

# Thesis Title

Thesis Subtitle

**CHAN, Yan-Chak Christopher**

**Supervised by:**

Prof. Dr. Hannes Taubenöck <sup>1</sup>  
Emran Alchikh Alnajar <sup>2</sup>

Masterarbeit submitted for the degree of  
Master der Naturwissenschaften  
in

Applied Earth Observation and Geoanalysis of the Living  
Environment (EAGLE)



Philosophische Fakultät (Historische, Philologische, Kultur- und  
Geographische Wissenschaften)  
Julius-Maximilians-Universität Würzburg

## **Forewords and Acknowledgements**

## **Declaration of Independent Work**

## **Figure list**

## **Abbreviations**

## **Abstract**

  Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur dignissim, quam maximus posuere cursus, magna justo rutrum erat, at mattis magna magna nec risus. Duis lacus lectus, condimentum a viverra eu, fermentum molestie lorem. Sed maximus, enim eu scelerisque dictum, sem erat mollis massa, in dictum ante libero a tortor. Cras nulla nisi, sollicitudin ac suscipit cursus, maximus non dui. Sed venenatis ligula id efficit imperdiet. Vivamus ut magna eleifend, rutrum ante facilisis, pulvinar turpis. Maecenas at interdum lorem. Duis vel varius ligula. Sed magna erat, egestas vitae varius id, cursus vitae neque. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Phasellus interdum lectus mi, a dapibus lorem tristique a. Phasellus molestie vestibulum metus a fringilla. Pellentesque at rhoncus nulla. Praesent posuere turpis nec leo fringilla egestas.

  Pellentesque auctor vel dolor eu viverra. Ut faucibus nunc orci, eu aliquam justo hendrerit vel. Proin auctor sed nisl non posuere. Vivamus orci orci, commodo eget semper nec, tempus at arcu. Nam eget leo cursus velit aliquam varius. Curabitur nisi dui, rutrum vitae elit a, mollis volutpat mauris. Suspendisse potenti. Nam convallis magna iaculis posuere aliquam. Quisque tristique rutrum placerat. Quisque ultricies molestie lacinia. Maecenas at nisi in neque dictum consequat.

# *Contents*

0.1	Forewords and Acknowledgements . . . . .	ii
0.2	Declaration of Independent Work . . . . .	iii
0.3	Figure list . . . . .	iv
0.4	Abbreviations . . . . .	v
<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Study Area of Interest . . . . .	5
1.1.1	Kalobeyei, Kakuma, Turkana, Kenya . . . . .	5
1.1.2	Dzaleka, Dowa, Malawi . . . . .	5
1.1.3	Research Questions . . . . .	6
<b>2</b>	<b>Literature Review</b>	<b>9</b>
2.1	Remote Sensing of Informal Settlements . . . . .	9
2.2	Deep Learning in Urban Remote Sensing . . . . .	9
2.2.1	Computer Vision and Building Segmentation . . . . .	9
2.2.2	Computer Vision and Convolutional Neural Networks	10
<b>3</b>	<b>Data and Methodologies</b>	<b>11</b>
3.1	Data . . . . .	11
3.1.1	Raster pre-processing . . . . .	11
3.1.2	Data Augmentation . . . . .	11
3.2	Architecture and hyperparameter selection . . . . .	12
3.2.1	The U-Net and U-Net variants . . . . .	12
3.2.2	Pre-trained weights on Deep Learning models . . . . .	12
3.3	Accuracy Assessment . . . . .	13
3.3.1	Precision, Recall, Sensitivity, and Specificity . . . . .	14
3.3.2	Overall Accuracy, Dice Score, and Intersection-over-Union . . . . .	15
3.4	Experimentation setup . . . . .	15
3.4.1	Project workflow . . . . .	15
<b>4</b>	<b>Findings</b>	<b>16</b>
4.1	Analysis . . . . .	16

<b>5</b>	<b>Discussion</b>	<b>17</b>
<b>6</b>	<b>Conclusion</b>	<b>18</b>
<b>7</b>	<b>Bibliography</b>	<b>19</b>
<b>A</b>	<b>Appendix</b>	<b>20</b>

# *Introduction*

The world's population is more urbanised than ever before. As of 2018, approximately 4 billion (55%) (UN DESA., 2018, Taubenöck et al., 2009) reside in urban areas, of which 60% reside in slums often located at the fringes of the city (Venables A., 2018). Urbanisation growth is expect to increase by 2.5 billions between 2018 to 2050, most of which will be in Asia and Africa (UN DESA., 2018). When population growth outpace development, slums became the supplier of significant housing stocks. These informal settlements are dynamic and represent a good reflection of cultural practices, access to resources, financial limitations and other socio-economic conditions. This means the informal settlement differs significantly between urban and rural settlements of roof covers, densities, and are subjected to different levels of access to resources and the types of resources.

