	0.47	0.05	0.43	0.06	0.44
Camp 20 Extension	0	0.76	0.02	0.75	0.06	0.72	0.05	0.7
Camp 2E	0.04	0.35	0.13	0.26	0.19	0.18	0.22	0.2
Camp 2W	0.06	0.33	0.1	0.29	0.19	0.16	0.23	0.2
Camp 3	0.24	0.21	0.13	0.32	0.25	0.14	0.32	0.21
Camp 4	0.18	0.98	0.18	0.98	0.24	0.84	0.31	0.92
Camp 4 Extension	0	0.5	0	0.49	0.05	0.41	0.09	0.45
Camp 5	0.09	0.53	0.1	0.52	0.22	0.37	0.24	0.39
Camp 6	0.08	0.28	0.09	0.27	0.17	0.17	0.19	0.19
Camp 7	0.12	0.6	0.17	0.54	0.27	0.39	0.32	0.45
Camp 8E	0.1	0.86	0.19	0.77	0.29	0.65	0.31	0.66
Camp 8W	0.09	0.69	0.16	0.62	0.3	0.47	0.3	0.48
Camp 9	0.1	0.55	0.19	0.46	0.29	0.33	0.32	0.36
Kutupalong RC	0.1	0.29	0.15	0.24	0.18	0.17	0.22	0.21

4.2 Maximum Likelihood Classification

Maximum likelihood classification has been performed over the four images from different dates in the study area. Following the same methodology as for the SVM classifier, the researched is interested to see the urbanization over the study area in Kutupalong Camp, Bangladesh , two different classes were used – Urban and Non-Urban. Training samples in form of polygons were collected for each class as follows. The MLC showed good performance in classifying very high spatial resolution images as it was not failing during the classification run process, unlike the SVM classification .

The classification process was lengthy, but if was finalized successfully. However, in order to fairly compare the two classifications, the we recreated the MLC classification from the scratch, with resampling the images to have the same size and resolution and the ones used in SVM classification. This way, we would fairly compare the classification performances.

The maps in *Figure 7* below are the results of supervised classification using a Maximum Likelihood Classification. The blue polygons indicate the classified urban class over the study area.

After classification is complete, the actual classified area in square kilometers for all four maps has been calculated, to understand the classification by calculating the geometry attributes of urban and non-urban classes.

Class/Date	24-12-17	12-02-18	31-07-18	24-09-18
Urban	3.2	3.6	5.25	7.8
Non-Urban	11.3	10.9	9.25	6.7

Figure 6 MLC: Representing number of square kilometres (km2) for each of the classes in each date

