Thesis Title

Thesis Subtitle

CHAN, Yan-Chak Christopher

Supervised by:
Prof. Dr. Hannes Taubenöck [1]
Emran Alchikh Alnajar [2]

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 efficitur 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

1 Introduction **3**
1.1 Study Area of Interest 3
1.1.1 Kalobeyei, Kakuma, Turkana, Kenya 3
1.1.2 Dzaleka, Dowa, Malawi 3
1.1.3 Research Questions 3

2 Literature Review **4**
2.1 Remote Sensing of Informal Settlements 4
2.2 Deep Learning in Remote Sensing 4
2.2.1 Computer Vision and Convolutional Neural Networks 4
2.2.2 Computer Vision in Building Segmentation 4

3 Data and Methodologies **5**
3.1 Data 5
3.1.1 Raster pre-processing 5
3.1.2 Data Augmentation 5
3.2 Architecture selection 6
3.3 Accuracy Assessment 6
3.3.1 Precision, Recall, Sensitivity, and Specificity 7
3.3.2 Overall Accuracy, Dice Score, and Intersection-over-Union 8
3.4 Experimentation setup 8
3.4.1 Project workflow 8

4 Findings **9**
4.1 Analysis 9

5 Discussion **10**

6 Conclusion **11**

7 Bibliography **12**

A Appendix **13**

Introduction

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, as defined as temporary settlement, many refugee camps

Study Area of Interest

Kalobeyei, Kakuma, Turkana, Kenya

Dzaleka, Dowa, Malawi

Research Questions

Literature Review

Remote Sensing of Informal Settlements

Deep Learning in Remote Sensing

Computer Vision and Convolutional Neural Networks

Computer Vision in Building Segmentation

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 throught 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...

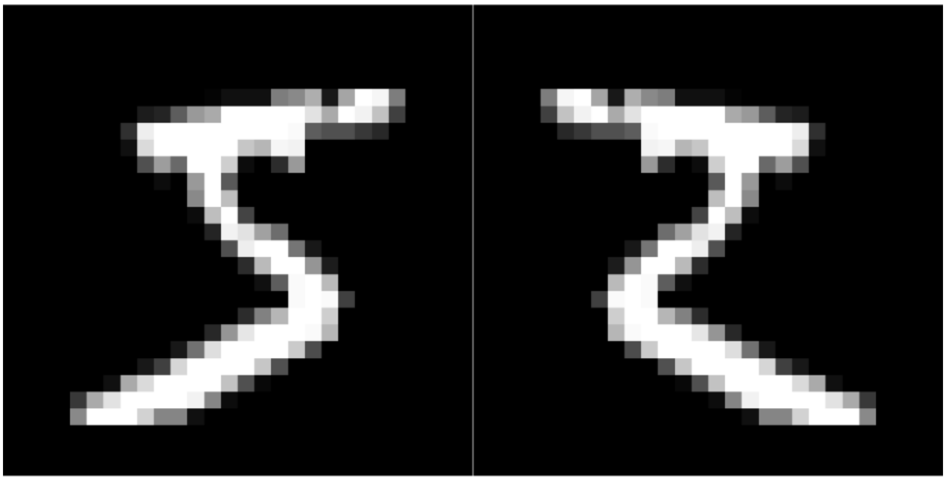

Figure 3.1: Perhaps geometric augmentation of horizontal flipping shall not be applied on the MNIST number of 5

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)

The U-Net and U-Net variants

The U-Net architecture was

Accuracy Assessment

Detail and scrutable 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 order 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 3.2)*

<table>
<tr><td colspan="2"></td><td colspan="2">True condition</td><td colspan="3"></td></tr>
<tr><td></td><td>Total population</td><td>Condition positive</td><td>Condition negative</td><td>Prevalence $= \frac{\Sigma \text{ Condition positive}}{\Sigma \text{ Total population}}$</td><td colspan="2">Accuracy (ACC) = $\frac{\Sigma \text{ True positive} + \Sigma \text{ True negative}}{\Sigma \text{ Total population}}$</td></tr>
<tr><td rowspan="2">Predicted condition</td><td>Predicted condition positive</td><td>True positive</td><td>False positive, Type I error</td><td>Positive predictive value (PPV), Precision = $\frac{\Sigma \text{ True positive}}{\Sigma \text{ Predicted condition positive}}$</td><td colspan="2">False discovery rate (FDR) = $\frac{\Sigma \text{ False positive}}{\Sigma \text{ Predicted condition positive}}$</td></tr>
<tr><td>Predicted condition negative</td><td>False negative, Type II error</td><td>True negative</td><td>False omission rate (FOR) = $\frac{\Sigma \text{ False negative}}{\Sigma \text{ Predicted condition negative}}$</td><td colspan="2">Negative predictive value (NPV) = $\frac{\Sigma \text{ True negative}}{\Sigma \text{ Predicted condition negative}}$</td></tr>
<tr><td colspan="2" rowspan="2"></td><td>True positive rate (TPR), Recall, Sensitivity, probability of detection, Power $= \frac{\Sigma \text{ True positive}}{\Sigma \text{ Condition positive}}$</td><td>False positive rate (FPR), Fall-out, probability of false alarm $= \frac{\Sigma \text{ False positive}}{\Sigma \text{ Condition negative}}$</td><td>Positive likelihood ratio (LR+) $= \frac{\text{TPR}}{\text{FPR}}$</td><td rowspan="2">Diagnostic odds ratio (DOR) $= \frac{\text{LR+}}{\text{LR−}}$</td><td rowspan="2">$F_1$ score = $2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision + Recall}}$</td></tr>
<tr><td>False negative rate (FNR), Miss rate $= \frac{\Sigma \text{ False negative}}{\Sigma \text{ Condition positive}}$</td><td>Specificity (SPC), Selectivity, True negative rate (TNR) $= \frac{\Sigma \text{ True negative}}{\Sigma \text{ Condition negative}}$</td><td>Negative likelihood ratio (LR−) $= \frac{\text{FNR}}{\text{TNR}}$</td></tr>
</table>

Figure 3.2: 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:

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

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:

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

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:

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

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

Project workflow

Findings

Analysis

Discussion

Conclusion

Bibliography

Appendix