Refugee camps are often the common or only way for displaced people to receive shelters and assistance. They are often setup in place of proximity to displaced population, whether that be from natural disasters, human caused disasters, or other reasons. Throughout history, refugee sites have provided haven to the world's most vulnerable population (UN, 2018, Turner S., 2016, UNHCR, 2021). However as of 2020, out of the 26.4 million refugees, only around 1.4 million have access to third country solution between 2016 to 2021 (UNHCR, 2021). Additionally, although officially defined as temporary settlement, many refugee camps have had longer than expected life cycle, some of them have even became "Secondary Cities" and therefore suffers similar problems of poor governance and rapid urbanisation which consequentially makes them unattractive as investment (Cities Alliance & AfDB., 2022). For the many refugee camps and informal settlements that have lasted well beyond their expected temporary role, there are in generally 3 ways of solving the issue: 1. Voluntary repatriation, 2. Reolocation to third country, 3. Local integration as outlined by the Global Compact on Refugees (UN, 2018), although actual implementations are often subjected to the wills of the host sovereign-state. Recent studies have suggested that local integration often have a net positive economical impact on the surrounding region (Alix-Garcia et al., 2018, Rummery A., 2019, IFC., 2018).

The United Nations Department of Economic and Social Affairs have set out a set of 17 Sustainable Development Goals (herein SDG) to be achieved as of 2030 (UN, 2015). Special attention are drawn to Goals 1 and 10 that are particularly relevant, of which this study hope to contribute to.

- *Goal 1: End poverty in all its forms everywhere*
  - *Target 1.1: By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than \$1.25 a day*
  - *Target 1.4: By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including microfinance*
  - *Target 1.b: Create sound policy frameworks at the national, regional and international levels, based on pro-poor and gender-sensitive development strategies, to support accelerated investment in poverty eradication actions*
- *Goal 10: Reduce inequality within and among countries*
  - *Target 10.1: By 2030, empower and promote the social, economic and political inclusion of all, irrespective of age, sex, disability, race, ethnicity, origin, religion or economic or other status*
  - *Target 10.7: Facilitate orderly, safe, regular and responsible migration and mobility of people, including through the implementation of planned and well-managed migration policies*

Thus, the overarching theme of this thesis is about reducing the geospatial data inequality of informal settlements (Herfort et al., 2021), thereby to improve future decision making in both humanitarian and non-humanitarian context. As the topic and data provider of this project, the Humanitarian OpenStreetMap (herein HOTOSM) have been at the forefront of using open and crowd sourced mapping data to support humanitarian causes from shorter term disaster response to longer epidemiology and microfinance campaigns (HOTOSM, 2021). Having up-to-date map is therefore paramount for short and long term humanitarian projects, and with the advent of Deep Learning in the last decade, AI-assisted mapping have became a major topic for innovation (e.g. Herfort et al., 2019, Kuffer et al., 2016, Wurm et al., 2021, Quinn et al., 2018).

## **Study Area of Interest**

### **Kalobeyei, Kakuma, Turkana, Kenya**

The Kakuma camp was first established in 1992, located in the North-West of Kenya in Turkana County. The camp was initially established to provide accommodation to the refugees fleeing the Second Sudanese Civil War as a temporary solution. However, as the conflict became drag out and followed by subsequent conflicts in the nearby region, the Kakuma camp have therefore been running for the past 30 years. As of 2020, Kakuma is home to 157,718 refugees with increasing number coming from the more recent Somali and Ethiopian-Eritrean conflict (IFC., 2018, UN-HABITAT, 2021). The Kalobeyei Integrated Settlement established in 2015 benefited from a much better spatial planning in order to facilitate inclusive socio-economic development (UN-HABITAT, 2021, UNHCR & DANIDA, 2019) (*see figure 1.1*).

