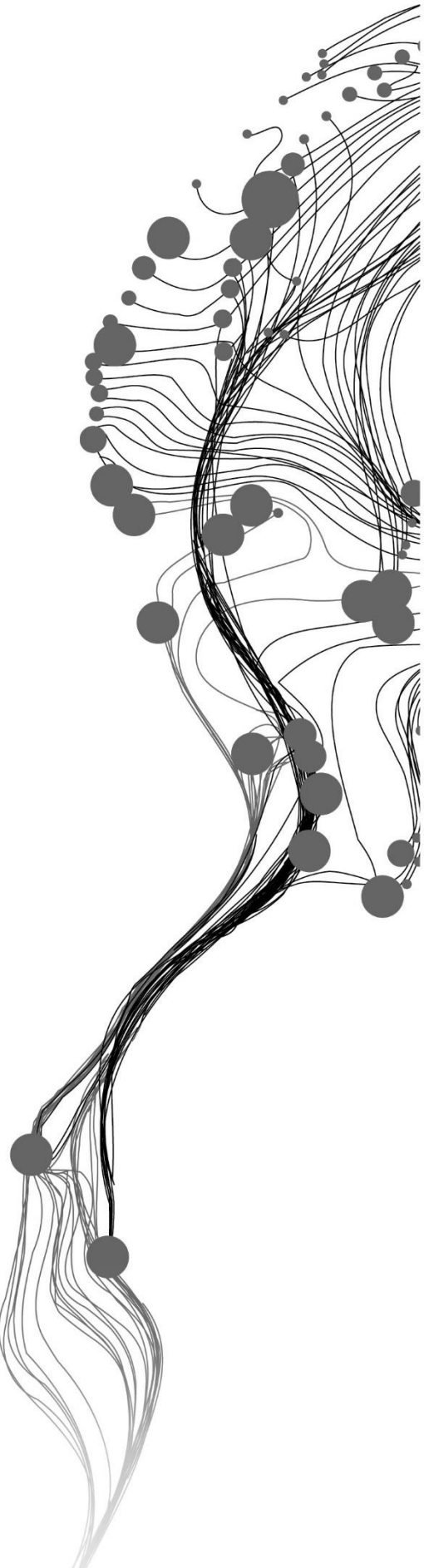


BUILDING DETECTION FROM COARSE AIRBORNE LASER SCANNER DATA

INTAN YULIA ANTASARI
February, 2019

SUPERVISORS:
dr.ir. S.J. Oude Elberink
dr. C. Persello



BUILDING DETECTION FROM COARSE AIRBORNE LASER SCANNER DATA

INTAN YULIA ANTASARI

Enschede, The Netherlands, February, 2019

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfillment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Geoinformatics

SUPERVISORS:

dr.ir. S.J. Oude Elberink

dr. C. Persello

THESIS ASSESSMENT BOARD:

prof.dr.ir. M.G. Vosselman

dr. M. Rutzinger (External Examiner, University of Innsbruck, Institute of Geography)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author and do not necessarily represent those of the Faculty.

ABSTRACT

In developing countries, urban areas have high building densities with complex building pattern. Therefore, detecting and classifying building data can be challenging as it is difficult to separate between the building and non-building area. The buildings object from coarse ALS data which have noisy content and complexity of urban area were obtained by using information provided by topographic map. The FCN has been used to accomplish this research. The planar segmentation has been conducted to detect the roof of the buildings object. The spatial information from 2D map was used to observe the location and the shape of the building. The spatial information from 2D map help to easily recognize the variation shape of the building. The points clipped using polygon from 2D map data to assign points that belong to building objects.

The trained network was tested into different region in Indonesia. The accuracy assessment calculated based on the F-Score. From RMSD in Lombok region, the cleaning data perform better than the noisy data with RMSD equal 1.03 %. The number of sites enhance the performance of the classification. The RMSD in another region that in Tanjung Lesung 20.33%, Tanggamus 30.57% and Makassar 19.70 %.

Key word: building detection, topographic map, ALS data, FCN

ACKNOWLEDGMENTS

I am really would like to thankful to my supervisor dr.ir. S.J. Oude Elberink for guidance, giving knowledge and understanding of this subject, supports and motivated me from the beginning till the end of this thesis. Also, to my second supervisor dr. C. Persello, thanks you very much for sharing knowledges and discussions, and make me more confident to develop this thesis.

Thanks to Aldino, Elyta and Dewi for the advice

Thanks to all my classmates in GFM 2017 department for the collaboration and knowledge exchange. And also, all my Indonesian colleagues Fajar, Ayu, Ratna, Afrin, Eko, Reza, Izzul, Rifat, Ayu Adi, Astria, Rizki, Aji, Aulia, and Yan for being a big new family.

Finally, last but not least, this thesis dedicated to half of my hearth, my husband and my children who struggling to keep full patient and survive during my absence in the home for period of 18 months. My blessing to my parents and my mother in-law for giving advices and moral support during the period of my study.

TABLE OF CONTENTS

1.	INTRODUCTION.....	7
1.1.	Motivation and Problem Statement	7
1.2.	Research Identification.....	8
1.3.	Project Set-up.....	9
1.4.	Method Adopted	9
1.5.	Thesis Structure	10
2.	LITERATURE REVIEW.....	11
2.1.	Building Detection	11
2.2.	Classification and Segmentation	11
2.3.	Fully Convolutional Network.....	12
2.4.	Fusion Topographic Map	13
3.	METHODOLOGY.....	14
3.1.	Dataset.....	14
3.2.	Methodology Overview.....	20
3.3.	Data Preprocessing	21
3.4.	Cleaning Training Data.....	21
3.5.	Label Preparation	23
3.6.	Point to image conversion	25
3.7.	Data Augmentation.....	27
3.8.	Network Training.....	28
3.9.	Network Testing.....	28
3.10.	Accuracy Assessment.....	29
4.	RESULTS AND DISCUSSION	30
4.1.	Results	30
4.2.	Accuracy Assessment.....	33
5.	CONCLUSION.....	35
5.1.	Conclusion.....	35
5.2.	The answer to Research Question.....	35
	APPENDIX I: DATASET VISUALISATION	40
	APPENDIX II: THE CLEANED DATA.....	48
	APPENDIX III PREDICTION RESULT	53
	APPENDIX IV THE PRECISION AND RECALL PERFORMANCE	59

LIST OF FIGURES

Figure 1 Point cloud data from ALS.....	8
Figure 2 Unclassified point cloud.....	9
Figure 3 The fully convolutional (segmentation) network transformation (Long et al., 2015).....	12
Figure 4 ALS data (left) and topographic map (right) of Mandalika, Lombok.....	14
Figure 5 The area of the dataset in Indonesia.....	15
Figure 6 The distribution of the training sites.....	16
Figure 7 Distribution sites for training and testing in Lombok.....	17
Figure 8 The visualization of ALS data.....	18
Figure 9 The location of the testing sites in Tanjung Lesung.....	18
Figure 10 The location of the testing sites in Tanggamus.....	19
Figure 11 The location of the testing sites in Makassar.....	19
Figure 12 Research workflow.....	20
Figure 13 The outliers in the ALS data.....	21
Figure 14 Clipped inside (left) and clipped outside (right).....	21
Figure 15 Segmentation result.....	22
Figure 16 The cleaning results shows point before cleaning (left) and after cleaning process (right).....	23
Figure 17 The combined of the points.....	23
Figure 18 The topographic map in vector format.....	24
Figure 19 The raster image of the corresponding label.....	24
Figure 20 The elevation image.....	25
Figure 21 The intensity image.....	26
Figure 22 The number of echoes.....	26
Figure 23 The height difference image.....	27
Figure 24 The overview classification steps.....	27
Figure 25 The point cloud visualized based on height that combined with polygon map.....	30
Figure 26 The corresponding label (left) and classification result (right), the building object shows in red colour.....	30
Figure 27 The classification results within same region.....	31
Figure 28 The comparison between noisy and cleaned training sample.....	31
Figure 29 Building object in the different region.....	32
Figure 30 Small buildings that have similar elevation with the neighbourhood.....	33
Figure 31 The classification result.....	33

LIST OF TABLES

Table 1 FCN layers.....	28
Table 2 Evaluation of the FCN performance in Lombok Region	34
Table 3 The performance for another region	34

1. INTRODUCTION

1.1. Motivation and Problem Statement

Building detection is a significant task in creating base maps. It is the process of detecting and classifying point cloud data which separates buildings from other objects in the scene. The building information can be used for urban planning, population estimation, sustainable development, risk management, and other applications. Commonly, the use of aerial photographs and high-resolution satellite images is the most effective data source for extracting building objects. However, manual digitizing of geometric buildings from remotely sensed images is costly and time consuming (Zhang, Yan, & Chen, 2006). Therefore, Airborne Laser Scanner (ALS) data that has height information can be utilized to extract the building object. Additionally, providing accurate and reliable base maps is needed to support government development program. Due to the complexity of urban scenes, the detection of building from ALS data remains a challenging task (Niemeyer, Rottensteiner, & Soergel, 2014).

In developing countries, urban areas have high building densities with complex building pattern. Therefore, detecting and classifying building data can be challenging as it is difficult to separate between the building and non-building area. Also, this classification process is constrained by the unavailability of training samples that can represent the region that have the complexity of the urban area the complexity of point cloud data depicted in Figure 1.

Building detection of ALS data is an active research in the last decade. In the recent time, a deep learning algorithm is used to improve the classification of ALS data into the ground, vegetation, and buildings. Several studies use deep learning for classification. Rizaldy, Persello, Gevaert, & Oude Elberink (2018) used Fully Convolutional Networks (FCNs) as FCNs architecture for ground classification. Persello & Stein (2017) explain deep FCNs that can learn a hierarchy of features associated with increasing levels of abstraction, from raw pixel values to edges and corners up to complex spatial patterns. Thus, this method is expected to be a solution for building detection in complex urban areas.

However, there are several limitations, not only for building detection but also in classification using deep-learning based. Some problems that can be distinguished for building detection are for example problems in finding information from the data, and problems caused by scene complexity (Elberink, 2008). Yang et al., (2018) identified several challenges among many deep learning-based classifications as well as in building detection. One of them was required a significant amount of training samples. Hence, a representative training sample is needed in the classification processes.

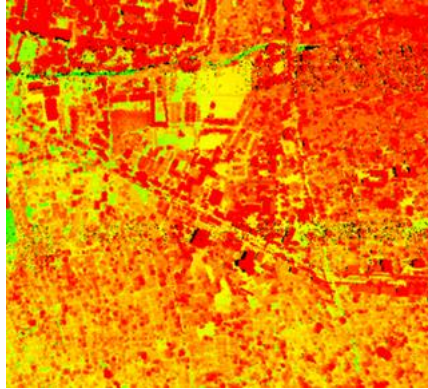


Figure 1 Point cloud data from ALS

This research explored and provided representative and diverse training samples from 2D map data for classification. We using deep learning approach in high density of buildings with a complex pattern urban area and investigate point feature information such as geometry information to improve classification. Therefore, the classification method can be applied to another region.

1.2. Research Identification

1.2.1. Research Objectives

The objective was to develop approaches for building detection from coarse ALS data which have noisy content and complexity of urban area by using information provided by topographic map. The sub-objectives that have to be conducted were as follows:

1. To investigate the use of 2D map data for the generation of training sample and the influence of point features such as geometry information, height, an intensity value and echo information for building detection.
2. To design deep learning architecture which can detect building and non-building
3. To investigate whether the proposed method can be applied to another region.

1.2.2. Research Questions

Sub-objective 1:

1. How can an existing spatial geoinformation such as 2D map be used to help to extract training samples?
2. How to collect the representative training samples?
3. What is point feature information of ALS data that can be used to improve detection?
4. How to employ those point features in deep learning methods for building detection?

Sub-objective 2:

How can a deep learning approach be used to detect building from coarse ALS data?

Sub-objective 3:

1. What are the accuracy and the performance of the proposed methods with completeness and correctness such as F1 score or kappa index?
2. How can the proposed method be employed in another region?

1.2.3. Innovation aimed at

Rizaldy, Persello, Gevaert, Elberink, & Vosselman, (2018) generate multi-class for instance building objects using deep learning approach with training and testing using AHN dataset in Netherlands that have different characteristic with Indonesia. Since representative training sample was significantly needed for classification in an urban area in Indonesia that has a high density of building with a complex pattern like in Figure 2. Therefore, this research proposed a method for detecting building from coarse ALS data by generating training sample in the complex urban area using spatial information such as 2D map and point feature information to improve classification results.

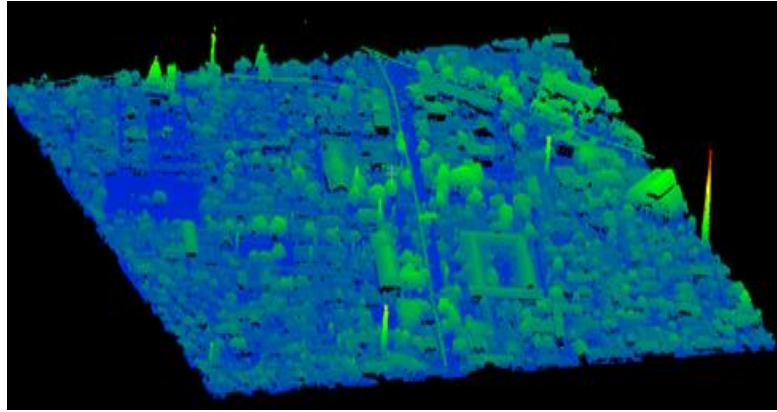


Figure 2 Unclassified point cloud

1.3. Project Set-up

In general, the workflow of the approaches are as follows:

- a. Data pre-processing
- b. Cleaning training data
 - a. Point in polygon information
 - b. Planar segmentation
 - c. Visual inspection
 - d. Combine ALS Data
- c. Label preparation
- d. Point to image conversion
- e. Data augmentation
- f. Network training
- g. Network testing
- h. Accuracy assessments

1.4. Method Adopted

The main idea is to detect building objects by using information from topographic map to improve the classification results. The image-based classification was employed by using FCN network architecture by Persello and Stein (2017). The point cloud was converted into single image and value of pixels representing value of point cloud within the highest point in order to detect building objects. The topographic map was used as corresponding labels. The results were building and non-building pixels.

1.5. Thesis Structure

This thesis contains of five chapters organized as follows:

- a. Chapter 1 presents introduction that explains motivation and problem statements, research identification, research objectives, research questions, innovation, project set-up, and adopted method.
- b. Chapter 2 provides literature reviews to support the research.
- c. Chapter 3 gives explanations about dataset and methodology used in this study.
- d. Chapter 4 shows the result and discussion.
- e. Chapter 5 contains conclusion, answer to research questions and recommendations.

2. LITERATURE REVIEW

2.1. Building Detection

Building detection is a classification for separating the building from another object such as vegetation (bushes and trees) and artificial ground (Vosselman, G., & Maas, 2010). Detecting building object is the initial process to generate building outlines and to reconstruct the 3D building object (Vosselman, G., & Maas, 2010). The main elements of detection buildings are that the objects are above the ground and contains planar segments.

Morphological operators have been used to separate ALS data into terrain and non-terrain (Bretar, Roux, Mallet, Heipke, & Soergel, 2011; Ghaffarian, Ghaffarian, El Merabet, Samir, & Ruichek, 2016; Mongus, Lukač, & Žalik, 2014; Weidner & Förstner, 1995; Zhang et al., 2003). The study about this method has been well explaining by Haralick & Sternberg (1987) that give the review about binary and grey scale, the operations of dilation, erosion, opening, and closing in image processing. Weidner & Förstner, (1995) use this method to obtain a normalized digital surface model (nDSM). It was generated by subtracting the original surface model with the DTM which is produced by using erosion followed by dilation. An initial building mask is obtained by using height thresholds to nDSM (Weidner, 1997).

Mallet, Chehata, & Bailly, (2016) explain the use of the feature of ALS data as an input that is intensity, heights, echoes, and local 3D geometry. The authors identify that a set of features of ALS data can be used for supporting the analysis of the ALS data. Hug & Wehr (1997) combining both surface reflectance of ALS data and 3D geometry to detect and to identify surface object automatically. This approach using surface reflectance, surface orientation and elevation to separate the detected surface object into a natural object(vegetation) and artificial object(building). The authors determine that artificial objects have continuous surface segments and specific characteristic. The characteristic has geometric regularities and simple geometric shapes like rectangular, triangles and circle.

The height-based features are one of the feature information that can be used for discriminating objects in the urban environment (Guo, Chehata, Mallet, & Boukir, 2011; Hug & Wehr, 1997). The authors explain that the elevation can separate high-rise, low-rise and ground elements such as to distinguish buildings and vegetation. Mallet, Chehata, & Bailly, (2016) determine that the height difference in its neighborhood between the point of interest and the lowest point can be used to discriminate the ground and non-ground.

2.2. Classification and Segmentation

Many classification techniques that have been done to classify feature from the ALS data. Generally, ALS data classification divided into two types (Charaniya, Manduchi, & Lodha, 2004) that are a classification of ALS data into terrain and non-terrain points or classification into features such as trees, buildings, etc.

The classification has been two main approaches (Bishop, 1995; Grilli, Menna, & Remondino, 2017) that are supervised classification and unsupervised classification. The supervised classification using data samples that contain labels. This data samples as an input to estimate classifier parameter called training dataset. The test dataset is a data that have been tested by the classifier.

Previous studies convert the point cloud into image data which is supervised pixel-based classification such as maximum likelihood (Bartel & Wei, 2000), Gaussian mixture modelling (Charaniya et al., 2004), neural network(Minh & Hien, 2011; Nguyen, Atkinson, & Lewis, 2005; Priestnall, Jaafar, & Duncan, 2000), and

rule-based classification (Huang, Shyue, Lee, & Kao, 2008) were employed classification depending on the purpose (Yan, Shaker, & El-Ashmawy, 2015).

Recently, a deep learning algorithm can be used to improve the classification not only from aerial images, remote sensing data but also from point cloud data. Some studies use a neural network for building detection from aerial images and remote sensing data (Bittner, Cui, & Reinartz, 2017; Davydova, Cui, & Reinartz, 2016; Yang et al., 2018; Yuan, 2016). Priestnall et al., (2000) propose classification from ALS data using an Artificial Neural Network to distinguish between buildings and trees which is using both topographic and spectral characteristics from ALS data. The elevation data from ALS data can be used as an additional source of information that can be used as input in the Hopfield neural network (HNN) (Nguyen et al., 2005). Minh & Hien, (2011) were using a neural network to process land cover classification. They were applied the intensity image combine with elevation data, panchromatic image, RGB image from ALS data. Also, Rizaldy et al. (2018) use FCN for ground classification which is they study can improve the performance of ground classification by converting all points into one image that can be more efficient.

To get an approximation of parameter, deep learning as supervised learners requires sufficient training samples (Erhan, Manzagol, Bengio, Bengio, & Vincent, 2009). The amount of training sample is necessary for the classification task (Yang et al., 2018) and more representative training data may improve the test result (Rizaldy1 et al., 2018). Moreover, Gevaert, Persello, Elberink, Vosselman, & Sliuzas, (2017) provide a labeling training sample using the spatial dataset to give the labels instead of manually labeling.

2.3. Fully Convolutional Network

Long, Shelhamer, & Darrell, (2015) use fully convolutional network (FCN) for pixel-wise segmentation task in image processing. Specific layers have been adjusted which implement the constructions of segmented output maps to create a network appropriate for pixel-wise segmentation, (Snuverink, 2017). The system has an end-to-end schema which uses an image, learns the features, and outputs labels for every pixel without additional works (Rizaldy, 2018). In Figure 3 depict a classification net generate the output a heatmap by transforming fully connected layers into convolution layers. The input of the classification network is an image with $b \times w \times d$, where spatial dimension is represented by $b \times w$ and the feature or channel dimension is represented by d (Long et al., 2015). In this study, the image-based classification was employed by using FCN network architecture by Persello and Stein (2017).

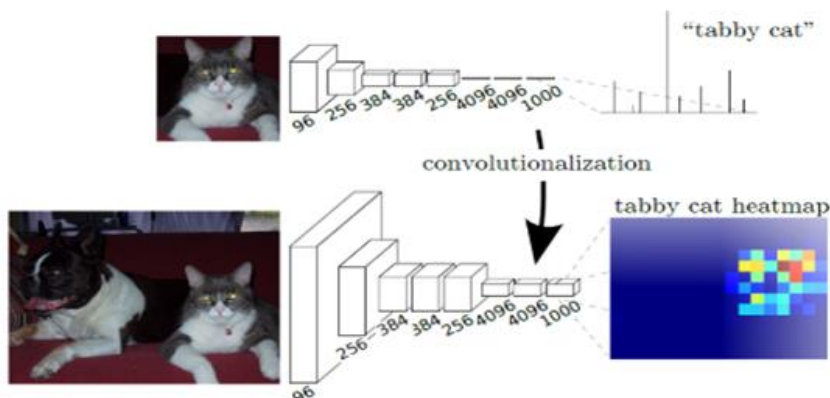


Figure 3 The fully convolutional (segmentation) network transformation (Long et al., 2015)

2.4. Fusion Topographic Map

The classification can be improved significantly by combining ALS data with 2D topographic map/ ground plan (Haala & Anders, 1997; Hofmann, Maas, & Streilein, 2002; Vosselman & Dijkman, 2001; Wang & Oude Elberink, 2016). The advantages of the ground plan are to detect the location and to define the orientation of the building (Suveg & Vosselman, 2001). The building outlines of the ground plans can provide the position of these building object (Vosselman & Dijkman, 2001). The maps have labeled polygons that have topographic classes like buildings, roads, vegetation, water and terrain (Wang & Oude Elberink, 2016).

