

MAPPING AND PREDICTING THE INTRA-URBAN DEPRIVATION DEGREES USING EO DATA

EQI LUO
June 2021

SUPERVISORS:
Dr. M.Kuffer
Dr. J.Wang



MAPPING AND PREDICTING THE INTRA-URBAN DEPRIVATION DEGREES USING EO DATA

EQI LUO

Enschede, The Netherlands, June 2021

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

Specialization: Urban Planning and Management

SUPERVISORS:

Dr. M.Kuffer

Dr. J.Wang

THESIS ASSESSMENT BOARD:

Prof.dr.ir. J.A. Zevenbergen (Chair)

Dr. S. Georganos (External Examiner, Université libre de Bruxelles)

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

The rapid global proliferation of slums is a major challenge in urbanisation. Most of the growing urban population in Low- to Middle-Income Countries (LMICs) is absorbed by slums and informal settlements (here called deprived areas). In the last decades, deprived areas have been identified and mapped to a great extent, given the increasing availability of very-high-resolution (VHR) satellite images and the development of machine learning (ML) techniques. Yet, most earth observation (EO) approaches only yield a binary delineation of deprived/non-deprived areas – an oversimplified understanding of urban deprivation that mostly built upon physical or morphological features, with little information inferred regarding the intensity, variation, and diversity of intra-urban deprivation. In this study, we attempt to explore the potential of using VHR EO-based data to predict the degrees of intra-urban deprivation in Nairobi, Kenya. This involves a two-step workflow of characterising and predicting a continuous index of deprivation degrees. First, a principal component analysis (PCA) is conducted to characterize the multi-dimensionality and intensity of deprivation as a set of continuous indices (i.e., the ‘*multi-deprivation portfolio*’), using 100m standard grids as analytical units. Next, a convolution neural network (CNN) based regression model is trained to directly predict the ‘*multi-deprivation portfolio*’, using only SPOT-7 images. The PCA results identify four major domains of deprivation, i.e., PC1: Poverty, accessibility to facilities, and maternal health support, PC2: Dense urbanization, absence of green space and waste management, PC3: Air and water contamination, and PC4: Transport infrastructure. Among these deprivation domains, PC2 is the most morphology-based domain and successfully captures the spatial configurations of slums in Nairobi. During the test of EO-based data for predicting the domains of deprivation, the best prediction of the proposed CNN regression model is also obtained in PC2, with an R^2 of 0.6543; whereas the CNN fails on other deprivation domains. Based on these results, this study confirms that urban deprivation is by nature a multi-dimensional, complex concept, and PCA is a useful tool to unpack and measure this multi-dimensionality in continuous scales. Most importantly, we demonstrate the potential of an EO-based method to directly capture the degrees of multiple deprivation with relatively high accuracy. We suggest scaling up this method to inter-city, national or even global level and produce larger-scale maps of deprivation degrees in LMICs cities in future studies.

Keywords: Deprivation, Slums, Earth Observation, Deep Learning, Low- to Middle-Income Countries

ACKNOWLEDGEMENTS

This is the last section that I wrote for my MSc Thesis – to arrive at this final manuscript, I spent nine months of dedication and efforts, along with numerous ups-and-downs and mood swings. I have to say it out loud now “Doing an MSc in ITC is definitely not easy, but I have achieved so many things never imaged before!”. When I look back on the two-year journey, everything feels surreal. Maybe I am being over-emotional or dramatic, but I am the type of people who like to expose their feelings. Now, with all these complex and genuine emotions, I would like to sincerely express my acknowledgements.

First and foremost, my deep, strong, overloaded gratitude to my MSc supervisors – Dr Monika Kuffer and Dr Jiong Wang, for your continuously warm support, professional guidance, enlightening feedback, and valuable critiques throughout the whole MSc supervision. It is my great pleasure to work with you, and I believe we have jointly built a cooperative, motivating team spirit. I remember there were sometimes when I felt frustrated and anxious about the research progress, yet, after our weekly meetings, you can always cheer me up with even stronger motivation. Your mentorship not only shapes me as a growing junior researcher equipped with increasing expertise but also inspires my deepening interests in the field of urban deprivation. Most importantly, your recognition helps me to further built up my confidence, while I also learn so much from you to stay modest. I hope we will still stay in touch even after graduation!

My special thanks to the IDEAMAPS project, SLUMAP project, for providing me with a large share of research data, without which this MSc thesis cannot be done. Also, I would like to thank the African Health and Population Research Centre for offering me an internship opportunity from which I developed lots of working skills. High appreciation to my internship supervisors – Dr Dana R. Thomason and Dr Caroline Kabaria, for your professional guidance and encouragement. Great thanks to Dr Jaap Zevenbergen for your valuable feedback on the proposal defence and mid-term report. Additionally, my strong thankfulness to the ITC Excellence Scholarship – I really appreciate the substantial final support in this two-year program. Also, my great thanks to all the staff in ITC for supporting us in such a difficult time under the covid-19 pandemic. A final special thanks in this paragraph for some of the amazing food in the ITC cafeteria, especially the spareribs randomly showing up on Friday.

Here, I would also express my overloaded gratitude to all the wonderful friends I have made in ITC. You have no ideas how much fun I had with you guys/girls and how amazing it is to stay in a foreign country, all speaking in English (well, I wish I knew other languages), coming in with seemingly totally different cultural backgrounds but ends up building such empathetic, equal, and supportive friendships with all of you – we really broke the socially constructed barriers and stereotypes. A special thanks to all the fellows in cluster 3-128, the room I have stayed in for nearly three months conducting my research analysis – you all together made this thesis journey much more fun. Lastly, my constant gratitude to all my lovely friends back in China – I miss you all in these two years – and thanks for always be there for me, no matter wherever and whenever I am.

My heaviest thankfulness to my incredible, unconditionally supportive, and cutest parent, whom I never stop missing. Thank you for always letting me follow my own willingness and encouraging me to obtain the MSc degree in the Netherlands from the beginning. I love you so much!

Lastly, I would like to spare some acknowledgements to myself, Eqi Luo, who has been devoting so many efforts and grown a lot during the past two years. As the MSc life approaching its end, I will soon start a new adventure somewhere in the globe. However, all the stories that happened in Enschede (a town I whining about nearly all the time but soon will miss) are life-long unforgettable, and I always appreciate all the people I have met here.

TABLE OF CONTENTS

1.	Introduction.....	1
1.1.	Background and justification.....	1
1.2.	Research problem.....	2
1.3.	Research objectives.....	3
1.4.	Conceptual framework.....	4
1.5.	Thesis structure.....	5
2.	Literature review.....	6
2.1.	Multi-dimensional deprivation.....	6
2.2.	Modelling deprivation.....	6
2.3.	Mapping deprivation.....	7
2.4.	Mapping deprivation through deep learning (CNN).....	8
2.5.	Configuration of the CNN model (classification vs regression).....	9
3.	Methodology.....	11
3.1.	Study area.....	11
3.2.	Overall Methodology.....	12
3.3.	Data.....	13
3.4.	Principal component analysis.....	20
3.5.	Deep CNN-based regression model.....	24
4.	Results.....	31
4.1.	Elimination of unsuitable indicators for PCA.....	31
4.2.	Decomposing deprivation dimensionality through the PCA component scores.....	31
4.3.	PCA-based multiple deprivation indices.....	33
4.4.	PCA results validation and interpretation.....	36
4.5.	CNN model implementation and optimization.....	42
4.6.	CNN prediction on the morphology-based deprivation index.....	43
4.7.	CNN prediction on other deprivation indices.....	44
5.	Discussion.....	46
5.1.	Multi-dimensionality nature of deprivation.....	46
5.2.	Measuring multiple deprivation.....	47
5.3.	Role of an EO-based method in deprivation mapping.....	48
5.4.	Limitations.....	49
6.	Conclusion and recommendation.....	50
6.1.	Conclusions.....	50
6.2.	Recommendations for further studies.....	50

LIST OF FIGURES

Figure 1 - The conceptual framework of multiple deprivation mapping, developed for this MSc study.	4
Figure 2 - An overall structure of Convolutional Neural Networks. Source: (Alom et al., 2019).	8
Figure 3 - Plots visualizing the difference between classification and regression outputs.	10
Figure 4 - The study area map of Nairobi. VHR image source: WorldView-3.	11
Figure 5 - The overall methodology of mapping multiple-deprivation degrees.	12
Figure 6 - The domains of multiple deprivation. Source: (Abascal et al., 2021)	13
Figure 7 - The deprived areas shown in VHR images. (a) Worldview-3, 2019. (b) SPOT-7, 2017.	18
Figure 8 - The land use typologies of urban areas in Nairobi, 2020. Base map: Worldview-3. Source: (Vanbrysse et al., 2021).	19
Figure 9 - The settlements extents of Nairobi. Source: (CIESIN & Novel-T, 2020).	20
Figure 10 - An illustration showing the inputs and outputs CNN-based regression model.	24
Figure 11 - The extracted image tiles for the CNN-based regression model. (a) examples showing the size of each image tile. (b) the total extracted image tiles for the whole study area.	25
Figure 12 - The sampling approach for extracting the training and validation datasets.	25
Figure 13 - The overall architecture of the proposed deep CNN-based regression model. BN means batch normalization, ReLU means rectified linear unit activation function.	26
Figure 14 - Spatial distribution of the extracted four sub-dimension deprivation indices: (a) PC1 - Poverty, accessibility to facilities, and maternal health support; (b) PC2 - Dense urbanization, absence of green space and waste management; (c) PC3 - Air and water contamination; and (d) PC4 - Transport infrastructure.	34
Figure 15 - Spatial distribution of the aggregate multi-deprivation index.	35
Figure 16 - A boxplot showing the distribution of PCA-based deprivation indices on slum and non-slum areas.	36
Figure 17 - The map overlying the slum boundary with PC1, PC3, PC4 and aggregate index.	36
Figure 18 - The map overlaying the slum boundary with PC2.	37
Figure 19 - The grouped boxplots showing the distribution of PCA results on different land use typologies.	38
Figure 20 - The visual assessment of PC2 on slum areas by comparing with VHR and street-view images. Source: (Mapillary, 2021).	39
Figure 21 - The visual assessment of PC2 on CBD area by comparing with VHR and street-view images. Source: (Mapillary, 2021).	40
Figure 22 - The visual assessment of PC2 on the atypical deprived areas by comparing with VHR and street-view images. Source: (Mapillary, 2021).	40
Figure 23 - The visual assessment of PC2 on the formal built-up areas by comparing with VHR and street-view images. Source: (Mapillary, 2021).	41
Figure 24 - Histogram of PC2 input samples.	42
Figure 25 - The density scatter plot of PC2 prediction on test dataset.	42
Figure 26 - The visual comparison between CNN prediction and the reference PC2 index.	43
Figure 27 - The CNN prediction on PC1 (Poverty, accessibility to facilities, and maternal health support).	44
Figure 28 - The CNN prediction on PC3 (Air and water contamination).	44
Figure 29 - The CNN prediction on PC4 (Transport infrastructure).	45
Figure 30 - The CNN prediction on the aggregate deprivation index.	45

LIST OF TABLES

<i>Table 1 – The list of candidate indicators for deprivation index formulation.</i>	<i>14</i>
<i>Table 2 – The summary of available VHR satellite images of Nairobi for this study.</i>	<i>18</i>
<i>Table 3 – A detailed summary of the proposed deep CNN-based regression model structure.</i>	<i>27</i>
<i>Table 4 – A summary of the hyper-parameter setting for the model initialization.</i>	<i>28</i>
<i>Table 5 – The list of hyper-parameters to be tuned for model optimization.</i>	<i>29</i>
<i>Table 6 – The summary of discarded indicators for PCA analysis after the quality check.</i>	<i>31</i>
<i>Table 7 – The retained principal components scores and the component loadings in the rotated matrix.</i>	<i>32</i>
<i>Table 8 – The mean of deprivation indices between slum and non-slum.</i>	<i>36</i>
<i>Table 9 – The optimal values after hyper-parameter tuning.</i>	<i>42</i>
<i>Table 10 – The CNN performance on the test datasets in predicting PC2.</i>	<i>42</i>

LIST OF ABBREVIATIONS

AMDI	Aggregate Multiple Deprivation Index
BN	Batch Normalization
CNN	Convolutional Neural Network
DHS	Demographic and Health Surveys
EO	Earth Observation
FC	Fully Connected
GIS	Geographic Information System
GRID3	Geo-Referenced Infrastructure and Demographic Data for Development
HR	High Resolution
IDEAMAPS	Integrated Deprived Area Mapping System
KMO	Kaiser-Mayer-Olkin
LMIC	Low- to Middle-Income Country
LULC	Land Use and Land Cover
MAE	Mean Absolute Error
MSE	Mean Squared Error
MAPU	Modifiable Areal Unit Problem
MDG	Millennium Development Goal
ML	Machine Learning
NDVI	Normalized Difference Vegetation Index
NTL	Night-Time Light
PCA	Principal Component Analysis
RCM	Rotated Component Matrix
ReLU	Rectified Linear Unit
RF	Random Forest
RMSE	Root Mean Squared Error
SDG	Sustainable Development Goal
SDI	Slum Dwellers International
SLUMAP	Remote Sensing for Slum Mapping and Characterization in sub-Saharan African Cities Project
SoVI	Social Vulnerability Index
SVM	Support Vector Machine
VGG	Visual Geometry Group
VHR	Very High Resolution

1. INTRODUCTION

1.1. Background and justification

Currently, more than half of the world's population lives in urban areas, with an estimated increase to 68% by 2050 (United Nations, 2019). As the world becomes more urbanised, many cities, especially in the Low- to Middle-Income countries (LMICs), are facing urbanisation problems like growing numbers of slum dwellers, lack of basic services and infrastructure, a rising level of inequalities and social exclusion etc. (Zhang, 2016). The rapid proliferation of slums is considered one of the most direct manifestations of urban poverty (Arimah, 2010). Therefore, the urge to upgrade slum conditions and reduce informality has been recognised as a main global challenge and compiled in many development agendas. For example, in the Millennium Development Goals (MDG) proposed by UN, 'to improve the lives of a minimum of 100 million slum dwellers by 2020' was established to address the rising expansion and severity of slum (United Nations, 2015), likewise in the Sustainable Development Goals (SDGs) target 11.1: "ensure access for all to adequate, safe and affordable housing and basic services and upgrade slums" (United Nations, 2018). Despite the efforts in upgrading slum conditions worldwide, the actual number of slum dwellers, however, increased from 807 million to 883 million between 2000 and 2014, of which the vast majority happened in LMICs, especially in Asia and Africa (United Nations, 2018).

The term 'slum' is widely used in urban and development studies but varies strongly across the world, which can be due to the inconsistency of slum definition itself, as well as the heterogeneity of slums in the sense of morphology and socio-economic status (Kohli et al., 2012). A broadly accepted definition of 'slum' refers to the household or a group of individuals which lack one or more of the followings: durable housing, sufficient living space, access to safe water, access to adequate sanitation and security of tenure (UN-Habitat, 2003). Nevertheless, this definition provided by UN-Habitat is household oriented, reflecting little information at the area level (e.g. lack of infrastructures, hazard-prone risk, low accessibility to facilities etc.) faced by slum dwellers living in the deprived areas (Lilford et al., 2019). In addition, the definition also casts multi-dimensional characteristics that may not be mutually inclusive, and therefore the characteristics of people living in slums may not fully be manifested by the morphology of substandard housing, but usually coupled with deprived conditions in other domains (Gilbert, 2007). One of the consequences is that the morphologies of slums vary significantly across or even within the same cities (Taubenböck et al., 2018). For instance, in Mumbai, the slums appear heterogeneously across space in terms of geometry, density, pattern, and environment (Taubenböck & Kraff, 2014), yet, few systematic studies have been performed to investigate such complexity (Kuffer et al., 2017).

Considering the multidimensionality and fuzziness in the characterisation of slums, the conventional dichotomy of "slums and non-slum areas" only provides an oversimplified understanding of deprivation, usually based upon their spatial configurations, locations and extents (Thomson et al., 2020). However, even within the same slum located in one city, there is a differing mixture of deprivation in terms of intensity and dimensions (social-economic, living conditions, ecological etc.) (Jankowska et al., 2011), whereas the authorities tend to decline the diversities of these slum-like regions (Baud et al., 2010). Similarly, the traditional aggregated census-based approach cannot reveal the inner variety in its study unit (usually neighbourhood), leading to few discussions on the diversity and cross-boundary clusters of deprivation (Kuffer et al., 2017). Another disadvantage is that once an area has been declared as a slum, it may bring unintended stigmatisation to its dwellers (Eksner, 2013). As such, it is more important to reveal the internal variation and heterogeneity of deprivations faced by slum dwellers (e.g., socio-economic

factors, environmental risks), which could be analysed later to underpin more comprehensive and contextualised slum upgrading plans.

Recently, more approaches were introduced to investigate slums via the lens of 'Multiple Deprivation', which were widely explored in previous census-based studies (Baud et al., 2009; Gill, 2015), as it enables to capture slum as a multi-dimensional manifestation resulted from not only the traditional aspects such as housing conditions, access to water and sanitation but also the socio-economic status, environmental, ecological factors of deprived areas and the dwellers (Ajami et al., 2019; Arribas-Bel et al., 2017; Kuffer et al., 2017; Thomson et al., 2020). In general, the multiple dimensions of deprivation faced by slum dwellers are not independent of each other. Instead, it is the interplay of such multiple facets of deprivation that characterising and shaping the diversity and complexity of the slums (Mahabir et al., 2016). Therefore, as a response to the SDG goals, it is important to increase our knowledge of the variation and diversity of deprivation within a city and to develop a generalised method identifying deprivation level, which could be transferred to other LMICs. With rapidly expanding deprived areas, the governments and policymakers require more detailed and contextual spatial information to formulate urban development plans and support decision-making regarding the pro-poor agenda (UCLG, 2018). However, most of the LMICs usually lack routinely updated and accurate census and geospatial data of deprived areas due to limited resources and technologies (United Nations, 2018), thus failing to target such issues and being trapped in the vicious circle of deprivation.

Conventionally, deprived areas were investigated via census or household surveys, considered as labour-intensive, costly, large-scale and easy to be outdated (Mahabir et al., 2016). In the last decade, using very-high-resolution (VHR) earth observation (EO) data to identify and map the spatial distribution of deprived areas has become one of the mainstreams in urban studies (Kuffer, Pfeffer, & Sliuzas, 2016), given the increasing availability of multi-spatiotemporal satellite image and the recognition that multiple deprivations partially manifest themselves on the physical morphologies in space (Duque et al., 2015; Taubenböck et al., 2009). Different EO-based methods and technologies have been applied to investigate deprived areas in terms of identification (Kit et al., 2012; Williams et al., 2020), temporal dynamics (Liu et al., 2019), and severity (Ajami et al., 2019; Arribas-Bel et al., 2017; Kuffer et al., 2020), ranging from local to the global level. In addition, the development of machine learning (ML) algorithms has opened a new gate for the EO community in image analysis. More advanced and efficient classifiers have been performed in slum-related studies, including traditional ML methods (i.e., using hand-crafted features) and deep learning models. For example, Leonita et al. (2018) applied two traditional ML algorithms, i.e., support vector machine (SVM) and random forest (RF), to identify deprived dwellings in Indonesia and compared the performances of two models. These powerful ML methods, given the availability of high-quality VHR data and increasing computational ability, could reach a high accuracy from 75% to 95% (Kuffer, Pfeffer, & Sliuzas, 2016). However, revealed by Mahabir et al. (2018), even though the number of such researches has been growing over time, most of them are still restrained on limited geographic scales, e.g. only small areas (particular neighbourhoods and blocks) within a city (Ma et al., 2017), and insufficient reflection on the multi-dimensionality of slums, as the EO images mainly reflect physical information of the land surface, resulting in a lack of understanding in deprived areas regarding their variety within the global context.

1.2. Research problem

Deprived areas have been effectively detected and mapped through EO techniques in the last decades, with more advanced methods being developed and achieving remarkable performance (Kuffer, Pfeffer, & Sliuzas, 2016; Lilford et al., 2019). Yet, very limited information in relation to multiple deprivations on deprived areas has been extracted from the satellite images, such as diversity, severity and dynamics etc.

Most EO-based methods generate binary delineations of the slums depended on their morphological characteristics (Kit et al., 2012; Kohli et al., 2016; Persello & Stein, 2017). In other words, the linkage between the spatial morphology (manifested in EO data) and multiple deprivations (usually ‘hidden behind the images’) has not been systematically explored. Thus, to build such a connection and support an in-depth and more holistic understanding of urban deprivation, this study unpacks multiple deprivations as a multi-dimensional, continuous spatial concept. Moreover, previous studies mostly focus on small urban parts or pre-delineated deprived pockets, covering areas of several km², rather than map deprivation at inter-/intra-city level (Ajami et al., 2019; Liu et al., 2019; Wang et al., 2019). However, we argue that depicting multiple deprivations at an intra-city level will surely provide more insights into the diversity of urban poverty and help the local government facilitate slum upgrading plans.

In this research, to avoid the imprecision and inconsistency of terminological discourses about ‘slums’, and explicitly stress the multi-dimensionality, as well as the continuous degree of deprivation, we, instead, decide to employ a more comprehensive and area-based term: ‘deprived areas’, which encompasses the multi-dimensional deprivation characteristics of slums and helps to unveil the marginalisation and socio-economic disparities of the deprived dwellers (Ajami et al., 2019; Arribas-Bel et al., 2017; Nolan, 2015; Thomson et al., 2020; Wurm & Taubenböck, 2018). Unlike conventional studies that identify areas as slums or non-slums (which, in essence, reflects only one dimension of deprivation), this research attempts to quantitatively measure the degree of multiple deprivations at standard gridded units within the entire urban area. More specifically, the ‘*multi-deprivation portfolio*’ – a set of continuous indices that indicate the degree of deprivation from multiple sub-domains and/or summarized domain – would be generated by this research. Afterwards, a convolutional neural network (CNN) model will be trained to directly estimate the degree of multiple deprivations, relying only on EO data. By doing this, the feasibility and effectiveness of leveraging the state-of-the-art method, combined with VHR EO images in predicting the degree of multiple deprivation would be examined.

1.3. Research objectives

1.3.1. General objective

The overall objective of this research is to characterise multiple deprivation by exploring the potential of an EO-based approach in capturing the multi-dimensionality of urban deprivation.

1.3.2. Specific objectives

The general objective can be further broken down into three sub-objectives, combined with research questions formulated at the operational level:

- 1. Characterise multiple deprivation and measure its spatial variation in continuous scales.**
 - a. What are the common key dimensions of multiple deprivations?
 - b. Which method is appropriate to unpack the multi-dimensionality of deprivation as a set of multiple deprivation indices?
 - c. What are the characteristics/diversity of deprivations within the study area?
- 2. Explore the potential of an EO-based method to predict the intra-urban continuous deprivation degrees.**
 - a. What are the criteria to divide the dataset for training, validation, and test?
 - b. How to train a CNN-based model to predict the continuous deprivation levels?
 - c. What are the suitable measures to evaluate the CNN-based model?
 - d. To what extent can the CNN-based model capture the degrees of multiple deprivations through VHR imagery?

3. Discuss the role of EO-based methods in deprivation mapping.

- What are the advantages of applying an EO-based model to directly capture deprivation degrees?
- What new insights of deprivation mapping does this research bring?
- Based on this research, how could EO-based methods contribute to deprivation mapping?

1.4. Conceptual framework

In this research, the key concept – ‘multiple deprivation’ is defined as a complex, multi-dimensional area-based manifestation dependent on various aspects. The word ‘multiple’ here underscores its multi-dimensionality but also somewhat contributes to the dissonance of which aspects should be included in defining multiple deprivation. In consequence, the definition of deprivation varies extensively from case to case. The goal of building this framework is not to exhaustively list all the existing definitions of multiple deprivation and/or their components and deliver a universal agreement, but rather provide a broad consensus in “what usually constitutes multiple deprivations?” and “How different methods capture the variation of deprivation”, thus underpinning the development of the research methodology.