The Kakuma refugee camp have fluctuated in population as a response to demand, however, a dramatic increase in population in 2013 to 2014 has culminated into the development of Kakuma 4 Camp and the Kalobeyei Settlement and the Kalobeyei Integrated Socio-Economic Development Plan (KISEDP) which is the local integration strategies. Both the Kakuma and Kalobeyei refugee camps have local integration as the targeted solution (UN-HABITAT, 2021, UNHCR & DANIDA, 2019). A comprehensive study of the formal and informal economy of Kakuma refugee camp conducted by the International Financial Corporation (IFC, 2018) suggests that that market catering for the refugees and surrounding towns is estimated at KES 1.7 billion (USD \$16.4 million). The economical vibrancy of local integration have improve significantly the impoverished Turkana county. However, challenges remain in integration into the wider Kenyan economy.

### **Dzaleka, Dowa, Malawi**

Originally an infamous prison camp under the Banda's Malawi Congree Party regime, the area was converted to become the Dzaleka Refugee Camp in 1994. Unlike the Kakuman and Kalobeyei camp, the Dzaleka Refugee Camp is located in the heart of Malawi, 45 km away from the capital Li-longwe, where it is home to around 52,000 refugees and receive on average 300 new residents every month. Most coming from the Great Lakes area, in particular, the Democratic Republic of Congo and Burundi. However, resurgence of past conflicts between the Republic of Congo and D.R. Congo have caused an increased of influx in recent years (UNHCR, 2014, Kaval E., 2016). Much of the infrastructure in the Dzaleka camp remain in rudimentary at best, and very little resources and statistics were available via the UNHCR and UNDP portals. The Northern extension to the Dzaleka main camp is known as the Katubza extension (*referred to as Dzaleka North by the rest of this thesis*), it is a well-planned plot of land consisting of 423 shelter shelters and were still inconstruction as of March 2021 (Gross G., 2021 & UNHCR, 2021) (*see figure 1.2*).