Haala & Anders, (1997) use a topographic map to process segmentation and object recognition. The authors use it to obtain the shape of the building object. This information can be helped to determine constraints by the planar surface intersection for reconstructing a building (Haala & Anders, 1997). The authors identify the result of segmentation from this approach better than image interpretation (Haala & Anders, 1997).

Hofmann et al., (2002) use a knowledge-based building detection based on ALS data and topographic map information which uses the pixel map as a source to obtain the pattern of the buildings in the ALS data. The pixel map not only presents the shape that can be used for detecting structures in the laser scanner data but also it can reduce the processing time(Hofmann et al., 2002). The authors using a GIS package for detecting building in segment regions. They use position, shape and attribute to analyze the segment(Hofmann et al., 2002). This approach gives an excellent result to identify the building object.

Wang & Elberink, (2016) segment ALS data using an existing topographic map to segment the ALS data. The authors use this data as a kind of background layer. The use of the topographic map can be used to group the point into the same group of objects. This approach can generate a point cloud that has corresponding information about the class, map object, a segment number, and a determination whether this point belongs the map class or not (Wang & Oude Elberink, 2016).

3. METHODOLOGY

3.1. Dataset

Datasets were collected from several areas in Indonesia i.e. Mataram (West Nusa Tenggara), Tanjung Lesung (Banten), Tanggamus (Lampung) and Makassar (South Sulawesi). The datasets contain 32 sites. The size of each site is 512 m x 512 m. Two primary data were used in this study. In Figure 4 an airborne laser scanning data and national topographic map are used to perform the proposed approach.

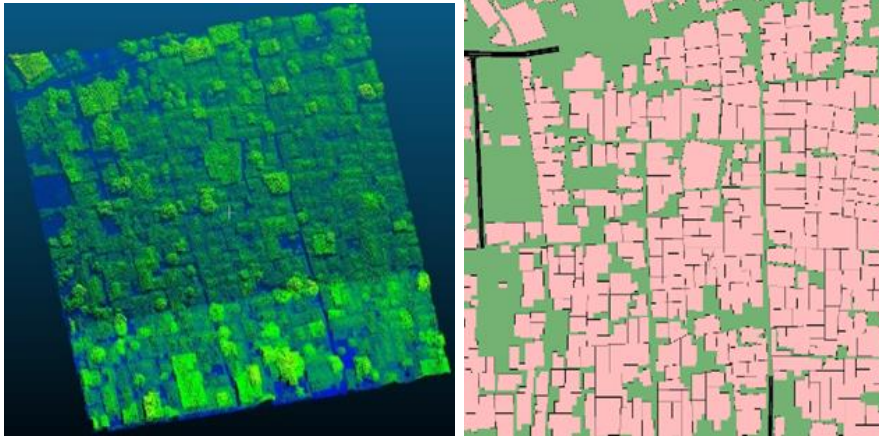


Figure 4 ALS data (left) and topographic map (right) of Mandalika, Lombok.

3.1.1. Airborne Laser Scanner Data

Point cloud data was acquired by airborne laser scanner using Leica ALS 70. The average point density is 4 – 12 points per square meter. The sites of Mataram and Makassar are projected in UTM zone 50S, and the sites of Tanjung Lesung and Tanggamus are projected in UTM zone 48S. All the sites use the WGS84 reference system. The ALS data is intended and designed as input data for base mapping at scale 1:5.000. All locations are in urban areas that relatively have a flat terrain. The landscapes have densely packed buildings that has quite a variance building pattern with vegetations among them. The datasets are available in LAS file. The data contains objects such as buildings, roads, trees, bridges and many other objects in urban areas. The ALS data has many information such as 3D coordinates in space, i.e. (x, y, z), intensity, echo number, number of echoes, and point source ID.

3.1.2. Topographic Base Map

The topographic base map is vector data in geodatabase format which was used as reference data at scale 1:5.000. It has geometric accuracy better than 5 meters. The base map consists of several layers for instance building, road, water body, bridge, and vegetation. These vector data manually delineated from ground orthophoto that were recorded at the same time as ALS data.

3.1.3. The Dataset Location Details

The ALS dataset was converted from las format into text format. The data was consisted of five columns. Three columns are coordinate information i.e. X, Y, and Z. The fourth column is intensity, and the fifth column is number of echoes. The dataset was divided into two categories i.e. training and testing. The

training data consisted of 20 sites in Mataram. The testing data consisted of five sites in Makassar, five sites in Tanggamus, and two sites in Tanjung Lesung. Figure 5 shows locations of the data over Indonesia region.



Figure 5 The area of the dataset in Indonesia

The sites have several characteristics i.e. densely packed small buildings, densely small building with vegetations among them, irregular shapes, and complex structures. Matikainen, Hyypä, & Kaartinen, (2004) mentioned that building objects can be roughly categorized into three types for example a large house area, a small-house area, and an industrial area.

Mataram is the capital city of West Nusa Tenggara Province, located in the west part of Lombok island. This location was chosen as the training data because this area has a different characteristic urban area. The place has relatively flat terrain. The sites can represent building shape and pattern. The distribution of the training sites in Mataram is depicted in Figure 6. It contains 20 sites of ALS data. These data were acquired in 2016. The training area was determined by the characteristic of buildings. The area has a densely large building with irregular shapes represented in sites 01, 02, 06, 08 and 09. Sites 03, 04, 07, 10, 11, and 14 have a densely small building. The area that has separately small building was represented by sites 16 and 19 (see Appendix 1 for detail visualization of ALS data). Figure 7 depicted the distribution of the training and testing data in Lombok region.

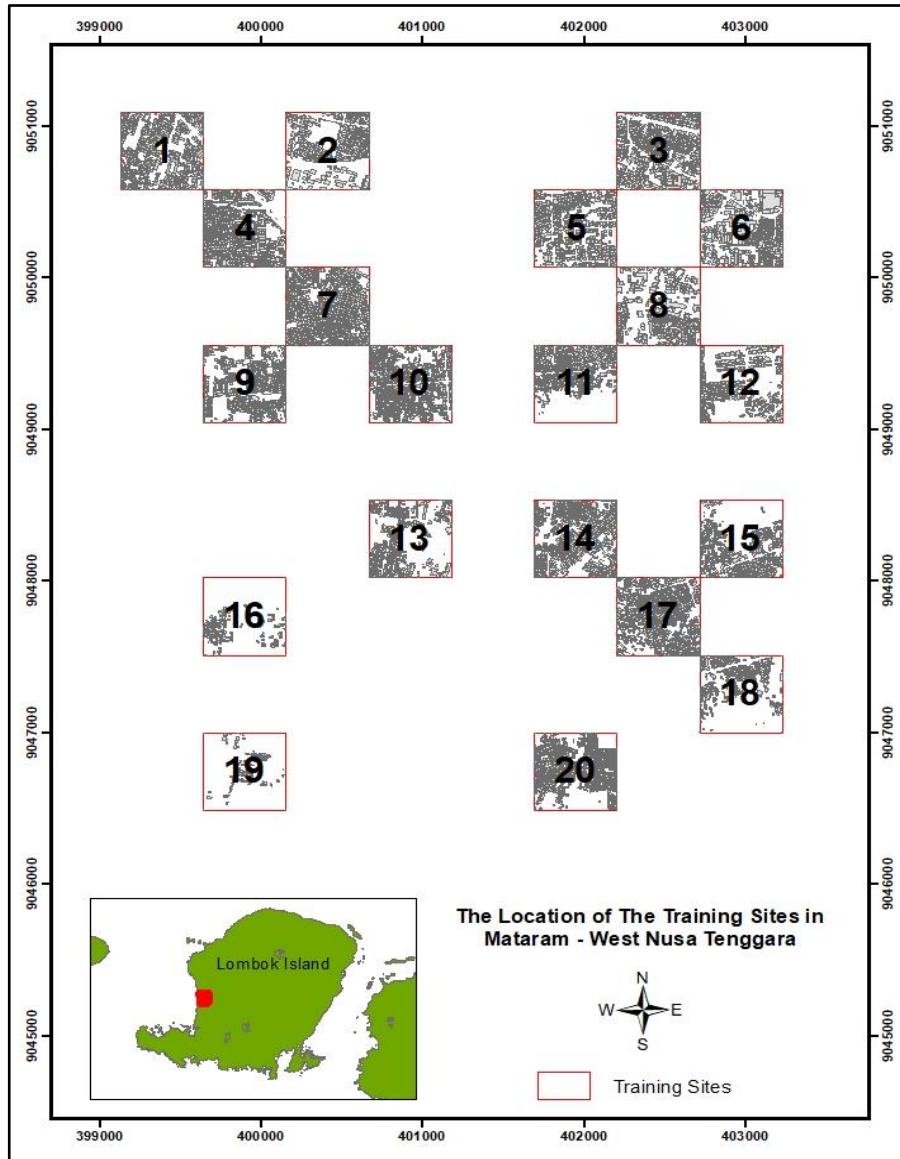


Figure 6 The distribution of the training sites

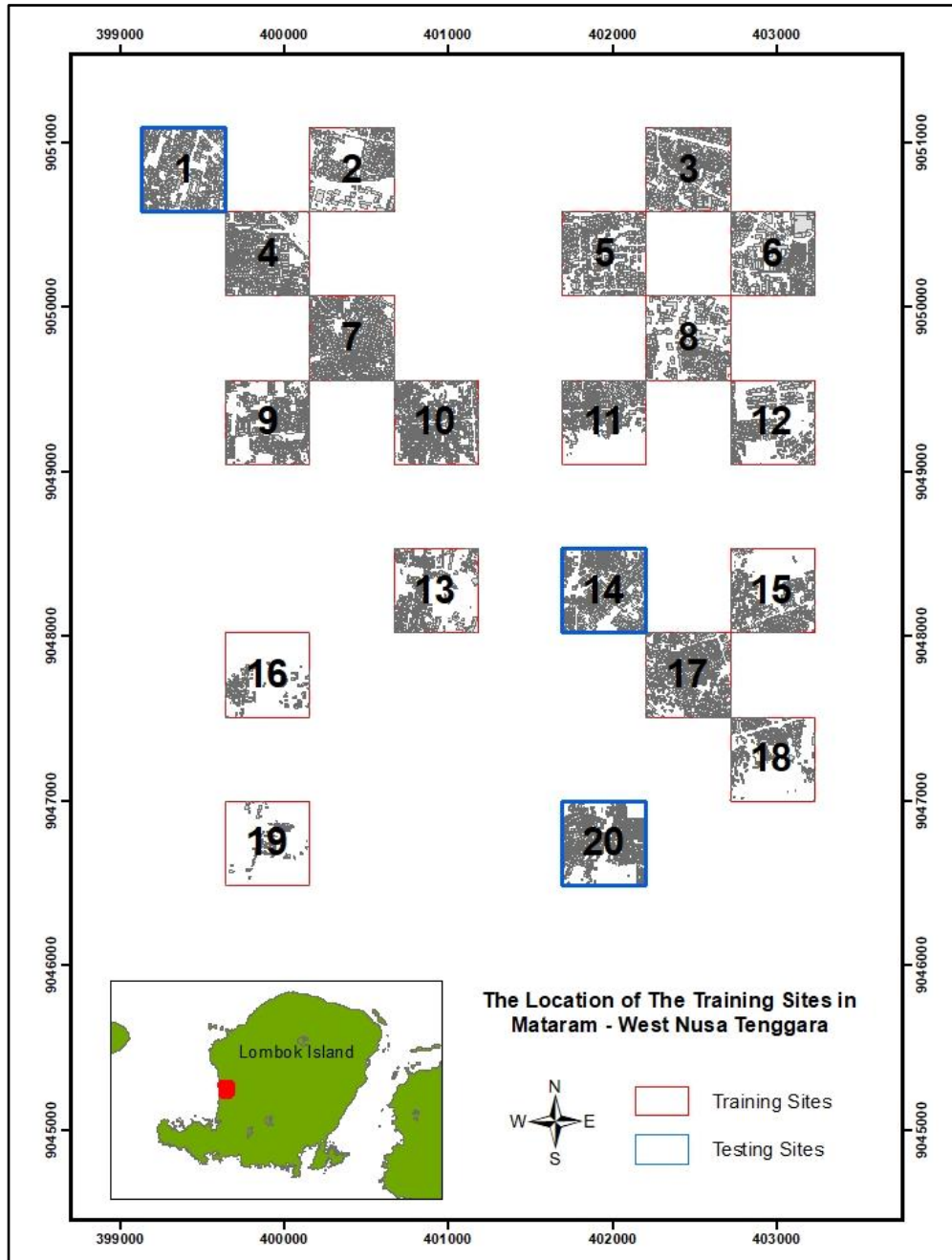


Figure 7 Distribution sites for training and testing in Lombok

Tanjung Lesung is located in Banten Province in the western part of Java Island. It was determined as the locations of two testing sites, namely site 21 and 22. This area has a densely small building with vegetations among the house. The distribution of the testing sites in Tanjung Lesung is depicted in Figure 9. Tanggamus is located in Lampung Province in Sumatera Island. As the testing sites, it has a densely small building that has rough areas. The sites have five sites namely 23, 24, 25, 26, and 27. The distribution of the testing sites in Tanggamus is depicted in Figure 10. Last region is Makassar. Makassar is the capital city in South Sulawesi Province. As an urban area, it has densely building with irregular patterns. Some area has a regular pattern such as residential area. The testing sites in this area are site 28, 29, 30, 31 and 32 (see Figure 11). The visualization of the point cloud is depicted in Figure 8.

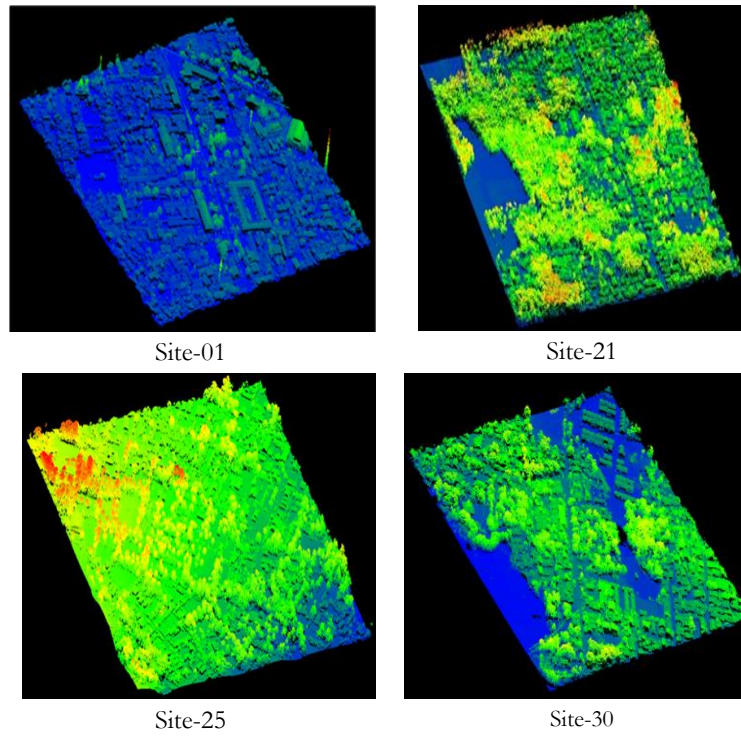


Figure 8 The visualization of ALS data

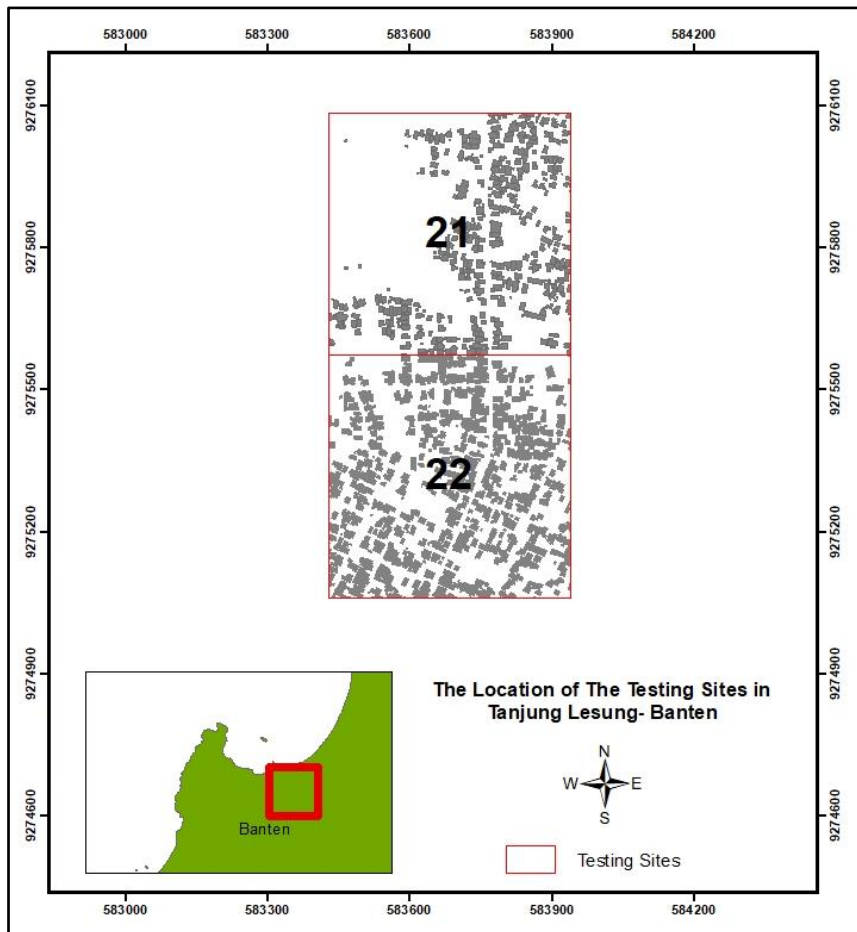


Figure 9 The location of the testing sites in Tanjung Lesung

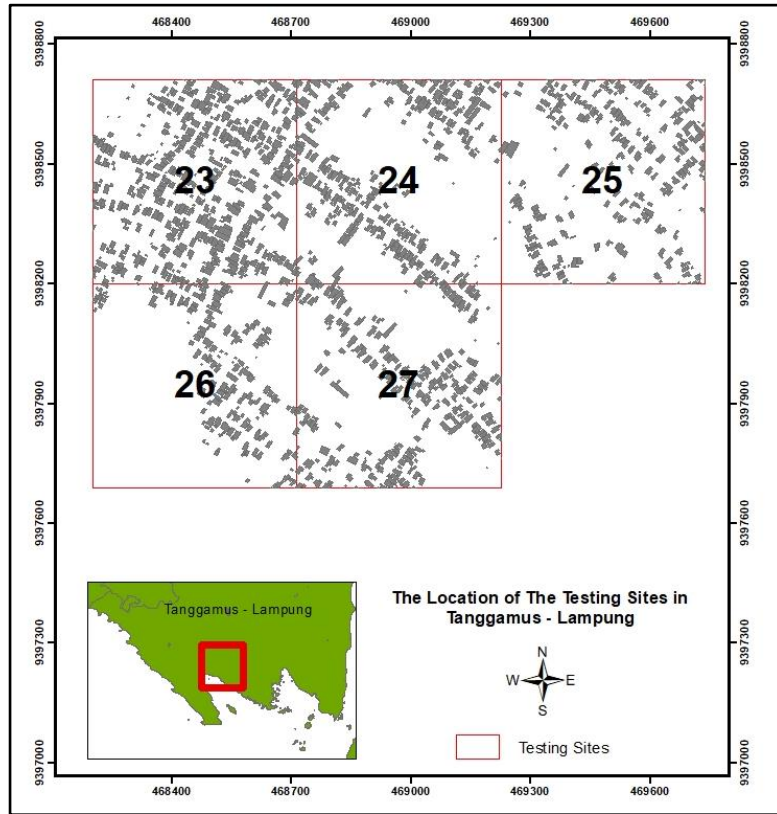


Figure 10 The location of the testing sites in Tanggamus

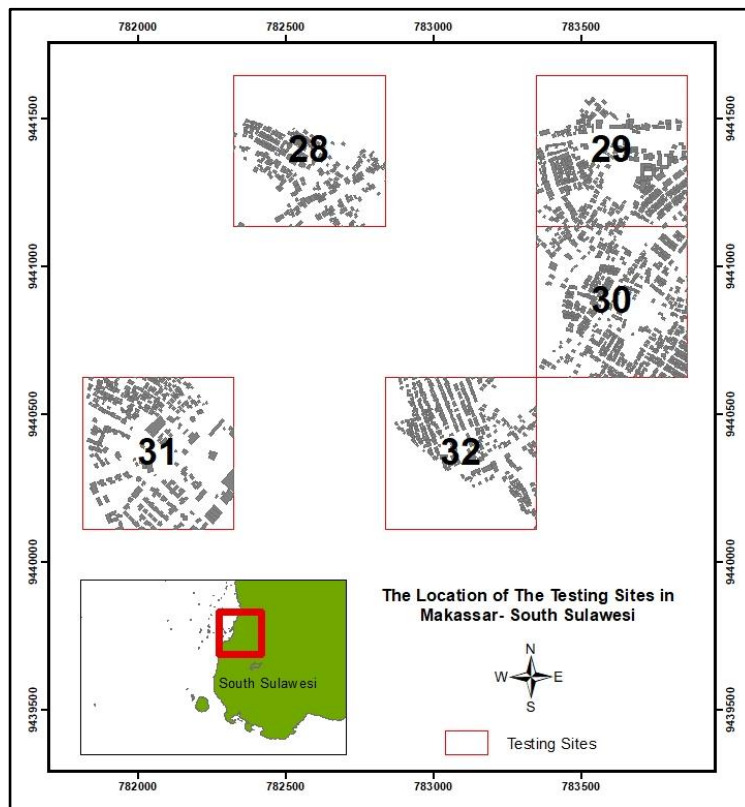


Figure 11 The location of the testing sites in Makassar

3.2. Methodology Overview

Image-based classification approach was proposed for detecting building from coarse point cloud data. In this study, a topographic map as an existing geospatial information was used to extract the building label and to detect the location of the buildings. The overview of the workflow can be seen in Figure 12.