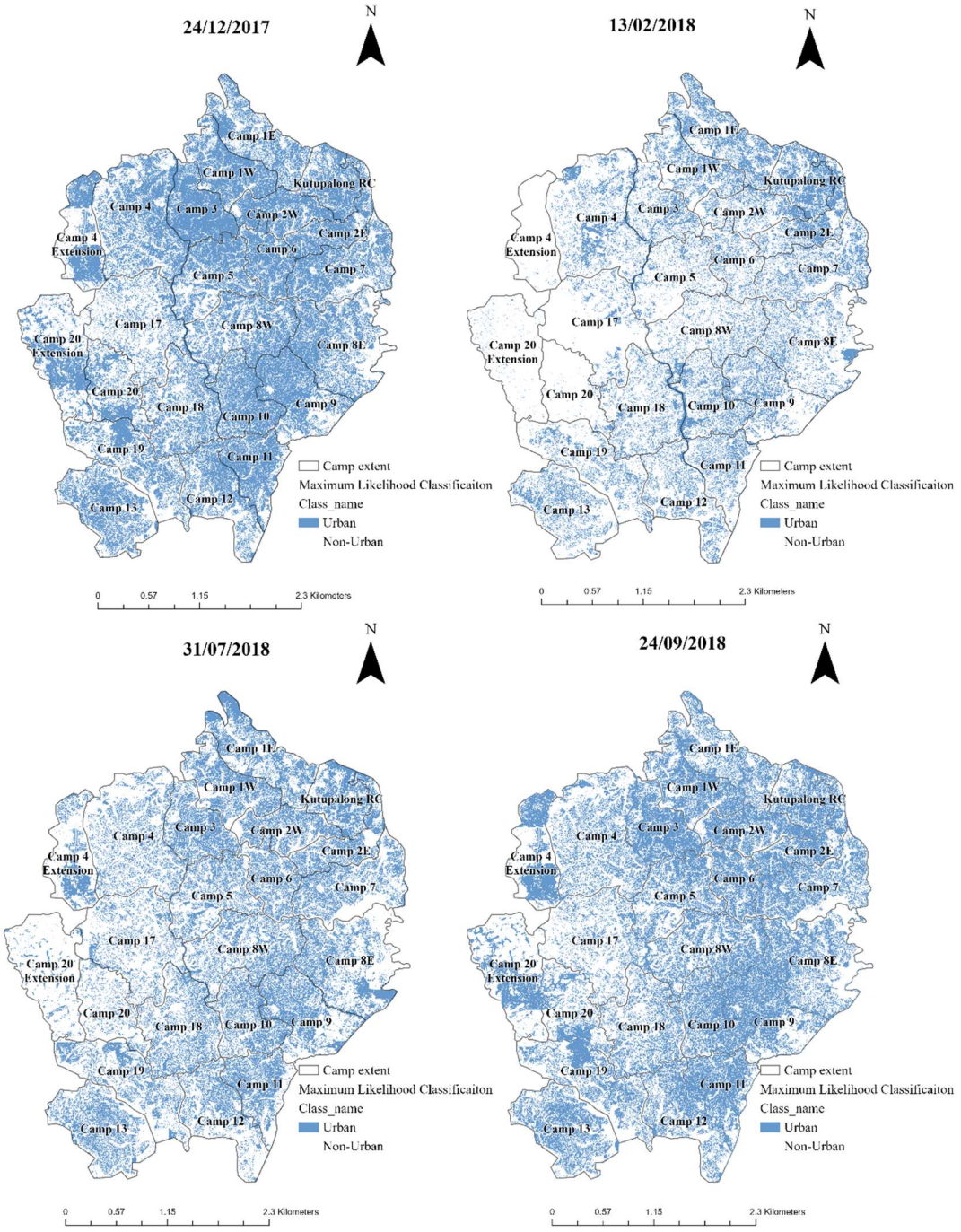

Figure 8 Maximum Likelihood Classification - Kutupalong Refugee Camp

Unlike in case of SVM, with MLC we can understand that the most significant change in terms of urban class expansion is between third and last image dates (31-07-18 to 24-09-18), which percentage wise gives increase from 36% of the total area on the third image, to 54% of the total area. In case of non-urban class, the same timespan (31-07-18 to 24-

09-18) give the highest decrease in terms of area from 9.25 km^2 to 6.7 km^2 (percentage wise this gives us decrease from 64% to 46%).

4.2.1 Change Detection – Maximum Likelihood Classification

Following the same mythology as for the SVM classification and in order to visualize the urban sprawl and detect changes between the dates, Change Detection Wizard of ArcGIS Pro has been utilized. The categorical change method of change detection has been configured over the 4 classified raster images. Processing was set to full extent and class configuration is as follows:

- From Classes: Non-Urban
- To Classes: Urban and Non-Urban

By overlaying the camp extent shapefile with classified raster , we would be able to detect the changes and camp-wise area statistics and percentage of the urbanized units for MLC classification. By grouping all the polygons for classes Urban and Non- Urban, we can calculate geometry attributes and percentage of presence withing a camp area. The *table 5* below presents this information for four (4) different dates.

The MLC classification output showed that camps with highest value of urban class for the first date of 24-12-17 are Camp 3 with 59%, Camp 1W with 40% , Camp6 with 33% and Camp 10 with 33%, while for the last date of 24-12-18 the camps with the highest Urban class value are Camp 3 with 71%, Camp 11 with 64% , Camp 10 with 62% and Camp 2W with 61%. The camps with the biggest urban sprawl in terms of percentage area, where the difference between first and last date are the highest are Camp 4 Extension with 47% increase of Urban class, Camp 11 and Camp 2E with 41% and Kutupalong RC with 33% of increase. Apart from Camp 4 Extension which is located in the north-west part of the study area, the camps 11, 2E and Kutupalong RC and spread across the east side of the border of study area, where the sprawl is noticeably high.