### **Research Questions**

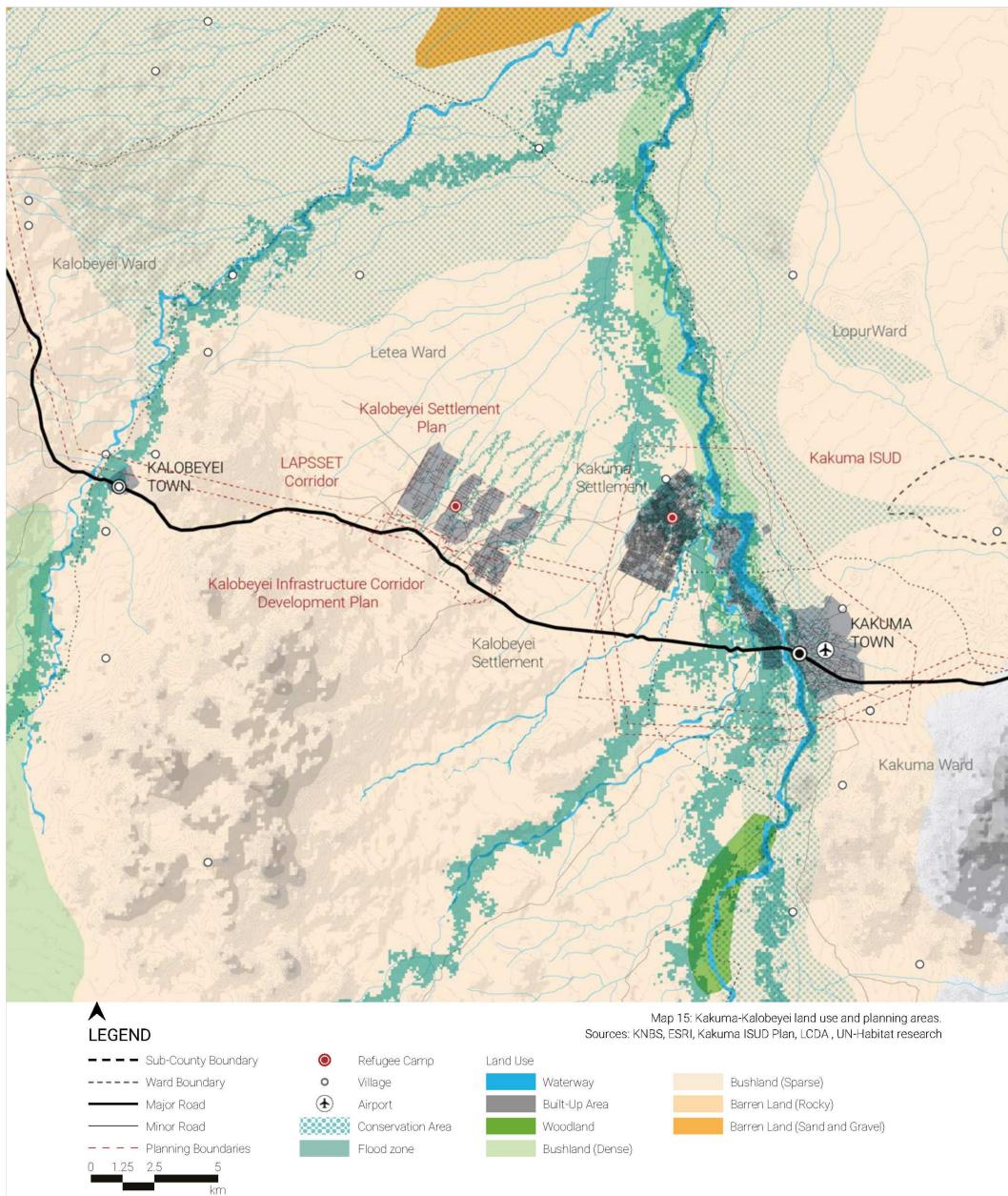


Figure 1.1: The Kakuma-Kalobeyei land use and planning areas (UN-HABITAT, 2018)

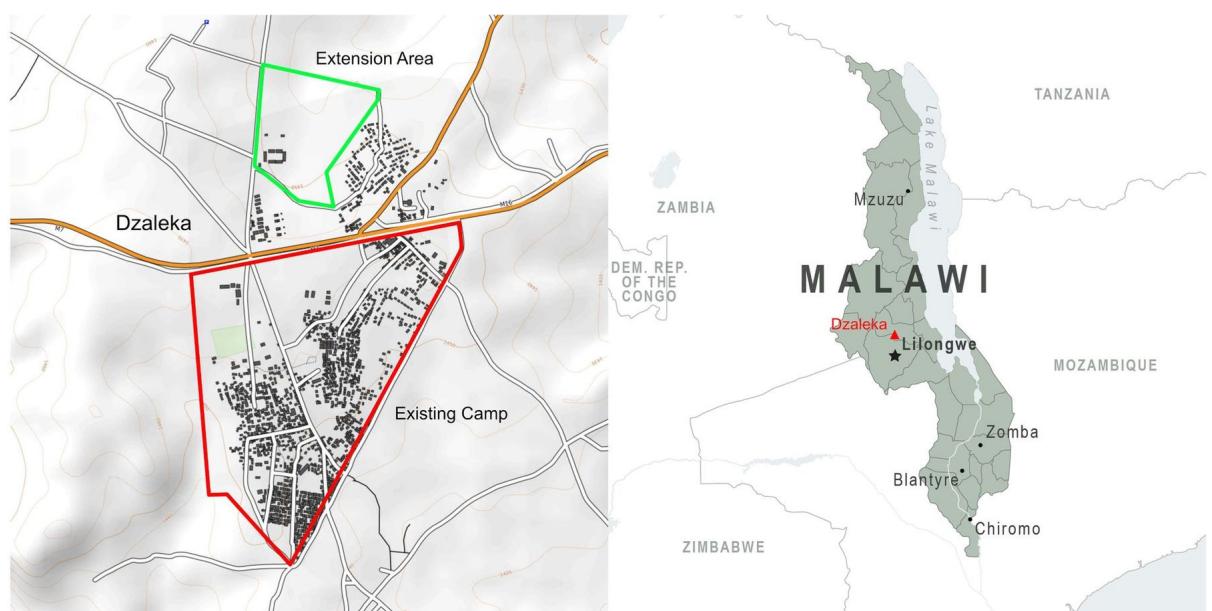


Figure 1.2: The main Dzaleka Refugee Camp and the Katubza extension plan (Dzaleka North) designed by Urban Design Advisor to the UNHCR Werner Schnellenberg (Gerhard G., 2021).