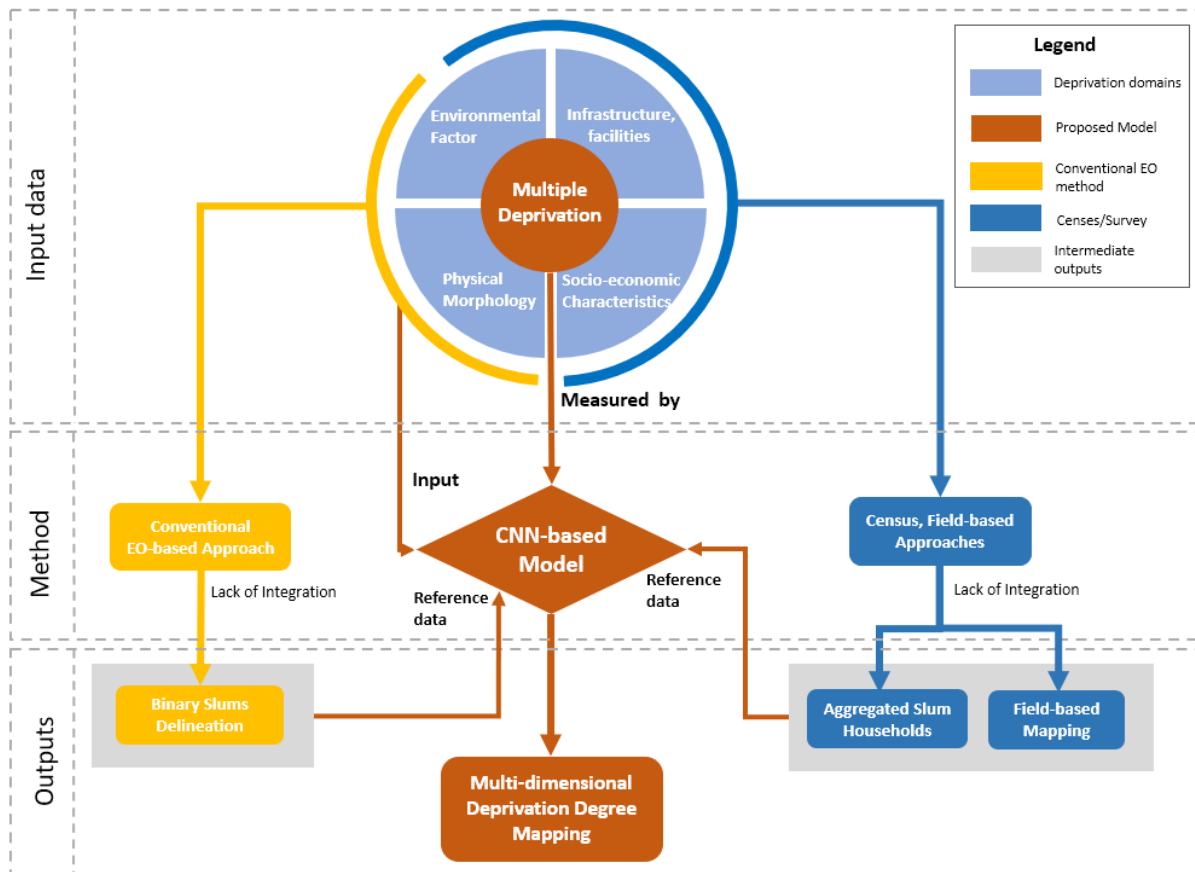


Figure 1 - The conceptual framework of multiple deprivation mapping, developed for this MSc study.

The multi-dimensionality of deprivation is commonly deconstructed into the following sub-domains, including but not limited to socio-economic status, physical morphology, environmental factors, infrastructures and facilities etc. Conventionally, the EO-based approaches cannot fully investigate the complexity of multiple deprivation, nor the survey- or filed-based methods. Hence, the outputs from these methods are usually siloed and lack integration. Therefore, this research proposes a novel third option – a CNN-based model that directly aims to quantitatively measure the complexity of multiple deprivation.

The crucial part of the designed framework is a CNN-based deep learning model enabling to capture the multi-dimensional deprivation from physical morphologies in the EO data. To attain the model objective, resultant deprivation mapping built upon different deprivation domains would be fed back into the CNN as training data, combined with VHR EO images. Thus, by training this model, the usefulness of using EO-oriented features to predict deprivation degrees and its multi-dimensionality will be examined, with a continuous data-driven multiple-deprivation index expected, thus increasing our present knowledge on how multidimensionality and variety of deprivation could be extracted from physical characteristics in the imagery data. Ultimately, the outputs from this proposed model would support unveiling the diversity of multiple deprivations at the citywide scale.

1.5. Thesis structure

The overall structure of this thesis is outlined as follows:

Chapter 1 introduces the background and rationales of this study and specifies the research problems and sub-objectives. A conceptual framework is also provided to clarify the inter-relationships of the main objects of interests in this study.

Chapter 2 provides a detailed review of previous studies, covering several crucial concepts in the field of urban deprivation and advanced mapping techniques – i.e., ‘multiple deprivation’, ‘multi-dimensional index formulation’, ‘deprivation mapping’; ‘convolutional neural network’ and ‘CNN-based regression model’. It allows to identify the research gap and help to propose the methodology.

Chapter 3 starts with a brief description of the study area, the reasons for choosing this study area, and an overview of the methodology; it then presents a summary of the input data and its pre-processing steps; lastly, it articulates the two major methods, i.e., the principal component analysis and convolutional neural network, applied in this study.

Chapter 4 presents the results obtained from all the analysis conducted in this research. It describes the PCA results and visualizes the multiple deprivation indices of Nairobi, followed by a validation section. Next, the CNN results are reported and compared with the PCA results by visual assessments.

Chapter 5 discusses the main finding obtained from the research and speaks out the advantages achieved by the proposed method in comparison to previous studies. A limitation discussion is also provided.

Chapter 6 finalizes the research paper by summarizing the major conclusions, highlights from this study, and further providing some potential directions for future studies.

2. LITERATURE REVIEW

2.1. Multi-dimensional deprivation

Deprivation and poverty were commonly measured as a one-dimensional phenomenon solely dependent on income or household consumption in previous urban poverty researches (Lucci et al., 2018). Such approaches usually identify poverty based on a single value or threshold (e.g. poverty line, Gini Index) to divide the population into two groups, i.e. poor and non-poor (Martínez et al., 2016), providing limited and often biased information to tackle poverty (Tigre, 2018). Parallely, related spatial analyses were also restricted on the “slum and non-slum” dichotomy whereby a place is classified as either slum or not (Patel et al., 2014). Following this dichotomy, researchers have generated a substantial amount of classification maps clearly presenting the spatial location and extent of slums (Kohli et al., 2016; Kuffer, Pfeffer, Sliuzas, et al., 2016; Persello & Stein, 2017; Williams et al., 2020), which could efficiently inform the urban planners and local communities to target the critically deprived regions. Yet, such dualism usually fails to further unveil the heterogeneity and variety of deprivations within and across slums, as the results are just formed of binary or multi-class categorical labels.

Recently, more studies have switched to the concept of ‘Multiple Deprivation’, which incorporates other multi-dimensional characters of human well-being beyond the monetary aspect. The measure of multiple deprivations is commonly conducted by generating a composite index. For example, Baud et al. (2008) designed a holistic framework based on the asset livelihoods approach, characterising deprivation as the interplay of physical, social, human and financial capitals and applied it to three Indian mega-cities. Likewise, Alkire et al. (2014) unpack deprivation into three domains, namely education, health and living standard for measuring acute global poverty. On the national level, the British government has a long history of monitoring deprivation via the UK indices of deprivation, comprised of several domains for more than 30 years (Gill, 2015). Although there is a slight variance in how researchers conceptualise multiple deprivations, it is widely recognised that deprivation should be investigated from more than only the financial aspect (Martínez et al., 2016).

2.2. Modelling deprivation

The most common way of modelling deprivation is by calculating a multivariate index that indicates the degrees of deprivation. Such indices are often named differently, e.g., deprivation index (Yuan & Wu, 2014), slum index (Duque et al., 2015; Engstrom et al., 2015), but all try to characterise deprivation. To present a brief outline, some common approaches in building the deprivation index were reviewed. In the early studies, different indicators were initially standardised (e.g., z-score method) into similar scales and then assigned with relative weights to formulate a single measure of deprivation (Dolk et al., 1995). Various criteria were applied to establish the weights, such as equal or arbitrary weights (Carstairs & Morris, 1990), expert opinions (Cabrera-Barona & Ghorbanzadeh, 2018) and previous literature. The output is quite straightforward and easy to interpret, so it can be reproduced across different regions and time (Allik et al., 2020). Later, a more refined approach is proposed where the scores of indicators representing the same domain are first combined to generate the indices of sub-deprivation and aggregated to a single index later. This composite measure allows to evaluate deprivation degree from different aspects individually and conduct inter-comparison. For example, Baud et al. (2008) calculated the deprivation index separately for four domains and summarised them with equal weight. However, these above methods are often criticised for the normative weighting subjective to value judgements and empirical perception from policymakers and researchers involved (Deas et al., 2003) and the inability to capture the intersection of multiple deprivation domains (Ipsum et al., 2015).

The third type of approaches, statistics-based methods, have become popular since no assumptions need to be pre-defined about the relative weights, thus argued to be more objective. Popular techniques include principal component analysis (Basu & Das, 2020; Vyas & Kumaranayake, 2006), multiple correspondence analysis (Ajami et al., 2019), factor analysis (Gill, 2015; Roy et al., 2020) and so on. Among the three approaches, PCA is suitable for quantitative indicators, while MCA works well on categorical data, and for factor analysis, it requires careful consideration of the communalities from the input. In the case of PCA, instead of manually determining a weight for each indicator, a set of linear combinations of variables will be derived based on a covariance or correlation matrix, which explains most of the variance of deprivation. It is widely used to reduce the large dimensionality of input data and then aggregate the retained components to generate a ‘data-driven’ index (Abdi & Williams, 2010).

2.3. Mapping deprivation

As stated before, the term ‘deprived areas’ was adopted to refer to slums, informal settlements, and other types of dwellings or settlements in slum-like conditions. In this section, previous studies that involved the detection of any aforementioned sub-categories of deprived areas were reviewed to present a brief summary of popular approaches and the state-of-the-art in deprivation mapping.

So far, various approaches have been performed to investigate deprived areas. Recently, a detailed review of deprivation mapping by Kuffer et al. (2020) has summarised four major methods widely applied in practice, namely, 1) census and household survey (e.g., Agarwal et al., 2018; Baud et al., 2009; Fink et al., 2012), 2) field-based mapping (e.g., Karanja, 2010; Makau et al., 2012), 3) visual interpretation of EO images by human (e.g., Anurogo et al., 2017; Gruebner et al., 2014) and 4) computational models using machine algorithms (e.g. Arribas-Bel et al., 2017; Engstrom et al., 2015; Mahabir et al., 2020). Yet, none of them is able to yield an integrated, scalable, frequently updated, and contextual result as each one has its own strengths and drawbacks. The census-based survey is the most traditional method in measuring deprivation. It often follows the definition of ‘slum’ by UN-Habitat, enabling to provide comprehensive information and conduct cross-city/country analysis (Basu & Das, 2020; Patel et al., 2014). However, one big disadvantage is that such surveys are mostly conducted at the household level and then aggregated into administrative units with irregular, variant spatial boundaries and scales. Through this aggregation, the statistics may be under-/over-estimated due to the Modifiable Areal Unit Problem (MAPU) (Nelson & Brewer, 2017), and some tiny slum pockets within an identified ‘rich’ area might also be overlooked (Christ et al., 2016; Subbaraman et al., 2012). Moreover, as census-based data are mostly collected at the household level, few area-level information of deprivation can be derived from it (Lilford et al., 2019). The field-based approach, on the other hand, is able to produce very local and contextual information with high validity and reliability by the local communities and NGOs, but hardly could be scaled up or generalised to other contexts (Kuffer et al., 2020). The visual interpretation approach requires intensive labour investment, and the criteria used to define slum boundary often vary among experts, thus also leading to high uncertainties on where to draw the boundaries of deprived areas (Kraff et al., 2020; Pratomo et al., 2017), albeit achieving relatively high accurate delineation.

Nevertheless, among all the methods, using ML algorithms to map deprivation from satellite images has drawn more attention in the EO community because of its high accuracy, the ability to be automatised, the increasing availability of HR/VHR data and the ability to cover large areas in principle (Kuffer, Pfeffer, & Sliuzas, 2016). Common ML-based approaches in deprivation mapping include support vector machine, logistic regression and random forest, which are sometimes employed jointly with object-based image analysis (OBIA) to better leverage the spectral and contextual features to achieve better prediction (Kuffer, Pfeffer, & Sliuzas, 2016). Yet, two major problems still restrict the performance of traditional ML methods: the first one is the access to sufficient training labels, i.e., the ground-truth delineation of

deprived and non-deprived areas; the second one is the designing of hand-crafted features and the selection of suitable features that can better capture the deprivation characteristics across various contexts, which are usually time-consuming and highly dependent on user expertise, and computational ability (Kuffer et al., 2020). Moreover, it is very expensive to retrieve VHR satellite images.

In general, with reference to Thomson et al. (2020), the research gaps of EO-based methods in relation to deprived area mapping could be described as follows: (1) limited scale of the study area, i.e., the detection only applied on small patches of slum-households and neighbourhoods, rather than at intra-/inter-urban scale. For example, Ajami et al. (2019) applied a CNN-based model to measure the deprivation degree only within the delineated slums in Bangalore, India, instead of upscaling to the whole city; (2) one-dimensionality. The traditional methods only take into account the physical or morphological characteristics of the urban fabric from VHR image to identify deprived areas, resulting in a limited inference on the socio-economic status of the inhabitants, which are actually the major concerns and target groups of pro-poor policies; (3) dichotomous detection, and by this we mean that common outputs of such researches only provide a binary classification of deprived and non-deprived areas, therefore failing to provide enough information for further exploration such as the internal gradient of multiple deprivations across space and dimensions.

2.4. Mapping deprivation through deep learning (CNN)

Convolutional Neural Network (CNN), as a subset of Deep Learning (DL) algorithms, has stepped forward compared to the conventional ML methods in the field of image analysis. An overall sketch of the CNN model is presented in *Figure 2*.

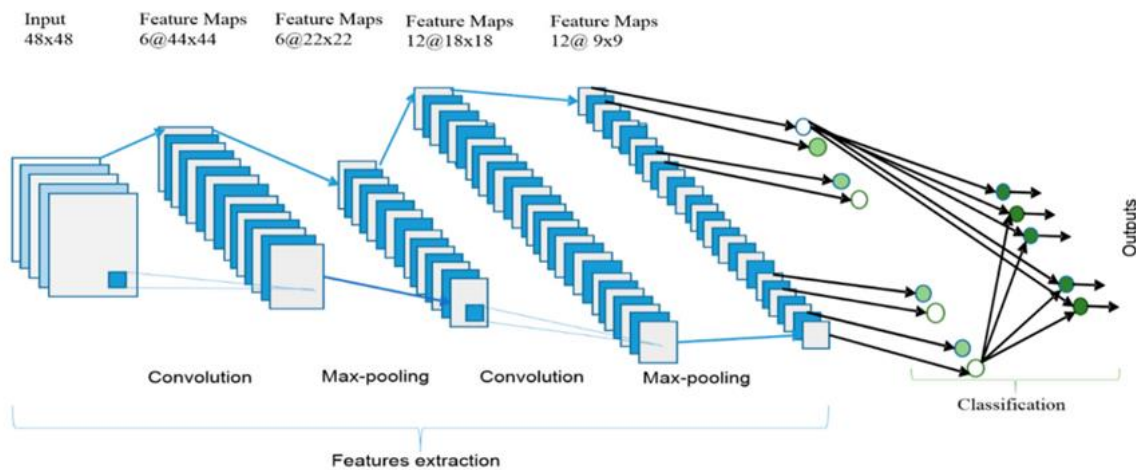


Figure 2 - An overall structure of Convolutional Neural Networks. Source: (Alom et al., 2019).

Similar to any kind of Neural Networks, the architecture of a standard CNN is also constructed in a multi-layer fashion, consisting of three parts: input, hidden layers, and output. As a supervised learning approach, CNN requires image tiles and their corresponding 'labels' as input. Within the hidden layers, there are three main types of layers, namely, convolutional layers, pooling layers and the fully connected layers, of which the first two types of layers learn and extract the spatial-contextual features from the input images (Alom et al., 2019). A typical CNN training process is achieved via back propagation (Hecht-Nielsen, 1989). First, the input maps are convolved with learnable kernels and followed by a liner or non-linear activation function. Then, such convolved output maps will be processed through a series of down-sampling operations to reduce the dimensionality of feature maps (O'Shea & Nash, 2015). After several stacked convolutional and pooling layers, the feature maps are fed into a set of fully connected layers in which the classifying decisions take place, with the prediction generated in the final layer (Alom et al.,

2019). Next, the error between model prediction and the ‘reference labels’ would be calculated by a specified loss function (e.g., sigmoid, SoftMax, Tanh) and then backpropagated to the network, and later on be minimized by the gradient descent method (Song et al., 2019). The weights of each neuron in the network are updated accordingly to minimize the loss error, and by doing these processes iteratively, the CNN model would finally converge.

In the latest years, CNN has attracted overwhelming preference from researchers in image analysis and computer vision fields due to its strong self-learning ability to automatically recognise and extract significant spatial-contextual features from the input data (Lecun et al., 2015). Various CNN-based architectures have been developed and widely applied in image classification tasks. To give some examples, classical CNN architectures include LeNet (LeCun et al., 1998), AlexNet (Krizhevsky et al., 2012), VGGNet (Simonyan & Zisserman, 2015), GoogLeNet (Szegedy et al., 2015), FractalNet (Larsson et al., 2016), ResNet (He et al., 2016) and so on.

By training a deep CNN model, deprived areas could be successfully detected, without the preparation of hand-crafted features as in conventional ML methods, while still yielding promising results of average accuracy over 80%, consuming less time and labour cost (Kuffer, Pfeffer, & Sliuzas, 2016). For example, Persello & Stein (2017) tested a series of CNN models with dilated convolution to distinguish informal settlements from other land use types, among which the best model achieved more than 85% accuracy compared to 77.01% accuracy by the SVM model. Besides, another crucial advantage of CNN is the transfer learning capability of employing a pretrained algorithm embedded with significant previously acquired knowledge, and fine-tune it to address similar tasks with high efficiency and accuracy. To give an example, Wurm et al. (2019) tested the transferability of CNN in semantic segmentation of slums by applying a deep FCN model trained on QuickBird images at 0.5m to a much courser imagery datasets of Sentinel-2 at 10m, where the positive prediction value showed remarkable improvement from 38% to 55%, further confirming the outstanding potential of transfer learning in deprivation mapping.

2.5. Configuration of the CNN model (classification vs regression)

In the EO community, CNN-based models have been largely exploited in various applications, most particularly in image classification problems, such as scene classification, object detection and object segmentation, due to its outstanding self-learning ability (Song et al., 2019). These classification tasks have been extensively dominating the applications of CNN models in the RS field, with a rising number of classification-wise studies being published. For instance, in deprivation mapping, existing studies mostly focus on binary classification of slum/non-slum (Liu et al., 2019; Mboga et al., 2017; Prabhu et al., 2021) or multi-classification between slums and other land use types, such as formal built-up, roads (Williams et al., 2020). In general, CNN-based approaches are able to achieve remarkable performance on class prediction, providing high accuracy in pixel-wise labelling of satellite images (Song et al., 2019).

Although the majority of the CNN applications are still dominated by classification tasks, an increasing number of studies have started to leverage the state-of-art deep regression techniques in EO image analysis, encouraged by its excellent performance on other domains, such as bone age assessment (Ren et al., 2019), object counting (Walach & Wolf, 2016), human pose estimation (Toshev & Szegedy, 2014). To give some examples, Pyo et al. (2019) trained a regression CNN model using hyperspectral images to estimate the concentration of phycocyanin and chlorophyll-a in waterbodies, achieving higher accuracy ($R^2 > 0.86$ and 0.73 , respectively) than conventional bio-optical methods. Li et al. (2020) developed an end-to-end deep regression approach for image registration, in which the corner displacement parameters of unaligned images can be accurately measured and then directly passed onto the project transformation matrix. These advanced applications again ascertain the need for more attention devoted to exploring CNN in performing regression tasks when classification labels could not provide enough information.

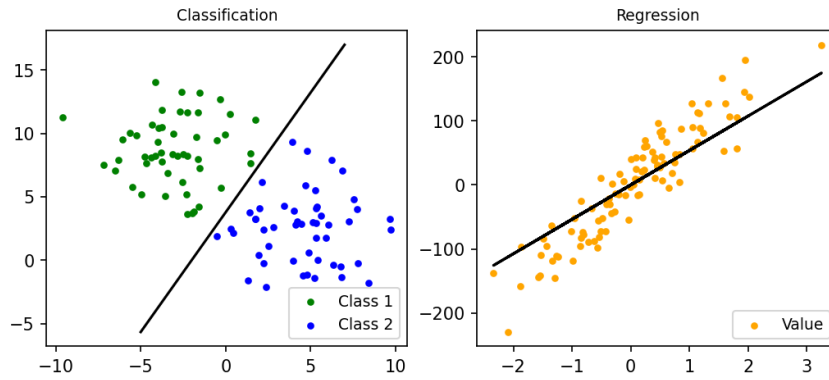


Figure 3 – Plots visualizing the difference between classification and regression outputs.

However, to the author’s best knowledge, few attempts have been made to investigate the potential of applying a CNN model on urban EO data to perform regression task, specifically for measuring deprivation degrees. In general, the scarcity of CNN-based regression application in deprived area mapping mainly results from the insufficiency of ‘suitable’ training data at a continuous scale. Here, the word ‘suitable’ underscores the primary requirement of the target variable in regression analysis, i.e., the dependent variable must be continuous quantities (Draper & Smith, 2014). In common regression problems, continuous numerical values are predicted as outputs from the model, unlike classification tasks where the outputs are discrete, categorical labels, as shown in Figure 3. Yet, in reality, most of the reference data about deprived areas are still delineated by dichotomous boundaries of deprived/non-deprived based on inconsistent varying definitions, along with several nominal, usually oversimplified descriptions like ‘high building density’, ‘overcrowding’ and ‘high pollution’ etc. (Lilford et al., 2019; Mahabir et al., 2016). To conclude, the lack of finer, detailed, and continuous measurement of deprived areas inhibits the CNN-based regression application for predicting deprivation degrees.

Most recently, the potential of applying CNN to capture deprivation level has been preliminarily explored by Ajami et al. (2019), where the authors first trained a deep CNN model to detect slums from formal built-up, and then modified the model architecture via changing the activation function from Log-likelihood into Euclidean loss in the final layer, so that it could perform as a regressor to predict the degree of deprivation. This pioneering study proved the feasibility of using a CNN-based regressor combined with GIS and/or hand-crafted features to quantitatively capture the variations of deprivation from EO data, especially regarding its socio-economic characteristics (Ajami et al., 2019). Nonetheless, the overall methodology is still built upon a prior binary-classification CNN model, rather than directly inferring the deprivation levels from satellite images due to the limited training data on deprived areas.

3. METHODOLOGY

3.1. Study area

In this research, Nairobi, the capital of Kenya, was selected as the study area. *Figure 4* presents the location and a general view of the study area, Nairobi. The boundary of the study area (682 km²) was delineated based on the coverage of available VHR satellite images. Existing slum extents are also visualized.

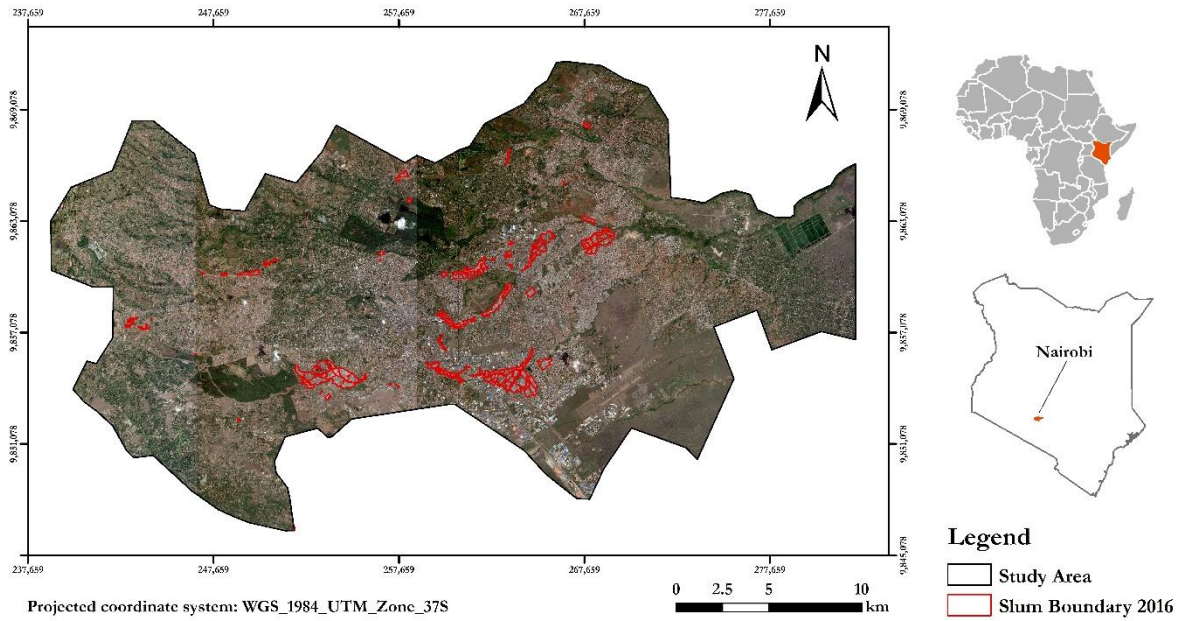


Figure 4 - The study area map of Nairobi. VHR image source: WorldView-3.