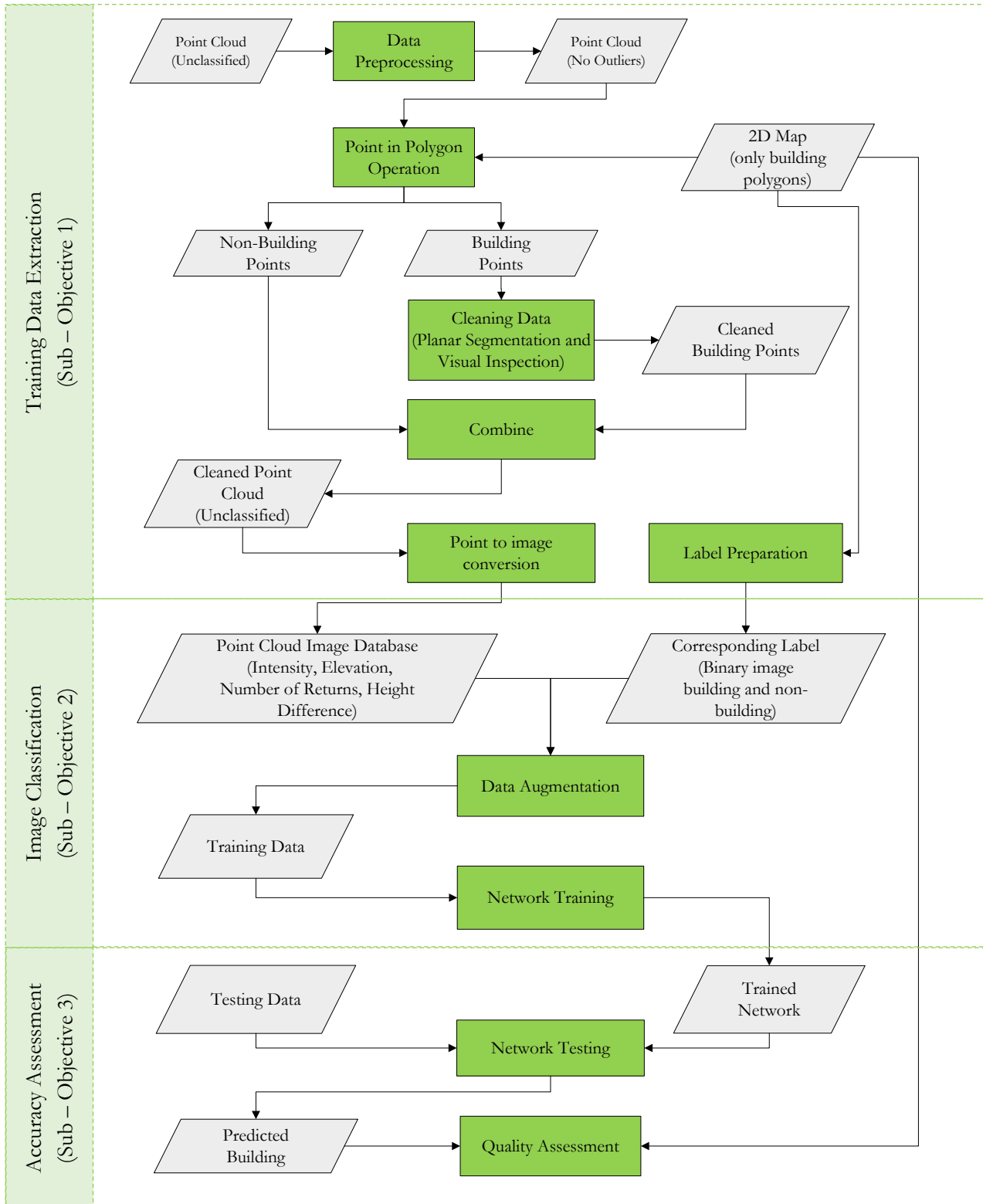


Figure 12 Research workflow

3.3. Data Preprocessing

The ALS data as an input have some outliers which can influence the performance of the classification. The outliers are the a measurement which do not correspond to the local surface and the local neighbourhood (Sotoodeh, 2006). The data pre-processing is a process to remove outliers from ALS data. This process had to be done to improve the quality of classification. It was removed by using Lidar360. Those points were removed automatically by determine the number of point in local neighbourhood and the standard deviation. The outlier data appear not only above the ground but also below the ground. The outliers in one of the ALS data are shown in Figure 13.

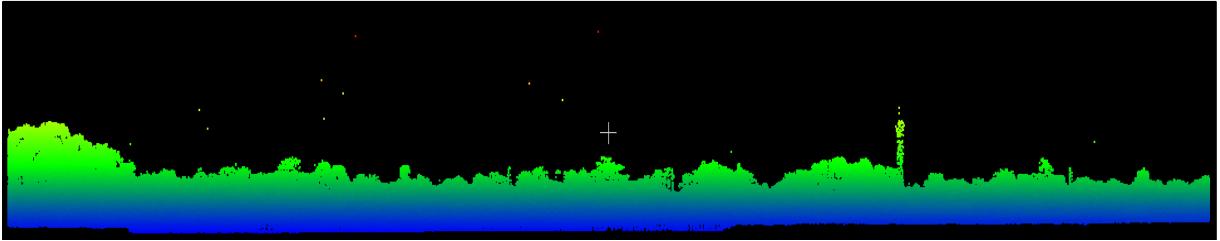


Figure 13 The outliers in the ALS data

3.4. Cleaning Training Data

3.4.1. Point in Polygon Operation

The cleaned points are unclassified. It means that the point does not assign into certain class. The points in polygon operation using clip operation. The purpose of this process is to obtain the points that belong to building objects. The polygons of the building objects were overlay with the points to specify the location and the shape. The points were obtained by group the point within polygon. The points clipped using polygon from 2D map data to assign points that belong to building objects.

In this study, the points were clipped inside and outside of the polygon. This process was done by using FME software. The clipped inside was used to detect the building objects. The clipped outside is the non-buildings objects. Clipped inside was clipped points within polygon that correspond with building object. Clipped outside was clipped points that lies outside the polygon. The clipped point cloud can be seen in Figure 14.

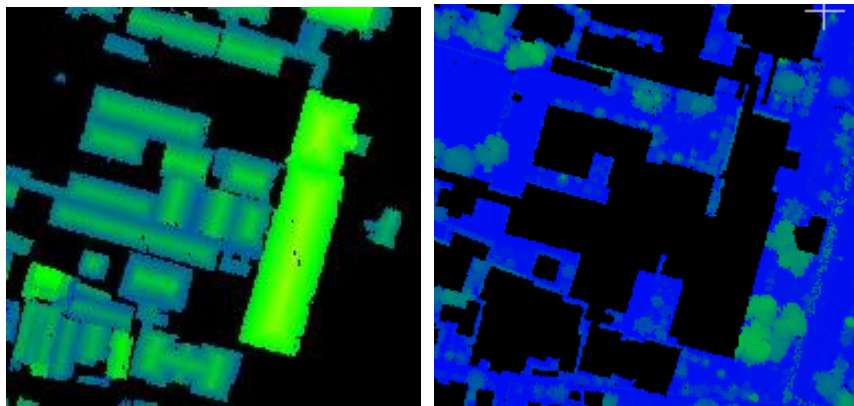


Figure 14 Clipped inside (left) and clipped outside (right)

3.4.2. Planar Segmentation

Segmentation is a approach to give label and to group point cloud based on the object categories and correspond to particular object (Elberink & Vosselman, 2009). The purpose of this process is to detect points that correspond to roof of the building. In general, the points on the roof might have points that represent another object for instance the points belong to trees. It can give influence of the height information of the building object.

The segmentation was done by using segmentation algorithm by Elberink and Vosselman, (2009). The Hough transform was used to detect planar points. The authors employed surface growing by select arbitrary points and then a few points in the neighborhood. Vosselman, G., Gorte, B. G. H., Sithole, G., & Rabbani, (2004) described surface growing algorithm. The points define as a seed surface if the plane fitting results have low residuals. The planar surface was expanded in the further points in the neighborhood. The seeds were determined based on local surface fitting and local smooth normal vector.

To detect the points that belong in the same type, the setting of parameter based on seed radius, grow radius, maximum distance growing, minimum segment size and flatness. The seed radius is the radius of the seed in the neighborhood. The size of the seed radius is 1 meter. The grow radius is the radius for growing the seeds. The grow radius size is 1 meter. The maximum distance assigned to 0.2 to 0.3 that is the distance from point to surface. The minimum segment is minimum segment remaining that has value 10 – 20. The flatness size is 0.4 – 0.5 is the measurement of the flatness. Figure 15 depicted the segmentation result. It shows that the building roofs were presented in the different color. The color represents the segment results of the roofs.

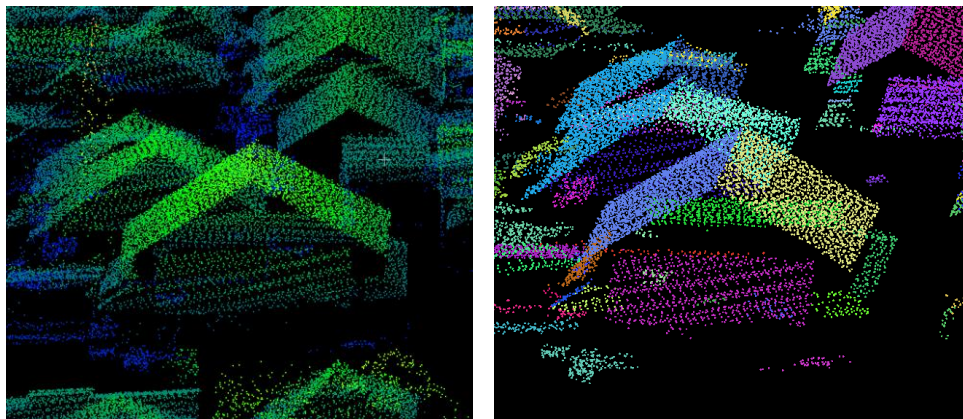


Figure 15 Segmentation result

3.4.3. Visual Inspection

The clipped and segmentation results still contain the points that are not belonging into building objects. The points are not belonging into building object can give the wrong information that can lead to misclassification. The points correspond to trees object and ground object can still remain in the segmentation results. So that, the visual inspection is necessary to check whether the points belong to correct object. This process was done by using Cloud Compare. The points are not corresponding to building object were selected and removed. Not only the points above the building but also the points that lies on the ground. Figure 16 depicted the results of the cleaning process by visual inspection. From the figure below, the points belonging to trees object with blue color and the ground areas colored in blue color. In means that, these points still remaining after clipped operation and segmentation process.

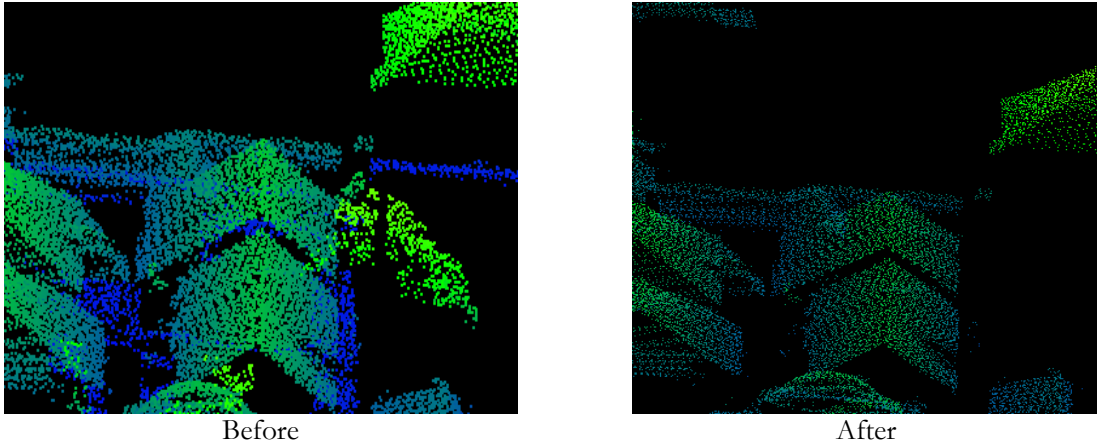


Figure 16 The cleaning results shows point before cleaning (left) and after cleaning process (right).

3.4.4. Combine ALS Data

After cleaning the data by segmentation and visual inspection, the cleaned clipped inside as building object were combined with the clipped outside as the non-building object. The aim is to group the points belonging into the building object that already cleaned with the points are not building object. This process using point cloud combine tools in FME. The combination of these points results is used as an input in the training network. The results of this step can be seen in Figure 17. The figure show that points belonging to trees and ground object were removed.

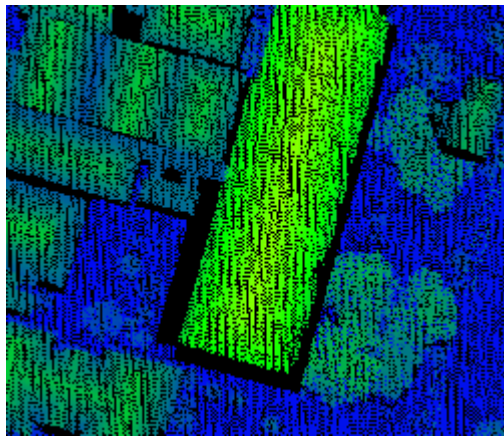


Figure 17 The combined of the points

3.5. Label Preparation

3.5.1. Generate the Corresponding Label

The corresponding labels were generated from topographic maps. The topographic maps were available in vector format. These maps contain several geometries based on points, lines, and polygons. They provide several layers for instance building, road, water body, vegetation, administrative boundary, and utility. The classes were aggregated into two categories i.e. building and non-building. The building layer which has polygon geometry was selected as a building label. Another class was determined as a non-building label. The results were used as reference maps to provide information about the building label. Figure 18 describe building polygons in vector format.

The attributes of the topographic map not only have information about building codes but also semantic information. The attribute of the reference map has to be fixed to provide an excellent corresponding label. The attribute was fixed by selection of database which still has an empty layer code. The empty attribute was corrected by filling an attribute code as building. Also, the database has semantic information. The code of building object was determined based on its function such as sports center, school, mosque, and graveyard. The code of building object was selected and altered into the same code as building to obtain a correct and consistent building label. This process was applied by using ArcMap software. The results of this process were a fixed database that contains a building object based on its geometry.



Figure 18 The topographic map in vector format

3.5.2. Polygon to Raster Conversion

The polygon was converted into a raster image. The pixel size is 1 x 1 meter so that it has the same size input as the point cloud image. The polygon to raster conversion was carried out using ArcMap. The value field that was used to convert the polygon was layer code which has label information. The result of this step was an image that consists of two labels i.e. building and non-building area.

Since training network cannot read zero values from the corresponding label, and the result of the conversion of polygon to raster still contains 0 value as a non-building, the pixel value was changed from 0 to 2 to support the network. So, it provided value 1 for building object and 2 for non-building object. Figure 19 show purple color represents building object while non-building object was represented as yellow. This image was used as a corresponding label image for classifications.

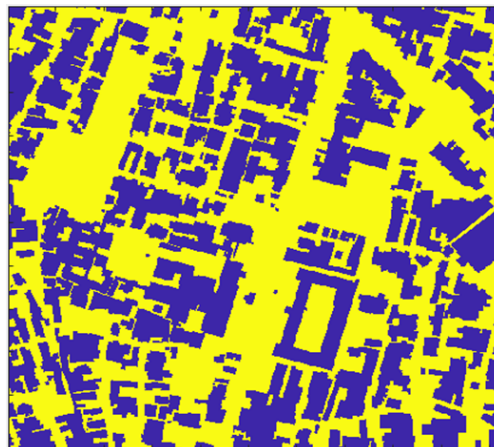


Figure 19 The raster image of the corresponding label

3.6. Point to image conversion

Input of the network was in image format (Long et al., 2015). Before classification could be performed, point cloud data should be converted into an image. The conversion was the initial process for providing input of the network. The ALS data contains a set of feature that can be used to classify an object such as intensity, height, return number, number of return and local 3D geometry (Clément Mallet, Chehata, & Bailly, 2016b). The features that were used to detect building in this study were intensity, height, and number of return as performing in previous research (Rizaldy et al., 2018). In order to detect building objects, the highest point within the pixel was used (Rizaldy et al., 2018). The building generally has high value than the ground object.

The pixel value was acquired from the value of the feature information. However, the limitation in this step was not all of feature information value can be shown in the pixel value. The pixel size of the image was 1 x 1 meter. It depends on the point density. There was at least one point that was expected within this pixel size (Rizaldy, 2018). The multi-dimensional image was generated from this step. Also, to obtain the same range of values as the channel image, the normalization of the image was processed.



Figure 20 The elevation image

The height feature can be one of the feature to differentiate objects in an urban environment (Guo et al., 2011). Figure 20 depicts the elevation information from the highest point. This image has a height value of the point cloud data. With this channel image, the object was distinguished based on the height. The building object has a high value than a non-building object.

Besides elevation, the intensity as feature information was also used as a channel image. The intensity feature was used to give an accurate object classification (Charaniya et al., 2004; Godin, Rioux, & Baribeau, 2005; Gressin, Mallet, Demantké, & David, 2013; Clément Mallet et al., 2016b). This feature was converted into an image from the fourth column in the matrix data. The result of the intensity image can be seen in Figure 21.

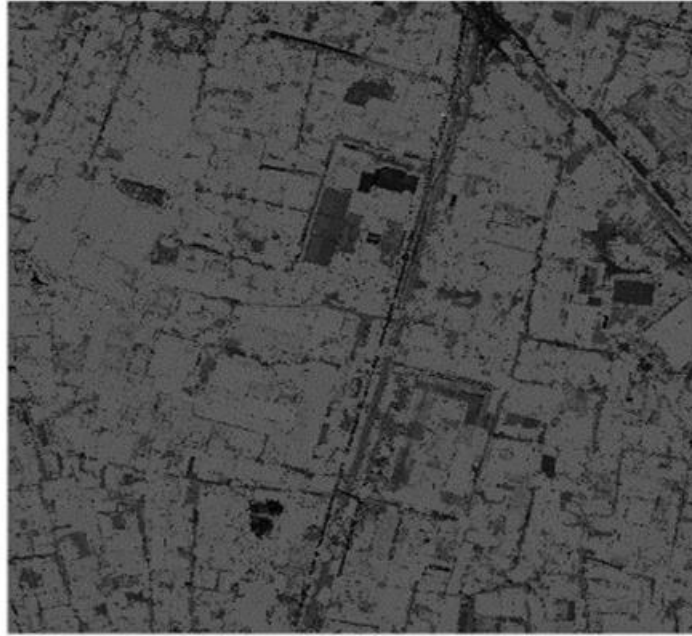


Figure 21 The intensity image

To distinguish between building and non-building objects, the number of returns was chosen as one of the channel images. The building object has only one number of returns than another object. Therefore, this feature can help to discriminate the building and non-building object. The number of returns image was created by converting the point cloud. This feature was transformed into the image from the fifth column in the matrix data (see in Figure 22).



Figure 22 The number of echoes

In the previous study, the height difference was used as additional information (C Mallet, Soergel, & Bretar, 2008; Niemeyer, Rottensteiner, & Soergel, 2012; Rizaldy et al., 2018). The height difference was generated by subtracting the elevation image with the lowest point in the particular window. By using 20 x 20 m horizontal window to capture conditions in the neighborhood that can help to discriminate the building object with another object. Figure 23 shows the height difference image.

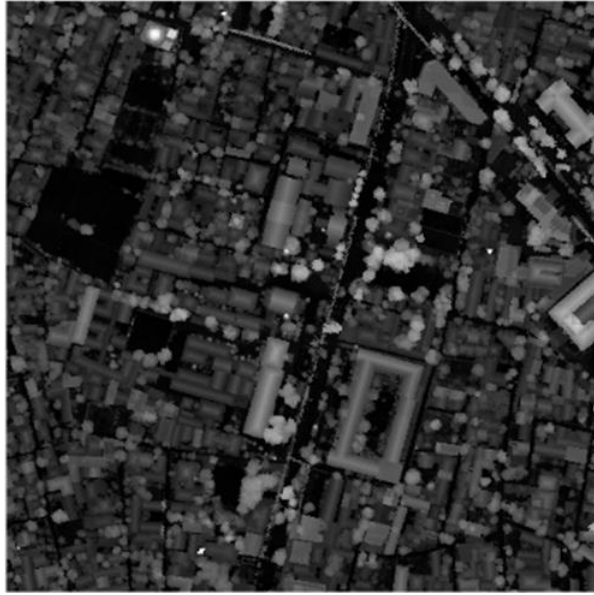


Figure 23 The height difference image

3.7. Data Augmentation

An overview of all design steps that consists of four phases can be seen in Figure 24. The training and validation datasets have to be defined to provide the input of the network. The training data was chosen 80% from the dataset, and testing data was chosen 20 % of the dataset. The input of the training network is a matrix that consists of feature information from point cloud data and corresponding label from 2D map data. Intensity, elevation, number of echoes and height differences was concatenate as multidimensional channel image. The corresponding label were obtained from 2D map information that labelled into building and non-building object. Increasing the number of training data is necessary to obtain more parameter to learn for the network. The object variation of the image can help the network more robust and more independent (Wardhana, 2018).