# *Literature Review*

## **Remote Sensing of Informal Settlements**

Informal settlement and slums mapping of developing countries require very high resolution (VHR) images which was unavailable until the turn of the century. The relatively new technology thus only began to gain traction within the last 2 decades. Particularly with the increase in the availability of VHR satellites. Increase in computational power had enabled novel techniques such as multi-layer machine learning, textural analysis, and novel geostatistical methods to emerge (Kuffer et al., 2016). Due to frequent repeat coverage of satellite's orbit, they can be used to fill in between periodic census that are costly and time consuming to conduct. Census also does not do a good job in capturing larger scale units and spatial patterns, potentially overlooking others socioeconomic determinants such as cropping cycles and infrastructure access etc. Availability of VHR sensors and publications on slums and remote sensing (Kuffer et al., 2016). Remote sensing of settlements largely falls under 2 categories, rural or urban. Due to the different make-up of socioeconomic context and urban morphology, sensing rural and urban settlements require different parameters. Additionally, there's no "one size fit all" way to generalise informal and formal settlement across the world, as physical geography, topography, cultural, and available resources often dictate the distribution, development, and settlement clusters pattern.

## **Deep Learning in Urban Remote Sensing**

### **Computer Vision and Building Segmentation**

Prior to the paradigm shift towards data-driven and ML based segmentation techniques, image segmentation were performed manually with the aid of a few algorithms (Pal & Pal, 1993). The image segmentation tasks often starts with acquiring less-noisy imagery or data, this is followed by applying multitudes of CV based edge detector (e.g. Sobel, Prewitt, Marr-Hildreth) or Grey-Level Co-occurrence Matrix kernel (e.g. Haralich Texture). This can be considered to be the pre-processing steps necessary to extract information

to select the parameters for the segmentation algorithms. (Pal & Pal, 1993, Blaschke T., 2010, Blaschke et al., 2014).

### **Computer Vision and Convolutional Neural Networks**

The issue with any Deep Learning Project is the high amount of data required (Tan et al., 2018, )

# *Data and Methodologies*

## **Data**

**Raster pre-processing**

**Normalisation**

**Cropping**

**Data Augmentation**

Data augmentation is perhaps one of the most crucial task in training a robust neural-network. It is an economical way of increasing generalisability without increasing model complexity, data augmentation achieve this through, firstly increasing the quantity of training and validation data, secondly encompassing a greater range of textural, geometrical, and colour variability through the creation of augmented pseudo-data (Shorten & Khoshgoftaar, 2019; Kinsley & Kukiela, 2020; Howard & Gugger, 2020; Zoph et al., 2019).

Data augmentation can generally be split into 3 categories: 1. Geometric/Affine distortion, 2. Colour distortion, and 4. Noise distortion. The application of which types of distortion to the *Train* and *Validation* dataset is highly dependent on the context of the semantic task. Therefore, care must be taken as to not introduce mislabelling (*see Figure 3.1*) (Ng A., 2018).

**Augmentation categories:**

- Geometric/Affine distortion
  - e.g. Fliping, Stretching, Rotation...
- Colour distortion
  - e.g. Colour Inversion, Solarise Colour, Greyscale...
- Noise distortion

- e.g. Blurring, Contrasting, Salt & Pepper...

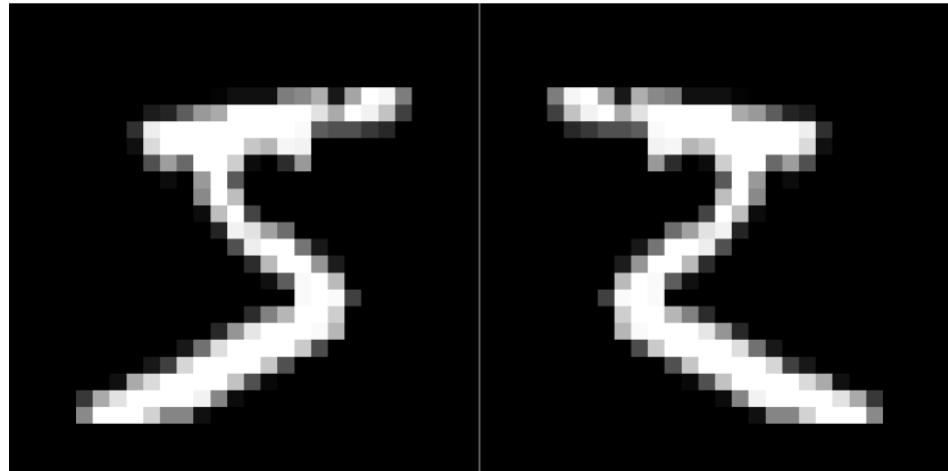


Figure 3.1: Perhaps geometric augmentation of horizontal flipping shall not be applied on the MNIST number of 5

## Architecture and hyperparameter selection

Model architecture and their associated hyperparameters selection is highly dependent on the computational resources and the task at hand (Ng A., 2018, Howard & Gugger, 2020). As this study aims to output a pixel-based binary segmentation which delineates building and non-building, and given the computational resources constraint, tried and tested architectures with relatively low number of training parameters is ideal.