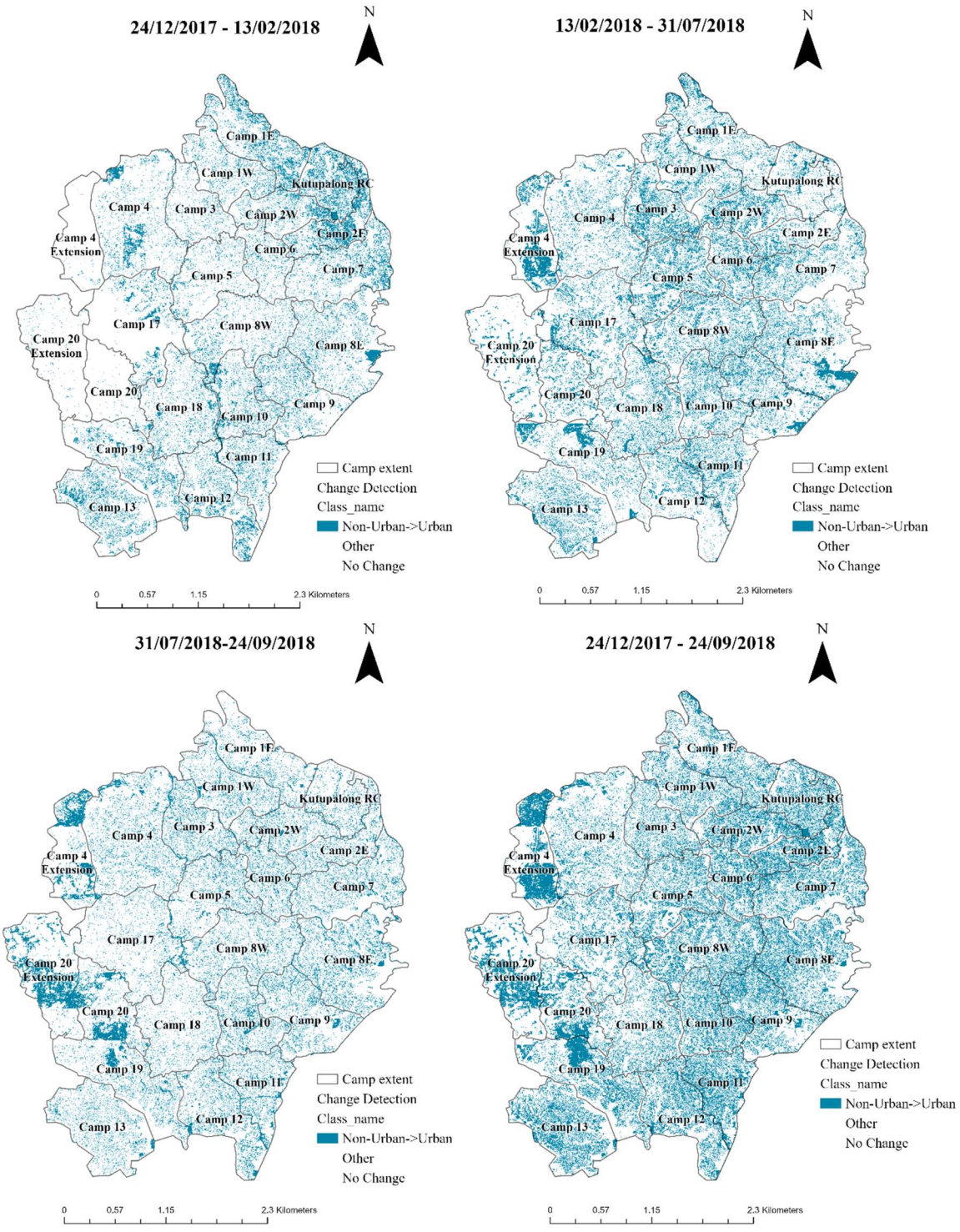

Figure 9 Change Detection Maps - Maximum Likelihood Classification

Table 6 Camp wise area statistics of the urbanized units. The MLC supervised classification was used to analyse the urbanization.