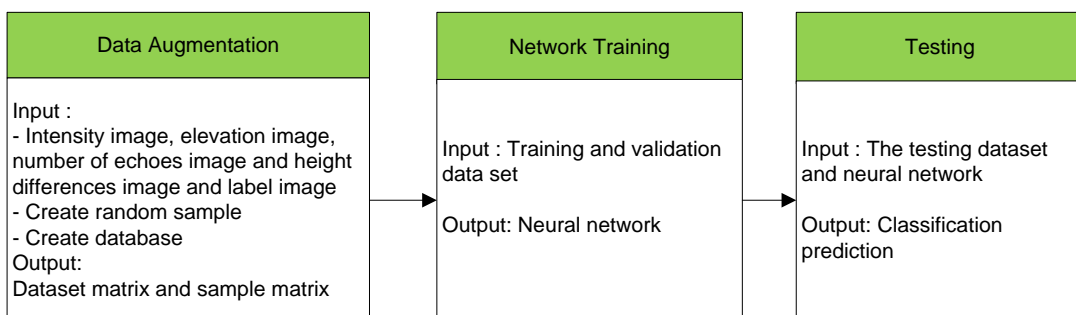


Figure 24 The overview classification steps

3.8. Network Training

In this study, for classification we used FCN_DK network by Persello and Stein, (2017). The network construction used no downsampling FCN architecture. It means the network can work without adding a more learnable parameter (Yu & Koltun, 2016). The filters used 5 x 5 kernels. The setting of the stride was one so that the filter moved without any gaps within image and avoided downsampling (Persello & Stein, 2017; Rizaldy, 2018). The dilation factors increased from one to six in order to handle the larger coverage (Persello & Stein, 2017); the layers can be seen in Table 1. The network was employed using MatConvNet.

Table 1 FCN layers

Input	Filter dimension (pixels)	Dilation	Stride	Pad
Convolution	5 x 5 x 3 x 16	1	1	2
Batch Norm				
ReLU				
Convolution	5 x 5 x 16 x 32	2	1	4
Batch Norm				
ReLU				
Convolution	5 x 5 x 32 x 32	3	1	6
Batch Norm				
ReLU				
Convolution	5 x 5 x 32 x 32	4	1	8
Batch Norm				
ReLU				
Convolution	5 x 5 x 32 x 32	5	1	10
Batch Norm				
ReLU				
Convolution	5 x 5 x 32 x 64	6	1	12
Batch Norm				
ReLU				
Convolution	1 x 1 x 64 x 2	1	1	0
Batch Norm				
Dropout				
Softmax				

3.9. Network Testing

The network testing is a process to test the trained network that can predict the building object from testing data. The input of the testing data is the data that are not used in network training. The input of this process is trained network and the test dataset. The trained network is the result from training network from previous step. The testing data have the same type with the training data. The result of this step is the prediction of the building object in tested data. The building objects were assigned class 1.

3.10. Accuracy Assessment

The quantitative of the result was assessed by comparing the result and the reference data. By comparing the area between the resulted objects and reference data, the completeness and correctness assessment were conducted (Bittner et al., 2017; Foody, 2002; Widyaningrum & Gorte, 2017). In this study, the number of true positive (TP), false positive (FP) and false negative (FN) were used to calculate the completeness and the correctness. True positive (TP) is the same area between reference data and the result. False Positive (FP) determines the only area in the reference data. False Negative (FN) defines the only area in the result data. The equation of the accuracy assessment can be seen below.

$$\textit{Precision} = TP / (TP + FP) * 100\%$$

$$\textit{Recall} = TP / (TP + FN) * 100\%$$

$$\textit{F1 score} = 2 * \frac{\textit{Precision} * \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

4. RESULTS AND DISCUSSION

4.1. Results

4.1.1. Cleaning the training sample

Based on information from 2D maps, the location and the shape of buildings can easily be recognized. The point cloud was depicted based on the height in Figure 25. The red lines represents building outlines. Blue lines show the area that has tree points within the polygon. These non-building points that lie within the polygon can give a wrong label that influences classification results. The points which were not corresponding to the building objects were removed manually from this review.

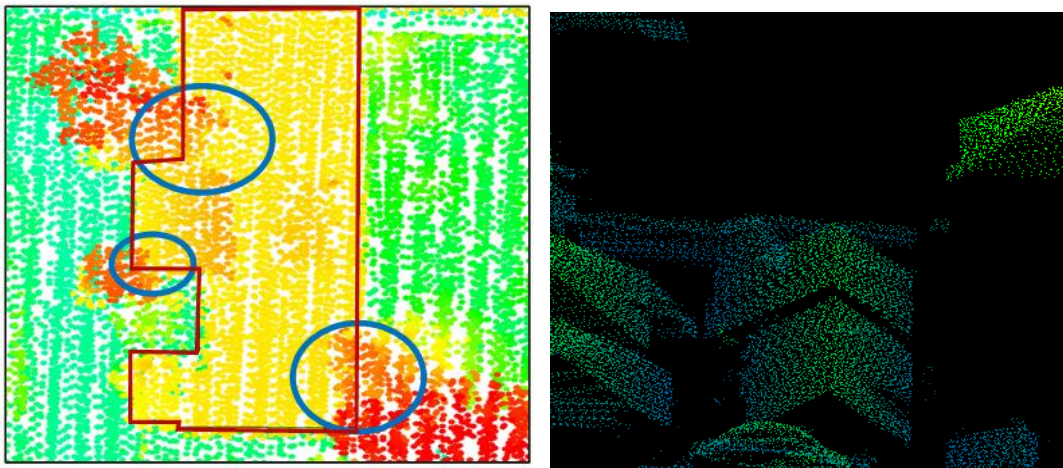
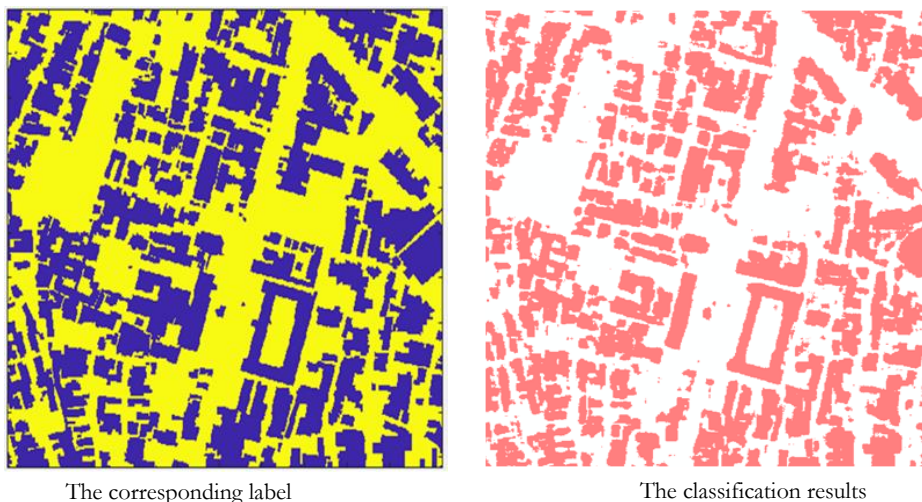


Figure 25 The point cloud visualized based on height that combined with polygon map

4.1.2. Training Result

This section provides the results that were generated by using the training sample from sites located in Lombok. Figure 26 shows the results that were generated by using the training sample in the same region. The training sample used 17 sites which were all sites located in Lombok except site 01, site 14 and sites 16. The sites were not using training sample that was used as testing data. The testing sites were chosen based on sites that have a densely built area.



The corresponding label

The classification results

Figure 26 The corresponding label (left) and classification result (right), the building object shows in red colour

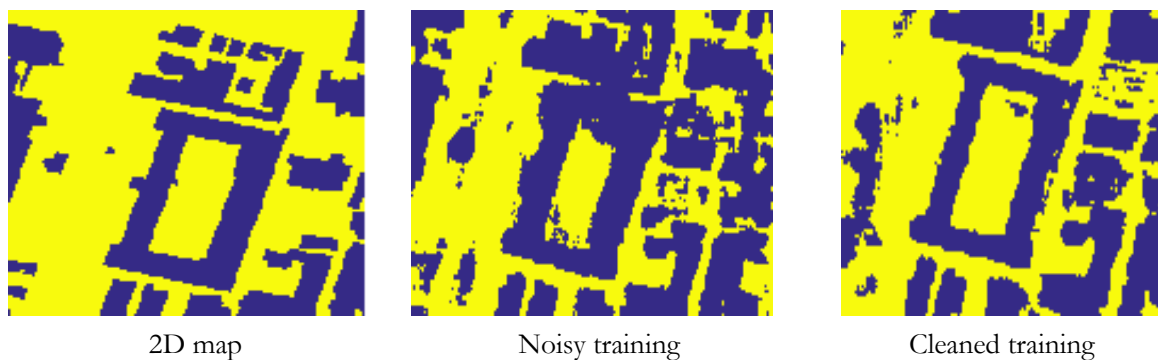
The building object mostly could be detected well. However, some areas were still unclassified and over-classified. The over-classified pixels were caused by the tree points located within building polygon. The network classified the tree object pixels as a building.



Site 01

Figure 27 The classification results within same region

The cleaning process was done for 20 sites in Lombok that used as training data. The cleaned training samples improved the results of the classification. By removing points that were not corresponding to the building objects within a polygon map reduced misclassification. Also, it minimized the overclassify points that did not correspond to building objects. presents comparison of classification results with reference map and classification results from noisy training samples and cleaned training samples see in Figure 28. The noisy results have over-segmented not only close to building but also the surrounding area because pixel predicted from the wrong information from noisy dataset. The clean data represented the building prediction almost similar with 2D map data. That means cleaning the data will strengthen the prediction result.



2D map

Noisy training

Cleaned training

Figure 28 The comparison between noisy and cleaned training sample

4.1.3. The Classification Result

Buildings as objects from another region were depicted in Figure 29. The classification used 20 sites in Lombok region. The training result was employed in another region to know the influences of training samples to the classification results. The building object in Site 21 could be detected well. However, over-classify and misclassification problems were still occurred. The misclassification was occurred in the buildings that were near or covered by another object like a tree. From the image below, we can see that Site 23 was poorly classified. Site 23 has a slope terrain and high vegetations, meanwhile, Lombok has a flat terrain. Therefore, the sloping terrain was misclassified as building in this site. Site 30 has a small building that has a similar size and low vegetations. In this situation, the classification result became overclassified. The other result in differences region can be seen in APPENDIX III PREDICTION RESULT

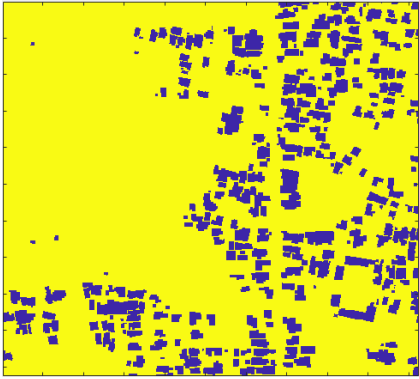
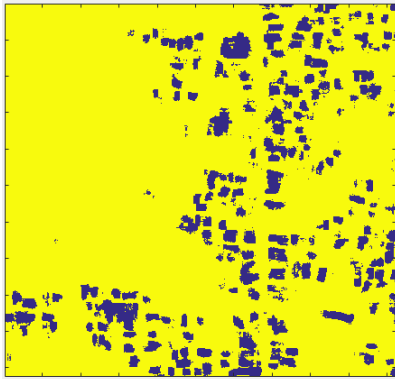
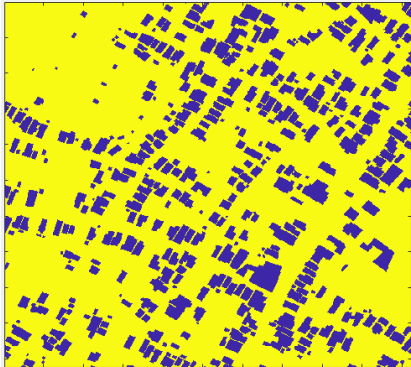
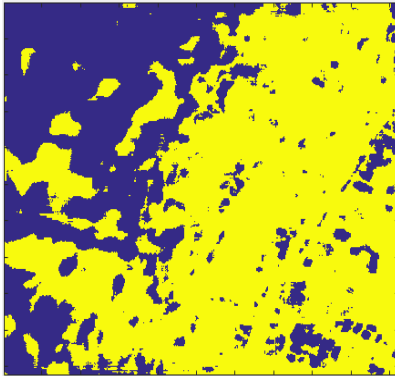
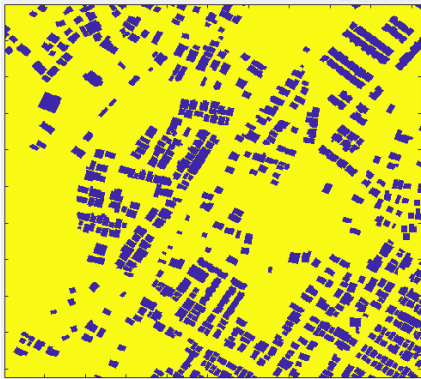
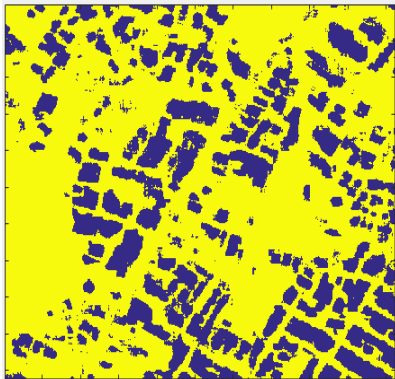
Site	Topographic map	FCN
Site 21		
Site 23		
Site 30		

Figure 29 Building object in the different region

In Figure 31 depicts the matching between classification results with a vector of a topographic map. The over-classifications were occurred outside the map polygons. Overclassified was represented by red color. From the visual review, there was some areas that have unclassified building objects. It could happened

because the building has a small size or it has a same elevation as its neighborhood, it can be seen in Figure 30.

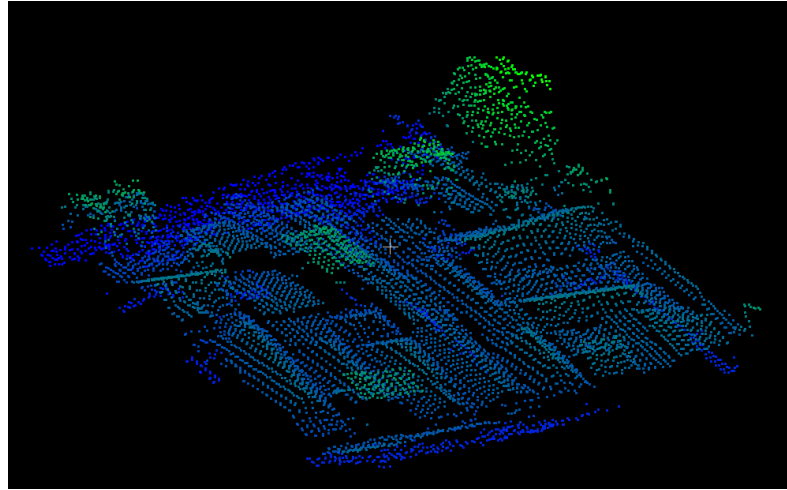


Figure 30 Small buildings that have similar elevation with the neighbourhood



Figure 31 The classification result

4.2. Accuracy Assessment

Accuracy assessment was assessed by estimating the overall accuracy and F-Score as shown in Table 2. The tested data divided into two region that in the same region and another region. The same region means that the tested data in Lombok using training data from Lombok as well. In Lombok region has 20 sites that divided into training dan testing data. The training data using 17 sites except site 01, 14, and 20 were used as test data the Lombok region. From RMSD the calculation of the F-score in Lombok region, the cleaning data perform better than the noisy data with RMSD equal 1.03 %.

Tanggamus, Tanjung Lesung and Makassar regions were tested using 17 sites and 20 sites in Lombok. From the accuracy assessment results, site 23 in Tanggamus that have steep terrain has lower F-score value. The

steep terrain has been classified as building in that region. The condition of the area in particular terrain condition influenced the accuracy results. The area that has less low vegetation like in site 21 has better accuracy than another region that has an obstacle. Table 3 present the performance of the trained network in another region. The table compare the accuracy from each region, it described the number of sites enhance the performance of the classification.

Table 2 Evaluation of the FCN performance in Lombok Region

Region	Sites	F-score (%)		
		Noisy	Cleaned	RMSD
Lombok	1	86.0367	87.5581	1.034115
	14	88.5315	89.2013	
	20	84.598	85.265	

Table 3 The performance for another region

Region	Sites	F-score (%)					
		17 Sites		RMSD	20 Sites		RMSD
		Noisy	Cleaned		Noisy	Cleaned	
Tanjung Lesung	21	81.8319	81.21952	27.24678	56.3315	82.9668	20.33271
	22	73.2322	34.7043		64.1024	74.9375	
Tanggamus	23	37.3173	42.4122	13.65823	16.5933	58.3018	30.57
	24	36.8932	41.3561		30.1685	50.7901	
	25	37.693	49.1162		21.7424	51.4634	
	26	32.0876	58.1342		23.4811	39.4064	
	27	49.3055	58.1342		22.8123	59.8369	
Makassar	28	68.4996	69.5552	10.03097	54.678	76.5218	19.70159
	29	66.587	75.3282		60.652	78.4257	
	30	71.6978	72.6843		68.8359	79.4819	
	31	59.4776	71.3201		59.534	75.29	
	32	42.6272	59.4902		37.7318	65.7696	

5. CONCLUSION

5.1. Conclusion

The FCN network was transferred to regions with point density around 4-12 points per square meter and does not have accurate training samples in Indonesia. By using spatial information from topographic maps such as corresponding labels can support the classification of ALS data using FCN. Based on the experiments, the proposed method which used an automatic approach could be used to detect buildings from a coarse point cloud data. Furthermore, by using feature information from point cloud data such as intensity, number of returns, elevation and height difference could support building detection.

The spatial information from 2D map help to easily recognize the variation shape of the building. The 2D map data was used as clip feature, corresponding label and reference data to validate the test result. The topographic map can be used as the corresponding label that can give information about the position of building objects. The polygon lines can help to perceive building objects. Not only location of the buildings but also the shape of the buildings.

The trained network was tested into different region in Indonesia. The accuracy assessment calculated based on the F-Score. From RMSD in Lombok region, the cleaning data perform better than the noisy data with RMSD equal 1.03 %. The number of sites enhance the performance of the classification. The RMSD in another region that in Tanjung Lesung 20.33%, Tanggamus 30.57% and Makassar 19.70 %.

By overlaying vector data with classification results, the misclassifications and overclassified occurred in the area that has small building and dense vegetation. The results present building objects that could be detected by using FCN network. The availability of training samples in the same region could improve the classification results.

5.2. The answer to Research Question

Sub-objective 1:

1. How can existing spatial geoinformation such as 2D map be used to help to extract training samples?

The topographic map can be used as the corresponding label that can give information about the position of building objects. The polygon lines can help to recognize building objects easily. Not only location of the buildings but also the shape of the buildings. However, from visual inspection, the 2D map still has an error like the points of trees and ground areas belong to the polygon building. The 2D map has several layers that can be used to extract another layer such road, water body and another land cover area. The spatial information from 2D map help to easily recognize the variation shape of the building. The 2D map data was used as clip feature, corresponding label and reference data to validate the test result. The points clipped using polygon from 2D map data to assign points that belong to building objects.

2. How to collect the representative training samples?

The training sample was collected and cleaned by using the 2D map data. The 2D map data was acquired at the same time with ALS data. It was manually digitized based on ground orthophoto to give a representation of building objects.

3. What is point feature information of ALS data that can be used to improve detection?

In Section 3.6 Point to image conversion discussed the feature information that can be used to improve the detection such as intensity, number of returns, elevation and height difference.

4. How to employ those point features in deep learning methods for building detection?
The point features were converted from point into image as a multidimensional channel such as intensity, number of returns, elevation and height difference. This multidimensional channel can be used as an input in the network. The conversion from point to image is explained in Section 3.6 Point to image conversion.

Sub-objective 2:

How can a deep learning approach be used to detect building from coarse ALS data?

This approach used a multidimensional channel image from point features and the topographic map as an input to detect the building. The output was an image that has two classes namely building and non-building.

Sub-objective 3:

1. What are the accuracy and the performance of the proposed methods with completeness and correctness such as F1 score or kappa index?
The classification resulted in accuracy assessment was depended on the condition of the region. The classification was done with two classes such as building and non-building. The accuracy in the same region was around 85%.
2. How can the proposed method be employed in another region?
The training sample in Lombok was employed in another region. Tanggamus has a steep terrain and has a low accuracy. However, this training sample gave better classification results in a flat terrain such as in Tanjung Lesung and Makasar.