Being one of the biggest cities in Africa, Nairobi covers an administrative area of in total 684 km², with an estimated population of 3 million (APHRC, 2014). In recent decades, Nairobi has been through rapid urbanisation and economic development. However, alongside this dramatic transition, the proliferation of deprived areas is still deeply rooted in the city, with nearly 50% to 60% of the total population residing in the slum or slum-like neighbourhoods (UN-Habitat, 2016). For instance, Kibera, located in the southwest of Nairobi, is the largest slum in Africa, whose dwellers have been enduringly suffering from poor housing conditions, overcrowding, and high pollutions, combined with lack of basic services (UN-Habitat, 2011). Apart from its large share of the deprived population, a high variation with the mixture of different deprivation dimensions also exist in Nairobi (Kraff et al., 2019), and yet, limited studies have been conducted to capture and quantify such variety, especially from an intra-city scale. Therefore, unveiling this ambiguous complexity becomes imperative to better inform the local planners. Another important reason to choose Nairobi is the data availability and richness. Fortunately, the authors have access to VHR satellite images at utmost 0.3m resolution and ground-truth of slum boundaries produced by local experts, given the linkage to other ongoing external projects, i.e., the SlumMap (SLUMAP, 2020) and the IDEAMAPS projects (Thomson et al., 2020). The retrieved VHR data covers the entire urban area of Nairobi, which is fundamental for this study to successfully train a deep-CNN model, as CNN usually requires thousands of input data. In addition, compared to other LMICs cities, slum-related data in Nairobi are more well-documented in terms of abundance, quality, and update frequency, due to the great presence of researchers, NGOs, and local initiatives actively investigating slums in Kenya. This data-rich environment of slums also helps the researchers to validate and compare the outputs from this study to previous knowledge.

3.2. Overall Methodology

As shown in the conceptual framework before (Figure 1), most of the traditional EO-based methods perform binary classification of deprived/non-deprived that built upon the physical features, while the survey- and filed-based approaches generate detailed information on the socio-economic status of deprivation, but mainly restricted to the household level. Such deprivation mapping products are usually siloed and lack a comprehensive reflection on the multi-dimensionality of deprivation. As such, this MSc research develops an integrated two-stepwise methodology to directly capture the diversity of deprived areas through predicting the intra-urban deprivation degrees from VHR images. The overall workflow of this research was visualised in Figure 5, comprised of two key technical parts: unsupervised learning and deep learning. The legend shows the sets of steps corresponding to three main research objectives.

Based on the literature review and the local context of Nairobi, the term ‘multiple deprivations’ is firstly conceptualised to inform the development of candidate indicators, covering all possible sub-domains of deprivation. All the collected input data are then transformed and resampled into 100m gridded raster layers, which was deemed as suitable units for this intra-city level study, also in line with other common global gridded datasets, e.g., WorldPop (Bondarenko et al., 2020). In the unsupervised learning process, principal component analysis (PCA) is applied to the deprivation indicators to generate a set of data-driven continuous indices indicating the degrees of multiple deprivation (i.e., the ‘*multi-deprivation portfolio*’); next, the results from PCA are iteratively refined by inspecting the PCA-model statistic metrics and validated/cross-discussed with previous studies and local knowledge.

Afterwards, the final ‘*multi-deprivation portfolio*’ is used as ‘labels’, combined with the VHR EO images to train a deep CNN-based regression model, aiming to estimate the deprivation degree from imagery features. Once trained, the CNN model will be applied to the whole urban area of Nairobi to predict the intra-urban degrees of deprivation. Finally, the gridded outputs will be validated and post-processed to discuss the utility of using an EO-based method to predict intra-urban deprivation degrees.

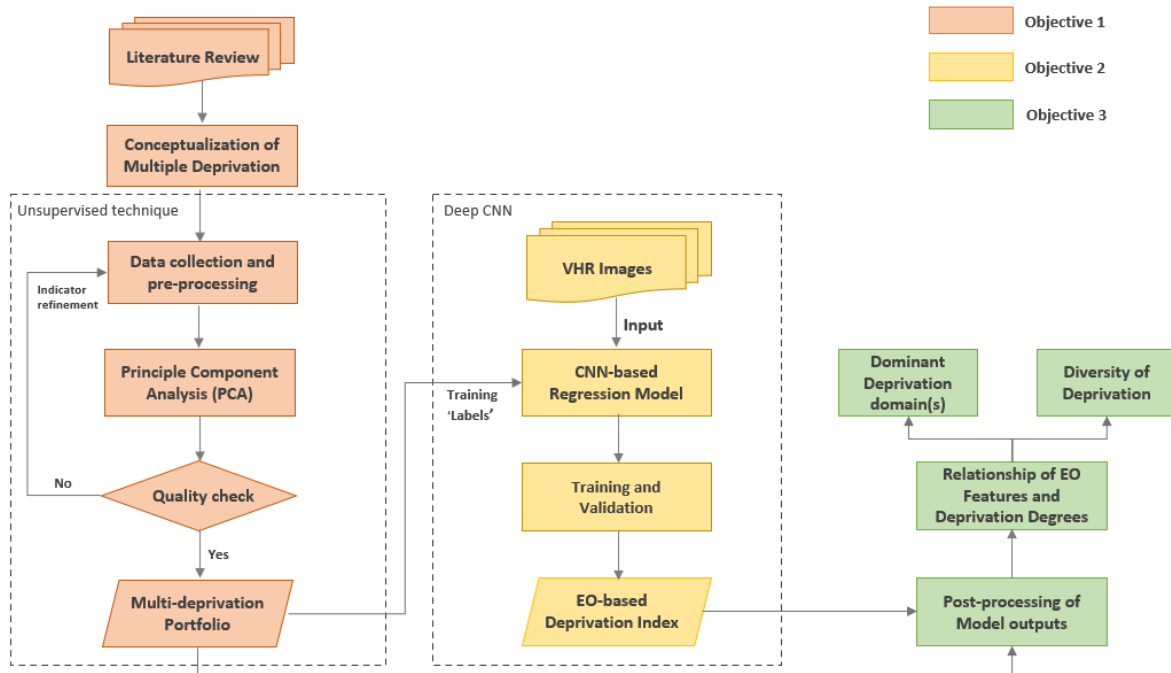


Figure 5 – The overall methodology of mapping multiple-deprivation degrees.

3.3. Data

In this research, the input data are divided into three major types: (1) deprivation-related spatial covariates data for characterising the ‘*multi-deprivation portfolio*’; (2) EO-based VHR satellite images for predicting the ‘*multi-derivation portfolio*’; and (3) the auxiliary datasets for the validation and discussion of the results.

3.3.1. Deprivation covariates data

3.3.1.1. Data requirements

In this study, considering the goal of building a continuous deprivation index at the intra-city level, a list of requirements was employed in the inclusion of candidate datasets. They are: (1) the data must be openly available so that this method could be replicated across other LMICs cities; (2) the data type should be numerical or quantifiable since PCA could only be applied on continuous numerical data; (3) the data must be spatial, i.e., the data format should be raster or could be rasterized in a sensible way; (4) the data should not be highly aggregated (e.g. census or survey data are usually collected at administrative units level), as the output index is set at 100m resolution; (5) the data needs to cover across the entire study area (Nairobi); (6) the data should fit in the local contexts of deprivation, in other words, the data should be representative of deprivation in the study area.

3.3.1.2. Candidate data collection

Recently, the Integrated Deprived Area Mapping System (IDEAMAPS) Network project has published a comprehensive, up to date, and complete scoping review based on existing deprivation studies, unpacking the multi-dimensionality of deprivation into three levels and nine sub-domains (Abascal et al., 2021). Therefore, the research adopts this conceptual framework (Figure 6) of multiple deprivation from the IDEAMAPS project as a reference to guide the searching and collection of all possible indicators for measuring the deprivation degrees.

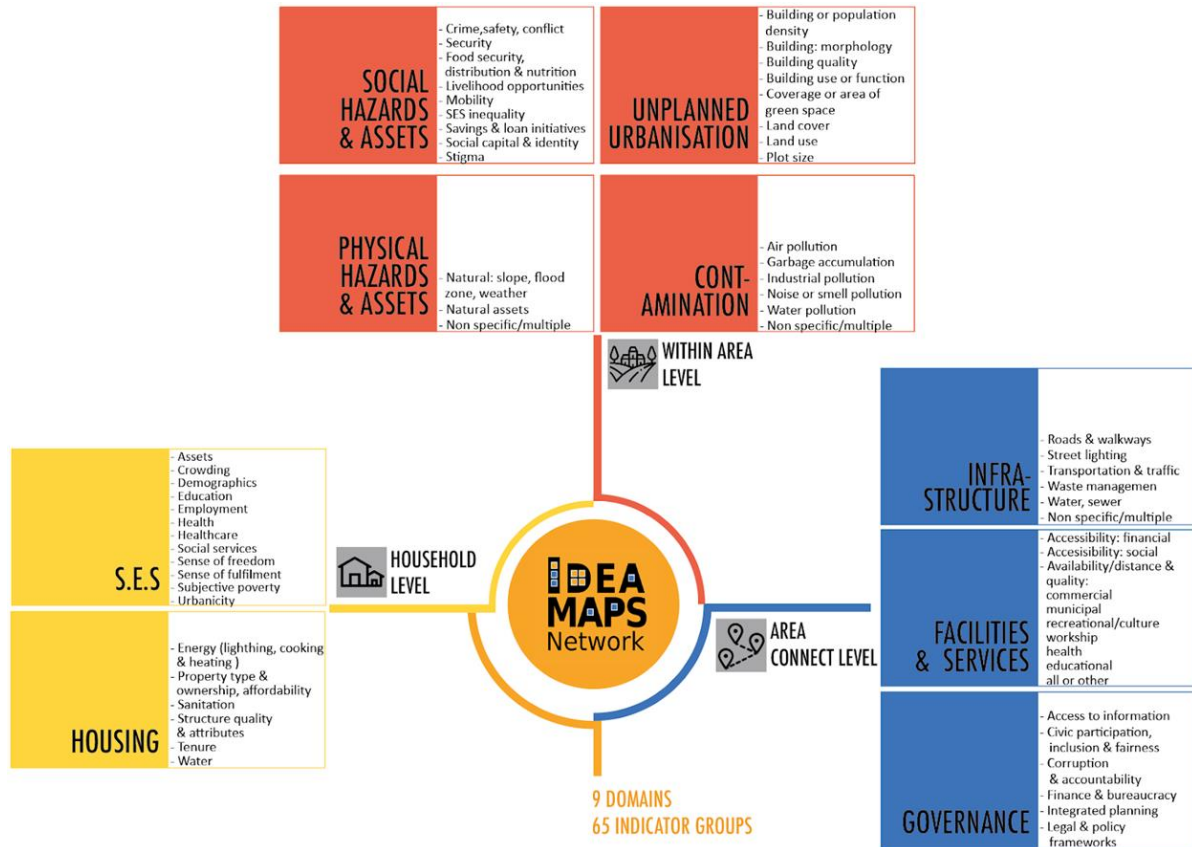


Figure 6 - The domains of multiple deprivation. Source: (Abascal et al., 2021)

Based on the adopted framework and the data requirements mentioned above, in total, 27 candidate indicators were preliminarily extracted from open databases. At least one indicator was collected to cover the sub-domains of multiple deprivation, except for the ones from physical hazards and assets, and governance domains. Indicators of physical hazards and assets are omitted, given limited data availability and local contexts, e.g., the slope indicator was discarded because the slum distribution is not really correlated with the slope in Nairobi. The exclusion of governance indicators is due to the nature of data and its limited availability - such data are usually qualitative and broad, which makes them hard to converted into the spatial format; also, they are more suitable for inter-city comparison. *Table 1* provides a summary of all the 26 candidate indicators, including the data description, justification for adopting the indicator in characterising deprivation, its effect on deprivation, i.e., whether the increase of the indicator value would contribute to or reduce the deprivation degree, and the original format and sources.

Table 1 – The list of candidate indicators for deprivation index formulation.

Candidate indicator	Effect	Description (year)	Rationale/Hypothesis	Original format	Data source
<i>Household Socio-economic Status</i>					
Skilled birth attendance	-	The estimated possibility of receiving skilled birth attendants during delivery. (2014)	Higher percent of receiving skilled birth attendants indicates less deprivation in maternal health care.	Raster tiff (~300m)	WorldPop Development and Health Indicators (Ruktanonchai et al., 2016)
Poverty	+	Estimated proportion of people per grid living in poverty, as defined by the Multidimensional Poverty Index. (2008)	Poverty rate is a key indicator in measuring deprivation. High poverty rates directly indicate more serious deprivation levels.	Raster tiff (~1000m)	WorldPop Development and Health Indicators (Tatem et al., 2013)
Female literacy	-	Estimated percentage of women aged 15–49 who are literate. (2014)	If more women are literate, the less deprivation in education level.	Raster tiff (~4700m)	The DHS model Surface (Burgert-Brucker et al., 2018)
Male literacy	-	Estimated percentage of men aged 15–49 who are literate. (2014)	If more men are literate, the less deprivation in education level.	Raster tiff (~4700m)	The DHS model Surface (Burgert-Brucker et al., 2018)
DT3 vaccination	-	Estimated percentage of children 12–23 months received a third dose of DPT vaccination. (2014)	The vaccination rates indicate the level of primary health care coverage for newborn children.	Raster tiff (~4700m)	The DHS model Surface (Burgert-Brucker et al., 2018)
Access to Insecticide-Treated Net (ITN)	-	Estimated percentage of the de facto household population who could sleep under an ITN if each ITN in the household were used by up to two people. (2014)	The use of ITN reduces the risk of malaria illness and severe disease caused by insects.	Raster tiff (~4700m)	The DHS model Surface (Burgert-Brucker et al., 2018)
Stunted Children	+	Estimated percentage of children under age 5 years stunted (below –2 SD of height–for–age according to the WHO standard). (2014)	The growth status of children indicates the level of deprivation.	Raster tiff (~4700m)	The DHS model Surface (Burgert-Brucker et al., 2018)
Unmet family planning	+	Estimated percentage of currently married or in-union women with an unmet need for family planning. (2014)	The unmet need for family planning contributes to deprivation faced by the household.	Raster tiff (~4700m)	The DHS model Surface (Burgert-Brucker et al., 2018)

Candidate indicator	Effect	Description (year)	Rationale/Hypothesis	Original format	Data source
<i>Housing</i>					
Improved housing	-	Estimated prevalence of improved housing (with improved water and sanitation, sufficient living area and durable construction). (2015)	Access to improved housing conditions reduces the level of deprivation.	Raster tiff (~4700m)	(Tusting et al., 2019)
Improved water source	-	Estimated percentage of the de jure population living in households whose main source of drinking water is an improved source. (2014)	With access to better water source, the deprivation level decreases.	Raster tiff (~4700m)	The DHS model Surface (Burgert-Brucker et al., 2018)
Open defecation	+	Estimated percentage of the population living in households using open defecation. (2014)	Households using open defecation are more deprived of sanitation.	Raster tiff (~4700m)	The DHS model Surface (Burgert-Brucker et al., 2018)
Pit latrines	+	Kernel density of the pit latrine locations in Nairobi, generated by a bandwidth of 1000m. (2015)	Households using pit latrines as defecation facilities are more deprived of sanitation.	Point vector	(Mahabir et al., 2020)
<i>Social Hazards and Assets</i>					
Armed conflicts	+	Kernel density of reported armed conflicts occurred in Nairobi, generated by a bandwidth of 1000m. (2019)	If an area is more exposed to armed conflicts, it is more deprived of the security level.	Point vector	The Armed Conflict Location Events Dataset project (ACLED, 2020)
<i>Contamination</i>					
PM 2.5	+	The annual concentrations (micrograms per cubic meter) of ground-level fine particulate matter (PM2.5) in Nairobi, with dust and sea salt removed. (2016)	High concentrations of PM 2.5 reduces the air quality.	Raster tiff (~1000m)	NASA SEDAC (Van Donkelaar et al., 2016)
Density of waterways	+	Kernel density of OSM waterways (river, stream, canal) in Nairobi, generated by a bandwidth of 1000m. (2020)	The water quality in urban Nairobi is heavily polluted due to the increased discharge of industrial, commercial, and domestic effluents (Mulei, 2012). Thereby, the proximity to rivers indicates deprivation in water quality.	Polyline vector	Open Street Map
Illegal dump sites	+	Kernel density of illegal trash dump sites in Nairobi, generated by a bandwidth of 1000m. (2017)	The presence of unplanned dump sites reflects poor waste management.	Point vector	(Ogutu et al., 2019)

Candidate indicator	Effect	Description (year)	Rationale/Hypothesis	Original format	Data source
<i>Unplanned Urbanization</i>					
Population density	+	Estimated population density per pixel at 100m resolution, adjusted to match the corresponding UNPD 2020 estimate of Nairobi. (2020)	Highly populated areas tend to be more deprived due to lack of living spaces and over-crowdedness.	Raster tiff (~100m)	WorldPop Grid Population Counts (Bondarenko et al., 2020)
Building density	+	The number of buildings per grid cell area in m ² . (2020)	High building density usually indicates the lack of sufficient living spaces.	Raster tiff (~100m)	WorldPop Urban Building Footprint (Dooley et al., 2020)
NDVI	-	Maximum Normalized difference vegetation index (NDVI) in Nairobi. (2019)	Low NDVI value indicates the absence of green space, which increases the deprivation in living environmental quality.	Raster tiff (~30m)	Climate Engine (Huntington et al., 2017)
<i>Infrastructure</i>					
Night-time light	-	The VIIRS night-time light value of Nairobi. (2016)	The presence of electrical lighting on the earth surface indicates street light infrastructure, also widely used as a proxy of income, development.	Raster tiff (~100m)	WorldPop Covariates (Lloyd et al., 2019)
Density of bus stations	-	Kernel density of informal bus stations in Nairobi, generated by a bandwidth of 500m. (2019)	High accessibility to public transport indicates less deprivation level.	Point vector	The DHS model Surface (Burgert-Brucker et al., 2018)
Distance to major roads	+	Distance of per grid centroid to the nearest OSM major road. (2020)	Proximity to major roads indicates less deprivation level in terms of access to transport.	Polyline vector	(OpenStreetMap, 2020)
Poor quality roads	+	Kernel density of road with poor quality (i.e., unpaved, unsurfaced, track, path, footpath, grave) in Nairobi, generated by a bandwidth of 1000m. (2020)	Areas dominated by roads with poor quality are deprived in transport infrastructure conditions.	Polyline vector	(OpenStreetMap, 2020)
<i>Facilities and Services</i>					
Distance to education facilities	+	Distance of per grid centroid to the nearest education facilities (schools, college, university etc.). (2020)	Areas with higher distance to education facilities are more deprived in accessibility.	Point vector	Humanitarian OpenStreetMap Team (HOTOSM, 2020)
Distance to health facilities	+	Distance of per grid centroid to the nearest health facilities (hospital, clinic, pharmacy etc.). (2020)	Areas with higher distance to health facilities are more deprived in accessibility.	Point vector	Humanitarian OpenStreetMap Team (HOTOSM, 2020)
Distance to financial facilities	+	Distance of per grid centroid to the nearest financial facilities (bank, ATM, bureau de change etc.).	Areas with higher distance to financial facilities are more deprived in accessibility.	Point vector	Humanitarian OpenStreetMap Team (HOTOSM, 2020)

3.3.2. Deprivation covariates data pre-processing

After establishing the pool of candidate indicators, several pre-processing procedures were conducted to prepare the input for the PCA model¹. The pre-processing steps include (1) resampling of the raster indicators, (2) rasterization of the vector layers and (3) Z-score standardization for all the pre-processed inputs. The following sections introduce the detailed procedures in each pre-processing task.

First, all the indicator layers with original formats in raster are reprojected to the Universal Transverse Mercator (UTM) Zone 37S coordinate system, the local coordinate system of Kenya, and then resampled into the standard 100m resolution grid. The resampling method used in the data preparation is ‘Cubic’ in ArcGIS, as recommended for interpolating continuous pixel values. The analysis scale for this research is set at 100m grid for four primary reasons: (1) 100m grid is the second finest resolution from all the input data, with only one indicator (i.e., NDVI) at 30m resolution. Thus, using 100m resolution would retain most of the information variance and richness from the datasets. (2) The 100m standard resolution is geographically coherent with existing global gridded datasets (e.g., WorldPop database), which may support the research approach and outputs to be linked, compared, and incorporated into other global studies/datasets or even generalised across cities in the future. (3) Considering this research is designed at an intra-city level, the size of analysing units should not be too large that masks off the diversity and gradient pattern of deprivation, nor it should be too small as it is usually costly to retrieve high resolution data, let alone its limited availability; (4) Last but not least, producing a deprivation degree map in high resolution would naturally highlight the most deprived areas, which usually scatter in and around the slum neighbourhoods. Even without any intention, in principle, such high-resolution deprivation mapping products may increase the stigmatization of slum dwellers.

Second, the vector datasets were converted into raster format at 100m grid resolution. For the rasterization of all indicators related to facilities and services domain as well as the indicator of major roads, Euclidean distance to the nearest facility was calculated in kilometres. The rest of the point and polyline vectors (except for the bus station point layer) were rasterized by calculating the kernel density with a bandwidth of 1 km, following the same setting of research conducted by Mahabir et al. (2020) about mapping slums in Nairobi. For the indicator of bus stations, 500m bandwidth was used to calculate the kernel density, considering it as a common walking distance to the bus station with good accessibility (Daniels & Mulley, 2013).

Finally, all the raster layers were standardized before entering the PCA model, given the large discrepancy of scale and measuring unit between individual indicators. PCA is well known as highly dependent on input scale because the input features with large variance and magnitude tend to dominate the retained components, thus leading to bias on the results (Baxter, 1995). To avoid so, the standardisation is achieved through rescaling the input features based on the standard score (Z-score) to obtain a normal distribution with a mean of zero and a standard deviation of one. The standardization was conducted using the following formula (Eq. 1).

$$x_{std} = \frac{x - \mu}{\sigma} \quad (1)$$

Where x_{std} is the standardized value of the observed indicators; x is the raw value of indicators; μ and σ refer to the mean and the standard deviation of the sample values, respectively.

¹ The pre-processing step was conducted in Python via Jupyter Notebook. The script is openly available at Github for the readers' reference (https://github.com/EquiLuo/MappingDeprivation/tree/main/PCA_Process).

3.3.3. VHR satellite image data

The available VHR EO imagery data for this research include Worldview-3 satellite image, and the SPOT-7 satellite image of Nairobi (Table 2), acquired via the collaboration with an ongoing project – SLUMAP (Remote Sensing for Slum Mapping and Characterization in sub-Saharan African Cities) (SLUMAP, 2020). Both of the VHR images cover the entire urban areas of Nairobi and can provide a very clear visual difference between slums and non-slum settlements even from just inspecting the imagery features.

Table 2 – The summary of available VHR satellite images of Nairobi for this study.

Name	Resolution	Producer	Year	Cloud coverage
Worldview-3	0.3m at panchromatic band. 1.2m at 8 multispectral bands: red, red edge, coastal, blue, green, yellow, near-IR1 and near-IR2	Maxar DigitalGlobe	2019	Several cloud patches scattered in the deprived areas in Nairobi.
SPOT-7	1.5m at 4 bands: red, green, blue and near-infrared.	Intellegience Airbus	2017	Very less

Although the Worldview-3 datasets have higher spatial resolution and more spectral bands, the SPOT-7 satellite images were selected as the final inputs for the CNN model due to the presence of several clusters of cloud in the Worldview-3 images, which will naturally introduce undesirable noises to the deep learning model. Figure 7 shows a snapshot zooming into one of the biggest slums – Kibera in Nairobi. In the left graph of Worldview-3, it could be clearly observed that a few cloud patches block the deprived settlements within Kibera, coupled with some areas darkened by the shadows. While for the SPOT-7 images in the right figure, all the physical morphology information is well retained without any cloud coverage. Most importantly, a well-trained CNN model requires the input data to contain as much as fewer noises, especially regarding the key target variables – in this case, the delineated slum extents, because they are the most direct morphological appearance of deprivation based on prior knowledge. In addition, removing the cloud coverage from EO images and filling up the missing information is often difficult to address, and inappropriate interpolation method may bring in extra noises to the model input, let alone this MSc research only has limited time availability.