Camp Name	24-12-17		13-02-18		31-07-18		24-09-18	
	Urban (km^2)	Non-urban (km^2)	Urban (km2)	Non-urban (km2)	Urban (km2)	Non-urban (km2)	Urban (km^2)	Non-urban (km^2)
Camp 10	0.16	0.33	0.19	0.31	0.24	0.26	0.31	0.19
Camp 11	0.11	0.36	0.14	0.32	0.22	0.24	0.3	0.17
Camp 12	0.09	0.54	0.17	0.46	0.2	0.43	0.29	0.34
Camp 13	0.13	0.63	0.2	0.56	0.29	0.46	0.35	0.41
Camp 17	0.05	0.9	0.08	0.87	0.21	0.74	0.22	0.73
Camp 18	0.18	0.57	0.18	0.57	0.28	0.47	0.27	0.48
Camp 19	0.08	0.69	0.15	0.62	0.23	0.54	0.3	0.47
Camp 1E	0.15	0.48	0.28	0.35	0.32	0.32	0.29	0.34
Camp 1W	0.21	0.32	0.19	0.34	0.26	0.28	0.3	0.23
Camp 20	0.06	0.43	0.02	0.47	0.08	0.41	0.17	0.32
Camp 20 Extension	0.06	0.7	0.01	0.75	0.06	0.71	0.24	0.53
Camp 2E	0.07	0.32	0.18	0.21	0.19	0.2	0.23	0.16
Camp 2W	0.12	0.27	0.12	0.27	0.19	0.2	0.24	0.15
Camp 3	0.27	0.19	0.15	0.31	0.25	0.2	0.32	0.13
Camp 4	0.37	0.78	0.28	0.87	0.31	0.85	0.4	0.76
Camp 4 Extension	0.01	0.49	0	0.49	0.14	0.36	0.24	0.26
Camp 5	0.16	0.45	0.1	0.51	0.23	0.38	0.28	0.33
Camp 6	0.12	0.24	0.1	0.26	0.17	0.19	0.22	0.14
Camp 7	0.18	0.53	0.23	0.48	0.27	0.44	0.34	0.37
Camp 8E	0.18	0.78	0.23	0.72	0.34	0.62	0.42	0.53
Camp 8W	0.13	0.64	0.17	0.6	0.31	0.46	0.38	0.39
Camp 9	0.17	0.48	0.2	0.45	0.3	0.35	0.37	0.28
Kutupalong RC	0.1	0.28	0.22	0.17	0.21	0.18	0.23	0.16

4.3 Accuracy assessment metrics

The main goal of the research is to evaluate the performance of the two classifiers over the different time periods, and to determine which of the classifiers produces superior results.

For each of the classified image (SVM and MLC) the random 100 points have been computed in order to get the accuracy results. The related UAV image for each date has been used for ground truth testing against the 100 random points produced. After the ground truth comparison with classifiers, confusion matrices have been computed to get overall accuracy, user accuracy, producer accuracy and Kappa Index of agreement.

The results of classified images of the four study time periods are shown in figure 7 and 8. The result indicate that both classifiers scored high overall accuracy and performed well when classifying UAV imagery in environment such as refugee camp settlement.

Class/Date	24-12-17		13-02-18		31-07-18		24-09-18	
	OA=85%		OA=90%		OA=94%		OA=83%	
	UA	PA	UA	PA	UA	PA	UA	PA
Urban	100%	48%	90%	69%	96%	91%	91%	67%
Non-Urban	83%	100%	90%	97%	94%	97%	79%	95%
Kappa	0.57		0.72		0.89		0.64	

Table 7 Confusion Matrix - Support Vector Machine

The overall accuracy yielded good results for the SVM classifier, having minimum value of 83% and maximum value of 85% . User accuracy and producer accuracy also showed favorable results, with exception of 48% of producer accuracy of urban class for the first date of SVM classified image. The kappa coefficient and the degree of agreement between categorization and truth values varies from 0.57 for the first image to 0.89 for the third image.

Class/Date	24-12-17		13-02-18		31-07-18		24-09-18	
	OA=81%		OA=86%		OA=87%		OA=85%	
	UA	PA	UA	PA	UA	PA	UA	PA
Urban	86%	54%	84%	68%	72%	90%	76%	90%
Non-Urban	79%	95%	87%	94%	95%	86%	93%	82%
Kappa	0.54		0.65		0.71		0.69	

Table 8 Confusion Matrix - Maximum Likelihood Classification

Maximum Likelihood Classification has produced similar results in comparison to SVM classified. The minimum overall accuracy is 81% for first image and varies to maximum 87% for the third image. User accuracy and producer accuracy have showed good scores however slightly less when in comparison to SVM classifier. The Kappa coefficient and degree of agreement between categorization and true values is varying from 0.54 for the first image to 0.71 for the third image.