### The U-Net and U-Net variants

The U-Net architecture was first developed by Ronneberger et al. (2015) for the task of cell segmentation in biomedical electronmicroscope images. The architecture feature a symmetrical Encoder-Decoder structure (*see figure ??*) and as with many other CNN, the architecture have transferred successfully well into the remote sensing domains (Höser & Künzer, 2020, Höser et al., 2020) (e.g. Jean et al., 2016, Xu et al., 2019)

### Pre-trained weights on Deep Learning models

Another major consideration of this study is to compare the performance of the architectures on various pre-trained weights. Due to the representation learning feature of CNN, studies have shown transfer learning on pre-trained

weights even across different domain dataset could result in better performance, especially on projects with less data availability (LeCun et al., 2015, Zhu et al., 2017, Tan et al., 2018). Pre-trained weights for this study are ImageNet weights and the drone-based building segmentation competition (OCC) weights. Thus, one of the objective of this study is to find out whether transfer training performance on various pretrained weights would outperform training from scratch.

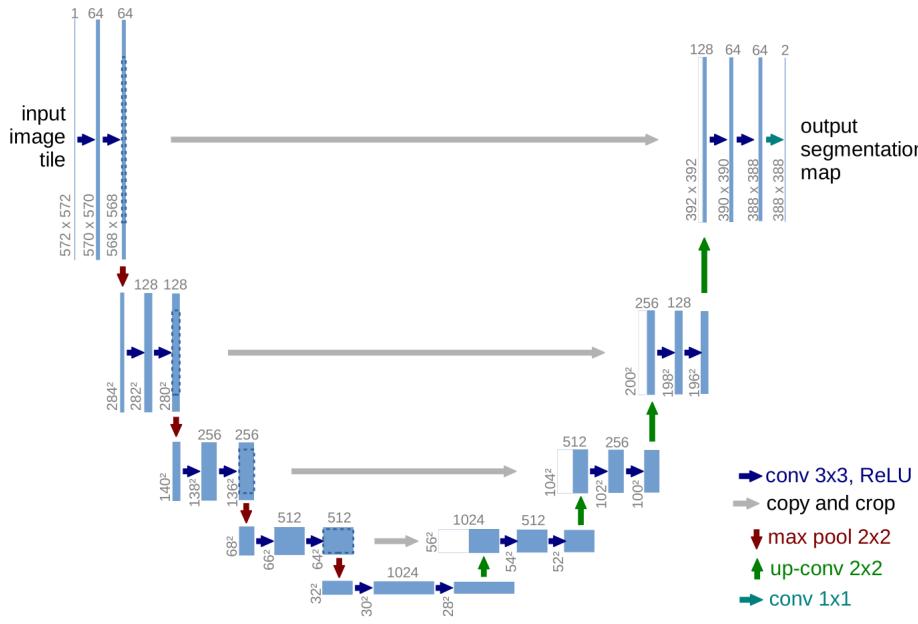


Figure 3.2: The Encoder-Decoder U-Net architecture (Ronneberger et al., 2015)

#### EfficientNet encoders

### Accuracy Assessment

Detail and scrutible accuracy assessments are fundamental towards any classification based analysis. This section will introduce and break down the various lower order and higher order class-based (thematic) accuracy assessment. By explaining the characteristics of each metrics, this will provide a much more granular nature of accuracy assessment in the findings of section 4. In general, accuracy assessment in remote sensing can be divided into 2 categories: 1. Positional Accuracy & 2. Thematic Accuracy. Of which, Positional Accuracy deals with the accuracy of the location while Thematic Accuracy deals with the labels or attributes accuracy (Congalton & Green, 2019 & Bolstad, 2019). The rest of this section will consider the lower or-

der and higher order accuracy metrics, with lower order metrics being more granular while higher order metrics more triturated but generalised.