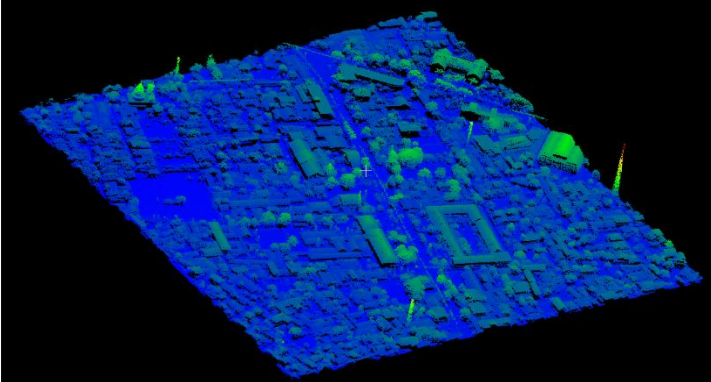
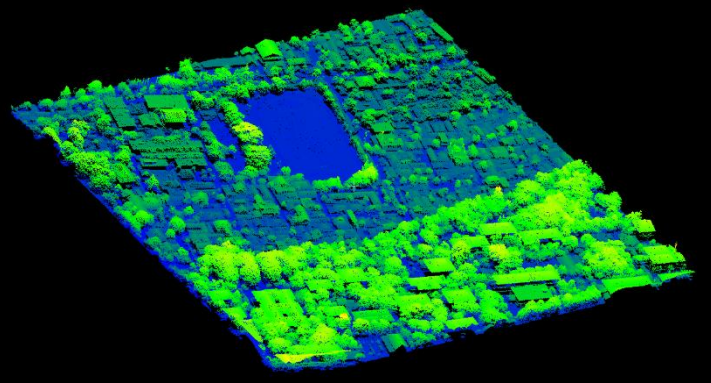
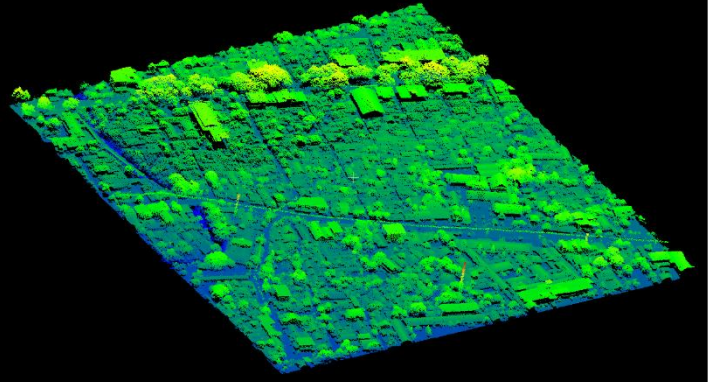
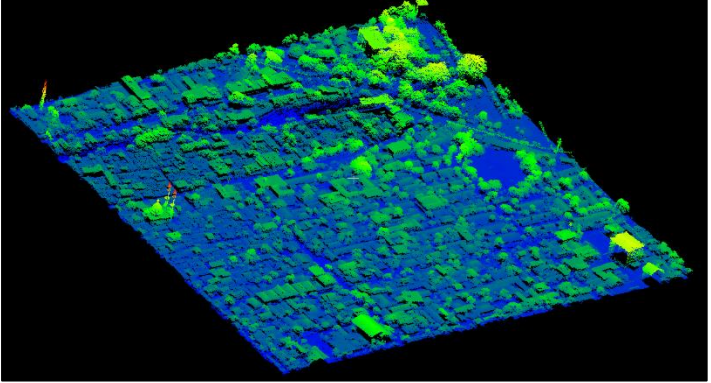
LIST OF REFERENCES

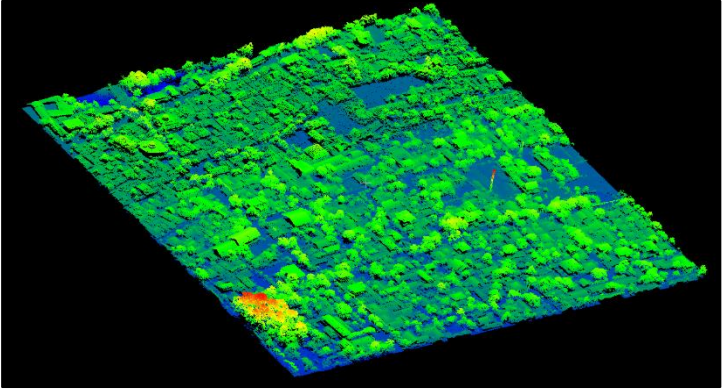
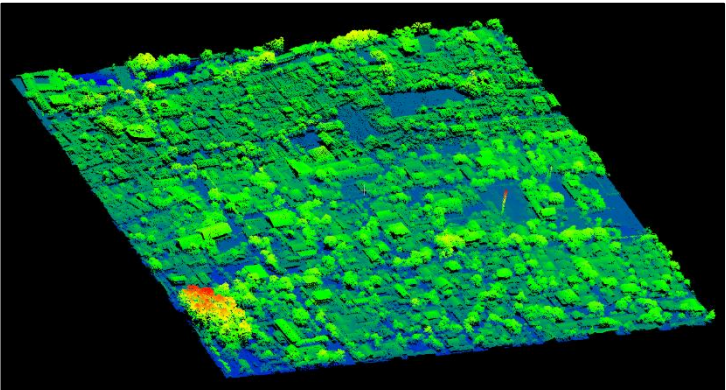
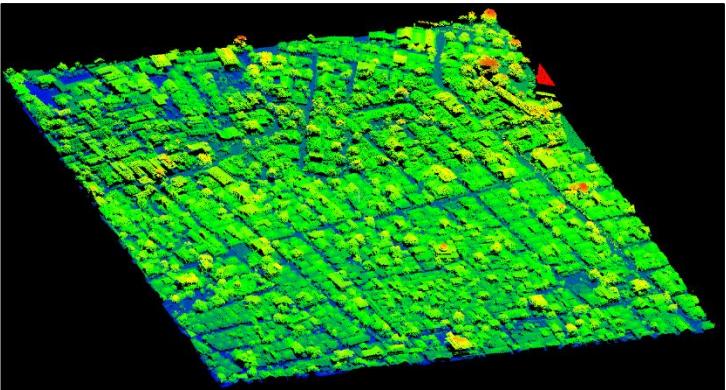
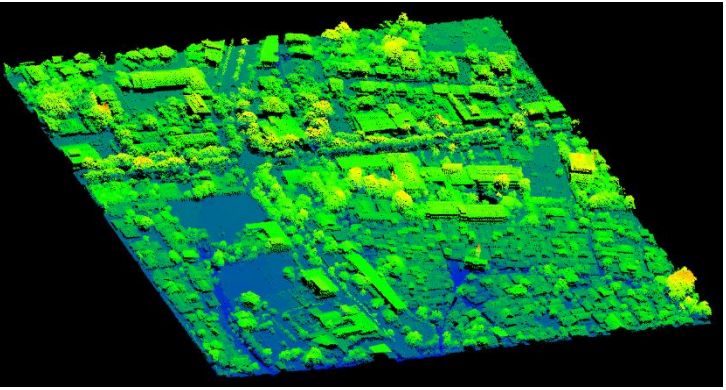
- Bartel, M. B., & Wei, H. (2000). Rule-based Improvement of Maximum Likelihood Classified LIDAR Data fused with co-registered Bands Marc, *162*(37), 4934–4937. Retrieved from <http://www.cvg.reading.ac.uk/projects/LIDAR>
- Bishop, C. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.679.1104&rep=rep1&type=pdf>
- Bittner, K., Cui, S., & Reinartz, P. (2017). Building Extraction From Remote Sensing Data Using Fully Convolutional Networks. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-1/W1*, 481–486. <https://doi.org/10.5194/isprs-archives-XLII-1-W1-481-2017>
- Bretar, F., Roux, M., Mallet, C., Heipke, C., & Soergel, U. (2011). Relevance assessment of full-waveform lidar data for urban area classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(6), S71–S84. <https://doi.org/10.1016/j.isprsjprs.2011.09.008>
- Charaniya, A. P., Manduchi, R., & Lodha, S. K. (2004). Supervised Parametric Classification of Aerial LiDAR Data 2. Background and Previous Work 3. Data Classification. *Conference on Computer Vision and Pattern Recognition Workshop*. <https://doi.org/10.1109/CVPR.2004.446>
- Davydova, K., Cui, S., & Reinartz, P. (2016). Building footprint extraction from digital surface models using neural networks. In *Image and Signal Processing for Remote Sensing XXII* (Vol. 10004, p. 100040J). <https://doi.org/10.1117/12.2240727>
- Elberink, S. O. (2008). Problems in Automated Building Reconstruction Based on Dense Airborne Laser Scanning Data. *Archives*, *37*(Part B3a), 93–98. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.157.187&rep=rep1&type=pdf>
- Elberink, S. O., & Vosselman, G. (2009). Building reconstruction by target based graph matching on incomplete laser data: Analysis and limitations. *Sensors*, *9*(8), 6101–6118. <https://doi.org/10.3390/s90806101>
- Erhan, D., Manzagol, P.-A., Bengio, Y., Bengio, S., & Vincent, P. (2009). The Difficulty of Training Deep Architectures and the Effect of Unsupervised. *International Conference on Artificial Intelligence and Statistics*, *5*, 153–160. Retrieved from <https://www.semanticscholar.org/paper/The-Difficulty-of-Training-Deep-Architectures-and-Erhan-Manzagol/5ecc56c3d60a2fa0f1c698d0ecbd280b9c50e240>
- Foody, G. M. (2002). Status of Land Cover Classification Accuracy Assessment. *Remote Sensing of Environment*, *80*(1), 185–201. [https://doi.org/10.1016/S0034-4257\(01\)00295-4](https://doi.org/10.1016/S0034-4257(01)00295-4)
- Gevaert, C. M., Persello, C., Elberink, S. O., Vosselman, G., & Sliuzas, R. (2017). Context-Based Filtering of Noisy Labels for Automatic Basemap Updating From UAV Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(8), 2731–2741. <https://doi.org/10.1109/JSTARS.2017.2762905>
- Ghaffarian, S., Ghaffarian, S., El Merabet, Y., Samir, Z., & Ruichek, Y. (2016). Automatic Building Roof Segmentation Based On PFICA Algorithm And Morphological Filtering From Lidar Point Clouds. *Asian Conference on Remote Sensing*. Retrieved from https://webapps.itc.utwente.nl/library/2016/conf/ghaffarian_aut.pdf
- Godin, G., Rioux, M., & Baribeau, R. (2005). Three-dimensional registration using range and intensity information. *Videometrics III*, *2350*(October 1994), 279–290. <https://doi.org/10.1117/12.189139>
- Gressin, A., Mallet, C., Demantké, J. Ô., & David, N. (2013). Towards 3D Lidar Point Cloud Registration Improvement Using Optimal Neighborhood Knowledge. *ISPRS Journal of Photogrammetry and Remote Sensing*, *79*, 240–251. <https://doi.org/10.1016/j.isprsjprs.2013.02.019>
- Grilli, E., Menna, F., & Remondino, F. (2017). A review of point clouds segmentation and classification algorithms. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(2W3), 339–344. <https://doi.org/10.5194/isprs-archives-XLII-2-W3-339-2017>
- Guo, L., Chehata, N., Mallet, C., & Boukir, S. (2011). Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(1), 56–66. <https://doi.org/10.1016/j.isprsjprs.2010.08.007>
- Haala, N., & Anders, K. H. (1997). Acquisition of 3D urban models by analysis of aerial images, digital surface models and existing 2D building information. In *SPIE Conference on Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III, SPIE Proceedings* (Vol. 3072, pp. 212–222). <https://doi.org/10.1117/12.281041>

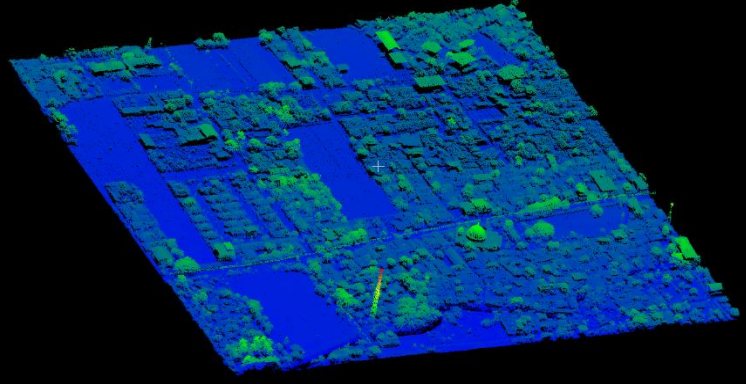
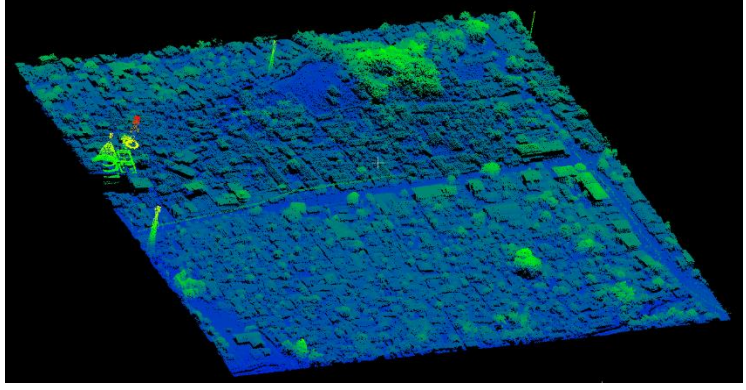
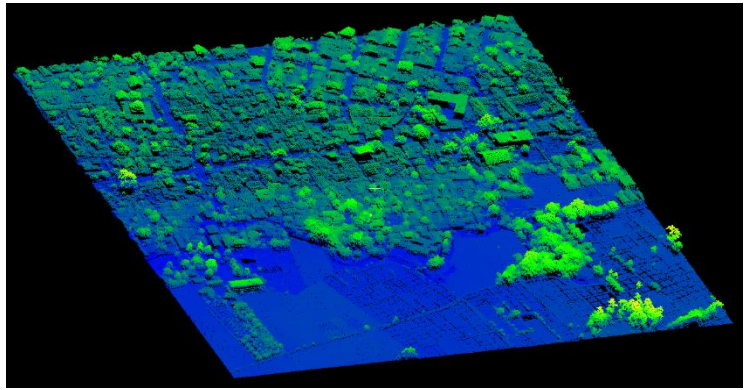
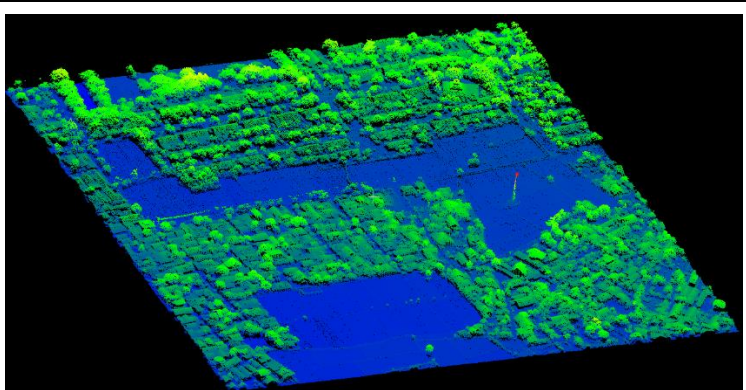
- Haralick, R. M., & Sternberg, S. R. (1987). Image Analysis Morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (4), 532–550. Retrieved from <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4767941>
- Hofmann, A. D., Maas, H.-G., & Streilein, A. (2002). Knowledge-Based Building Detection Based On Laser Scanner Data And Topographic Map Information. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34(3/A), 169–174. Retrieved from https://tu-dresden.de/bu/umwelt/geo/ipf/photogrammetrie/ressourcen/dateien/forschung/publikationen/pubdocs/2002/2002_Hofmann_Maas_Streilein_ISPRSComIII2002.pdf?lang=en
- Huang, M.-J., Shyue, S.-W., Lee, L.-H., & Kao, C.-C. (2008). A Knowledge-based Approach to Urban Feature Classification Using Aerial Imagery with Lidar Data. *Photogrammetric Engineering & Remote Sensing*, 74(12), 1473–1485. <https://doi.org/10.14358/PERS.74.12.1473>
- Hug, C., & Wehr, A. (1997). Detecting and identifying topographic objects in imaging laser altimeter data. *International Archives of Photogrammetry and Remote Sensing*, 32(3 SECT 4W2), 19–26. Retrieved from http://www.ifp.uni-stuttgart.de/publications/wg34/wg34_hug.pdf
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 07–12–June, 3431–3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- Mallet, C., Chehata, N., & Bailly, J.-S. (2016a). Airborne LiDAR Data Processing. *Optical Remote Sensing of Land Surface*, 249–297. <https://doi.org/10.1016/B978-1-78548-102-4.50006-5>
- Mallet, C., Chehata, N., & Bailly, J. S. (2016b). Airborne LiDAR Data Processing. In *Optical Remote Sensing of Land Surface: Techniques and Methods* (pp. 249–297). <https://doi.org/10.1016/B978-1-78548-102-4.50006-5>
- Mallet, C., Soergel, U., & Bretar, F. (2008). Analysis of Full-Waveform Lidar Data for Classification of Urban Areas. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII, 85–92. Retrieved from <http://www.csie.ntu.edu.tw/~cjlin/>
- Matikainen, L., Hyyppä, J., & Kaartinen, H. (2004). Automatic Detection Of Changes From The Laser Scanner And Aerial Image Data For Updating Building Maps. *LAPR/SIS, Vol. 35-B2*, (January), 434–439. <https://doi.org/10.1007/978-4-431-54108-0>
- Minh, N. ., & Hien, L. P. (2011). Land cover classification using LiDAR intensity data and neural network. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, 29(November 2016), 429–438. <https://doi.org/10.7848/ksgpc.2011.29.4.429>
- Mongus, D., Lukač, N., & Žalik, B. (2014). Ground and building extraction from LiDAR data based on differential morphological profiles and locally fitted surfaces. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 145–156. <https://doi.org/10.1016/j.isprsjprs.2013.12.002>
- Nguyen, M. Q., Atkinson, P. M., & Lewis, H. G. (2005). Superresolution Mapping Using A Hopfield Neural Network With LIDAR Data. *IEEE Geoscience and Remote Sensing Letters*, 2(3), 366–370. <https://doi.org/10.1109/LGRS.2005.851551>
- Niemeyer, J., Rottensteiner, F., & Soergel, U. (2012). Conditional Random Fields For Lidar Point Cloud Classification In Complex Urban Areas. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1–3. Retrieved from <https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/I-3/263/2012/isprsannals-I-3-263-2012.pdf>
- Niemeyer, J., Rottensteiner, F., & Soergel, U. (2014). Contextual Classification Of Lidar Data And Building Object Detection In Urban Areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, 152–165. <https://doi.org/10.1016/j.isprsjprs.2013.11.001>
- Persello, C., & Stein, A. (2017). Deep Fully Convolutional Networks for the Detection of Informal Settlements in VHR Images. *IEEE Geoscience and Remote Sensing Letters*, 14(12), 2325–2329. <https://doi.org/10.1109/LGRS.2017.2763738>
- Priestnall, G., Jaafar, J., & Duncan, A. (2000). Extracting urban features from LiDAR digital surface models. *Computers, Environment and Urban Systems*, 24(2), 65–78. [https://doi.org/10.1016/S0198-9715\(99\)00047-2](https://doi.org/10.1016/S0198-9715(99)00047-2)
- Rizaldy, A. (2018). *Deep Learning-Based Dtm Extraction From Lidar Point Cloud*. The University of Twente.
- Rizaldy, A., Persello, C., Gevaert, C., Elberink, S. O., & Vosselman, G. (2018). Ground and multi-class classification of Airborne Laser Scanner point clouds using Fully Convolutional Networks. *Remote Sensing*, 10(11), 1723. <https://doi.org/10.3390/rs10111723>
- Rizaldy1, A., Persello, C., Gevaert, C. M., & S.J. Oude Elberink. (2018). Fully Convolutional Networks For Ground Classification From Lidar Point Clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2(1). <https://doi.org/10.5194/isprs-annals-IV-2-231-2018>