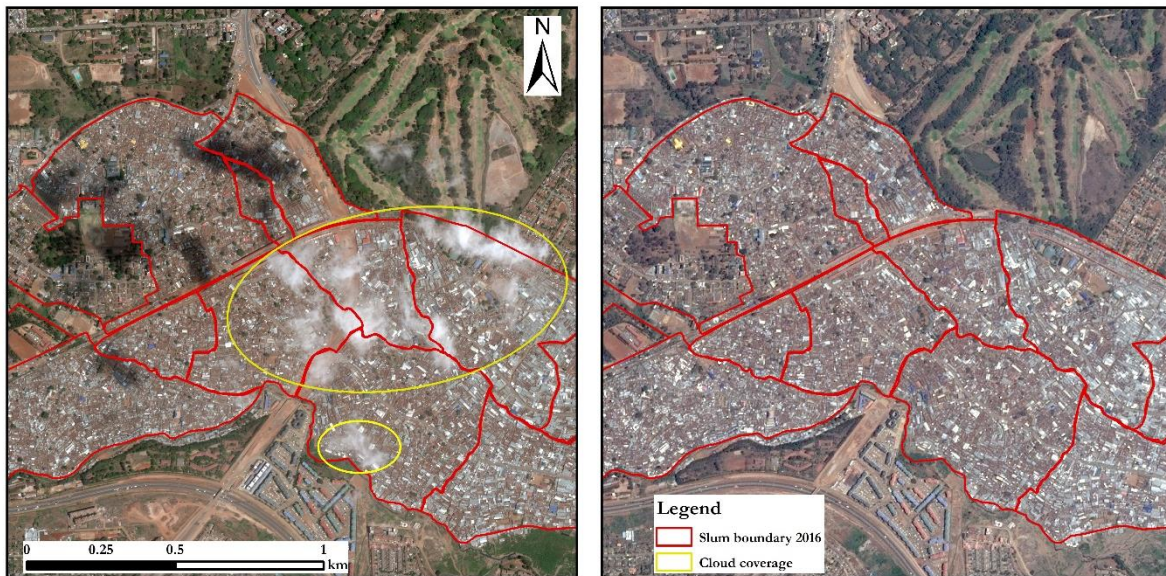


Figure 7 – The deprived areas shown in VHR images. (a) Worldview-3, 2019. (b) SPOT-7, 2017.

Indeed, it is agreed that the richness of EO information would be slightly reduced by adopting the 1.5m data from SPOT-7, rather than the 1.2m image of Worldview-3. However, the trade-off of compromising on spatial resolution is still reasonable to avoid bringing extra noises and bias to the model and to prioritize that the research problems and objectives could be first systematically explored.

3.3.4. Auxiliary data

In this research, several auxiliary datasets were used to provide detailed information about the contexts and help to support the validation and cross-discussion of the model outputs. The word ‘auxiliary’ here emphasize the supplementary and referencing functions of these datasets, although they are not directly used as raw inputs for the models in this study. The following sections briefly introduce the supplementary datasets and their benefits brought into this study.

One of the most important datasets is the reference slum boundaries of Nairobi (see *Figure 4*). This is because, unlike traditional classification problem where the outputs can be directly validated with the ground-truth label, in this study, the characterization of multiple deprivations is designed to be measured in continuous values. Nonetheless, in theory, it should be foreseeable that the areas with a high estimate of deprivation degree would somewhat correlate (more specifically, geographically overlapping) with the slum region. Therefore, in the absence of ‘continuous labels’ for deprivation in Nairobi, the previously delineated slum extents become compulsory for supporting the validation of model outputs. The reference slum boundaries were collaboratively delineated by the local communities and NGOs in 2016, under the guidance and support from Slum Dwellers International (SDI). These established slum maps were also further validated by the local experts and hence should be regarded as reliable ground truth data. In total, there are 161 delineated slum polygons, accounting for 11.15 km².

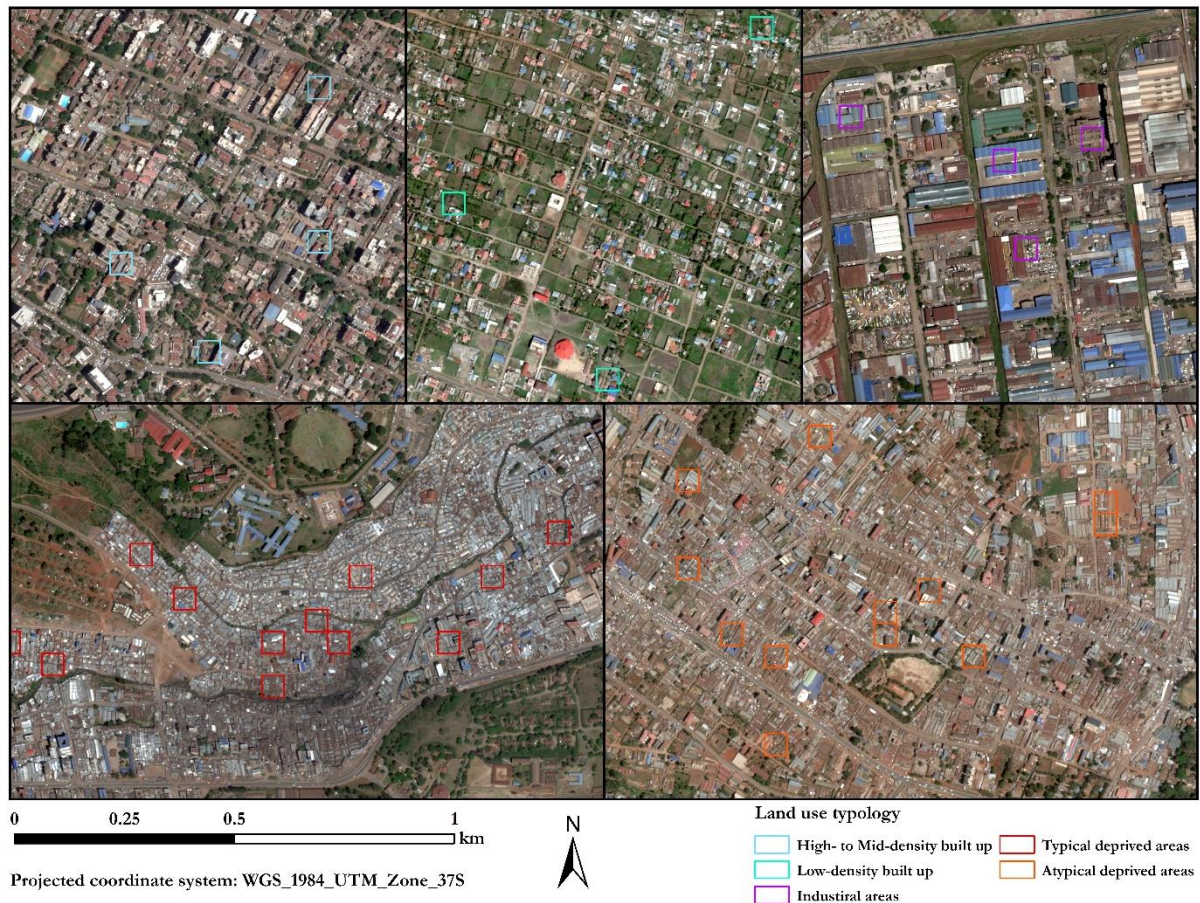


Figure 8 - The land use typologies of urban areas in Nairobi, 2020. Base map: Worldview-3. Source: (Vanbrysse et al., 2021)

In addition to the slum boundaries, this research reaches out to a very recent output produced by the SLUMAP team (Vanhuyse et al., 2021), which contains more detailed information on land-use/land-cover (LULC) typologies within Nairobi. This dataset collected 500 in-situ samples at 50m grid and classified the land use into different typologies, including high- to mid-density built up, low-density built up, industrial area and urban deprived areas (Figure 8). The major highlight of the SLUMAP dataset is that it further divides the deprived areas in Nairobi into two sub-classes, namely, typical deprived urban areas and atypical deprived urban areas (Vanhuyse et al., 2021). As shown in Figure 8, the typical deprived areas are usually dominated by highly dense, compact clusters of small settlements with irregular road networks; on the other hand, the atypical deprived areas accommodate compact to mid-dense building clusters, and the size of buildings are usually slightly bigger than the typical deprived settlements, coupled with a bit more organized road networks. In general, introducing this dataset would enrich the insights and understanding from this research and help to validate the results.

Lastly, the Kenya Settlements Extents Database, produced by the Geo-Referenced Infrastructure and Demographic Data for Development (GRID3) initiative (CIESIN& Novel-T, 2020), was also included as an auxiliary dataset. This dataset contains the spatial boundary for three types of settlements, i.e., built-up areas, small settlements areas and hamlets (Figure 9). These extents are automatically derived from the Maxar/ECOPia building footprint (Maxar, 2021) and further classified based on the building density of the settlement agglomerations. This dataset provides a clear classification of built-up and non-built-up for the city of Nairobi and thus helps to confirm which types of land use should be included for this analysis.

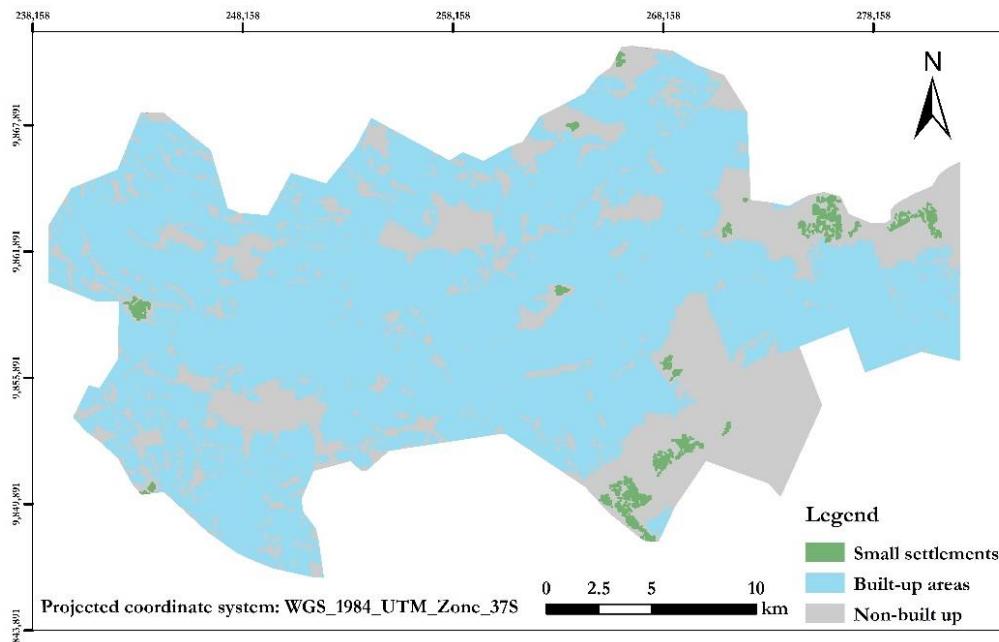


Figure 9 - The settlements extents of Nairobi. Source: (CIESIN & Novel-T, 2020)

3.4. Principal component analysis

Principal component analysis (PCA), as one of the most popular multivariate statistic methods, was first invented by Pearson (1901) and further elaborated by Hotelling (1933). It intends to derive a set of uncorrelated principal components (PCs) through an orthogonal transformation of the original variables that enables to retain most of the variance from the original datasets (Jolliffe & Cadima, 2016). Due to its outstanding performance in transforming data with large dimensionality while still minimising the information loss, PCA has been extensively applied by researchers to investigate multiple deprivation, specifically in identifying its variability and formulating the multiple deprivation index, ranging from

neighbourhood to area scale (Aungkulanon et al., 2017; Basu & Das, 2020; Mishra, 2018). Additionally, PCA is a non-parametric, automated statistic approach, independent from any presumption of the probability distribution for the original datasets, which makes it relatively easy and straightforward to implement (Abdi & Williams, 2010). In this sense, PCA could be considered as an unsupervised learning technique since it automatically explores the similarities and differences among the variables within the datasets and then derives the PCs that together explain the maximum variance (Hastie et al., 2009). Lastly, another advantage of PCA is avoiding the subjectivity from manually or semi-manually assigning weights (there are always personal judgements involved, which are inconsistent across people and contexts) to different variables when aggregate the deprivation indicators into a composite index (Deas et al., 2003). Given above all considerations, PCA is chosen to generate the multiple deprivation indices.

In this analysis, all the 26 collected indicators (refer to *Table 1*) are firstly used as inputs for running the PCA. The inputs are standardized before entering the model to avoid bias from varying measurement units. The PCA was conducted via the software Statistical Package for the Social Sciences (SPSS). And prior to the application, several assumptions related to the fitness of the data on the PCA model need to be fulfilled. The Pearson correlation coefficient matrix was calculated to make sure that highly correlated (correlation coefficient $p > 0.8$) pairs of indicators were not included and at least two or more correlation coefficients are above 0.3 or more (Tabachnick & Fidell, 2007). The Kaiser-Mayer-Olkin test (KMO) assesses the sampling adequacy of the data and to what degrees does the input fit the PCA model (Hutcheson & Sofroniou, 1999; Kaiser, 1960). The KMO value is suggested to be more than 0.6 as an acceptable threshold (Hair, 2006). Lastly, Bartlett's test of sphericity was checked as statistically significant, with returned value $p \leq 0.001$. All the 26 candidate indicators satisfied the above criteria and thus intactly retained for the PCA analysis.

3.4.1. PCA-based index construction

To achieve a comprehensive, transparent and quantitative insight towards the diversity and variation of multiple deprivation in Nairobi, this study implements a PCA-based technique to formulate a set of continuous deprivation indices – called the '*multi-deprivation portfolio*'. The calculation method of the multi-deprivation indices is mainly based upon the 'SoVI recipe' (Cutter et al., 2003), added with slight modification inspired by the work from (Abson et al., 2012). Here, the word 'SoVI' refers to 'Social Vulnerability Index', and the 'SoVI recipe' is a classical index aggregation approach widely used in social vulnerability studies (Sherbin & Bardy, 2016; Fekete, 2009; Frigerio et al., 2016; Jankowska et al., 2011). Although the target variable here has changed from 'social vulnerability' to 'multiple deprivation', it is still sound and reasonable to transfer this 'SoVI recipe' as a manual to calculate the aggregate deprivation index, because, essentially, both of them are multi-dimensional, complex concepts that indicate the status of people being deprived or disadvantaged in one or several domains (Whelan & Maître, 2005). Nonetheless, it is not enough to just provide a single aggregate index, but more importantly, to capture the independent, unique facets of multiple deprivation. As such, the individual, explicit sub-dimension (i.e., the PC scores) of multiple deprivation are also retained and quantified as the sub-dimensional deprivation indices (Abson et al., 2012). The expected '*multi-deprivation portfolio*' contains two products: an aggregate multiple deprivation index and a set of indices measuring the deprivation degree in individual sub-domains. To conclude, this tailored PCA-based approach enables to not only present a summarized final score of the multiple deprivation level but also deconstruct the multi-dimensionality and complexity of deprivation into individual, less inter-correlated but more pure facets of multi-deprivation. The remaining sections would articulate the calculation procedures in detail.

Suppose there is a dataset with n samples and p variables, which is equivalent to a $n \times p$ matrix X . The matrix consists of n p -dimensional vectors, denoted as $X = (X_1, X_2, \dots, X_p)$ (Jolliffe & Cadima, 2016). The PCA aims to derive a set of linear combinations of the original vectors (i.e., the columns of the matrix

X) that can maximize the remained variance as much as possible. In essence, this linear combination of original variables is what usually referred to as the ‘Principal Component’. For example, the first component could be denoted via the following formula (Eq. (2)):

$$PC_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \quad (2)$$

where PC_i , subscripted by i ($i \leq p$) to indicate the order of the PCs (in this example, $i = 1$), is the i_{th} PC scores; a ($a_{i1}, a_{i2}, \dots, a_{ip}$) is a set of PC-specific loadings employed on the corresponding variables to calculate the linear sum. The first PC is calculated in a way to make itself accounting for the maximum variance from the original datasets, followed by the second PC explaining the largest share of the remaining variance, and so on.

One important parameter to decide in PCA analysis is the number of retained PCs, because it directly relates to the total variance explained. This study follows the classic rule of thumb proposed by Kaiser (1960) that if the Eigenvalue of a PC is greater than 1, it will be retained, combined with the scree plot of the Eigenvalue. Next, the extracted PCs are rotated using the varimax orthogonal rotation method such that the rotated PCs are uncorrelated among each other. This ‘uncorrelatedness’ is extremely significant because it ensures that each extracted PC measures the different dimension of deprivation. Meanwhile, the rotated component matrix (RCM) is generated to help the researchers interpret the components by identifying the dominant variables (i.e., variables with high factor loadings) within that component. In other words, the sub-domain/facet of multiple deprivation reflected by each PC is determined by the significant loadings of the deprivation indicators on that component (Abson et al., 2012). Additionally, all the sub-dimensional deprivation indices are normalized to the scale between 0 to 1, where 0 represents the least deprived and 1 indicate the highest level of deprivation.

As mentioned before, this study also seeks to generate an overall aggregate measure of the multiple deprivation degrees, so the retained PC scores (not normalized) are ultimately combined together to calculate the aggregate multiple deprivation index. The aggregation approach follows the ‘SoVI recipe’ and is denoted in Eq. (3):

$$AMD I = \sum_{i=1}^q \sigma * PC_i \quad (3)$$

Where the Aggregate Multiple Deprivation Index is denoted as $AMD I$; q is the number of retained PCs from the previous analysis; σ is the directional adjustment coefficient, either equal to -1 (negative directionality) or +1 (positive directionality), dependent on the influence of the PC to the deprivation.

More specifically, the positive directionality (+1) would be employed on those PCs that contribute to the multiple deprivation level, i.e., the increase of the values in those indicators would lead to the increase of deprivation degree. Note that just adding the directional adjustment does not import any weight on the PCs; instead, in this formula, each PC is regarded as equally important in contributing to the final deprivation score. This equal weight aggregation method is chosen due to the absence of *a priori* knowledge about the relative importance of different deprivation sub-domains in Nairobi. In fact, such knowledge can only be generated via very detailed field surveys and require validation from local experts’ judgements (Schmidtlein et al., 2008). Yet, given the COVID-19 pandemic outbreak, any field data collection with physical contact with people is hardly possible. Therefore, this research decides to give each PC equal weight as the best option in order to avoid potential bias from subjectively assigning weights but also ensure transparency and interpretability of the results. In addition, the $AMD I$ value is also normalized into the range between 0 to 1 due to the same reason given above.

3.4.2. PCA-based index refinement

All the 26 candidate indicators were first included in the PCA analysis to generate preliminary results. Then, the results are statistically checked based on the data suitability criteria stated above and also geographically visualized in the 100m standard grids to be inspected on the VHR satellite images and slum boundaries. In addition, the data fitness was carefully checked with a specific focus on ‘to what extent is the indicator appropriate for this intra-urban 100m grid analysis?’, as some of the original datasets have a relatively coarse resolution. The refinement was running iteratively until the result fulfils the quality requirements and its spatial patterns stay geographically inherent and stable with the auxiliary datasets.

In the beginning, all the indicators were collected to cover the entire urban areas (i.e., contain both the urban built-up and non-built-up) of Nairobi, as the research aims to provide a complete measure for the intra-urban deprivation degrees. However, during the PCA analysis, the researchers realise that, by nature, the characteristics/intrinsic patterns of multiple deprivation vary substantially between urban built-up and urban non-built-up, let alone there is large heterogeneity even within the deprived built-up regions (Kuffer et al., 2017). For instance, common indicators for measuring the social-economic deprivation (e.g., population density, income level, food expenditure) or morphological deprivation (e.g., building density, building material, structure) would become naturally unapplicable and meaningless if applied to non-built-up areas due to the absence of settlements and residents. As such, including the non-built-up areas in this analysis may introduce extra noises and inconsistency to the PCA model so that the important patterns of multiple deprivation on urban built-up cannot be recognized and captured. Additionally, it should be emphasized again that, in this context, one of the fundamental objectives is to not only provide a mapping product of multiple deprivation level but also know how to use it in order to better inform the local governments and stakeholders in helping the people in need. Given the above, this research decided to only focus on the multiple deprivation of urban built-up areas – where the people in need live in. Additionally, several candidate indicators were finally discarded to avoid potential bias on the results. A summary of the discarded indicators and the justification is provided in the results section 4.1.

3.4.3. PCA-based index validation

As introduced before, PCA is essentially an unsupervised learning technique from which the results are solely data driven. In this case, the magnitude and number of input indicators decide how much information can actually be drawn to capture the multi-dimensionality and variety of the deprivation level in Nairobi. Also, the quality of input data directly influences the reliability and quality of the outputs. Therefore, it is important to validate the ‘*multi-deprivation portfolio*’ using independently verified datasets and also compare it with local knowledge. The validation process includes two major parts: (1) statistical analysis combined with visual assessments; (2) local expert validation.

Firstly, the auxiliary datasets are used for PCA results validation and/or cross-discussion via performing some basic exploratory statistical analysis. The major datasets used for validation are the slum boundaries and the land use typologies from SLUMAP project. Moreover, a detailed visual assessment is applied on the VHR satellite images and street-view images (Mapillary, 2021) for better comparing the PCA results with the ground truth and providing a more ‘*in-situ*’ understanding of deprivation in Nairobi.

Secondly, a local validation is also conducted via an online semi-structured interview with the local experts. The local experts are selected from researchers with experience in geo-spatial data analysis and adequate knowledge of the local deprivation characteristics in Nairobi to provide more detailed, contextual insights and potential improvements for this approach. During the one-hour semi-structured interview, the presenter first shows the mapping products of ‘*multi-deprivation portfolio*’, ask several fixed questions (e.g., “Is the deprivation pattern in this map reflect the local situation?”), and lead to open discussion. The key feedback, suggestions and comments from the interview are noted for the results and discussion.

3.5. Deep CNN-based regression model

As described in the overall methodology (Figure 1), after establishing the ‘multi-deprivation portfolio’ from PCA using all the selected indicators, a CNN-based regression model would be trained to directly predict those PCA-based indices using only EO data (Figure 10). All of the output indices from PCA are at the standard 100m grid. Hence, the PCA-based indices (i.e., the ‘multi-deprivation portfolio’) can be considered as a set of measures for the overall level of multiple deprivation within a square of 100m by 100m. In this sense, the task of the CNN-based regression model could be further specified as ‘using EO data to predict the overall level of the multiple deprivation within a 100m grid’.

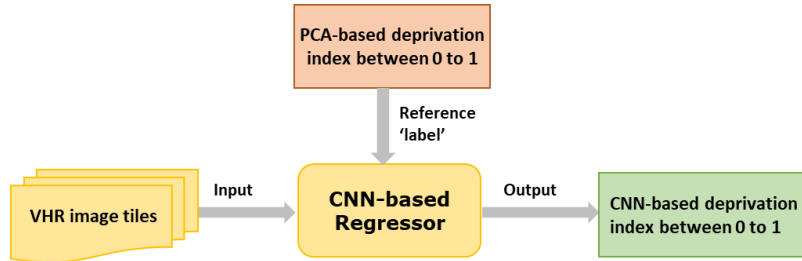


Figure 10 - An illustration showing the inputs and outputs of the CNN-based regression model.

The remaining sections articulate the input preparation and sampling method for this model, the configuration of the proposed CNN regression model, the model optimization, and finally the evaluation metrics for the model performance.

3.5.1. Input preparation

3.5.1.1. Image tiles extraction

Before training the CNN-based regression model, the first thing is to prepare the input images for the model. The selected VHR EO image for the CNN model is SPOT-7, given its very low cloud coverage. The SPOT-7 image has four bands (RGB + near-infrared) with a spatial resolution of 1.5m, covering the entire study area of Nairobi. As specified above that the multiple deprivation degrees were measured at the 100m grid scale, it is natural and intuitive to also use the same size of EO imagery information for predicting this index. Considering this, the SPOT-7 image needs to be clipped into a series of small image tiles, of which the size is geographically equivalent to the 100m standard grid. Additionally, the clipped image tiles are also required to be spatially coincident with the PCA-based indices grids to ensure each of the image tiles corresponding to the correct reference values of multiple deprivation indices.

The stepwise procedures of generating the image tiles are described as follows. Firstly, a polygon vector of 100m-by-100m fishnet was created covering the whole study area and then snapped into the PCA-based indices layer. Next, all the fishnet grids intersected with the non-built-up areas were eliminated. Afterwards, a centroid layer of all the retained fishnet grids was generated. The converted centroid layer is especially significant for extracting the image tiles, because: although the size of the image tile is required to be as much as close to 100m grid, due to the 1.5m resolution of SPOT-7 image, it is hardly possible to clip an image tile of the exact same size (since 100m cannot be divided by 1.5m with no remainders). Additionally, a standard CNN model also requires a fixed input size of a square patch. Therefore, the centroid point was used as an alternative to creating the square sampling area by expanding the pixel in which the centroid locates with 32 strides. So, the size of each sampling area becomes 65 by 65 pixels, which is geographically equal to a 97.5m grid. Next, the square sampling areas were vectorized into ‘sampling grids’ polygon layer (which are exactly in alignment with the raster pixels). Finally, the sampling grids were used to divide the input image tiles for the CNN model.