4.4 Shannon's Diversity Index

Shannon's Diversity Index shows us species diversity in the community, in the context of this research, it is giving us an understanding of the urban sprawl composition and class richness and evenness over time (Bourne and Conway 2014). The diversity index of the urban for MLC classification for dates 24-12-2017, 13-02-2018, 31-07-2018 and 24-09-2018 was respectively 0.53 , 0.56, 0.65 and 0.69 while for SVM classifier 0.4 , 0.5 , 0.65 and 0.62 following the same order. The trend of results of Shannon's diversity index for both classifiers shows lower diversity between the first two image dates, however it has an exponential growth in value for the third and fourth date. This has shown that the urban sprawl in the Kutuaplong refugee camp has increased, due to the influx of refugees and urban expansion as a need for more housing. If the urban class is unevenly distributed throughout space, the increased value of Shannon's Diversity index reflects that.. (K. Madhavi Lata et al., 2009).

Date/Classifier	24-12-17	12-02-18	31-07-18	24-09-18
MLC	0.5288482	0.5609598	0.6557419	0.69000136
SVM	0.4038027	0.500757	0.6580772	0.62542934

Table 9 Evolution of Shannon's Diversity Index for MLC and SVM Classifier

4.5 Discussion

In terms of the computing power, MLC has an advantage in analyzing very high spatial resolution imagery. MLC classification has successfully analyzed original 10 cm resolution images, while the SVM has been failing to do so. Finally, to create the fair comparison, the researcher decided to reduce and change the spatial resolution of a raster datasets and sets rules for aggregating or interpolating values across the new pixel size to 0.5.

The Kappa coefficient corrects standardized measure of agreement between two categorical scores produced by the two rates. Based on Landis and Koch measurement of observer agreement The Kappa interpretation of SVM classification gives us understanding that agreement is substantial for values of 0.57, 0.72 and 0.64 and almost perfect agreement for 0.89. The values for MLC classification have similar trend of values where classification of images 1-4 have values of 0.54 , 0.65 , 0.71 and 0.69 respectively. In comparison of the two classifications , the Kappa coefficient for MLC classifier shows higher agreement with exception of last-date image where MLC yields better results.

These results answer to the research question, indicating that the SVM classifier is superior and gives better performance in classifying urban class.

When it comes to calculating urbanization, the research indicates that there has been an exponential expansion of urban class from 24-12-17 to 24-09-18 from 2.01 km2 to 5.37 km2 for SVM and from 3.2 km2 to 7.8 km2 for MLC. The urban class however reduced from 12.58 km2 to 9.95 km2 for SVM and from 11.3 km2 to 6.7 km2 for MLC classifier.

The results found in the research are relevant for urban sprawl analysis in refugee camp settlement and for Humanitarian actors.

The evolution and increase in the values of Shannon's Diversity Index indicates that there is an increase in urban sprawl and development tend to be more dispersed over a period of time. This indicates rapid increase of urban sprawl. The results of this index give us the idea of spatio-temporal patters of urban growth in Kutualong Refugee camp.

5. Conclusions

The thesis demonstrated application of remote sensing classification techniques using 4 UAV images from different categorical dates in order to identify and calculate the urban sprawl in Kutupalong Refugee Camp, Bangladesh which us under great urban expansion due to the influx of Rohingya refugees from the neighboring Myanmar. The Rohingya emergency was one of the biggest crises in 2017 which has severely affected the change of physical land scape of the host community in Bangladesh.

The research analyzed the expansion of the refugee camp from 2017 to 2018. The objective of the research was to understand which of the techniques yield better results. The research was conducted to understand and evaluate performance and agreement of two different machine learning classifiers – Support Vector Machine and Maximum Likelihood Classification.

To answer the research question which machine learning classifier technique yields better performance in urban sprawl classification in Refugee camp context both of the classifiers' performance were similar in terms of overall accuracy for both of the classes. In terms of overall accuracy, the advantage has been given to SVM classifier as it produced slightly better results better results.

The research provides an overview of urban sprawl over time and direction of which the expansion is growing . Both classifiers showed effective method with high overall accuracy for classification in refugee camp, which is a complex environment. This is especially important for the camp planning purposes, policy makers , but as well for other elements like which have importance in refugee camp context like forestry, water bodies, road infrastructure, land slide risks, etc. Secondly, this approach can help support the implementation of a structured analysis of the refugee camp settlements, camp and shelter planning, inventory of housing and logistical activities.