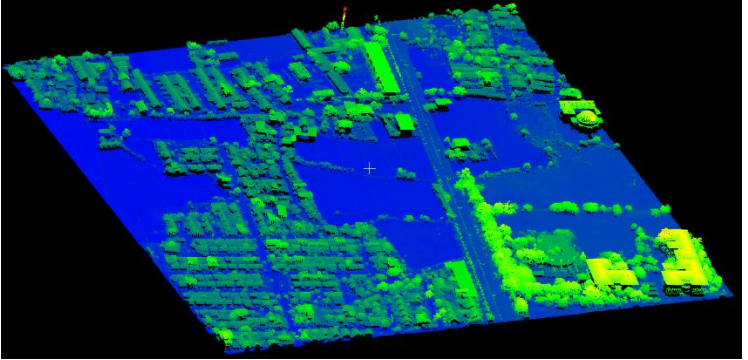
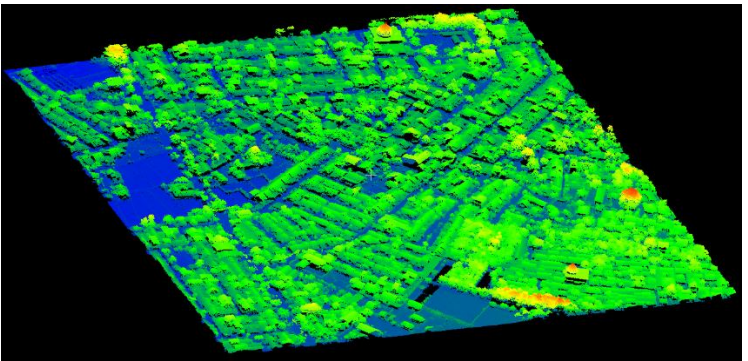
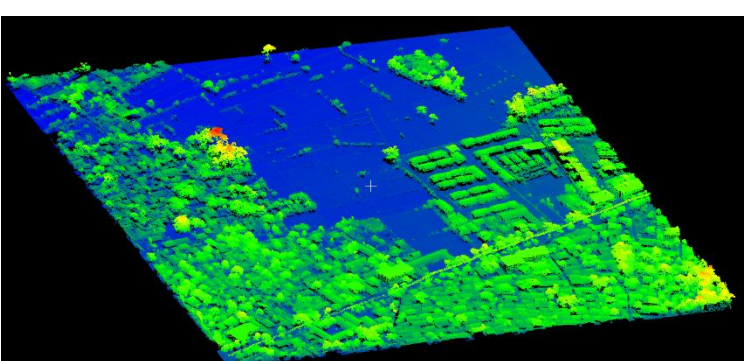
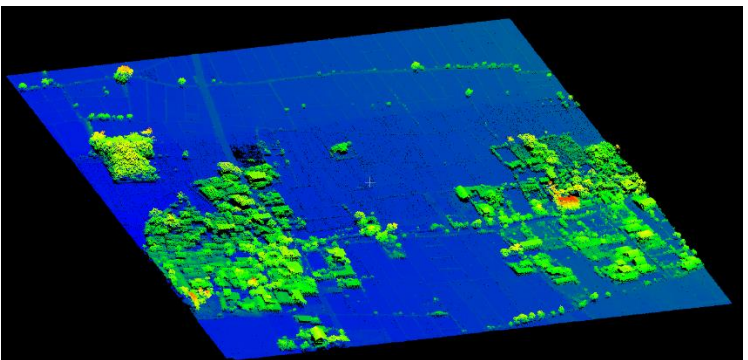
- Snuverink, I. A. . (2017). *Deep Learning for Pixelwise Classification of Hyperspectral Images*. The Delft University of Technology.
- Sotoodeh, S. (2006). Outlier Detection in Laser Scanner Point Clouds. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVI(Part 5), 297–302. Retrieved from <http://www.toposys.com/>
- Suveg, I., & Vosselman, G. (2001). *Map based building reconstruction from laser data and images*. In: *Proceedings of Automatic Extraction of Man- Made Objects from Aerial and Space Images (III)*. Ascona: Balkema Publishers.
- Vosselman, G., & Maas, H.-G. (2010). Airborne and terrestrial laser scanning. In *Airborne and terrestrial laser scanning* (Vol. 4, pp. 183–184). Whittles Publishing. <https://doi.org/10.1080/17538947.2011.553487>
- Vosselman, G., Gorte, B. G. H., Sithole, G., & Rabbani, T. (2004). Recognising structure in laser scanner point clouds. *Remote Sensing and Spatial Information Sciences*, 32(5), 33–38. <https://doi.org/10.1002/bip.360320508>
- Vosselman, G., & Dijkman, S. (2001). 3D building model reconstruction from point clouds and ground plans. *Int. Arch. of Photogrammetry and Remote Sensing*, XXXIV, 37–43. Retrieved from <http://www.isprs.org/proceedings/XXXIV/3-W4/pdf/Vosselman.pdf>
- Wang, Y., & Oude Elberink, S. J. (2016). Map Based Segmentation Of Airborne Laser Scanner Data. (p. 74). <https://doi.org/10.3990/2.424>
- Wardhana, G. (2018). *Automatic Segmentation and 3D Reconstruction of Liver and Tumor*. University of Twente.
- Weidner, U. (1997). Digital Surface Models for Building Extraction. *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, (September). <https://doi.org/10.1007/978-3-0348-8906-3>
- Weidner, U., & Förstner, W. (1995). Towards automatic building extraction from high-resolution digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4), 38–49. [https://doi.org/10.1016/0924-2716\(95\)98236-S](https://doi.org/10.1016/0924-2716(95)98236-S)
- Widyaningrum, E., & Gorte, B. G. H. (2017). Challenges and opportunities: One stop processing of automatic large-scale base map production using airborne lidar data within gis environment case study: Makassar City, Indonesia. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* (Vol. 42, pp. 365–369). <https://doi.org/10.5194/isprs-archives-XLII-1-W1-365-2017>
- Yan, W. Y., Shaker, A., & El-Ashmawy, N. (2015, March 1). Urban land cover classification using airborne LiDAR data: A review. *Remote Sensing of Environment*. Elsevier. <https://doi.org/10.1016/j.rse.2014.11.001>
- Yang, H. L., Yuan, J., Lunga, D., Laverdiere, M., Rose, A., & Bhaduri, B. (2018). Building Extraction at Scale Using Convolutional Neural Network: Mapping of the United States. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(8), 2600–2614. <https://doi.org/10.1109/JSTARS.2018.2835377>
- Yu, F., & Koltun, V. (2016). Multi-Scale Context Aggregation by Dilated Convolutions. *Proc. Int. Conf. Learn. Represent*, 1–13. <https://doi.org/10.16373/j.cnki.ahr.150049>
- Yuan, J. (2018). Learning Building Extraction in Aerial Scenes with Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(11), 2793–2798. <https://doi.org/10.1109/TPAMI.2017.2750680>
- Zhang, K., Chen, S. C., Whitman, D., Shyu, M. L., Yan, J., & Zhang, C. (2003). A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(4 PART I), 872–882. <https://doi.org/10.1109/TGRS.2003.810682>

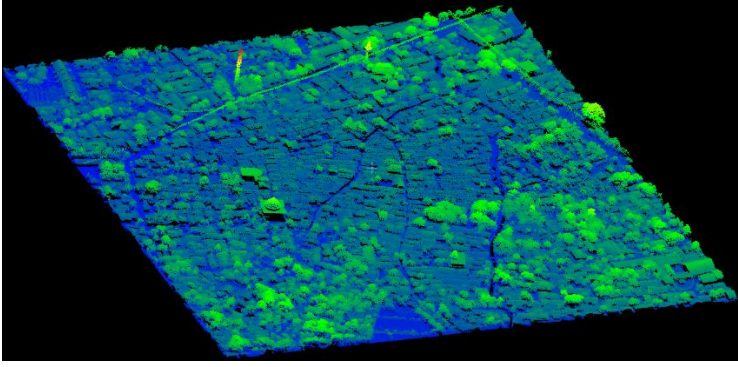
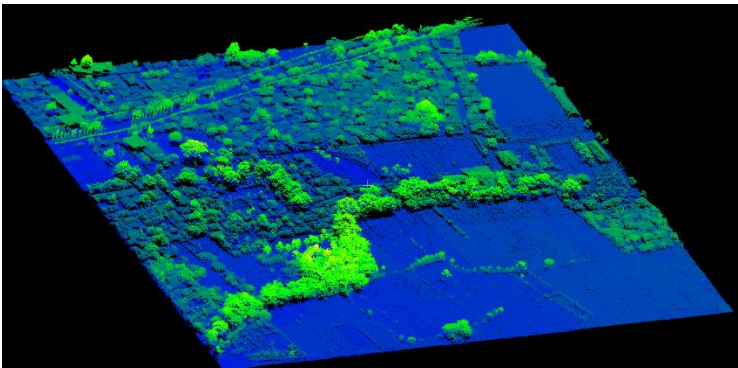
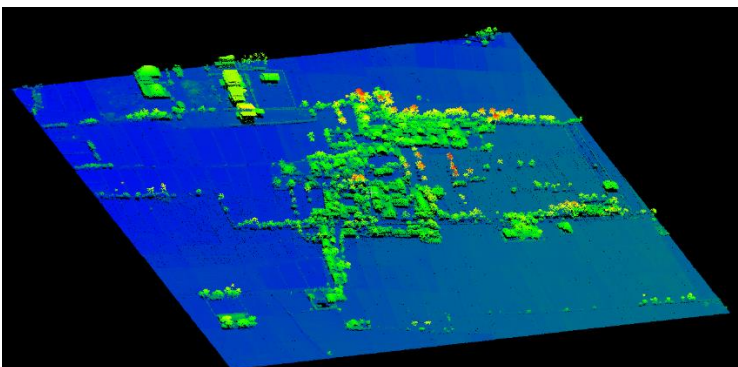
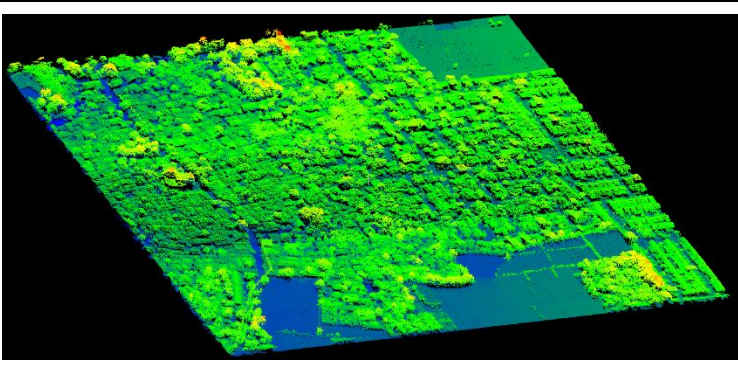
APPENDIX I: DATASET VISUALISATION

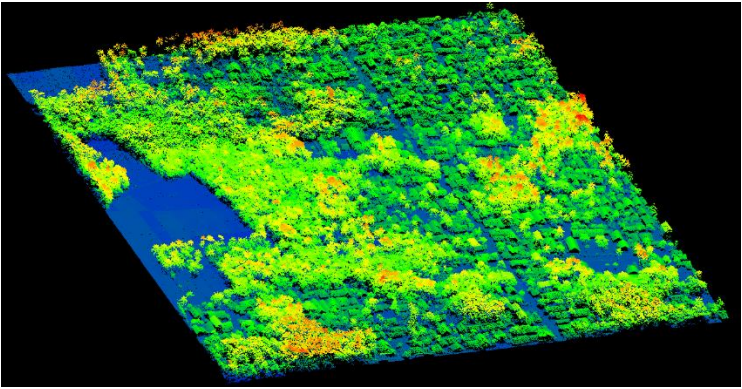
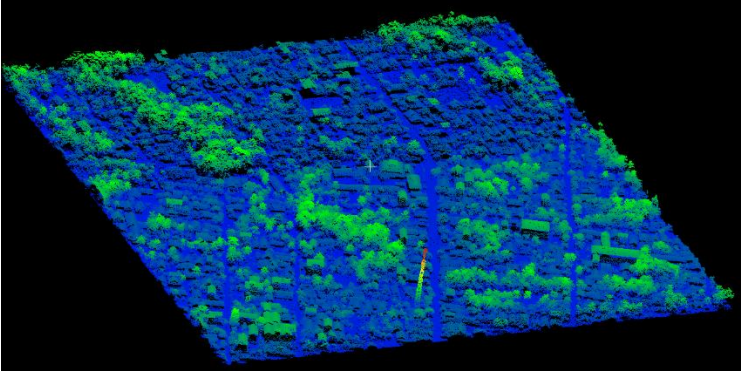
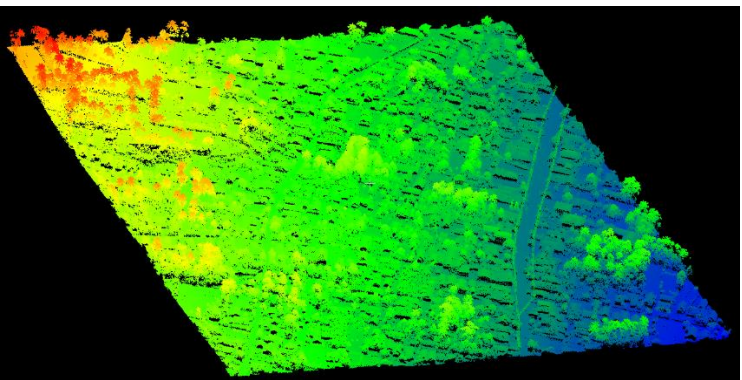
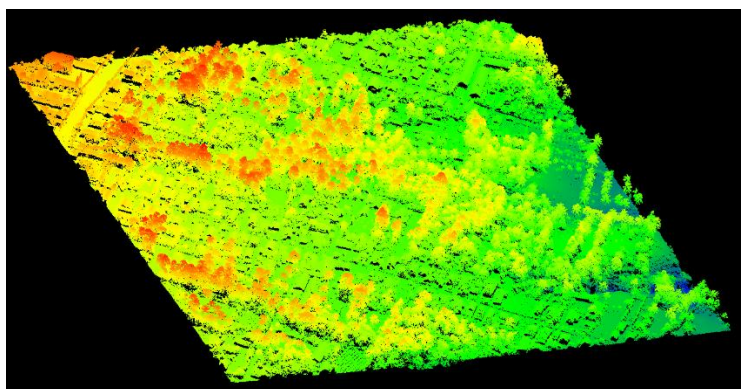
Site	Characteristic	Figure
Site 01	A large building, irregular shape building	
Site 02	A large building, irregular shape building with vegetation between them	
Site 03	Densely packed building	
Site 04	Densely small building with the irregular shape	

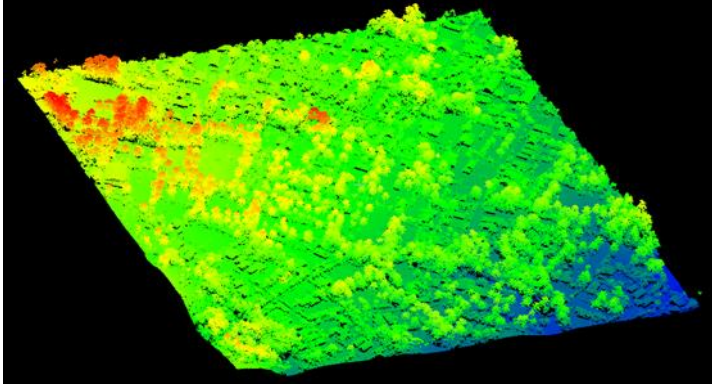
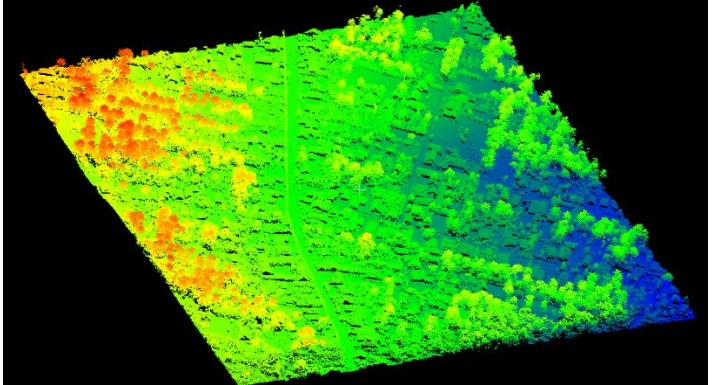
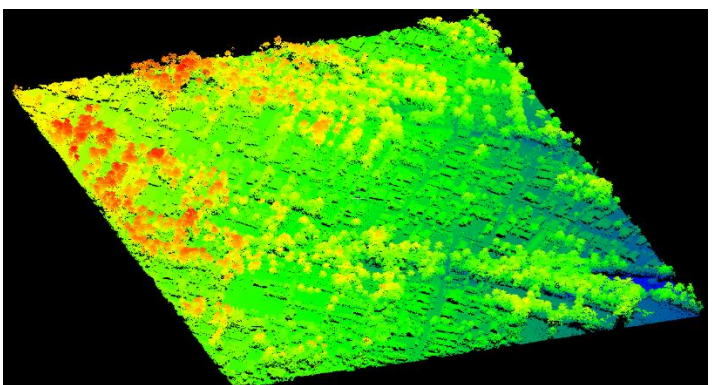
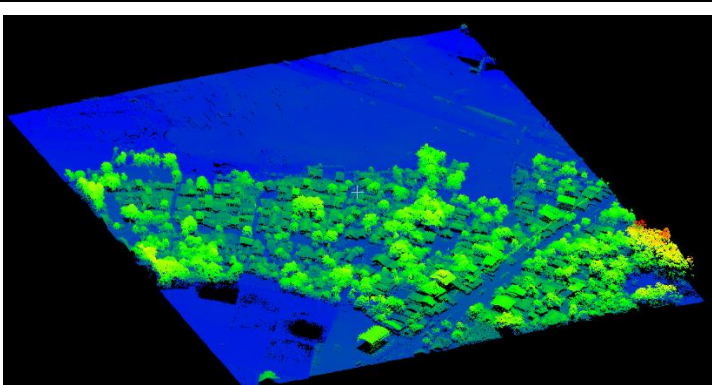
Site	Characteristic	Figure
Site 05	Densely small building	
Site 06	Densely small building	
Site 07	Densely small building	
Site 08	Densely large building	

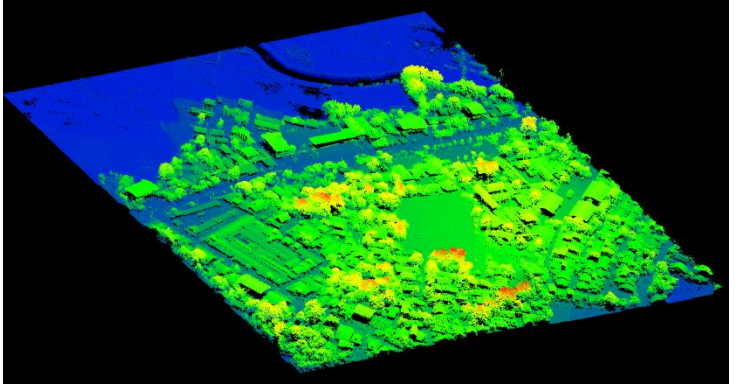
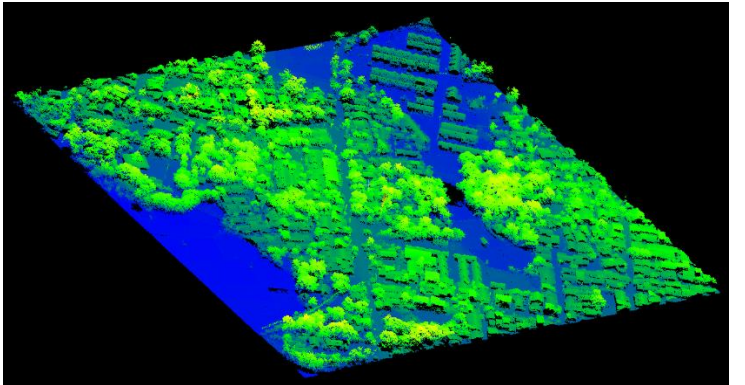
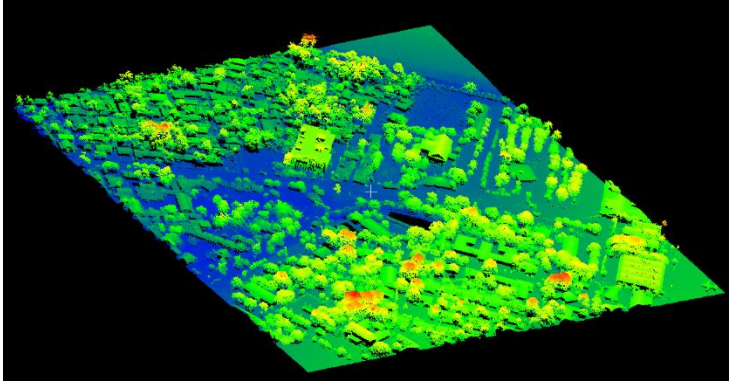
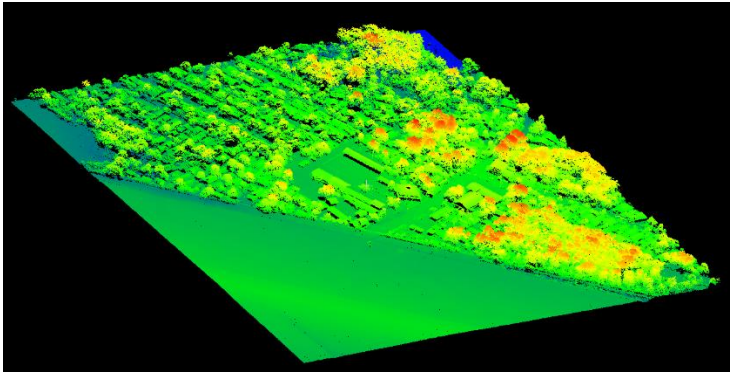
Site	Characteristic	Figure
Site 09	Densely small building	
Site 10	Densely small building, irregular shape	
Site 11	Densely small building, irregular shape	
Site 12	Densely small building, irregular shape	

Site	Characteristic	Figure
Site 13		
Site 14	The densely packed building, irregular shape building	
Site 15	The densely packed building, irregular shape building	
Site 16	Rural area, rarely small building	