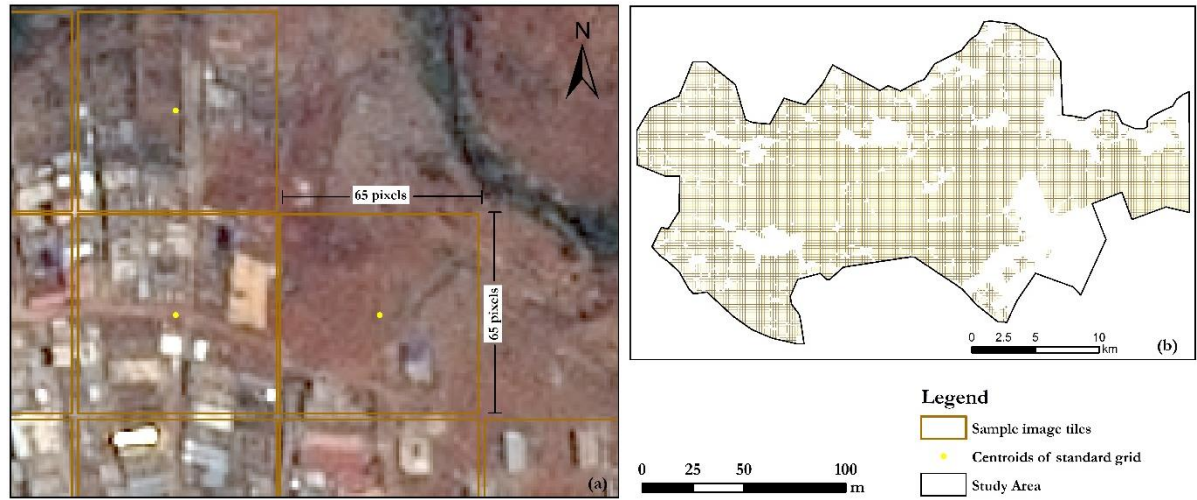


Figure 11 - The extracted image tiles for the CNN-based regression model. (a) examples showing the size of each image tile. (b) the total extracted image tiles for the whole study area.

Although the established size of the image tile (97.5×97.5m) is slightly smaller than the standard 100×100m grid, however, for the ease of data preparation and model implementation, this size is finally employed to divide the VHR image. In total, 45559 image tiles of 65×65 pixels were extracted from the original SPOT-7 image and will be used as inputs for the CNN-based regression model (Figure 11).

3.5.1.2. Sampling of the training, validation, and test dataset

As now the total 45,559 sample sets are well prepared, another crucial preparation task before entering the CNN model is to find a suitable way in dividing the overall samples into training, validation and test datasets. For the sampling process, two primary questions need to be answered: (1) what is a suitable proportion to divide training, validation and test datasets? (2) which sampling method should be applied to ensure that the subsets contain sufficient and comprehensive information from the original data?

There are a lot of ways to divide the total samples into subsets for model implementation. The common rule of thumb follows the proportion such as 70%, 15%, 15% or 80%, 10%, 10% for dividing the training, validation and test dataset, respectively. It is particularly suggested to adopt those ratios when the number of total samples is relatively limited (Xu & Goodacre, 2018). However, given the total number of 45559 samples, which can already be claimed as ‘large geo-dataset’, the training-validation-test splitting ratio does not necessarily need to adhere to the rule of thumb. Instead, this study specifically intends to train the CNN model using relatively small datasets because one of the research motivations lies on that if the model can be trained with only limited samples but still yields fairly accurate result in predicting multiple deprivation degrees, then this advanced approach could be transferred to other LMICs cities to inform local planners of urban deprivation with less cost in preparing training samples. In the end, the authors decide to take 7000 (around 15%) samples for training and validation, while the remaining 38559 samples would not be involved in any steps of the model development but only used to test the CNN model. Within the 7000 samples, 4900 (70%) is used for training, and 2100 (30%) is used for model validation, following the common rule of thumb.

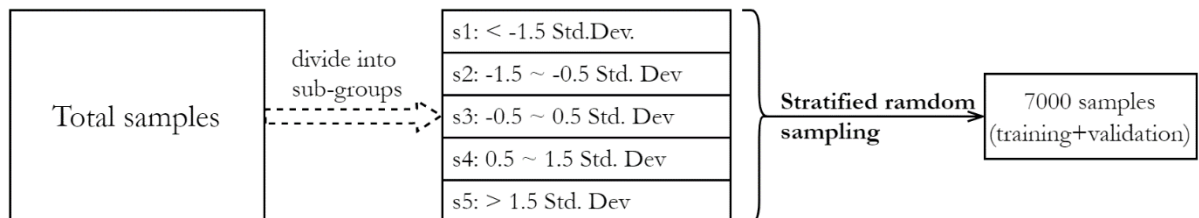


Figure 12 - The sampling approach for extracting the training and validation datasets.

The stratified random sampling method is employed to extract the 7000 samples (Figure 12). Compared to random sampling, stratified random sampling ensures that all the ‘sub-groups’ within the total samples can be well represented by the extracted subsets. In other words, the distribution pattern of deprivation level in the original dataset would still be retained in the extracted samples. The total samples were first divided into 5 sub-groups using the interval of 1 standard deviation so that each sample within the same sub-group shares an identical label. Then, a package called ‘sampling tool’ developed by Buja & Menza (2013) was applied to conduct the stratified sampling based on the previously classified label.

3.5.2. Regression CNN configuration

In this study, a deep CNN-based regression model is proposed for extracting the spatial-contextual features from the VHR image to predict the PCA-based multiple deprivation indices in Nairobi. The proposed CNN regression model is built upon the VGG-16 (Simonyan & Zisserman, 2015) with some modifications to fit in the regression purpose and research context. The most important innovation of this proposed CNN architecture is the insertion of a regression layer (i.e., the Sigmoid activation layer) to replace the conventional ‘SoftMax’ layer in the classification model, such that the outputs would be a set of continuous values between 0 to 1, instead of categorical, discrete labels (Babu et al., 2016).

It is also worth mentioning that the input image size for CNN has changed. Recall that when dividing the image tiles, the individual image size was decided on 65×65 pixels. However, as the proposed model is modified from VGG-16, it usually requires the input size to be an even number. Therefore, the actual size of the input image entering the CNN model is set at 64×64 pixels ($96m \times 96m$ grid), with the last column and row left outside. A more detailed description of the implemented CNN model is provided below.

3.5.2.1. Architecture of the CNN regression model

The overall architecture of the proposed CNN regression model is visualized in Figure 13. In general, this deep regression model consists of two major parts, namely, the feature extraction and the regression, each with several sublayers. A more detailed summary of the model structure is also shown in Table 3.

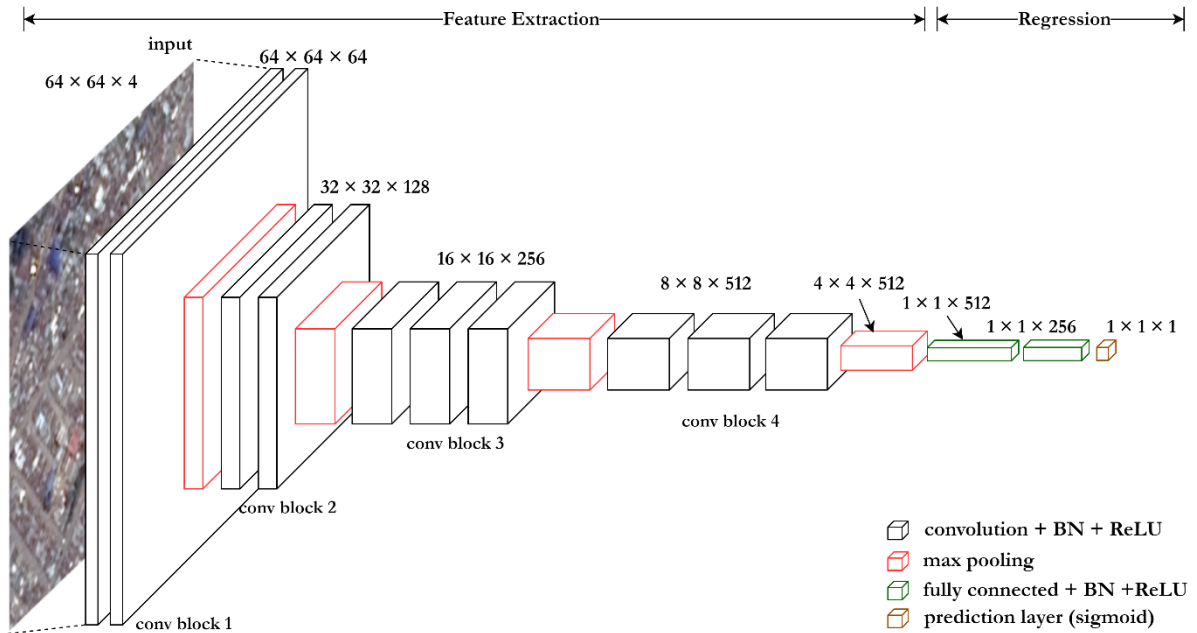


Figure 13 - The overall architecture of the proposed deep CNN-based regression model. BN means batch normalization, ReLU means rectified linear unit activation function.

Table 3 – A detailed summary of the proposed deep CNN-based regression model structure.

Stage	Layer name	Number of filters	Size of feature map (H × W × D)	Filter Size
Conv block 1	Input layer		$64 \times 64 \times 4$	
	Conv1-1	64	$64 \times 64 \times 64$	3×3
	BN + ReLU			
	Conv1-2	64	$64 \times 64 \times 64$	3×3
	BN + ReLU			
	Max Pooling		$32 \times 32 \times 64$	2×2
Conv block 2	Conv2-1	128	$32 \times 32 \times 128$	3×3
	BN + ReLU			
	Conv2-2	128	$32 \times 32 \times 128$	3×3
	BN + ReLU			
	Max Pooling		$16 \times 16 \times 128$	2×2
Conv block 3	Conv3-1	256	$16 \times 16 \times 256$	3×3
	BN + ReLU			
	Conv3-2	256	$16 \times 16 \times 256$	3×3
	BN + ReLU			
	Conv3-3	256	$16 \times 16 \times 256$	3×3
	BN + ReLU			
Conv block 4	Max Pooling		$8 \times 8 \times 256$	2×2
	Conv4-1	512	$8 \times 8 \times 512$	3×3
	BN + ReLU			
	Conv4-2	512	$8 \times 8 \times 512$	3×3
	BN + ReLU			
	Conv4-3	512	$8 \times 8 \times 512$	3×3
FC stage	BN + ReLU			
	Max Pooling		$4 \times 4 \times 512$	2×2
	1st fully connected		512	
	BN + ReLU			
Regression	2nd fully connected		256	
	BN + ReLU			
	Sigmoid activation		1	

In total, 4 convolution blocks comprised of 11 individual convolutional layers are built to extract the features for multi-deprivation index estimation. In this analysis, a ‘convolution block’ is defined as a series of convolutional layers, with each followed by a batch normalization (BN) layer and a rectified linear unit (ReLU) layer and placed with a max-pooling layer at the end of the block. Throughout the coevolution blocks, all the convolutional layers are set at the same filter size of 3×3 to keep the output feature map unchanged. The convolutional layer applies filters (also called kernels) in a moving window fashion to extract the contextual information from the input image. The batch normalization layer is applied right after the convolutional layer and before the activation layer in order to reduce the internal covariance shift and fasten the training speed (Ioffe & Szegedy, 2015). Then, ReLU is chosen as the activation function due to its lower computation cost and the ability to avoid vanishing gradient problems compared to other common activation function like Sigmoid or Tanh (Glorot et al., 2011). At the end of each convolution block, the max-pooling layer is applied to reduce the size of the feature map while still retaining the high-level features (Kalchbrenner et al., 2014). After passing through all the convolutional blocks, the extracted

feature map from the feature extraction phase is flattened and directly fed into the fully connected layers (the FC stage and regression in *Table 3*). In the FC stage, two fully connected hidden layers are applied to further process the feature vectors, with each of the neuron nodes set as 512 and 256, respectively. Finally, the regression layer is applied to predict the multiple-deprivation index. To work as a regressor and produce a continuous value between 0 to 1, the activation function in the final output layer is set as Sigmoid. Thus, a numerical value within the range of 0 to 1, which indicates the level of multiple deprivation degree, would be the final expected output from this proposed deep CNN regression model.

In total, this proposed CNN-based regression model contains 11,969,153 trainable parameters to learn.

3.5.2.2. Model training initialization setting

In this study, considering that the ‘*multi-deprivation portfolio*’ provides more than one measure of the multiple deprivation degrees in Nairobi, thus, before training the CNN model, the ‘target variable’ needs to be decided, i.e., which measure from the ‘*multi-deprivation portfolio*’ should be first be predicted by the model? As a starting point, the authors decide to use the sub-dimensional deprivation index that most related to morphological deprivation as the input ‘label’ to explore the potential of using the EO method in estimating deprivation degree. It is expected that to some extent, the variance of morphology-based deprivation should be partially captured by the proposed CNN regression model, as the input imagery data carries substantial morphology information.

As stated in section 3.5.1.2, in total, 7000 image tiles of 64×64 pixels were selected as the training and validation datasets. Among them, 4900 tiles (70%) are used for model training, whereas the remaining 2100 tiles (30%) for model validation. To initialize the model training process, several hyper-parameters are set based on the knowledge from previous deep regression studies (Ajami et al., 2019; Lathuiliere et al., 2020; Pyo et al., 2019; Ren et al., 2019) and empirical experience. A summary of the hyper-parameter setting for model initialization is provided in *Table 4*.

Table 4 – A summary of the hyper-parameter setting for the model initialization

Hyper-parameter	Value
Optimizer	Adam
Loss function	Mean Squared Error (MSE)
Batch Size	32
Learning rate (lr)	lr = 0.001, with reducing learn rate on plateau.
Epochs	80

Adam is applied as the model optimizer because of its high computational efficiency and generally stable performance on various applications (Kingma & Ba, 2015). The batch size is empirically set at 32. The learning rate is set as 0.001, along with the call back function of ‘reducing learning rate on plateau’ to help the model reach the minima once the learning stagnates. The number of epochs is empirically set at 80. Lastly, the loss function for the model initialization follows the common choice applied in regression tasks, the mean squared error (MSE). The proposed CNN model is trained by minimizing the MSE loss between the regression output y_i and the reference deprivation index produced by PCA \hat{y} . The MSE loss function is defined as:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad \text{Eq. (4)}$$

where N is the number of training samples.

The CNN model is implemented using Keras deep learning library and the Python programming language via Jupyter Notebook ². To address the high computational cost in processing large geo-datasets, the whole CNN simulation is performed in the modelling platform of ITC Centre of Expertise in Big Geodata Science (CRIB, 2021) with the NVIDIA Jetson AGX 8-core GPU and 32 GiB memory, which greatly helps to save the training time and fasten the model optimization. In general, one complete training process takes about 20 minutes, which is not very time-consuming.

3.5.3. Model hyper-parameter optimization

After the first CNN training initialization finished, the model is optimized via fine-tuning some of the hyper-parameters. Table 5 provides a list of the hyper-parameters to be explored for the model optimization. To find the optimal combination in a systematic and efficient way, a stepwise hyper-parameter tuning strategy is established, where each hyper-parameter will be tuned in a specific order while holding the other hyper-parameters still. For example, in this study, the batch size is firstly checked by only changing its values, with the other hyper-parameters following the default setting, and then compare the model performances to select the optimal value. This updated optimal value will be passed onto the next hyper-parameter tuning but staying still to find the optimal value for the second hyper-parameter, and so on.

Table 5 – The list of hyper-parameters to be tuned for model optimization.

Order	Hyper-parameter	Explored values
1	Batch Size	(32, 64, 128, 256, 512)
2	Learning rate	(0.0001, 0.0005, 0.001, 0.005)
3	Loss function	(MSE, MAE, Log cosh, Attention loss)

Only three key hyper-parameters are explored in this study. The batch size and learning rate are usually considered as one of the most important hyper-parameters influencing the model training process, so they are chosen as the first two parameters to be tuned. Additionally, the CNN configuration is usually very context-dependent, meaning that the CNN model settings always need to be dynamically adjusted according to its application context (e.g., if there is any imbalance in input distribution; what is the subject of high interest, etc.). Therefore, other loss functions are also tested in case they may fit better in this regression task. The ‘attention loss’ is included as one of the explored options because it is proved to be particularly useful in dealing with the imbalanced distribution of input values in regression task (Ren et al., 2019). This imbalance distribution may also occur in the multi-deprivation degrees of Nairobi, as only limited areas were delineated as slums (morphologically deprived areas) based on prior knowledge. The attention loss (AMAE) is denoted in Eq. (5) as adopted from (Ren et al., 2019):

$$\alpha_i = \left| \frac{y_i}{\hat{y}_i} - 1 \right|$$

$$AMAE = \frac{1}{n} \sum_{i=1}^n \alpha_i^k |\hat{y}_i| \quad \text{Eq. (5)}$$

where α_i measures the relative difference between model prediction y_i and reference value \hat{y}_i ; k is the parameter specifying how many ‘attentions’ should be employed on the ‘hard’ subjects (the ones with a large discrepancy between prediction and reference) so that the model could automatically down-weight the dominant majority of easily predicted subjects (Lin et al., 2017).

² The complete Python script of the CNN simulation is available at GitHub for the readers’ reference (https://github.com/EquLuo/MappingDeprivation/tree/main/CNN-based_regression_model).

3.5.4. Performance evaluation

To comprehensively evaluate the performance of the proposed CNN-based regression model, three common metrics for regression task were adopted. The evaluation metrics include mean absolute error (MAE), root mean squared error (RMSE) and the coefficient of determination (R^2). The MAE calculates the average of the absolute difference between the prediction and the actual value, thus indicating the average magnitude of the error from the predictions. The RMSE measures a squared magnitude of the average error from the predictions. The RMSE tend to give bigger weight to large errors, so it is useful for detecting the outliers in predictions. These MAE and RMSE are calculated by Eq. (6) and Eq. (7), respectively.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad \text{Eq. (6)}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad \text{Eq. (7)}$$

Where y_i is the predicted deprivation index; \hat{y}_i is the reference deprivation index (i.e., the deprivation index generated by the PCA); and n is the total number of samples included.

The R^2 provides an overall assessment of to what degrees does the regression model fits the actual data. It measures the proportion of the variance of the target variable explained by the independent variables of the regression model. The R^2 ranges between 0 to 1, and the closer to 1 indicates the better fitness of the model. The calculation of R^2 is shown in Eq. (8).

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)}{\sum_i (\hat{y}_i - \bar{y})} \quad \text{Eq. (8)}$$

Where y_i is the predicted deprivation index; \hat{y}_i is the reference deprivation index; \bar{y} is the mean of \hat{y}_i .

In addition to the statistical metrics, the CNN outputs are also geographically visualized and compared with the PCA reference data, i.e., the '*multi-deprivation portfolio*'. The visual assessments are conducted for each of the deprivation indices produced by the CNN model.

4. RESULTS

4.1. Elimination of unsuitable indicators for PCA

Since PCA is a very data-dependent approach, it requires the input data with good suitability to retain meaningful variance from the original datasets. To achieve this, a set of PCA-based indices produced by different subsets of the inputs were visualized and iteratively compared with the reference data.

In the end, 11 candidate indicators were discarded, including all the DHS model surface datasets (7 in total) (Burgert-Brucker et al., 2018), the improved housing prevalence (Tusting et al., 2019), the pit latrines (Mahabir et al., 2020) and the armed conflicts (ACLED, 2020). Detailed justification for discarding those indicators is given in Table 6. Although the information richness of multiple deprivation would shrink to some level after the removal of several indicators, the authors still argue that it is more important to ensure the indices are produced from relatively fine-resolution datasets with high quality and balanced distribution, especially for this intra-city analysis.

Table 6 – The summary of discarded indicators for PCA analysis after the quality check.

Removed datasets	Justification
The DHS model surface in Kenya 2014; The improved housing prevalence	Unsuitable spatial scale: The original resolution of these indicators is around 4.7km, which may be too coarse for this 100m grid intra-urban analysis. Thus, using such ‘big scale’ data would distort and mask off the spatial heterogeneity of other finer indicators and bring unwanted spatial autocorrelation from the 4.7km grid (Abson et al., 2012).
Pit latrines	Imbalanced distribution: the locational data of pit latrine was highly concentrated on existing slums because of the greater presence of non-government organizations (e.g., local initiatives, research institutes) conducting investigations in and around the deprived areas (Mahabir et al., 2020). Thus, the data outside slums might be under-reported.
Armed conflicts	Imbalanced distribution: The Armed conflicts and events dataset are mostly retrieved from media-based sources (Eck, 2012). After a visual assessment, it is observed that the locational crime data tend to be largely underreported, especially outside the central business district, because the occurred events may not be reported on media.

4.2. Decomposing deprivation dimensionality through the PCA component scores

From the PCA results, four principal components are retained, with each of the PC’s eigenvalue scoring more than 1. A KMO test value of 0.862 is also achieved, which can be considered good for this analysis (Hair, 2006). In total, the first four PCs together explain 63.8% of the total variance from the original 15 variables included in the PCA. The varimax rotation is applied for a better interpretation of each component. Table 7 details the four extracted PCs and their effects on deprivation, eigenvalue, and percentage of explained variance, along with the dominant loadings bolded and highlighted in different colour shade. Note that each component is interpreted into different facets of multi-deprivation according to its heavily loaded factors. The effect column indicates whether the retained PCs contribute to or reduce the degrees of deprivation.

Table 7 – The retained principal components scores and the component loadings in the rotated matrix.

Component Name	Effect	Eigenvalue	% variance explained	Rotated Component Matrix				
				Dominant variables	Component loadings			
					PC1	PC2	PC3	PC4
Poverty, accessibility to facilities, and maternal health support	+	4.010	26.73	Estimated % of people living in poverty	0.865	-0.239	-0.064	0.088
				Distance to health facilities	0.844	-0.088	-0.277	-0.014
				Distance to education facilities	0.779	-0.135	0.225	-0.146
				Distance to financial facilities	0.723	-0.063	-0.399	0.127
				% of skilled birth attendance in delivery	-0.709	0.186	0.272	-0.081
				Night-time light value	-0.619	0.346	0.309	-0.314
Dense urbanization, absence of green space and waste management	+	2.651	17.67	Population density	-0.201	0.800	0.083	0.118
				Building density	-0.032	0.794	0.042	0.007
				Maximum NDVI value	0.118	-0.653	-0.396	0.271
				Density of illegal dumpsites	-0.192	0.606	0.145	0.045
Air and water contamination	+	1.599	10.66	Annual concentrations of pm 2.5	-0.165	0.160	0.856	-0.079
				Density of waterways	-0.196	0.290	0.466	0.298
Transport infrastructure	+	1.311	8.74	Density of roads with poor quality	-0.231	0.348	-0.028	0.640
				Density of bus station	-0.385	0.336	-0.024	-0.598
				Distance to OSM major roads	0.417	-0.030	-0.019	0.456
Total Variance explained			63.804					

The first component (PC1) accounts for 26.7% of the total variance. PC1 is heavily dominated by the following six indicators: population living in poverty, distance to health facilities, distance to education facilities, distance to financial facilities, percentage of skilled birth attendance and night-time light (NTL) value. Among these, the poverty rate and NTL value reflect the level of poverty – NTL is a very popular proxy for measuring the poverty level (Jean et al., 2016; Kuffer, Pfeffer, et al., 2018); while the three distance indicators explain the accessibility to common facilities and the skilled birth attendance indicates maternal health care level. Thus, PC1 is titled “Poverty, accessibility to facilities, and maternal health support”.

The second component (PC2) explains 17.67%, the second largest share from the total variance. PC2 is heavily loaded on four indicators, including population density, building density, maximum NDVI value and the density of illegal dumpsites. The combination of these indicators tends to reflect those areas with highly dense buildings where the large population reside, also coupled with the lack of vegetation coverage and poor waste collection. As such, PC2 is termed as “Dense urbanization, absence of green space and waste management”.

Regarding the third component (PC3), it includes two dominant factors – the annual concentration of pm2.5 and water density – which together account for 10.66% of the total variance. As mentioned earlier that most of the waterways have been seriously polluted in Nairobi, the spatial density of waterways then becomes a proxy to indicate the water pollution degree. Therefore, PC3 can be titled “Air and water contamination”.

The last retained component (PC4) accounts for 8.74% of the total variance. PC4 is largely dominated by three indicators – density of roads with poor quality, the density of bus station and distance to major roads, with each pertaining to transport condition and road infrastructure. Based on this, PC4 is labelled as “Transport infrastructure”.

Lastly, it is worth stressing that the naming convention employed on the retained PCs is to help unpack the multi-dimensionality of urban deprivation into inter-independent sub-domains, as well as highlighting the unique characteristics and patterns within the identified sub-domains of multi-deprivation.