Bibliographic References

Bello, Olalekan Mumin, and Yusuf Adedoyin Aina. 2014. “Satellite Remote Sensing as a Tool in Disaster Management and Sustainable Development: Towards a Synergistic Approach.” *Procedia - Social and Behavioral Sciences* 120.

Blaschke, Thomas et al. 2014. “Geographic Object-Based Image Analysis – Towards a New Paradigm.” *ISPRS Journal of Photogrammetry and Remote Sensing* 87.

Boccardo, Piero. 2013. “New Perspectives in Emergency Mapping.” *European Journal of Remote Sensing* 46(1).

Bourne, Kirstin S., and Tenley M. Conway. 2014. “The Influence of Land Use Type and Municipal Context on Urban Tree Species Diversity.” *Urban Ecosystems* 17(1).

Braun, Andreas Braun. 2020. “Spaceborne Radar Imagery – An under-Utilized Source of Information for Humanitarian Relief.” *Journal of Humanitarian Engineering* 8(1).

Braun, Andreas, and Volker Hochschild. 2017a. “A SAR-Based Index for Landscape Changes in African Savannas.” *Remote Sensing* 9(4).

———. 2017b. “Potential and Limitations of Radar Remote Sensing for Humanitarian Operations.” *GI_Forum* 1.

Braun, Andreas, Stefan Lang, and Volker Hochschild. 2016. “Impact of Refugee Camps on Their Environment A Case Study Using Multi-Temporal SAR Data.” *Journal of Geography, Environment and Earth Science International* 4(2).

Burgess, S. F. 2018. “Military Intervention in Africa : French and US Approaches Compared. Remember to Check Citations for Accuracy before Including Them in Your Work Results Per Page:” : 5–25.

Chang, Anjin et al. 2011. “Canopy-Cover Thematic-Map Generation for Military Map Products Using Remote Sensing Data in Inaccessible Areas.” *Landscape and Ecological Engineering* 7(2).

Charbonneau, Bruno. 2017. “Intervention in Mali: Building Peace between Peacekeeping and Counterterrorism.” *Journal of Contemporary African Studies* 35(4).

Chen, J., Zhou, Y., Zipf, A., & Fan, H. “Deep Learning from Multiple Crowds: A Case Study of Humanitarian Mapping.” *IEEE Transactions on Geoscience and Remote Sensing* : 1713–22.

Corbane, Christina et al. 2021. “Convolutional Neural Networks for Global Human Settlements Mapping from Sentinel-2 Satellite Imagery.” *Neural Computing and Applications* 33(12).

Denis, Gil et al. 2016. “The Evolution of Earth Observation Satellites in Europe and Its Impact on the Performance of Emergency Response Services.” *Acta Astronautica* 127.

Faye, Malang. 2021. “A Forced Migration from Myanmar to Bangladesh and beyond: Humanitarian Response to Rohingya Refugee Crisis.” *Journal of International Humanitarian Action* 6(1).

Friedl, Mark A. et al. 2010. “MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of New Datasets.” *Remote Sensing of Environment* 114(1).