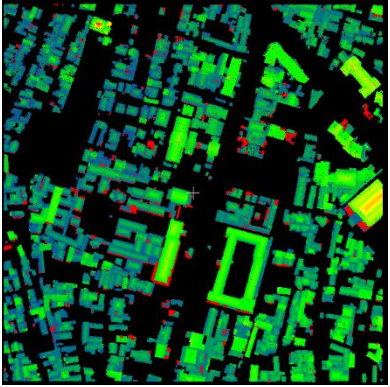
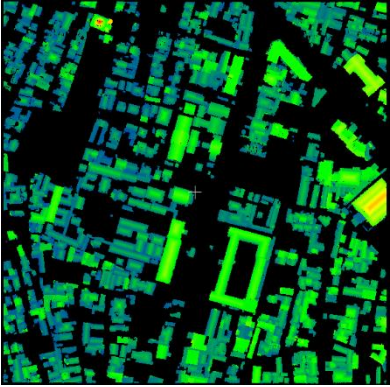
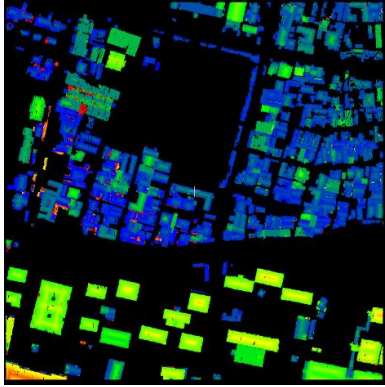
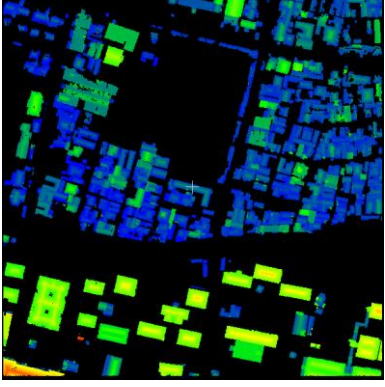
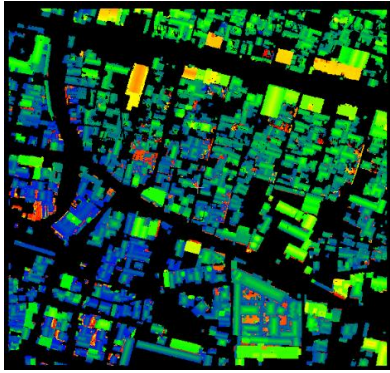
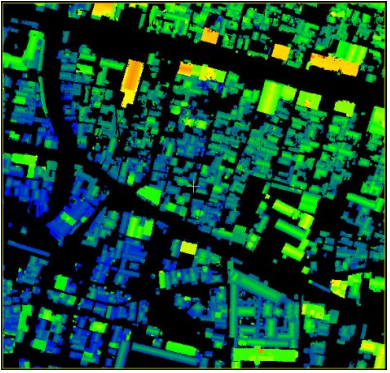
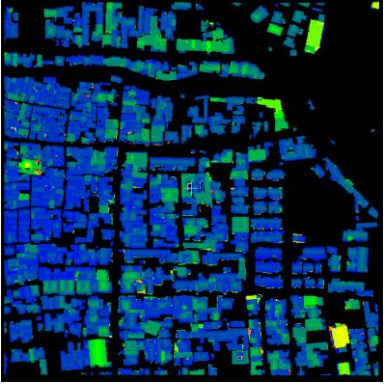
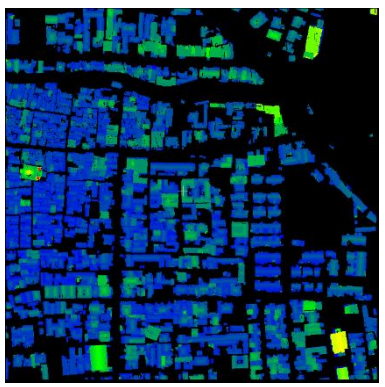
Site	Characteristic	Figure
Site 17	Densely packed building with the irregular shape	
Site 18	Rural area, dense small building	
Site 19	Rural area, rarely small building	
Site 20	Densely small building with vegetation between them.	

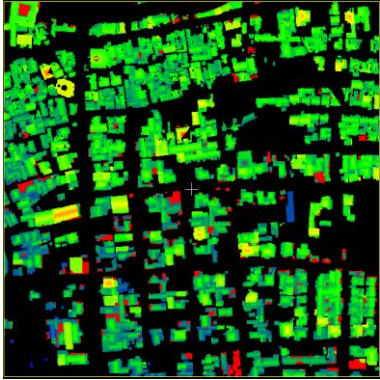
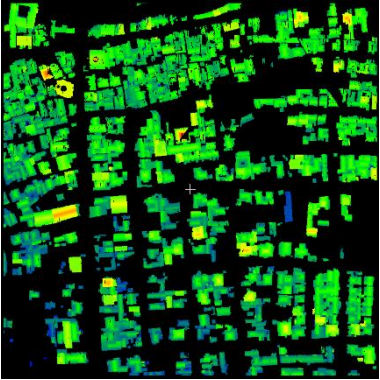
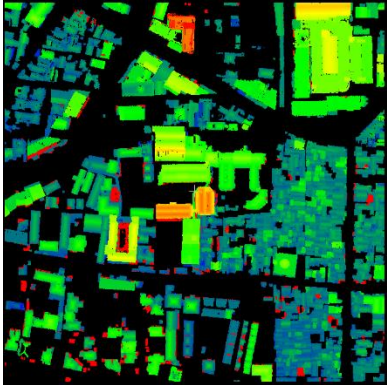
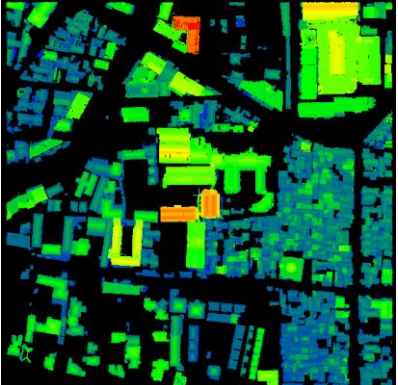
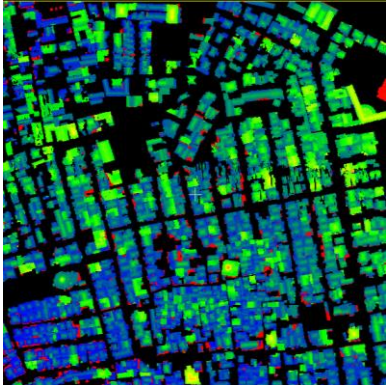

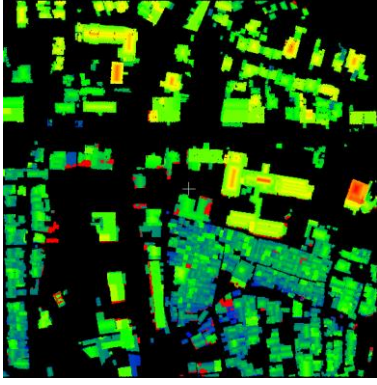
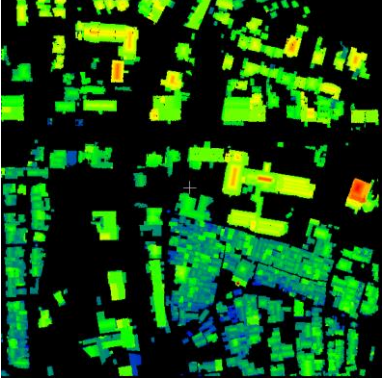
Site	Characteristic	Figure
Site 21	Densely small building with vegetation between them.	
Site 22	Densely small building with vegetation between them in a flat area	
Site 23	Densely small building with vegetation between them in the slope area	
Site 24	Densely small building with vegetation between them in the slope area	

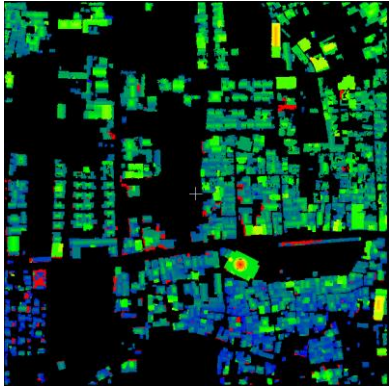
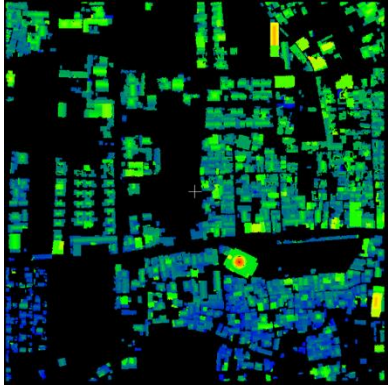
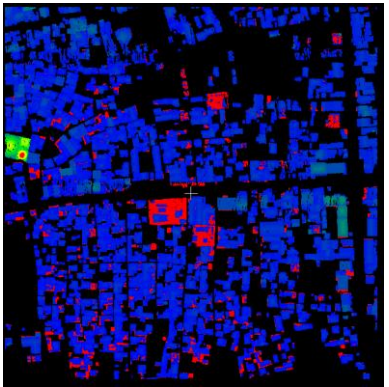
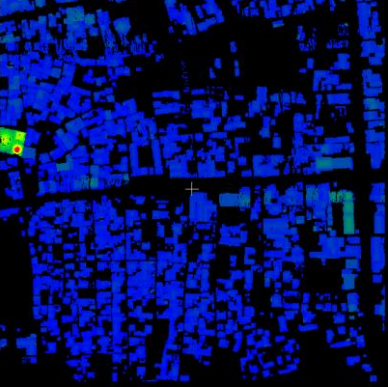
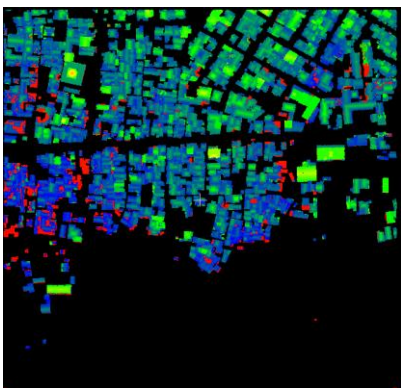
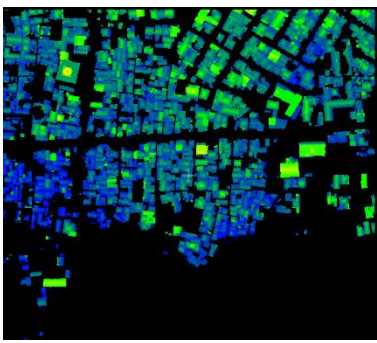
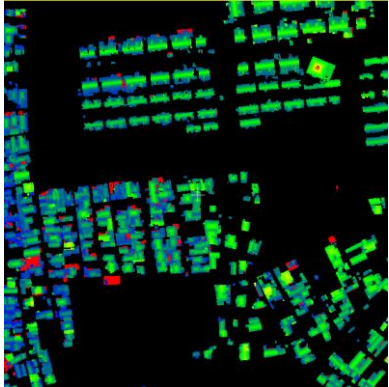
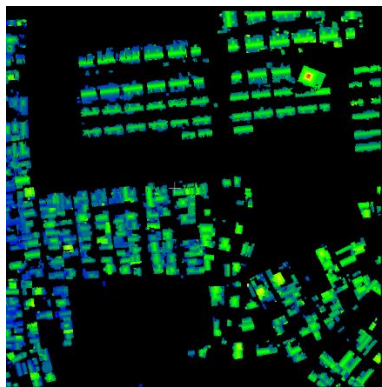
Site	Characteristic	Figure
Site 25	Densely small building with vegetation between them in the slope area	
Site 26	Densely small building with vegetation between them.	
Site 27	Densely small building with vegetation between them.	
Site 28	Rural area, small building	

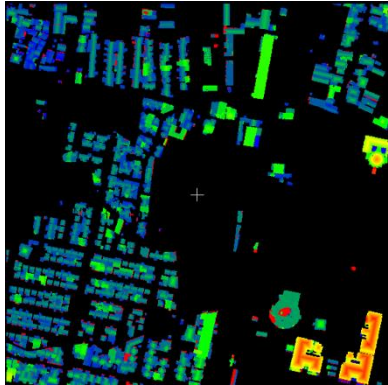
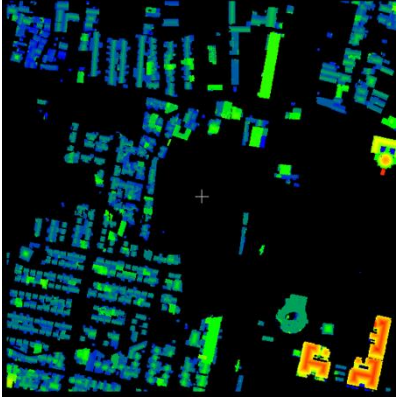
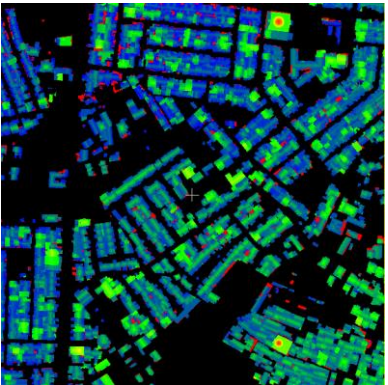
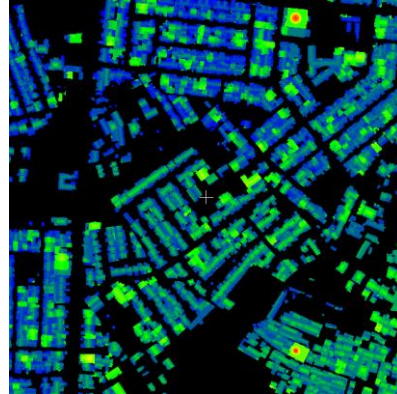
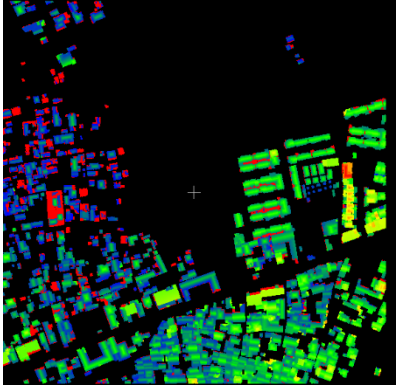
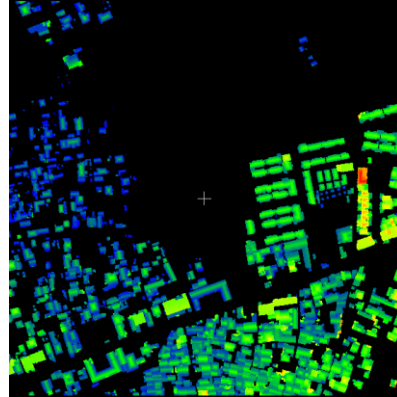
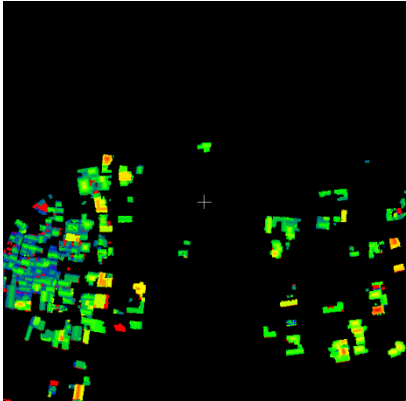
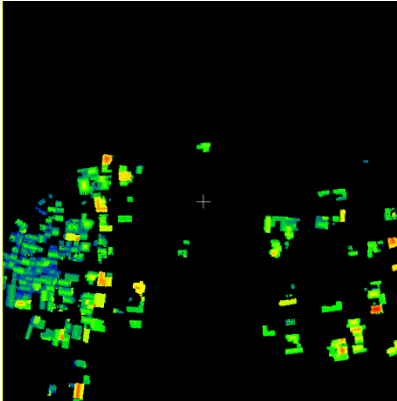
Site	Characteristic	Figure
Site 29	Densely small building with vegetation between them.	
Site 30	Densely small building with vegetation between them.	
Site 31	Densely small building with vegetation between them.	
Site 32	Densely small building with vegetation between them.	

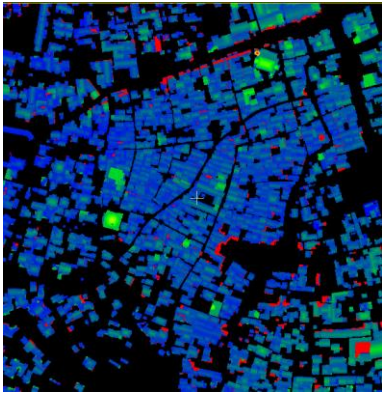
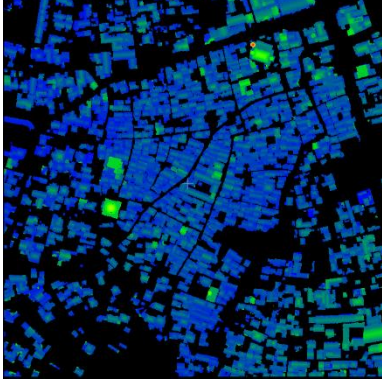
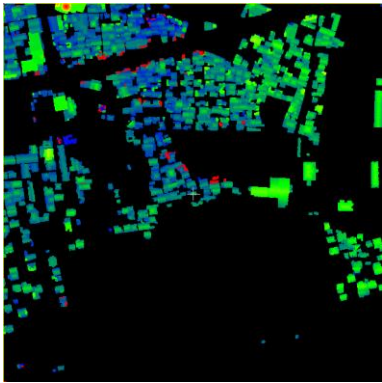
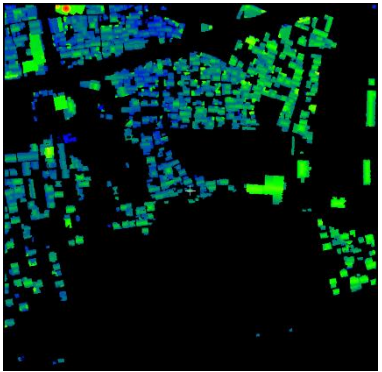
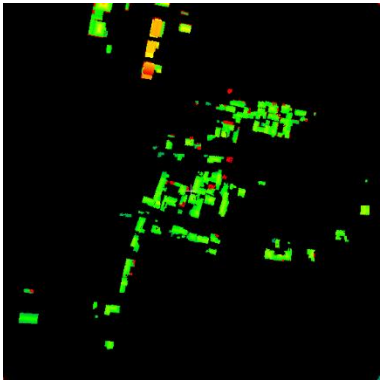
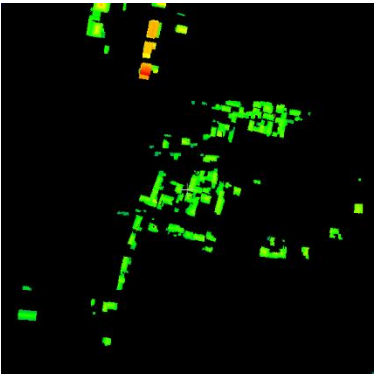
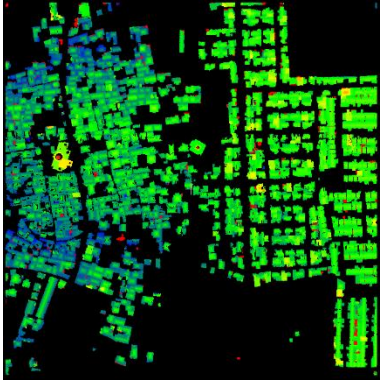
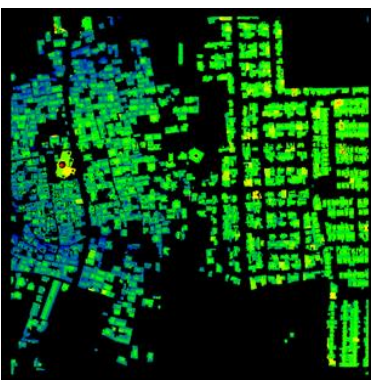
APPENDIX II: THE CLEANED DATA

Sites	Clipped Result with Errors	Cleaned Data
01	 Aerial view of a city block with a grid of buildings. The image is overlaid with a color-coded data map. The colors range from dark blue to bright yellow. There are several small red and orange spots scattered throughout the grid, indicating errors or specific data points.	 The same aerial view of the city block as in the 'Clipped Result with Errors' column. The color-coded data map is now cleaner, with the red and orange spots removed, leaving only the blue, green, and yellow colors.
02	 Aerial view of a city block with a grid of buildings. The image is overlaid with a color-coded data map. The colors range from dark blue to bright yellow. There are several small red and orange spots scattered throughout the grid, indicating errors or specific data points.	 The same aerial view of the city block as in the 'Clipped Result with Errors' column. The color-coded data map is now cleaner, with the red and orange spots removed, leaving only the blue, green, and yellow colors.
03	 Aerial view of a city block with a grid of buildings. The image is overlaid with a color-coded data map. The colors range from dark blue to bright yellow. There are several small red and orange spots scattered throughout the grid, indicating errors or specific data points.	 The same aerial view of the city block as in the 'Clipped Result with Errors' column. The color-coded data map is now cleaner, with the red and orange spots removed, leaving only the blue, green, and yellow colors.
04	 Aerial view of a city block with a grid of buildings. The image is overlaid with a color-coded data map. The colors range from dark blue to bright yellow. There are several small red and orange spots scattered throughout the grid, indicating errors or specific data points.	 The same aerial view of the city block as in the 'Clipped Result with Errors' column. The color-coded data map is now cleaner, with the red and orange spots removed, leaving only the blue, green, and yellow colors.