4.3. PCA-based multiple deprivation indices

Based on the PCA analysis, the ‘*multi-deprivation portfolio*’ is established to provide a comprehensive and continuous characterisation for the multiple deprivation degrees in Nairobi. As mentioned earlier, the ‘*multi-deprivation portfolio*’ is a compilation of the individual sub-dimensional deprivation indices (i.e., the extracted PC scores) and the aggregate multi-deprivation index (i.e., the additive score of all the PCs with equal weight). In the following paragraphs, a detailed description of the ‘*multi-deprivation portfolio*’ and its spatial distribution in Nairobi is presented.

4.3.1.1. The sub-dimensional deprivation index

The spatial distribution of the sub-dimensional deprivation indices is visualized in *Figure 14* at the standard 100m grid level. All the deprivation values are divided into 10 sub-class with the same interval (0.1), for the ease of visual interpretation and inter-comparison between PCs. It could be shortly noticed that the emerging spatial distribution pattern of the deprivation index varies strongly across the PCs, with each scoring high or low values on different areas within Nairobi. In spite of this large discrepancy, a level of spatial correlation could still be observed between some of the PCs, e.g., both PC2 and PC4 detects similar highly deprived areas; the PC1 and PC2 generally shows an east-west division pattern of deprivation.

Regarding PC1, the most deprived areas can be founded on the east outskirts of Nairobi, especially the right-upper corner highlighted by dark red colour. Contrarily, the pan-central areas score the lowest deprivation degree of PC1 on average. In general, it can be observed that as the distance to the city core increases, the deprivation index of PC1 also tends to rise, showing an outwardly ascending pattern. Therefore, the PC1 result map indicates that, compared to the city centre, most of the urban periphery in Nairobi is more deprived in terms of income level, accessibility to common facilities and the chance to receive maternal health support.

In terms of PC2, the vast majority of the deprivation values stays in a very low range, as shown in *Figure 14. (b)*, where most of the areas in Nairobi are covered by darker green colours. Nonetheless, some highly deprived spots can still be recognized from the study area, although only accounting for relatively small patches. There are three major deprived clusters identified in Nairobi, with the first one at the lower mid-left part, the second one (like a triangle shape) located at the opposite side of the first one and the third one, also the biggest one, located in the upper mid-right part of Nairobi. According to the map, it can be claimed that the average level of PC2 deprivation is low, but some small pockets (i.e., the highlighted deprived spots) are in severe deprivation resulted from dense urbanization, absence of green spaces and poor waste management. Additionally, it is worth underlining again that PC2 mainly reflects the urban morphology aspect of multi-deprivation and a detailed discussion about it is provided later.

In terms of PC3, the spatial distribution indicates a strong concentration of high deprivation values on the mid-east part of Nairobi, especially the bottleneck-shaped region that connects east Nairobi to the main urban body. The mid-lower part of Nairobi is also facing above mediate deprivation level. On the other hand, the lower deprivation level of PC3 is observed at the east edge, coupled with the northwest corner. Generally, the majority of Nairobi is deemed as relatively deprived of good air and water quality.

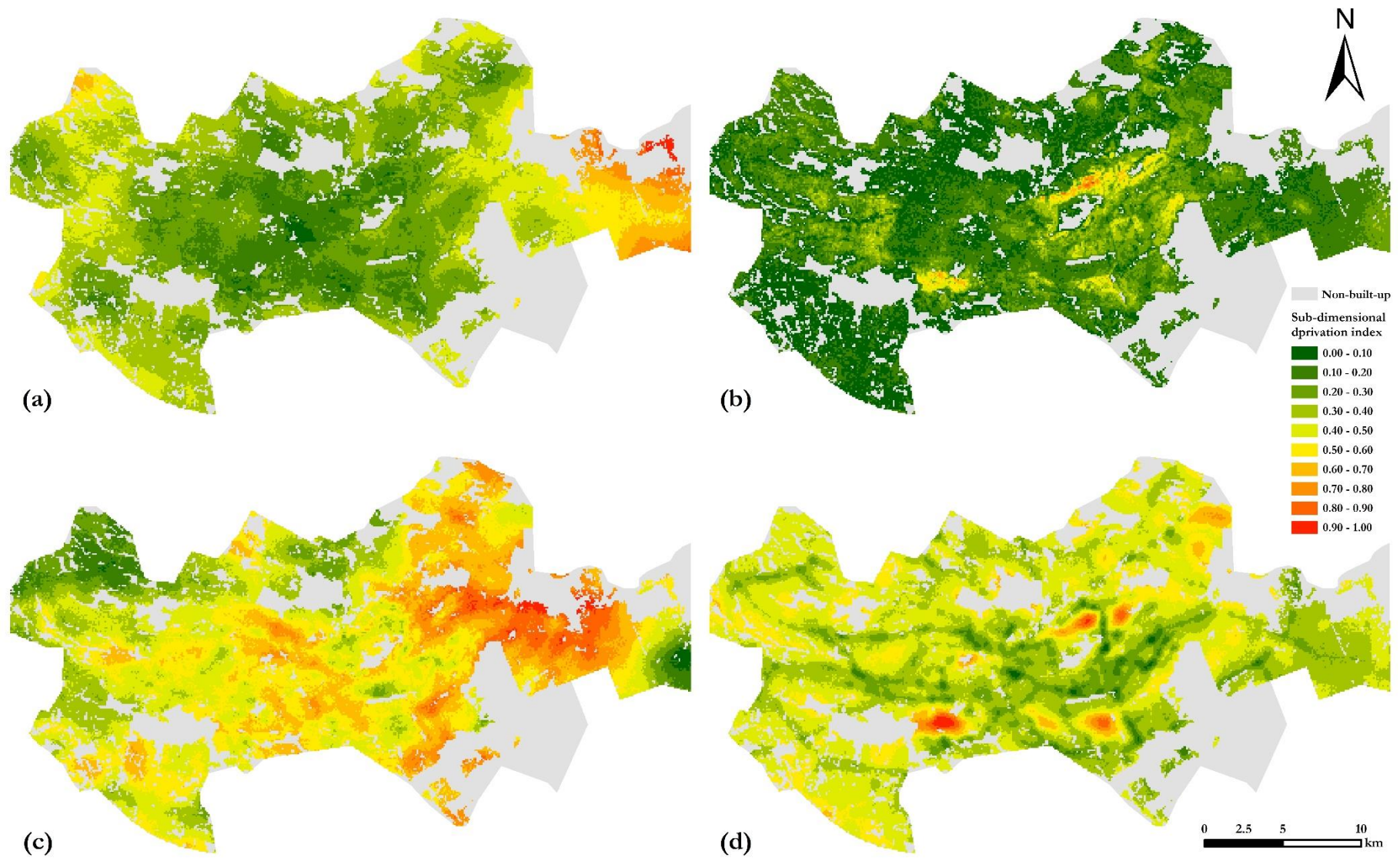


Figure 14 – Spatial distribution of the extracted four sub-dimension deprivation indices: (a) PC1 - Poverty, accessibility to facilities, and maternal health support; (b) PC2 - Dense urbanization, absence of green space and waste management; (c) PC3 - Air and water contamination; and (d) PC4 - Transport infrastructure.

As for PC4, it can be apparently seen that the areas of low deprivation level follow a distribution pattern similar to the traffic roads, with some nodes of even lower deprivation value scattered along the paths. Yet, some highly deprived 'islands', which are surrounded by lower values meanwhile standing alone among each other, still appear on the map. It is interesting to note that the locations of such deprived islands in PC4 are somewhat spatially overlapping with the highly deprived spots in PC2. Overall, regarding deprivation level of transport infrastructure, the city centre and the roads extending outward to the east and west are the least deprived, and as the distance to them increases, the deprivation level also tends to rise, but there are also some islands in between that are characterized by high deprivation level.

4.3.1.2. The aggregate multi-deprivation index

Figure 15 visualizes the spatial distribution of the aggregate multi-deprivation index, calculated via an additive, equal-weight aggregation method (Eq. (3)). It can be clearly observed that there is a significant variation of the aggregate index within Nairobi, which results from the aggregation of those individual sub-dimensional indices.

Generally speaking, based on the map, the entire urban areas tend to show an east-west division of multiple deprivation degrees, where most areas in eastern Nairobi have a relatively higher level of multi-deprivation on average compared to the rest of the city. This spatial concentration of high deprivation values is in alignment with the spatial pattern of PC1 and PC3, where the east of Nairobi is also characterised as a highly deprived region. Moreover, it can be observed that the city centre and its neighbouring areas are less deprived, except that some highly deprived spots and patches, which have already been recognized in the maps of PC2 and PC4, are still appearing on the aggregated index map. Regarding the westside of Nairobi, the overall multi-deprivation level can be described as low-to-medium, with only some areas scoring medium values in the middle and the lower edge.

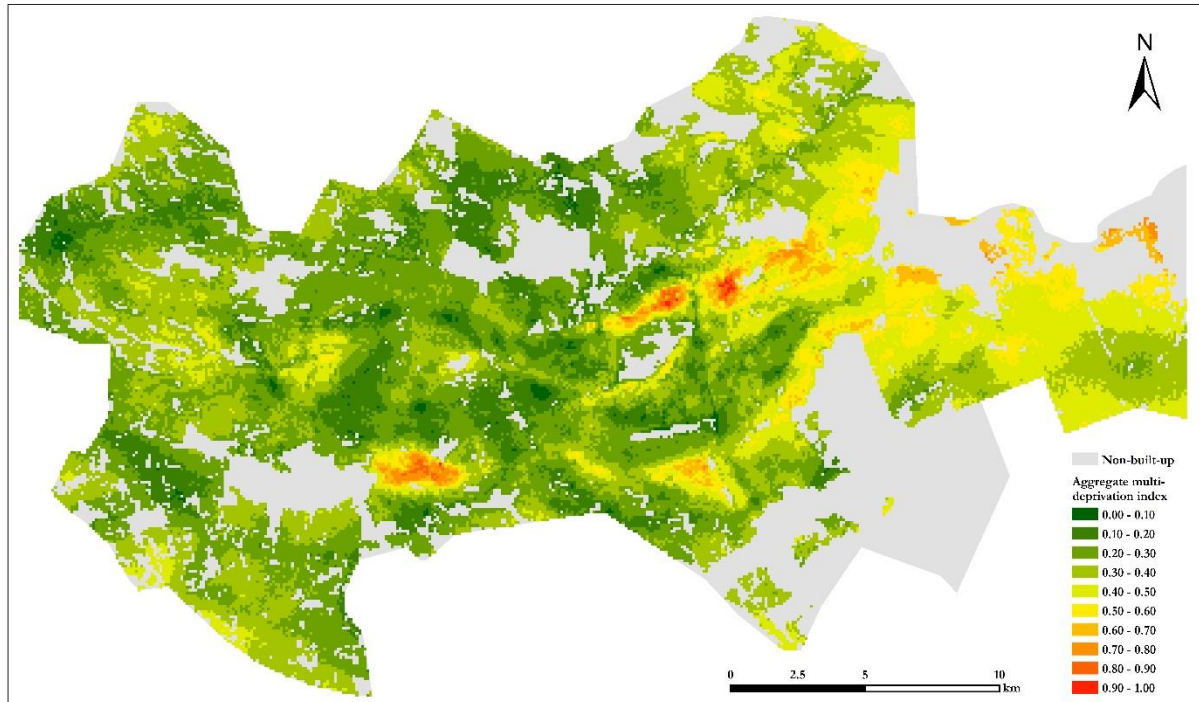


Figure 15 – Spatial distribution of the aggregate multi-deprivation index.

4.4. PCA results validation and interpretation

4.4.1. Comparison with slum boundary 2016

The PCA results are first checked using the slum boundary dataset. All the built-up areas that do not fall within the slum extents are considered non-slums. The distribution of PCA-based deprivation indices on slums and non-slum areas is shown in *Figure 16*. A summary of the mean of deprivation indices and its difference between slum and non-slum is also provided in *Table 8*.

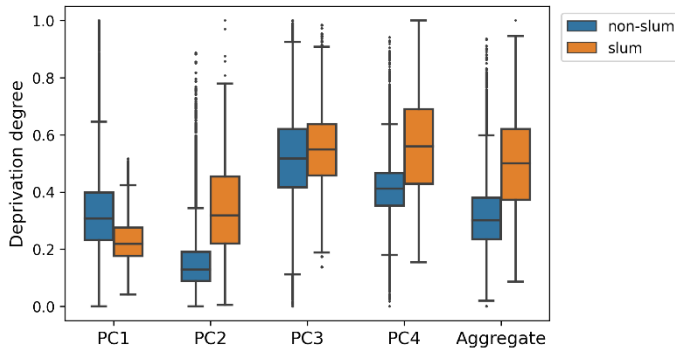


Table 8 – The mean of deprivation indices between slum and non-slum.

	Mean value		
	Non-slum	Slum	Difference
PC1	0.330	0.239	-0.091
PC2	0.150	0.340	0.190
PC3	0.515	0.564	0.049
PC4	0.408	0.568	0.160
Aggregate	0.315	0.503	0.188

Figure 16 – A boxplot showing the distribution of PCA-based deprivation indices on slum and non-slum areas.

Based on the boxplot, it is clear that except for PC1, the majority of slum areas always have higher deprivation degrees than the majority of non-slum areas. Surprisingly, the deprivation level in PC1 shows an opposite feature, i.e., most of the non-slum areas are more deprived compared to the defined slum regions, with a difference of -0.091 between the mean of non-slum and slum (*Table 8*). This discrepancy indicates that non-slum areas in Nairobi are more deprived in one or more of the following aspects: income level, accessibility to facilities and maternal health support. Regarding PC3, the deprivation level does not differ a lot, with slums scoring slightly higher on average. And for PC4 and the aggregate index, both the slum regions outweigh non-slums, confirming that slums are more deprived in general.

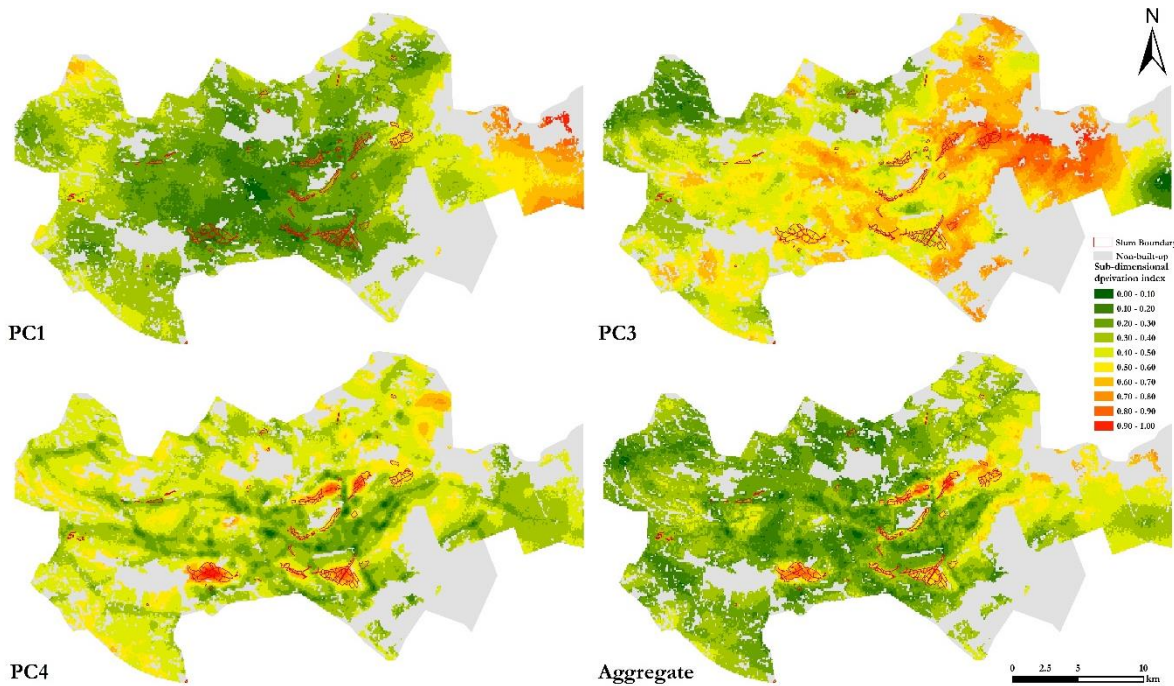


Figure 17 – The map overlaying the slum boundary with PC1, PC3, PC4 and aggregate index.

Figure 17 allows for spatially overlaying the slum boundary with the deprivation indices for better visual interpretation. Corresponding to the description above, the slum boundary does not fall within the highly deprived areas in PC1, nor does it exclusively scatter in more deprived areas in PC3. Nevertheless, most of the areas of high deprivation degrees in PC4 and aggregate index are geographically overlapping with the slum boundary, meaning that there is a distinct difference between slums and non-slums regarding them.

Among all the facets of multiple deprivation, PC2 shows the strongest variation of average deprivation level between slum and non-slum (0.190, from Table 8). Apart from this, Figure 18 further reveals an obvious spatial pattern, where nearly all the slum extents are geographically overlapping with the highly deprived spots identified by PC2. Based on the previous categorization of each PC, these highly deprived spots from PC2 are characterised as dense urbanization, absence of green space and poor waste management, among which the first two characteristics can be considered as morphology-based (Table 7). Given this inherence of morphological features in PC2, plus its spatial coincidence with the defined slum extents, it can be claimed that the spatial manifestation of the slum in Nairobi is mostly captured or reflected by PC2. Therefore, in this case, PC2 is regarded as the morphology-based deprivation domain, which is highly likely to spatially emerge as slum entities, and its continuous value characterises the degree of morphology-based deprivation. Additionally, previous studies have proved that the EO-based methods have the unique strength in detecting and capturing the physical/morphological features with high accuracy, covering large scales (Kuffer et al., 2020; Kuffer, Pfeffer, & Sliuzas, 2016). Given the above considerations, PC2 will be the primary object of interest to be further investigated and explained in this study.

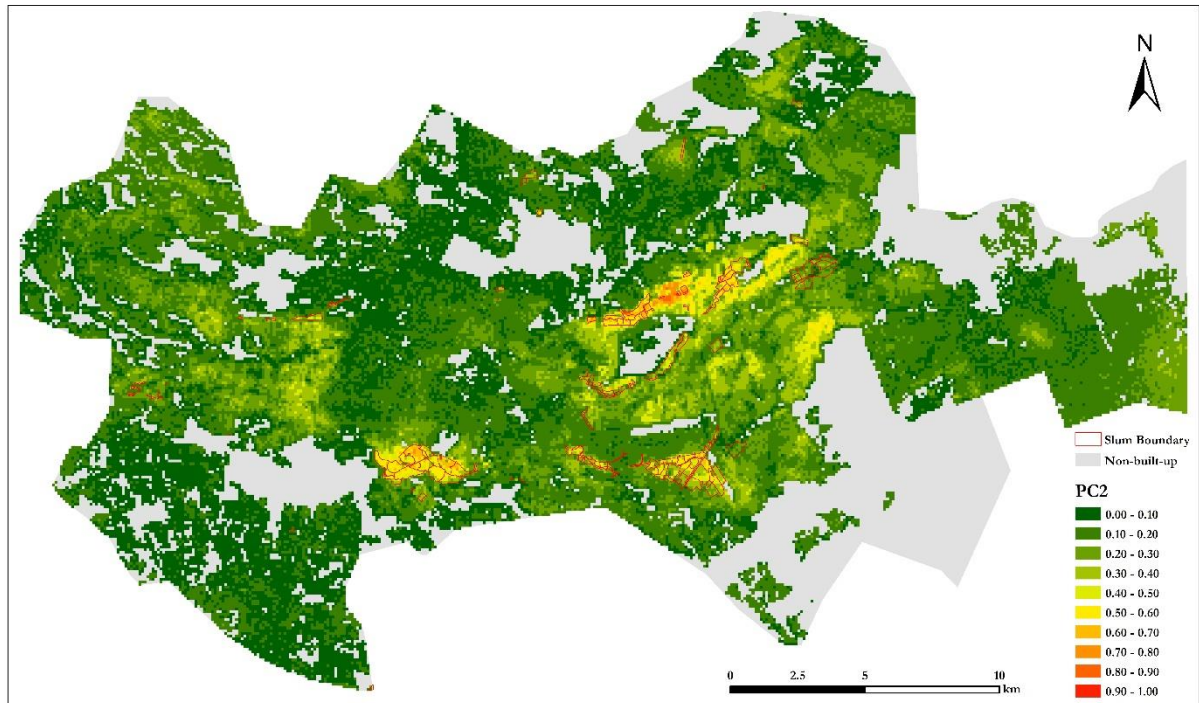


Figure 18 – The map overlaying the slum boundary with PC2.

4.4.2. Variance of LULC typologies within deprivation

The LULC typologies data are also included for validating the PCA results. To compare the distribution of each PCA-based index across different land use types, multiple boxplots are visualized in *Figure 19*. In total, five land use classes are compared, namely, high to mid-density built area, low density built area, industry/large structure, and two types of deprived areas, as shown in the legend of *Figure 19*. Note that here, class 4 (Type 1 deprivation) refers to the typical deprived areas with highly dense building density and unstructured road networks (see *Figure 8* the lower-left picture), whereas class 5 (Type 2 deprivation) represents the atypical deprived areas of mid-dense building clusters and less irregularity in the road networks (see *Figure 8* the lower-right picture).

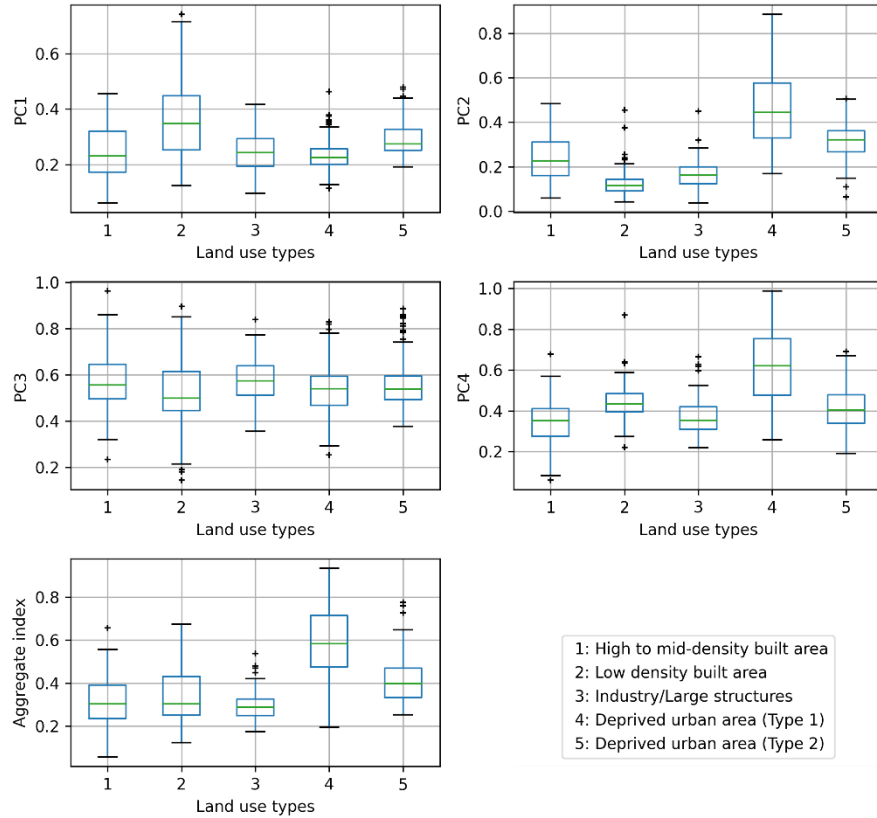


Figure 19 – The grouped boxplots showing the distribution of PCA results on different land use typologies.

In PC1, the highest level of deprivation occurs in the low-density built areas, which usually located in the sub-urban regions of Nairobi with a relatively long distance to the city centre. The lower level of deprivation in PC1 is, however, mainly distributed in the high to mid-density built-up areas, as well as the highly-dense deprived areas, both of which mostly located in the central urban areas.

With regards to PC2, the morphology-based deprivation domain, the highest level of deprivation degrees can be clearly observed in the typical deprived areas, closely followed by the atypical deprived areas. On the other hand, the low-density built areas score the lowest in morphological deprivation degrees, while the more compact built areas still show a relatively high deprivation level on average. In short, there is a strong correlation between the PC2 deprivation degrees and the spatial form and density of settlements.

Moving to PC3, a relatively even distribution of deprivation degrees can be noticed among the land use types. However, it is worth noting that, on average, the industrial areas and/or the large structures are the most deprived in PC2, meaning that they suffer more from air and water contamination. This pattern intuitively makes sense because industrial production is usually the major source leading to air and water pollution. Additionally, the average lowest level of PC3 deprivation occurs in the low-density built areas.