Hansen, Matthew C. et al. 2008. “A Method for Integrating MODIS and Landsat Data for Systematic Monitoring of Forest Cover and Change in the Congo Basin.” *Remote Sensing of Environment* 112(5).
Hassan, Farhad et al. 2020. “Urbanization Change Analysis Based on SVM and RF Machine Learning Algorithms.” *International Journal of Advanced Computer Science and Applications* 11(5).
Heisbourg, François. 2013. “A Surprising Little War: First Lessons of Mali.” *Survival* 55(2).
Hogland, John, Nedret Billor, and Nathaniel Anderson. 2013. “Comparison of Standard Maximum Likelihood Classification and Polytomous Logistic Regression Used in Remote Sensing.” *European Journal of Remote Sensing* 46(1).
Honest, M. et al. “An Investigative Environmental Impact Assessment for Kutupalong Refugee Camp and Surroundings, Bangladesh.” *Swedish Civil Contingencies Agency*.
Horsman, Graeme. 2016. “Unmanned Aerial Vehicles: A Preliminary Analysis of Forensic Challenges.” *Digital Investigation* 16.
K. Madhavi Lata, v. Krishna Prasad, K. V. S. Badarinath, and v. Raghavaswamy. 2009. “Measuring Urban Sprawl: A Case Study of Hyderabad.” *Geospatial World*.
Karimi, Firoozeh, Selima Sultana, Ali Shirzadi Babakan, and Shan Suthaharan. 2019. “An Enhanced Support Vector Machine Model for Urban Expansion Prediction.” *Computers, Environment and Urban Systems* 75.
Labib, S.M, and N Hossain. 2018. “Environmental Cost of Refugee Crisis: Case Study of Kutupalong Balukhali Rohingya Camp Site A Remote Sensing Approach.” In *Conference: 26th Annual GIScience Research UK*,.
Lambert, Jonas et al. 2013. “Monitoring Forest Decline through Remote Sensing Time Series Analysis.” *GIScience & Remote Sensing* 50(4).
Lang, S et al. 2015. “Humanitarian Emergencies: Causes, Traits and Impacts as Observed by Remote Sensing.” *Remote Sensing of Water Resources, Disasters, and Urban Studies*: 483–512.
Lang, Stefan et al. 2020. “Earth Observation Tools and Services to Increase the Effectiveness of Humanitarian Assistance.” *European Journal of Remote Sensing* 53(sup2).
Luque, Amalia, Alejandro Carrasco, Alejandro Martín, and Ana de las Heras. 2019. “The Impact of Class Imbalance in Classification Performance Metrics Based on the Binary Confusion Matrix.” *Pattern Recognition* 91.
Maxwell, Aaron E., and Timothy A. Warner. 2020. “Thematic Classification Accuracy Assessment with Inherently Uncertain Boundaries: An Argument for Center-Weighted Accuracy Assessment Metrics.” *Remote Sensing* 12(12).
MONTCLOS, M.-A. P. D., and P. M. KAGWANJA. 2000. “Refugee Camps or Cities? The Socio-Economic Dynamics of the Dadaab and Kakuma Camps in Northern Kenya.” *Journal of Refugee Studies* 13(2).
Quinn, John A. et al. 2018. “Humanitarian Applications of Machine Learning with Remote-Sensing Data: Review and Case Study in Refugee Settlement Mapping.” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376(2128).
Rimal, Bhagawat et al. 2019. “Effects of Land Use and Land Cover Change on Ecosystem Services in the Koshi River Basin, Eastern Nepal.” *Ecosystem Services* 38.

Rohi, Godall, O'tega Ejofodomi, and Godswill Ofualagba. 2020. "Autonomous Monitoring, Analysis, and Countering of Air Pollution Using Environmental Drones." *Heliyon* 6(1).

Rotte, Ralph. 2016. "Western Drones and African Security." *African Security Review* 25(1): 85–94.

Sahana, Mehebub, Selim Jahangir, and MD. Anisujjaman. 2019. "Forced Migration and the Expatriation of the Rohingya: A Demographic Assessment of Their Historical Exclusions and Statelessness." *Journal of Muslim Minority Affairs* 39(1).

Sisodia, Pushpendra Singh, Vivekanand Tiwari, and Anil Kumar. 2014. "Analysis of Supervised Maximum Likelihood Classification for Remote Sensing Image." In *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*, IEEE.

Tiede, Dirk, Pascal Krafft, Petra Füreder, and Stefan Lang. 2017. "Stratified Template Matching to Support Refugee Camp Analysis in OBIA Workflows." *Remote Sensing* 9(4).

UNHCR. 2020. "Kutupalong: El Campo de Refugiados Más Grande Del Mundo." *https://eacnur.org/blog/kutupalong-el-campo-de-refugiados-mas-grande-del-mundo-tc_alt45664n_o_pstn_o_pst/*.

Villa, Tommaso et al. 2016. "An Overview of Small Unmanned Aerial Vehicles for Air Quality Measurements: Present Applications and Future Prospectives." *Sensors* 16(7).

Wang, Shifeng, Emily So, and Pete Smith. 2015. "Detecting Tents to Estimate the Displaced Populations for Post-Disaster Relief Using High Resolution Satellite Imagery." *International Journal of Applied Earth Observation and Geoinformation* 36.

van Westen, C.J. 2013. "3.10 Remote Sensing and GIS for Natural Hazards Assessment and Disaster Risk Management." In *Treatise on Geomorphology*, Elsevier.

Yaacoub, Jean-Paul, Hassan Noura, Ola Salman, and Ali Chehab. 2020. "Security Analysis of Drones Systems: Attacks, Limitations, and Recommendations." *Internet of Things* 11.

UNIGIS PT

NOVA IMS Information Management School