The metrics described in this section form part of the larger family of accuracy assessment metrics that can be constructed from the confusion matrix (*see Figure 3.3*)

		True condition			
		Total population		Prevalence	Accuracy (ACC) =
Predicted condition	Total population	Condition positive	Condition negative	$= \frac{\sum \text{Condition positive}}{\sum \text{Total population}}$	$\frac{\sum \text{True positive} + \sum \text{True negative}}{\sum \text{Total population}}$
	Predicted condition positive	True positive	False positive, Type I error	Positive predictive value (PPV), Precision = $\frac{\sum \text{True positive}}{\sum \text{Predicted condition positive}}$	False discovery rate (FDR) = $\frac{\sum \text{False positive}}{\sum \text{Predicted condition positive}}$
Predicted condition negative	Predicted condition negative	False negative, Type II error	True negative	False omission rate (FOR) = $\frac{\sum \text{False negative}}{\sum \text{Predicted condition negative}}$	Negative predictive value (NPV) = $\frac{\sum \text{True negative}}{\sum \text{Predicted condition negative}}$
	True positive rate (TPR), Recall, Sensitivity, probability of detection, Power = $\frac{\sum \text{True positive}}{\sum \text{Condition positive}}$	False positive rate (FPR), Fall-out, probability of false alarm = $\frac{\sum \text{False positive}}{\sum \text{Condition negative}}$	Positive likelihood ratio (LR+) = $\frac{\text{TPR}}{\text{FPR}}$	Diagnostic odds ratio (DOR) = $\frac{\text{LR+}}{\text{LR-}}$	$F_1$ score = $\frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$
		False negative rate (FNR), Miss rate = $\frac{\sum \text{False negative}}{\sum \text{Condition positive}}$	Specificity (SPC), Selectivity, True negative rate (TNR) = $\frac{\sum \text{True negative}}{\sum \text{Condition negative}}$	Negative likelihood ratio (LR-) = $\frac{\text{FNR}}{\text{TNR}}$	

Figure 3.3: The Confusion Matrix

## Precision, Recall, Sensitivity, and Specificity

### Precision, Recall, and Specificity

**Precision** and **Recall**, aka. Positivie-Predictive-Value and Sensitivity/True-Positive-Rate Respectively. The two metrics are often used together, another common denomination especially in remote sensing literature are User's Accuracy and Producer's Accuracy (Congalton & Green, 2019 & Wegmann et al., 2016). To avoid further confusion in nomenclature, **Precision** and **Recall** will be used from hereon.

**Precision** is the measure of correctly predicted Positive class (True Positive) against all positive prediction assigned to that class (True Positive + False Positive) i.e. Given the predicted results, of those that are predicted as positive, what proportion were True. It can be expressed mathematically as:

$$\text{Precision} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})} \quad (3.1)$$

Meanwhile, **Recall** measures the correctly predicted Positive class (True Positive) against both the correct and incorrect predicton on the Positive reference class (True Positive + False Negative) i.e. Given the predicted

results, of those that are referenced as positive, what proportion of those were True. It can be expressed mathematically as:

$$Recall = \frac{True\ Positive}{(True\ Positive + False\ Negative)} \quad (3.2)$$

**Specificity**, aka. True-Negative-Rate measures correctly predicted Negative class (True Negative) against the correct and incorrect prediction on the Negative reference class (False Positive + True Negative) i.e. Given the predicted results, of those that are referenced as negative, what proportion of those were True. It can be expressed mathematically as:

$$Specificity = \frac{True\ Negative}{(False\ Positive + True\ Negative)} \quad (3.3)$$

Therefore, higher **Recall** suggests the model is better at identifying positives and vice-versa higher **Specificity** suggests the model is better at identifying negatives. Since this is an exercise that aim to maximise the positive prediction as a binary building segmentation classifier, emphasise will be placed on maximising **Precision** and **Recall**.

## Overall Accuracy, Dice Score, and Intersection-over-Union

### Experimentation setup

Each network architecture and their associated weights will be trained on 2 data setup. The first setup consist of only the Kalobeyei, Kakuma camp where the labels include drone motion artefacts and rooftops are relatively homogeneous. The second setup consist of data from the Kalobeyei camp and also the rest of Dzaleka, Dowa camp. The second setup will introduce imperfection in labelling and complex heterogeneous rooftops and morphologies. The two data setup will allow comparison between the models response of each class-based accuracy assessment metrics.

To truly assess the performance of the architecture output, a combination of the validation loss, class-based accuracy assessments, and

### Project workflow

## *Findings*

### Analysis

## *Discussion*

## *Conclusion*

## *Bibliography*

## *Appendix*