As for PC4, the dense deprived areas score the highest value on average, with a relatively large difference compared to the rest land use types. This may indicate that the typical deprived areas are strongly deficient in transport infrastructure in terms of road quality, adequate bus stations coverage and proximity to near major roads.

In general, from an aggregate perspective, in Nairobi, the highest deprivation level still falls within the deprived urban areas, with the typical dense deprived areas scoring first, followed by the atypical, mid-dense deprived areas. This pattern is retained mainly due to the variation from PC2 and PC4. Contrarily, the formal built-ups have lower values in aggregate multi-deprivation degrees.

4.4.3. Visual assessment of PC2 with the VHR and street-view images

In this section, a visual assessment is conducted on PC2 – the most morphology-dominated dimension, by comparing it with the VHR satellite images. Additionally, the street-view images of Nairobi are also included for visual assessment to add more ground information. The street-view images are retrieved from the Mapillary website, an online platform for accessing street-level imagery data (Mapillary, 2021).

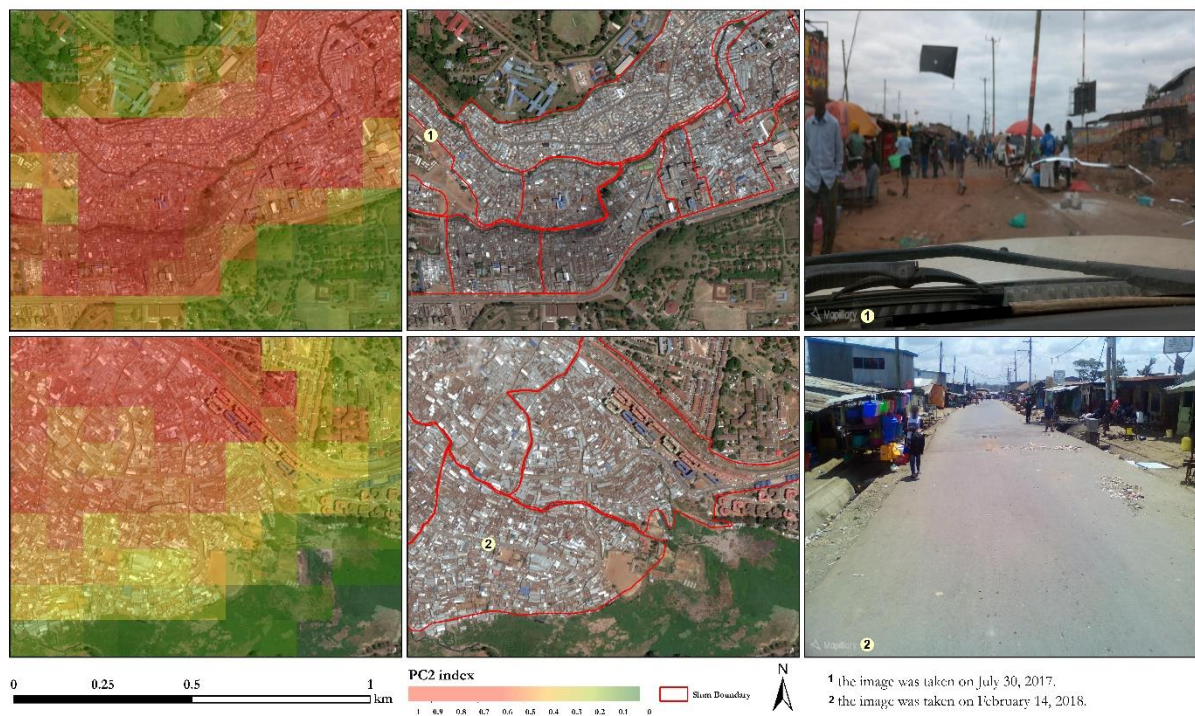


Figure 20 - The visual assessment of PC2 on slum areas by comparing with VHR and street-view images. Source: (Mapillary, 2021).

First, the defined slum areas are visually checked. Figure 20 selects the two most representative areas to reveal the huge spatial variation of morphological deprivation degrees in Nairobi. Apparently, nearly all the orange-to-red grids (i.e., the highly deprived areas as identified by PC2) are spatially overlapping the slum boundaries such that even on the gridded map of PC2, those highly deprived grids form a boundary that explicitly distinguishes themselves from other land uses. When the urban fabric changes from dense slums to green vegetation and/or more structured built-up areas, the PC2 deprivation index starts to drop. This pattern is confirmed by Figure 20, where the regions covered by vegetation or mixed with several sparsely scattered buildings have the lowest deprivation value. Two street-level photos of typical slums are purposely selected as representative ground view images. Photo-1 (Mathare slum) shows a relatively high density of pedestrians walking on unpaved roads. Several garbage tiles can be found on the street. Photo-2 (Kibera slum) shows a better infrastructure condition with paved roads but still dominated by the dense concentration of low-height, unorganized sub-standard housings, coupled with trashes on the street.

Figure 21 presents the visual inspection of PC2 deprivation degrees in the central business district (CBD) area. Normally, CBD is one of the most developed areas in a city, with a mixture of different land-use types and great accessibility to jobs, facilities, and infrastructures. Overall speaking, the CBD has a quite low level of deprivation in PC2, although the upper-left part is more deprived compared to the rest, which may result from its higher building density. Nonetheless, as further unveiled by the street-view image, the CBD in Nairobi is generally a very well structured and modernized urban areas.

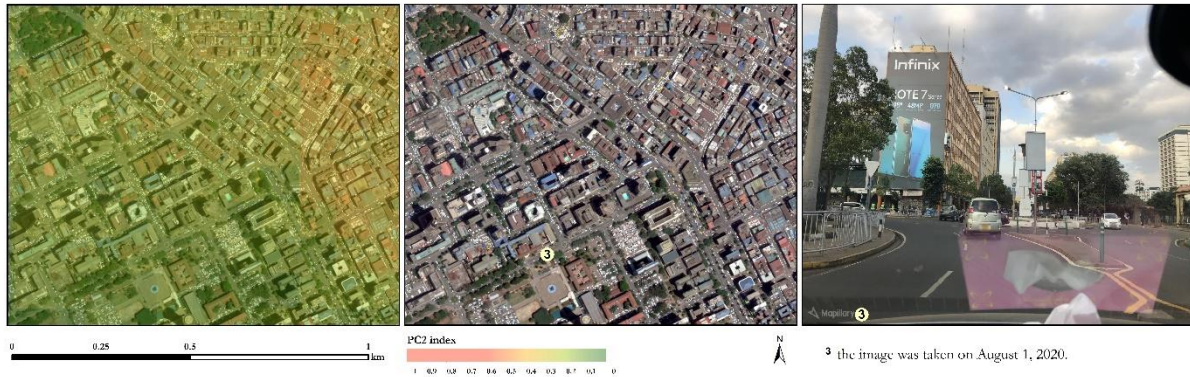


Figure 21 – The visual assessment of PC2 on CBD area by comparing with VHR and street-view images. Source: (Mapillary, 2021).

The atypical deprived areas are also carefully checked via visual evaluation (Figure 22). A distinct gridded boundary that differentiates the highly deprived patches from less deprived patches can be found on both of the selected sites from the first subgraphs. Such gridded boundaries also roughly overlap with the boundaries that actually divide the atypical deprived areas and formal built up, as shown in the second subgraphs in Figure 22. Therefore, it can be concluded that using the proposed PCA approach enables to capture or identify the mid-dense deprived areas from formal areas. In addition, the two street-level photos show a lower building density and less waste accumulated on the streets in the atypical deprived areas in comparison to the slum.



Figure 22 – The visual assessment of PC2 on the atypical deprived areas by comparing with VHR and street-view images. Source: (Mapillary, 2021).

Lastly, two sites dominated by well-organized, low-density buildings along with some green coverages are visually checked. Unsurprisingly, the two areas both have very low deprivation level in PC2, especially for the grids that mainly comprised of vegetation. The two street-view images further affirm this non-deprivation status. As shown in the ground photos, such areas often have nicely paved and clean roads, sufficient green space and a low density of settlements.



Figure 23 – The visual assessment of PC2 on the formal built-up areas by comparing with VHR and street-view images. Source: (Mapillary, 2021).

4.4.4. Local expert validation on PCA results

The local validation of the PCA results was conducted via an online semi-structured interview with a local expert, who is a researcher from the African Health and Population Research Centre (APHRC) with strong experience and substantial local knowledge on deprivation in Nairobi.

From the interview, the expert further confirmed that PC2 best resembles to the local patterns of deprivation at the ground, as perceived by the local people. In general, the highly deprived areas in PC2 are located within and around the *de facto* deprived neighbourhoods, and the ‘well-off’, non-deprived areas have low deprivation values (Figure 18). This may imply that in Nairobi, the indicators that strongly differentiate the deprived from non-deprived areas are mostly related to the physical morphology, green space coverage and waste management. The PC4 also somewhat performs well in capturing the deprived slum region, in line with the local knowledge that most of the slums suffering from poor road quality (e.g., unpaved, unstructured). In terms of PC3, the expert suggested that the overall spatial pattern of contamination is relatively valid, as the east part of Nairobi locates a lot of industry, and the waterways crossing through the city centre are highly polluted. The validity of PC1 is a bit controversial because large parts of the urban outskirts are predicted as deprived by PC1, but according to ground knowledge, those sub-urban areas are usually low-density neighbourhoods whose residents tend to be middle-upper class. This counter-local deprivation pattern may result from the dominance of ‘distance’ indicators and night-time light in PC1 because urban peripheries are often far from facilities and have lower building density. Yet, this may not necessarily indicate that the sub-urban is more deprived of accessibility because the residents usually have greater access to automobiles.

4.5. CNN model implementation and optimization

Through the validation and visual assessment of PCA results, PC2 is recognized as the primary subject given its strong morphological characteristics and spatial coincidence with slum extents. Therefore, as a starting point, PC2 – the morphology-based deprivation index – is first fed into a deep CNN regression model to explore the potential of leveraging EO data in predicting deprivation degrees.

As mentioned early, a stepwise hyper-parameter tuning is conducted for model optimization. For the detailed records of performed tuning experiments, please refer to Annex 1 attached in the Appendix. *Table 9* provides the final results of hyper-parameter tuning. Overall, three hyper-parameters were tested, namely batch size, learning rate and loss function. Based on the result, the optimal value for batch size is found at 256, and the one for learning rate is 0.001. Here, the attention loss is specifically employed for the PC2 simulation because the original distribution of PC2 is highly skewed – i.e., the vast majority of the training samples have a very low value of less than 0.2 in PC2, as shown in *Figure 24*. As such, it is necessary to ensure the model identify and be able to predict the high-value samples. The optimized model achieves an MAE of 0.422, RMSE of 0.0585 and R^2 of 0.6637 on the validation dataset (Annex1).

Table 9 – The optimal values after hyper-parameter tuning.

Order	Hyper-parameter	Explored values	Optimal value
1	Batch Size	(32, 64, 128, 256, 512)	256
2	Learning rate	(0.0001, 0.0005, 0.001, 0.005)	0.001
3	Loss function	(MSE, MAE, Log cosh, Attention loss)	Attention loss (k = 1.15)

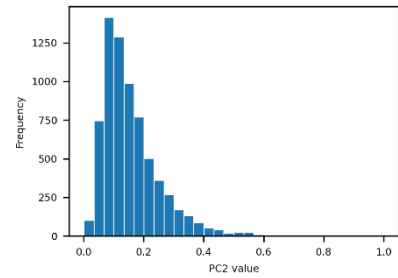


Figure 24 - Histogram of PC2 input samples.

After the model optimization, the final configured CNN-based regression model is implemented on the unseen test dataset, i.e., the remaining 38559 sample grids (around 85% of the total) that are independent of the model training and validation process. By doing this, the generalization ability of the CNN model in predicting the intra-city deprivation degrees for the entire Nairobi, given only limited training data, is tested. The performance of the proposed CNN-based regression model on the test dataset is reported in *Table 10*. The MAE and RMSE value on the test dataset is 0.0421 and 0.0582, respectively. The proposed model achieves an R^2 of 0.6543 on the test dataset, indicating the trained model can moderately explain the variance of the PC2 index. In fact, from all the experiments conducted for the PC2 simulation, the resultant R^2 values always fall within a stable range of around 0.65, with very small variances. Therefore, the reported R^2 value can be considered as a reasonable assessment of the model performance.

The density scatter plot comparing the CNN prediction with the reference values on the test dataset is visualized in *Figure 25*. The vast majority of the test samples concentrates on lower values of PC2. It can be observed that in the lower range, the difference between model prediction and reference data is relatively small and stable. However, as the value increases over 0.4, the error between the prediction and reference data also starts to become larger, with more samples being underestimated by the CNN model.

Table 10 – The CNN performance on the test datasets in predicting PC2.

	MAE	RMSE	R^2
PC2	0.0421 ± 0.0402	0.0582	0.6543

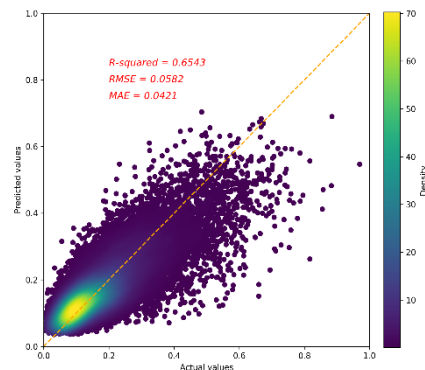


Figure 25 – The density scatter plot of PC2 prediction on test dataset.

4.6. CNN prediction on the morphology-based deprivation index

To provide a more intuitive and straightforward understanding of the outputs, the CNN predicted morphology-based deprivation index is projected back to the 100m standard grids and visually compared with the PC2 deprivation map (Figure 26). Note that here the symbology for visualization of PC2 values is adapted using different intervals to underscore the highly deprived areas, considering the large skewness in PC2 value distribution. In this sense, a value more than 0.5 can already be regarded as highly deprived.

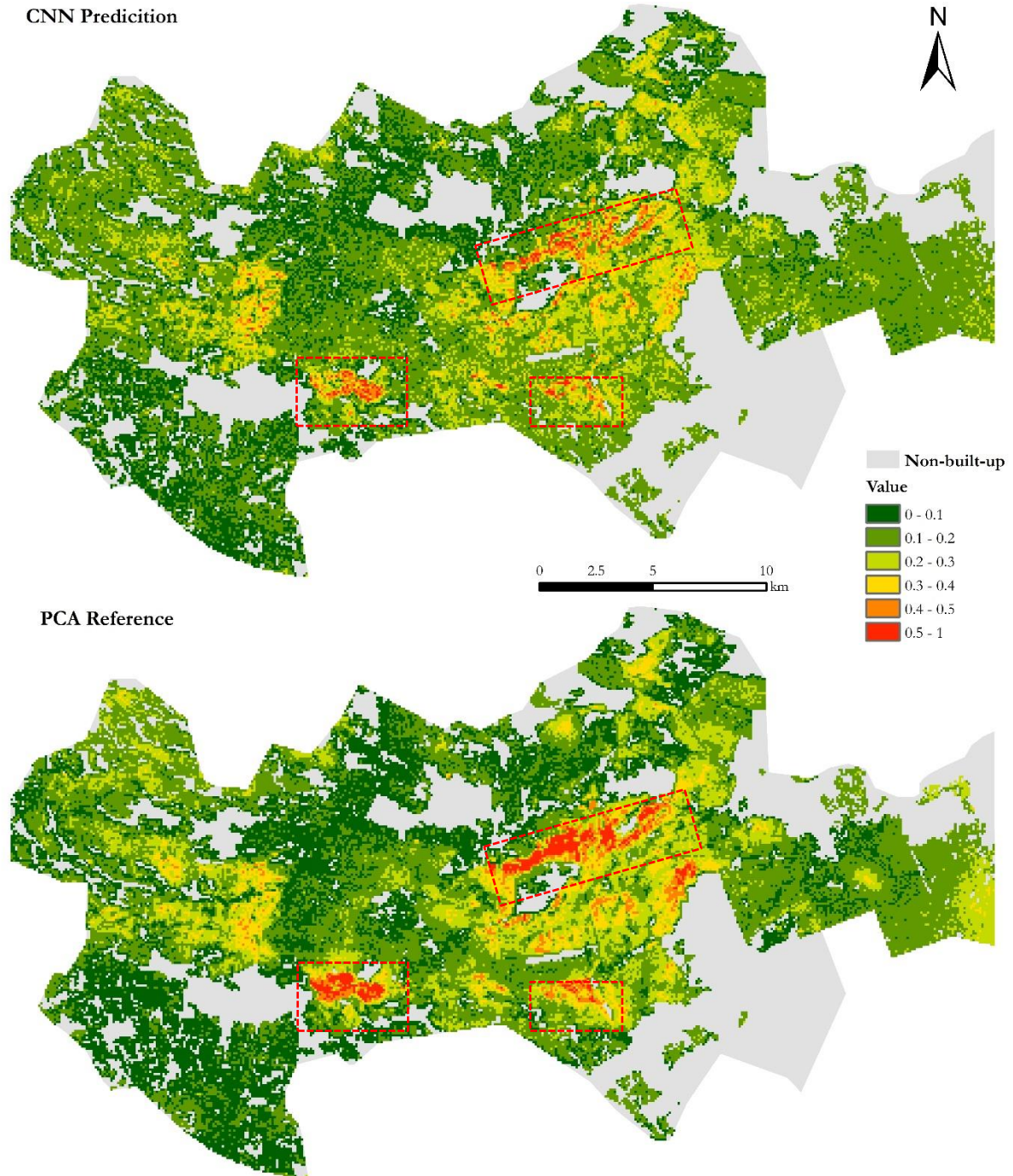


Figure 26 - The visual comparison between CNN prediction and the reference PC2 index.

Based on the map, the overall spatial pattern of the CNN prediction is similar to the PCA-based deprivation degree. It is clearly shown that the highly deprived areas are identified by the proposed model

within the roughly same spatial locations as in the PCA results. For example, the lower-left deprived spot – i.e., Kibera, the biggest slum area in Nairobi, as highlighted by the red dashed boxes – also appears in the CNN output map and characterised as one of the most deprived areas. Similar spatial coincidences of the highly deprived patches can also be observed in other highlighted dashed boxes. However, the severity of deprivation degrees from CNN output tends to be lower than the PCA-based degrees, as the size of deprived areas in CNN output is smaller, combined with also lighter colours. In addition, regarding the areas with low deprivation level, the proposed CNN model generates a noisier pattern in contrast to the spatial continuum observed in the PCA map.

Nonetheless, based on the reported R^2 of 0.6543 (Table 10) combined with the similar spatial patterns of PC2 between the prediction and reference shown in Figure 26, it can be claimed that a deep CNN-based regression model is able to partially capture the variation of morphological deprivation degrees.

4.7. CNN prediction on other deprivation indices

After testifying the potential of using deep learning technique in directly predicting PC2, the CNN model is also experimented on other deprivation indices. The CNN implementation on other deprivation indices follows the same optimized hyper-parameter setting as PC2, except the loss function adapted to MSE. In total, PC1, PC3, PC4 and the aggregate index have all been tested by the CNN model. The remaining sections describe the CNN model performance on each deprivation index.

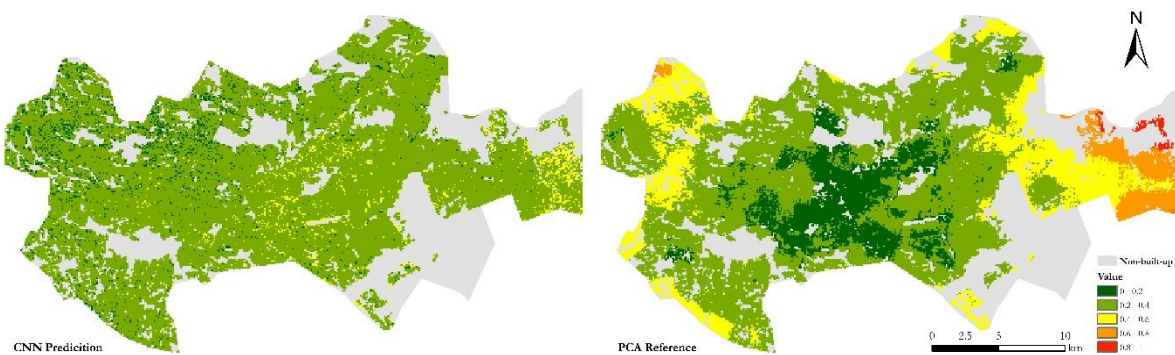


Figure 27 – The CNN prediction on PC1 (Poverty, accessibility to facilities, and maternal health support).

PC1 prediction: As presented in Figure 27, by only relying on the satellite images, the CNN model fails to predict the deprivation degrees in PC1. The CNN results show very little spatial variation across the city, with most of the areas dominated by deprivation level between 0.2 to 0.4. This even distribution pattern obtained in the PC1 prediction cannot correspond to the PCA-based index.

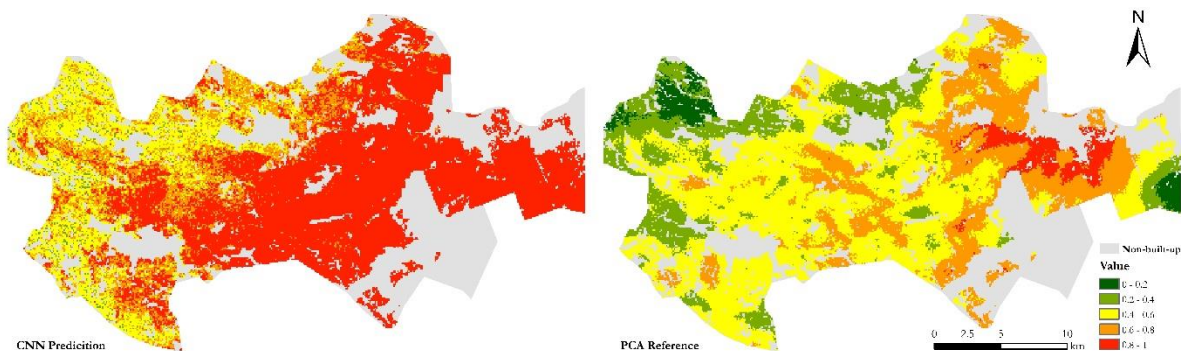


Figure 28 – The CNN prediction on PC3 (Air and water contamination).

PC3 prediction: The model outputs on PC3 are largely overestimated compared to the reference PCA values, as the vast majority of Nairobi is highlighted by the dark red colour (*Figure 28*). Overall, the CNN model also fails to explain the variance of deprivation degrees in PC3. However, it is worth noting that a general spatial pattern, where the east part of Nairobi is more deprived in PC3 compared to the west, can still be captured even within such noisy results.

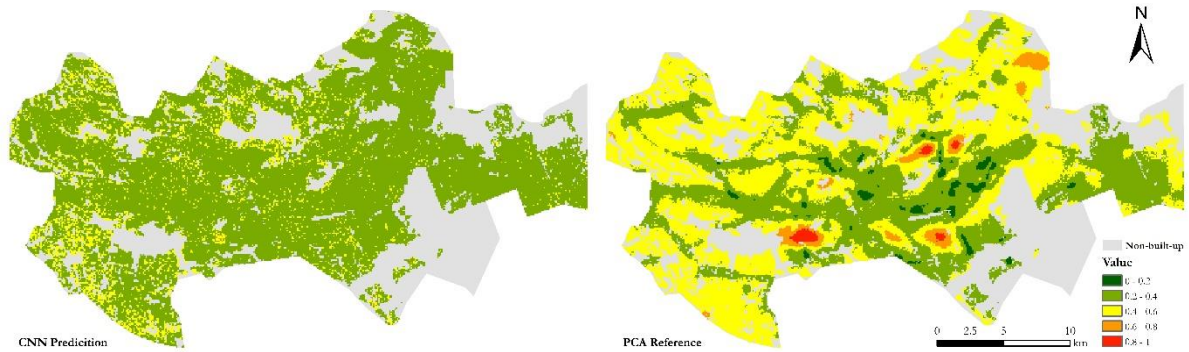


Figure 29 – The CNN prediction on PC4 (Transport infrastructure).

PC4 prediction: Similar to PC1, the CNN prediction on PC4 also shows an evened spatial pattern where large homogeneity of the PC4 distribution can be observed (*Figure 29*). Based on the model prediction, nearly all urban built-up areas are mistakenly characterised as non- or low-deprived, coupled with even those high deprivation patches being uncaptured on the map.