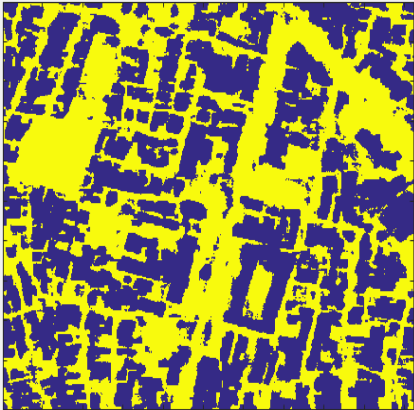
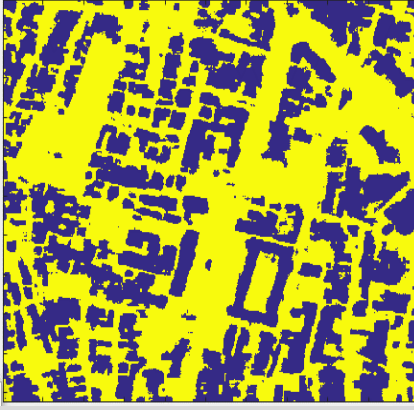
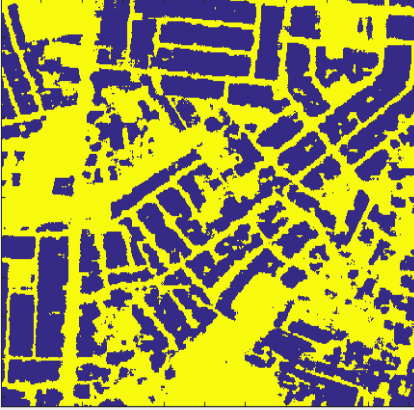
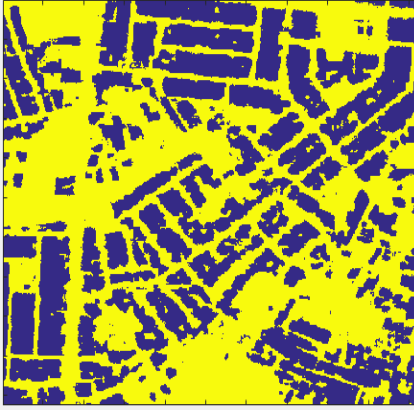
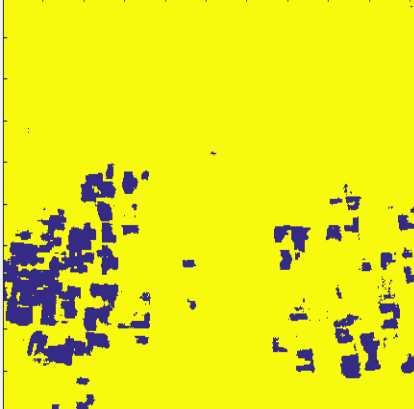
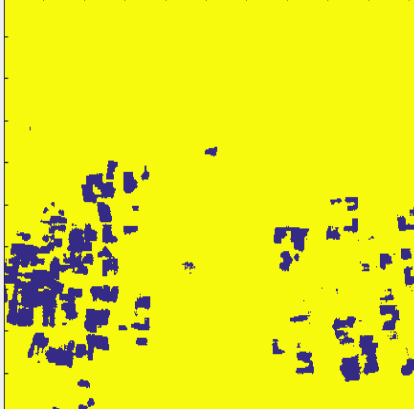
Sites	Clipped Result with Errors	Cleaned Data
05		
06		
07		
08		

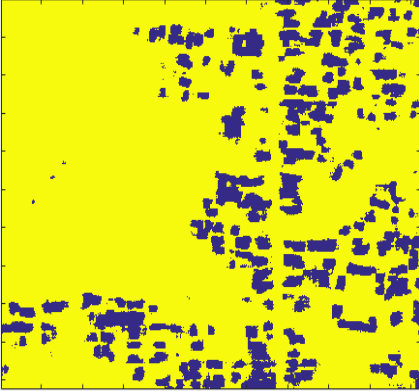
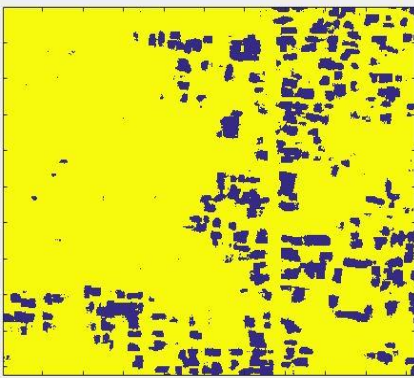
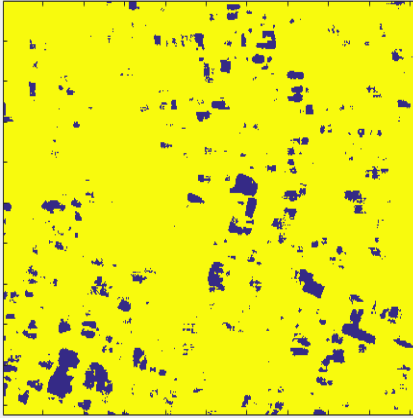
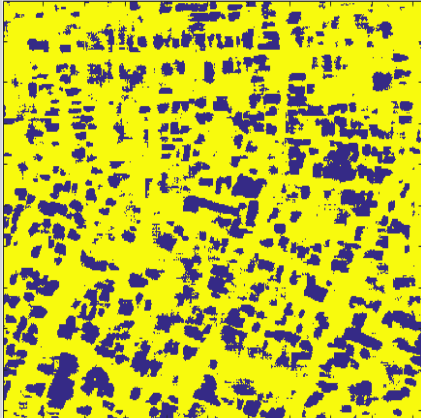
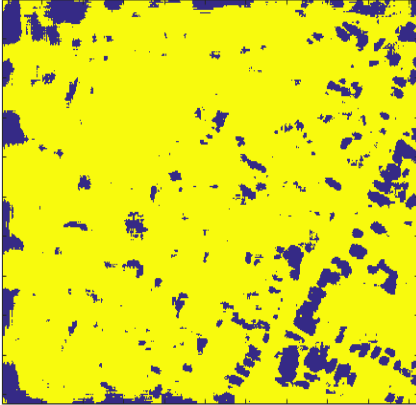
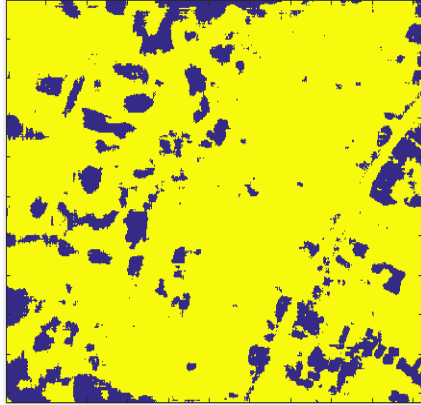
Sites	Clipped Result with Errors	Cleaned Data
09		
10		
11		
12		

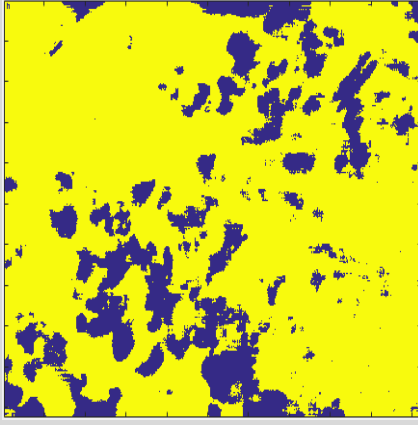
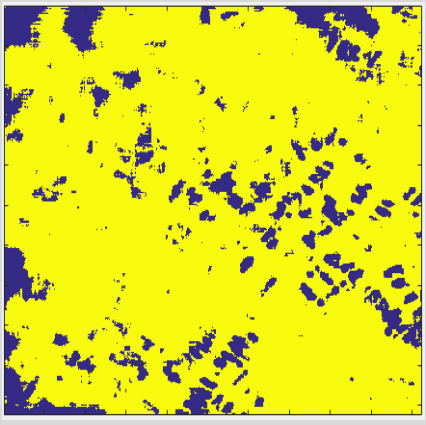
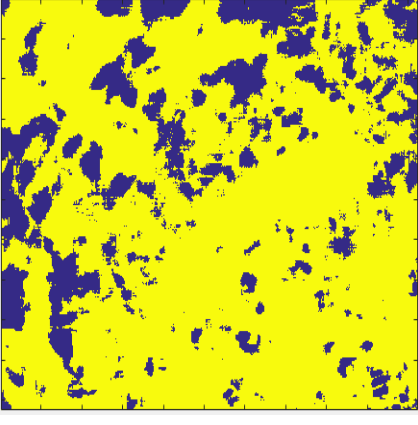
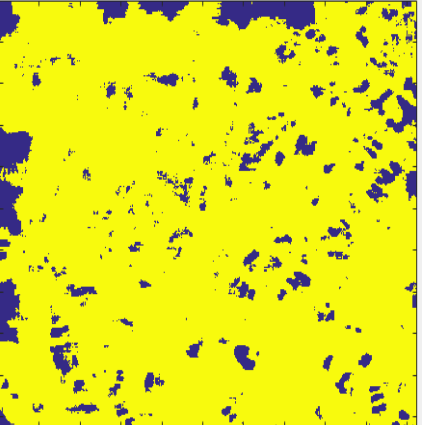
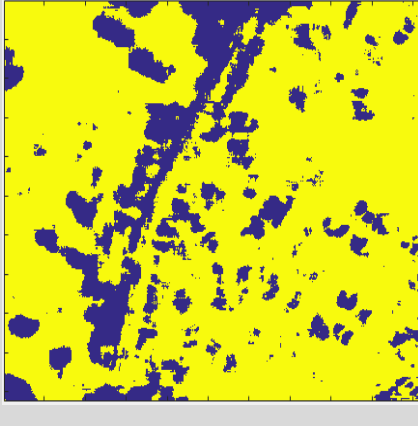
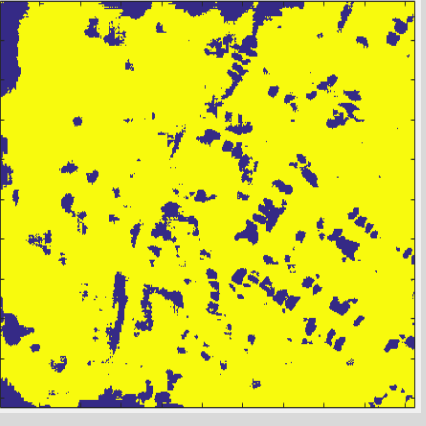
Sites	Clipped Result with Errors	Cleaned Data
13	 <p>This image shows a satellite view of a city block with a grid of buildings. The buildings are overlaid with a color-coded data layer. The colors range from blue to red, with a high concentration of red and orange pixels, indicating significant errors or noise in the data. A white crosshair is centered in the image.</p>	 <p>This image shows the same city block as the previous image, but with the data layer cleaned. The colors are now predominantly blue and green, with very little red or orange, indicating that the errors have been removed. A white crosshair is centered in the image.</p>
14	 <p>This image shows a satellite view of a city block with a grid of buildings. The buildings are overlaid with a color-coded data layer. The colors range from blue to red, with a high concentration of red and orange pixels, indicating significant errors or noise in the data. A white crosshair is centered in the image.</p>	 <p>This image shows the same city block as the previous image, but with the data layer cleaned. The colors are now predominantly blue and green, with very little red or orange, indicating that the errors have been removed. A white crosshair is centered in the image.</p>
15	 <p>This image shows a satellite view of a city block with a grid of buildings. The buildings are overlaid with a color-coded data layer. The colors range from blue to red, with a high concentration of red and orange pixels, indicating significant errors or noise in the data. A white crosshair is centered in the image.</p>	 <p>This image shows the same city block as the previous image, but with the data layer cleaned. The colors are now predominantly blue and green, with very little red or orange, indicating that the errors have been removed. A white crosshair is centered in the image.</p>
16	 <p>This image shows a satellite view of a city block with a grid of buildings. The buildings are overlaid with a color-coded data layer. The colors range from blue to red, with a high concentration of red and orange pixels, indicating significant errors or noise in the data. A white crosshair is centered in the image.</p>	 <p>This image shows the same city block as the previous image, but with the data layer cleaned. The colors are now predominantly blue and green, with very little red or orange, indicating that the errors have been removed. A white crosshair is centered in the image.</p>

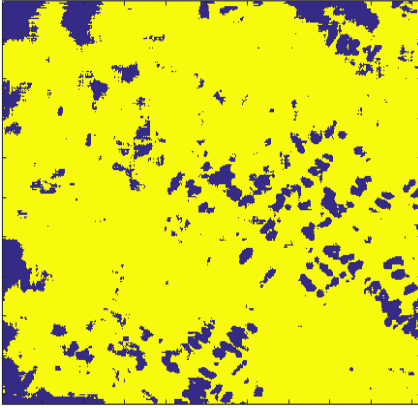
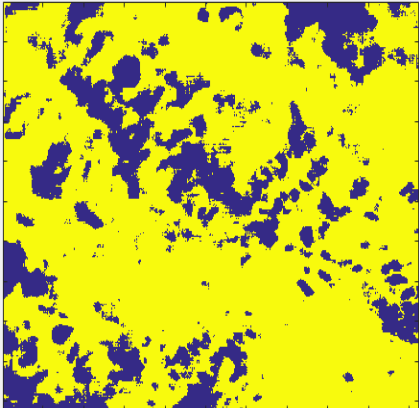
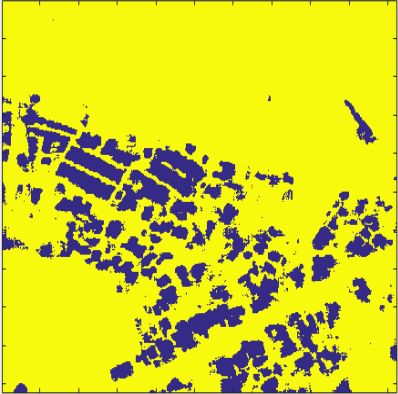
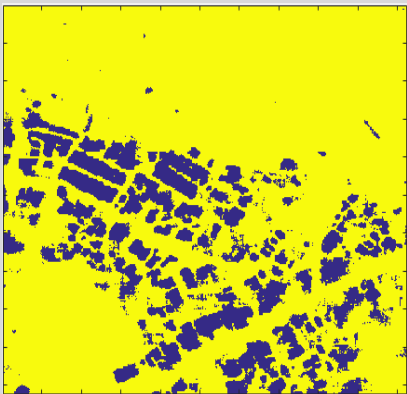
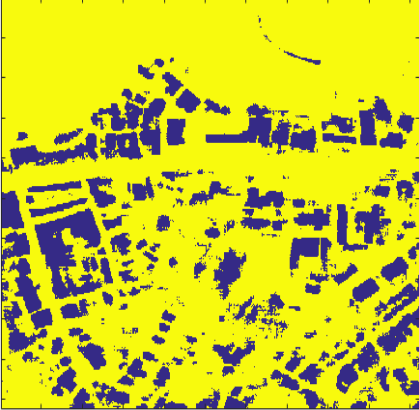
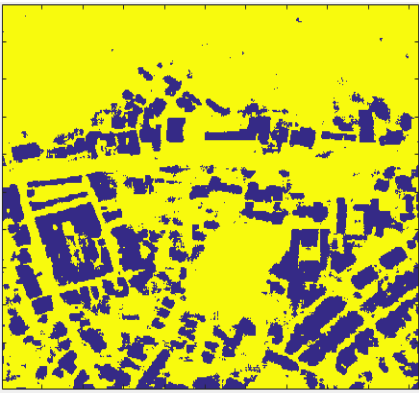
Sites	Clipped Result with Errors	Cleaned Data
17	 <p>This image shows a dense urban grid with a high density of small, irregularly shaped polygons. The colors are primarily blue and green, with some red and yellow spots scattered throughout, indicating errors or noise in the data.</p>	 <p>This image shows the same urban grid as the 'Clipped Result with Errors' but with a significantly reduced number of polygons. The remaining polygons are more uniform in shape and color, representing the cleaned data.</p>
18	 <p>This image shows a dense urban grid with a high density of small, irregularly shaped polygons. The colors are primarily blue and green, with some red and yellow spots scattered throughout, indicating errors or noise in the data.</p>	 <p>This image shows the same urban grid as the 'Clipped Result with Errors' but with a significantly reduced number of polygons. The remaining polygons are more uniform in shape and color, representing the cleaned data.</p>
19	 <p>This image shows a dense urban grid with a high density of small, irregularly shaped polygons. The colors are primarily blue and green, with some red and yellow spots scattered throughout, indicating errors or noise in the data.</p>	 <p>This image shows the same urban grid as the 'Clipped Result with Errors' but with a significantly reduced number of polygons. The remaining polygons are more uniform in shape and color, representing the cleaned data.</p>
20	 <p>This image shows a dense urban grid with a high density of small, irregularly shaped polygons. The colors are primarily blue and green, with some red and yellow spots scattered throughout, indicating errors or noise in the data.</p>	 <p>This image shows the same urban grid as the 'Clipped Result with Errors' but with a significantly reduced number of polygons. The remaining polygons are more uniform in shape and color, representing the cleaned data.</p>

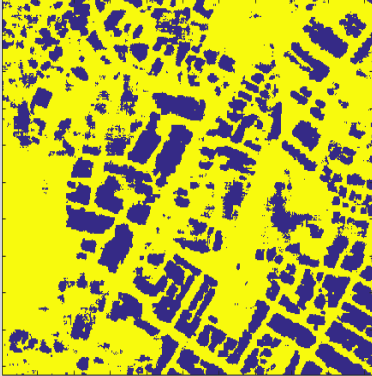

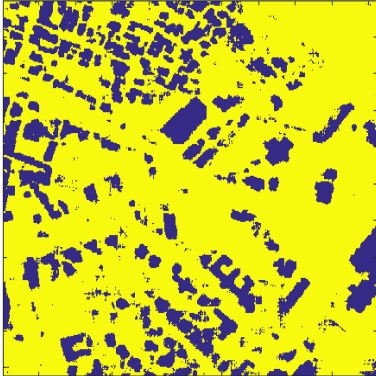
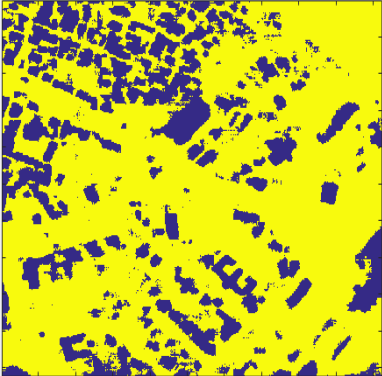
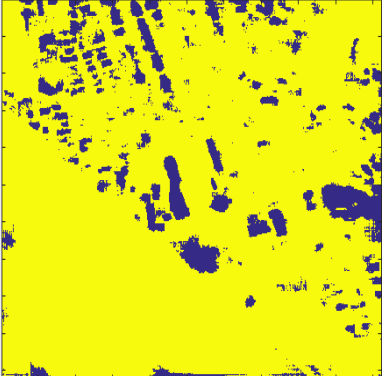
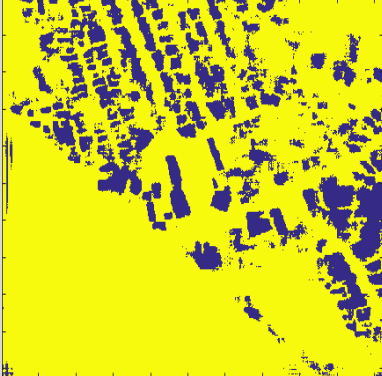
APPENDIX III PREDICTION RESULT

Area	Sites	Noisy Image	Cleaned Image
Lombok	1		
	14		
	20		

Area	Sites	Noisy Image	Cleaned Image
Tanjung Lesung	21		
	22		
Tanggamus	23		

Area	Sites	Noisy Image	Cleaned Image
	24		
	25		
	26		

Area	Sites	Noisy Image	Cleaned Image
	27		
	28		
	29		

Area	Sites	Noisy Image	Cleaned Image
	30		
	31		
	32		

APPENDIX IV THE PRECISION AND RECALL PERFORMANCE

Region	Sites	Precision (PA)			Recall (UA)				
		Noisy	Cleaned	Noisy	Cleaned	Noisy	Cleaned		
Lombok	1	92.361	85.8736	89.5512	90.527	80.523	89.3101	88.0631	86.937
	14	90.2941	87.493	88.7083	90.284	86.8364	90.9777	91.6116	91.795
	20	84.8264	82.0489	81.2902	82.692	84.3707	88.7436	91.6892	89.634
Tanjung Lesung	21	82.4675	76.4129	76.5624	76.5405	81.2061	86.6714	44.5576	90.571
	22	71.0601	21.5243	82.0996	64.9466	75.5414	89.5199	52.557	88.5611
Tanggamus	23	31.5241	33.4314	11.2756	78.6157	45.7192	57.9905	31.4033	46.3303
	24	44.5159	39.4524	26.6852	72.4121	31.4994	43.4528	34.6978	39.1145
	25	53.5438	48.2519	15.9316	77.6681	29.0834	50.0119	34.226	38.4803
	26	56.6623	60.1663	19.1876	77.0027	22.3809	56.2349	30.2499	26.4784
	27	68.7788	60.1663	15.2327	81.1257	38.426	38.4357	45.4058	47.3987
Makassar	28	89.049	91.8504	85.4683	89.9263	55.6561	55.9696	40.1969	66.595
	29	71.2012	90.6755	84.2902	86.8843	62.5345	64.424	47.3682	71.468
	30	85.3378	92.0499	86.2781	92.3528	61.8172	60.0507	57.2601	69.7597
	31	82.5367	81.0241	78.2531	82.4272	59.9865	63.6919	48.0419	69.2902
	32	41.0085	76.1228	82.3029	68.641	44.3788	48.8226	24.4766	63.1288