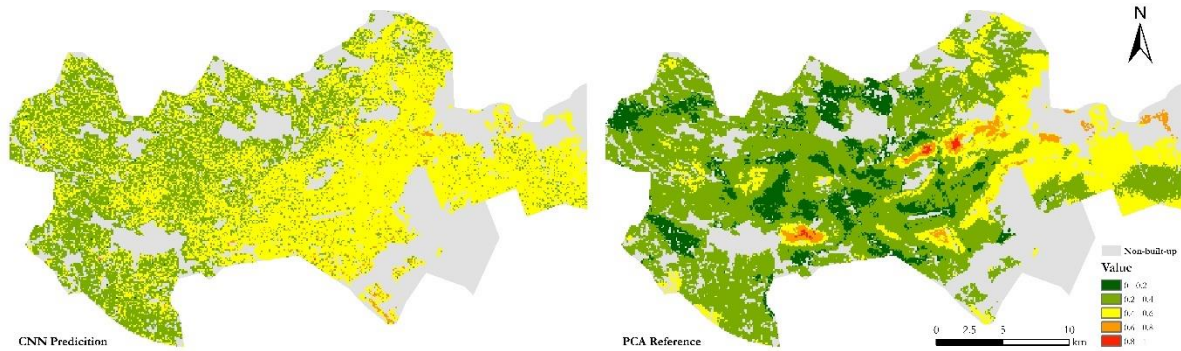


Figure 30 – The CNN prediction on the aggregate deprivation index.

Aggregate deprivation index prediction: Lastly, the model is trained to predict the aggregate deprivation index in Nairobi. According to *Figure 30*, the proposed CNN model, again, fails to explain the variance of the aggregate deprivation degrees. A lot of noises can be observed on the prediction map, especially on the west side of Nairobi, and the overall severity of intra-urban deprivation is over-estimated to some extents. In short, the CNN model cannot yield meaningful prediction on the aggregate deprivation index.

Given the above model mis-performances on other deprivation indices, it can be concluded that by just using EO-based imagery data, the proposed CNN-based regression model is unable to capture the non-morphological deprivation degrees in Nairobi.

5. DISCUSSION

In this study, we designed a two-step method that enables to first characterise the degrees of multiple deprivation as continuous indices and then examine the potential of an EO-based method to directly capture the multiple deprivation degrees. We used two key techniques, i.e., PCA for characterising the multi-dimensionality of deprivation, and a deep CNN-based regression model for predicting the multi-deprivation degrees. The discussion part is framed in terms of the multi-dimensionality nature of deprivation, the measuring of multiple deprivation, the role of an EO-based method in deprivation mapping, and lastly, the limitations of this study.

5.1. Multi-dimensionality nature of deprivation

In the beginning, we would like to point out again that, by nature, deprivation is a multi-dimensional, complex concept consisting of different but also inter-connected components, e.g., socio-economic status, physical morphology, environmental quality (Abascal et al., 2021; I. Baud et al., 2010; Thomson et al., 2020). This intrinsic complexity and intertwined nature of deprivation inhibit the researchers, local governments and any other stakeholders involved from arriving at a consensus on its definition, leading to the following investigation and characterisation of multiple deprivation lack consistency across or even within the LMICs cities. Therefore, the initial conceptualisation of multiple deprivation and the development of the related indicators directly determine the extent, magnitude, and dimensions of the resultant characterisation of deprivation.

As stated before, this study does not aim to exhaustively enumerate all the components and/or indicators related to deprivation, nor to deliver a universal definition that works omnipotently across any cities – both are impossible given the complexity of reality. However, we do attempt to propose a general conceptual paradigm, which includes the major domains of multiple deprivation, especially in LMICs, with the flexibility to be adapted to different contexts via allowing users to discard and/or insert some indicators based on the local knowledge. This objective has been obtained by leveraging the ‘IDEAMAPS Domains of Deprivation Framework’ (Abascal et al., 2021) to establish the inventory of multi-deprivation indicators for Nairobi, with some adaptation made to underscore the local contexts in (*Table 1*).

Despite the difficulty in conceptualising deprivation, another great challenge is the uncertainties in the input data for characterising deprivation. Aiming to provide a transferrable and scalable approach, we purposely establish the indicators from open databases, which inevitably bring some levels of uncertainties to the deprivation mapping results. In this study, though we initially collect 26 candidate indicators, only 15 of them are included in the deprivation characterisation after the quality check (*Table 6*). Some indicators (e.g., the DHS datasets) are removed due to their coarser resolution that leads to high uncertainties of the distribution at the grid level. Discarding unsuitable indicators helps to reduce the level of uncertainties and resultant biases, but meanwhile, the information richness of the model outputs regarding deprivation degrees also shrinks (in this case, mainly in the social-economic status). Therefore, we could only interpret and discuss the deprivation results based on the original magnitude of the input data and also realise that deprivation can only be partially measured due to the input data constraint. Moreover, high uncertainty is an unavoidable risk of using open data; nevertheless, with more open data becoming available and the development of advanced data mining techniques, deprivation mapping can still largely benefit from this data-rich environment (Mahabir et al., 2020), as long as the scope and uncertainty of input data are well concerned and informed by the researchers.

5.2. Measuring multiple deprivation

To characterise the multi-dimensionality of deprivation, this study employed an unsupervised PCA-based approach to produce a set of data-driven continuous measures of deprivation degrees. We named such PCA-based mapping products the ‘*multi-deprivation portfolio*’ (see *Figure 14* & *Figure 15*) to underscore the comprehensiveness and richness of information provided by the continuous indices. Most of the traditional approaches generate binary or multi-categorical mapping products of urban deprivation, ranging from pixel-, grid- to administrative-level (Agarwal et al., 2018; Kuffer et al., 2017; Kuffer, Pfeffer, Sliuzas, et al., 2016; Mahabir et al., 2020). Such products with discrete labels are very straightforward and communicable to the local planners and stakeholders; however, it often fails to further unveil the variation and severity within the deprivation classes, coupled with the lack of reflection on deprivation diversity. Therefore, we suggest that our proposed ‘*multi-deprivation portfolio*’ could fill in this gap by detailing deprivation on a continuous scale so that the value directly represents the deprivation degree. Besides, the continuous index is also versatile to be categorised back into different classes (e.g., deprived/non-deprived) if the users require a general binary mapping of deprived areas.

This research also reveals that measuring urban deprivation is sensitive to the analysing scale and thus needs to be carefully decided on (in this study, 100m grids). During PCA refinement, we first encountered difficulties in interpreting the mapping results because the indicators with broader spatial scales tend to mask off or distort the information from a relatively finer scale. This problem has been solved by retaining only the indicators with relatively fine scales. Yet, this does not necessarily mean that PCA is unsuitable for large scale analysis; instead, it is the inclusion of indicators that should always adapt to the study scale, e.g., for a national deprivation mapping, the discarded DHS indicators (*Table 6*) may be highly valuable. Besides, another key finding is that when we narrow to the urban built-up areas, the results significantly improve as more of the slum boundaries are captured. This improvement may imply that mapping deprivation at a relatively fine grid level needs to first distinguish the built-up areas.

Moreover, we would also like to stress the important role of the 100m standard grid in adding flexibility to product application, as well as the transferability and scalability of the PCA-based deprivation mapping method. Firstly, the gridded outputs can strongly support governmental decision-making activities by overlaying the grid cells of deprivation index with other crucial spatial layers. For example, the users may overlay the gridded deprivation index with flood hazard maps to identify the target areas exposed to floods, meanwhile also suffering from serious deprivation, thus prioritizing them to receive more investments and policy intervention. Secondly, as highlighted by the IDEAMAPS project (Kuffer et al., 2020), a gridded mapping system can work flexibly across different spatial scales, ranging from small area units, intra-city towards national or even global scale, by sensibly adjusting the size of the grid or aggregating the finer grid-level information into a larger scale. In this way, the gridded results can be easily integrated into other layers with varying scale, hence more transferable among LMICs cities. Moreover, a gridded deprivation index does not present a sharp boundary that clearly differentiates deprived and non-deprived areas as in the traditional administrative-units-based analysis, therefore avoiding the stigmatisation of deprived neighbourhood.

On the other hand, through the visual assessment of the ‘*multi-deprivation portfolio*’, we also realise the ‘trade-offs’ of information richness between the aggregate multi-deprivation index (*Figure 15*) and the sub-dimensional deprivation indices (*Figure 14*). By delivering a single aggregate map of multi-deprivation, the policymakers can quickly identify the urban areas facing high deprivation level, yet which specific domains contribute to the high deprivation degree among different areas still remains unclear. Thus, we suggest the users should be more cautious about the underlying reasons of the highly deprived areas on the aggregate map and always refer back to the sub-dimensional, easily interpretable indices. Additionally, in the index aggregation, another common challenge is “which weighting system should be applied?”. For this study,

we decided on equal weights considering the absence of local knowledge on the priority of different deprivation domains. Nonetheless, we still suggest that, given reliable local experts knowledge and engagements from the stakeholders, the researchers are encouraged to adopt different weights to highlight the most crucial deprivation dimension(s). For example, after the local validation with expert, we may suggest adding more weights on PC2 in the index aggregation to better inform the local planner, as PC2 is identified as the most accurate domain that represents the slum patterns in Nairobi.

5.3. Role of an EO-based method in deprivation mapping

In the second part of this study, we aim to train a deep CNN-based regression using only EO data to predict the continuous deprivation indices. Prior to the model training, the ‘*multi-deprivation portfolio*’ needs to be thoroughly interpreted to identify which PC should be first fed into the model, as the satellite images only contain physical and morphological features. The identification of target PC requires adequate knowledge of the multi-dimensional deprivation, which can be achieved via local validation, statistical analysis, and visual assessments with reference data. From the validation, we notice that PC2 is the most morphology-based deprivation domain that spatially captures the vast majority of slum areas in Nairobi, thus used as the first input. The CNN model yielded an R^2 of 0.6543, indicating a moderate-to-good performance in explaining the variance of PC2. Besides, the visual comparison also displayed a similar, consistent spatial pattern between the CNN prediction and PC2 index map (Figure 26). Therefore, we testify that an EO-based method has the potential to directly predict morphological deprivation degrees. This may open a new gate to the deprivation mapping community in generating scalable, transferrable, and ‘end-to-end’ mapping products of the continuous morphological deprivation index.

However, when we tried to train the proposed CNN on other deprivation indices and the aggregate index, none of the experiments was able to generate good results. We suggest that, by only relying on satellite images, a CNN-based regression model cannot explain the variation of other non-morphological deprivation domains in Nairobi. This finding is in line with the previous studies that EO-based approaches could only capture the physical and morphological domains of deprivation, and need to build a remote sensing proxy to model other aspects of deprivation (Kuffer et al., 2020; Taubenböck et al., 2018). A potential way to improve the model performance on non-morphological deprivation may be by introducing other supplementary data. For example, Ajami et al. (2019) added hand-crafted and GIS features to the trained CNN model to predict deprivation degrees of slum settlements in Bangalore, and the R^2 increased from 0.67 (using only EO data) data to 0.75. Notwithstanding, we may still argue whether adding additional hand-crafted feature layers is meaningful for an ‘end-to-end’ learning model like CNN because it may compromise the convenience and efficiency of the state-of-the-art approach.

Even though the EO-based methods fail to directly capture non-morphological deprivation, they still have the irreplaceable strengths that traditional survey- or field-based approaches cannot fully achieve, e.g., covering large areas, high efficiency and accuracy, scalability, and transferability (Kuffer et al., 2020; Kuffer, Pfeffer, & Sliuzas, 2016; Kuffer, Wang, et al., 2018). The important thing to further explore is “How can we use EO-derived information in helping multi-deprivation mapping?”. A clear yet not largely explored direction is to build the physical proxies that allow using image-based features to directly estimate deprivation degrees (Duque et al., 2015; Georganos et al., 2019; Jankowska et al., 2011; Stoler et al., 2012). This method is based on the assumption that deprived status would physically manifest itself in the imagery features in a distinct manner (Kuffer, Wang, et al., 2018). To bridge the gap between EO-based features and multiple deprivation, we suggest investigating more discriminant and robust EO variables that better reflect deprivation levels and jointly introducing ground-based data to state the uncertainties in the EO-based mapping products.

5.4. Limitations

The first limitation lies on the difficulties in providing a complete characterisation of deprivation, due to the inconsistency and quality of deprivation covariates data collected. The original spatial resolutions of the included indicators vary from 30m to nearly 1000m. Although all the input were resampled into the standard 100m resolution, the down-sampling of courser data would inevitably introduce some level of bias. Also, the output year of the indicators is not very consistent, among which the oldest one was the poverty information of 2008. However, given the limited data availability and the exploratory purpose of this study, the included datasets should be regarded as acceptable.

Secondly, the analysing unit of 100m grid itself decides that the output cannot further reveal the variation smaller than a grid, which may be argued by some researchers that such resolution is too coarse in mapping intra-urban deprivation. We indeed observed that the grided map of PC2 cannot really provide a sharp, explicit boundary between slum and non-slum (which is not our research intention) but rather shown in an obfuscating, dark-to-light spectrum. We also noticed that for those grids located around the edge of the slum boundary, the predicted deprivation levels are often underestimated due to the presence of other land covers with non-deprived features (e.g., roads, green space), thus leading to higher errors occurring on the grids intersected with the slum boundaries.

6. CONCLUSION AND RECOMMENDATION

6.1. Conclusions

The main objective of this research is to characterise and measure multiple deprivation by exploring the potential of an EO-based approach to directly capture the multi-dimensionality of urban deprivation. To achieve this, we proposed a two-step methodology that first characterises multiple deprivation as a set of continuous degrees and later on uses only VHR EO data to predict the established continuous deprivation degrees. The final conclusions are drawn by summarizing the key findings of each sub-objectives.

Regarding the first sub-objective of characterising multiple deprivation and measuring its spatial variation in continuous scales, we achieved it by first identifying the common deprivation domains using the IDEAMAPS framework (Abascal et al., 2021) and then applying PCA on 15 deprivation indicators at 100m standard grids. Through the PCA, we delivered a collection of gridded deprivation mapping products called the '*multi-deprivation portfolio*', each a continuous index between 0 to 1, indicating the degrees of deprivation in different domains. After the validation, we conclude that given suitable and sufficient data, PCA has its unique advantages on deconstructing the multi-dimensionality of deprivation into inter-independent, explicit sub-domains with few assumptions, as well as providing a data-driven continuous measure of deprivation level with more flexibility to support multiple user needs. In addition, PC2 was recognized as the most morphology-based deprivation domain, spatially capturing most of the slum configurations, thus recognised as the dominant deprivation pattern in Nairobi.

The second sub-objective seeks to explore the potential of EO-based methods to directly capture the intra-urban continuous deprivation degrees. To our best knowledge, this is the first study that attempts to apply a CNN-based regression model to directly estimate the deprivation degrees at an intra-urban level from VHR imagery features. Among the experiments, we observed promising performance of CNN in estimating PC2, with an R^2 of 0.6543. However, regarding other non-morphology-based deprivation domains, the proposed CNN model failed. Therefore, we conclude that an EO-based method has the ability to capture deprivation degrees in Nairobi, although only restricted to the morphological dimension.

The last sub-objective is to discuss the role of EO-based methods in deprivation mapping. Based on this research, we confirm that EO-based methods have unique strengths in capturing physical and morphological deprivation domains compared to traditional methods (e.g., census, filed-based mapping). In general, we suggest that the major roles of EO-based methods in deprivation mapping could focus on two aspects: (1) directly capturing the morphological deprivation degrees across different cities and scales; (2) developing more advanced and flexible physical proxies to infer or support the multi-deprivation characterisation.

6.2. Recommendations for further studies

Looking forward, as more spatial data related to urban deprivation, VHR EO images become available across LMICs, along with the development of advanced modelling techniques, we see a great potential of providing more accurate, routinely updated, and scalable mapping products of urban deprivation. Tiptoeing from this research, we would like to list a few potential directions for further studies:

1. We set the study scale of this research at an intra-city level, using 100m grids as analytical units. However, producing only city-level deprivation maps is not enough to support SDG11.1 and facilitate international cooperation. Therefore, we suggest future studies may scale up this approach into inter-city, national or even global scale by sensibly expanding the grid size.

2. It is worth applying this approach to other LMIC cities with varying context not only to test the transferability of our approach, but also to compare the deprivation pattern across cities. Given the large heterogeneity between deprived areas, the deprivation domains established by PCA may also differ strongly across cities. Such information can help us better understand the commonality and diversity of deprived areas, which is important to build a global deprivation inventory.
3. As a pilot study in leveraging the EO-based method to directly predict deprivation degrees, we did not really go in-depth to train and optimize an ‘excellent’ CNN model due to limited time and resources, which means that there is surely room to improve the model performance. Therefore, we strongly encourage researchers to further explore more advanced techniques and model architectures to improve prediction accuracy.

LIST OF REFERENCES

- Abascal, Á., Rothwell, N., Shonowo, A., Thomson, D. R., Elias, P., Elsey, H., Yeboah, G., & Kuffer, M. (2021). "Domains of Deprivation Framework" for Mapping Slums, Informal Settlements, and other Deprived Areas in LMICs to improve urban planning and policy: A Scoping Review. <https://doi.org/10.20944/preprints202102.0242.v2>
- Abdi, H., & Williams, L. J. (2010). Principal component analysis. In *Wiley Interdisciplinary Reviews: Computational Statistics* (Vol. 2, Issue 4, pp. 433–459). John Wiley & Sons, Inc. WIREs Comp Stat. <https://doi.org/10.1002/wics.101>
- Abson, D. J., Dougill, A. J., & Stringer, L. C. (2012). Using Principal Component Analysis for information-rich socio-ecological vulnerability mapping in Southern Africa. *Applied Geography*, 35(1–2), 515–524. <https://doi.org/10.1016/j.apgeog.2012.08.004>
- ACLED. (2020). *The Armed Conflict Location & Event Data Project*. <https://acleddata.com/data-export-tool/>
- Agarwal, S., Jakes, S., Essex, S., Page, S. J., & Mowforth, M. (2018). Disadvantage in English seaside resorts: A typology of deprived neighbourhoods. *Tourism Management*, 69, 440–459. <https://doi.org/10.1016/j.tourman.2018.06.012>
- Ajami, A., Kuffer, M., Persello, C., & Pfeffer, K. (2019). Identifying a Slums' Degree of Deprivation from VHR Images Using Convolutional Neural Networks. *Remote Sensing*, 11(11), 1282. <https://doi.org/10.3390/rs11111282>
- Alkire, S., Adriana, C., & Seth, S. (2014). Multidimensional poverty index 2014: Brief methodological note and results. *Oxford Poverty and Human Development Initiative (OPHI)*, June, 1–18. <http://www.ophi.org.uk/wp-content/uploads/Global-MPI-2014-Brief-Methodological-Note-and-Results.pdf?0a8fd7>
- Allik, M., Leyland, A., Ichihara, M. Y. T., & Dundas, R. (2020). Creating small-area deprivation indices: A guide for stages and options. *Journal of Epidemiology and Community Health*, 74(1), 20–25. <https://doi.org/10.1136/jech-2019-213255>
- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Van Essen, B. C., Awwal, A. A. S., & Asari, V. K. (2019). A State-of-the-Art Survey on Deep Learning Theory and Architectures. *Electronics*, 8(3), 292. <https://doi.org/10.3390/electronics8030292>
- Anurogo, W., Lubis, M. Z., Pamungkas, D. S., Hartono, & Ibrahim, F. M. (2017). A Spatial Approach to Identify Slum Areas in East Wara Sub-Districts, South Sulawesi. *IOP Conference Series: Earth and Environmental Science*, 98(1), 12030. <https://doi.org/10.1088/1755-1315/98/1/012030>
- APHRC. (2014). Population and Health Dynamics in Nairobi's Informal Settlements: Report of the Nairobi Cross-Sectional Slums Survey (NCSS) 2012. *Nairobi: APHRC*.
- Arimah, B. C. (2010). *The Face of Urban Poverty*. 2010(22).
- Arribas-Bel, D., Patino, J. E., & Duque, J. C. (2017). Remote sensing-based measurement of Living Environment Deprivation: Improving classical approaches with machine learning. *PLoS ONE*, 12(5). <https://doi.org/10.1371/journal.pone.0176684>
- Aungkulanon, S., Tangcharoensathien, V., Shibuya, K., Bundhamcharoen, K., & Chongsuvivatwong, V. (2017). Area-level socioeconomic deprivation and mortality differentials in Thailand: Results from principal component analysis and cluster analysis. *International Journal for Equity in Health*, 16(1), 1–12. <https://doi.org/10.1186/s12939-017-0613-z>
- Babu, G. S., Zhao, P., & Li, X. L. (2016). Deep convolutional neural network based regression approach for estimation of remaining useful life. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9642, 214–228. https://doi.org/10.1007/978-3-319-32025-0_14
- Basu, T., & Das, A. (2020). Formulation of deprivation index for identification of regional pattern of deprivation in rural India. *Socio-Economic Planning Sciences*, 100924. <https://doi.org/10.1016/j.seps.2020.100924>
- Baud, I., Kuffer, M., Pfeffer, K., Sliuzas, R., & Karuppannan, S. (2010). Understanding heterogeneity in metropolitan india: The added value of remote sensing data for analyzing sub-standard residential areas. *International Journal of Applied Earth Observation and Geoinformation*, 12(5), 359–374. <https://doi.org/10.1016/j.jag.2010.04.008>
- Baud, I. S. A., Pfeffer, K., Sridharan, N., & Nainan, N. (2009). Matching deprivation mapping to urban governance in three Indian mega-cities. *Habitat International*, 33(4), 365–377. <https://doi.org/10.1016/j.habitatint.2008.10.024>

- Baud, I., Sridharan, N., & Pfeffer, K. (2008). Mapping urban poverty for local governance in an Indian mega-city: The case of Delhi. *Urban Studies*, 45(7), 1385–1412. <https://doi.org/10.1177/0042098008090679>
- Baxter, M. J. (1995). Standardization and Transformation in Principal Component Analysis, with Applications to Archaeometry. *Applied Statistics*, 44(4), 513. <https://doi.org/10.2307/2986142>
- Bondarenko, M., Kerr, D., Sorichetta, A., Tatem, A., & WorldPop. (2020). *Census/projection-disaggregated gridded population datasets, adjusted to match the corresponding UNPD 2020 estimates, for 51 countries across sub-Saharan Africa using building footprints*. University of Southampton. <https://eprints.soton.ac.uk/443943/>
- Buja, K., & Menza, C. (2013). *Sampling design tool for ArcGIS: Instruction manual.[for ESRI ArcGIS 10.0 Service Pack 3 or higher]*. NOAA/National Centers for Coastal Ocean Science. <https://coastalscience.noaa.gov/project/sampling-design-tool-arcgis/>
- Burgert-Brucker, C. R., Dontamsetti, T., & Gething, P. W. (2018). The DHS Program’s Modeled Surfaces Spatial Datasets. *Studies in Family Planning*, 49(1), 87–92. <https://doi.org/10.1111/sifp.12050>
- Cabrera-Barona, P., & Ghorbanzadeh, O. (2018). Comparing Classic and Interval Analytical Hierarchy Process Methodologies for Measuring Area-Level Deprivation to Analyze Health Inequalities. *International Journal of Environmental Research and Public Health*, 15(1), 140. <https://doi.org/10.3390/ijerph15010140>
- Carstairs, V., & Morris, R. (1990). Deprivation and health in Scotland. *Health Bulletin*, 48(4), 162–175.
- Center for International Earth Science Information Network (CIESIN) Columbia University, & Novel-T. (2020). *GRID3 Benin Settlement Extents Version 01, Alpha*. <https://doi.org/10.7916/D8-3TN0-1686>
- Christ, K., Baier, K., & Azzam, R. (2016). Slums and informal housing in India: a critical look at official statistics with regard to water and sanitation. *Water International*, 41(2), 308–324. <https://doi.org/10.1080/02508060.2016.1139656>
- CRIB. (2021). *The ITC Geospatial Computing Portal*. <https://crib.utwente.nl/>
- Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly*, 84(2), 242–261. <https://doi.org/10.1111/1540-6237.8402002>
- Daniels, R., & Mulley, C. (2013). Explaining walking distance to public transport: The dominance of public transport supply. *Journal of Transport and Land Use*, 6(2), 5–20. <https://doi.org/10.5198/jtlu.v6i2.308>
- de Sherbin, A., & Bardy, G. (2016). Social vulnerability to floods in two coastal megacities: New York City and Mumbai. *Vienna Yearbook of Population Research*, 1(1), 131–165. <https://doi.org/10.1553/populationyearbook2015s131>
- Deas, I., Robson, B., Wong, C., & Bradford, M. (2003). Measuring Neighbourhood Deprivation: A Critique of the Index of Multiple Deprivation. *Environment and Planning C: Government and Policy*, 21(6), 883–903. <https://doi.org/10.1068/c0240>
- Dolk, H., Mertens, B., I Kleinschmidt, Walls, P., Shaddick, G., & Elliott, P. (1995). A standardisation approach to the control of socioeconomic confounding in small area studies of environment and health. *Journal of Epidemiology and Community Health*, 49(SUPPL. 2). https://doi.org/10.1136/jech.49.Suppl_2.S9
- Dooley, C., Boo, G., Leasure, D., Tatem, A., Bondarenko, M., & WorldPop. (2020). *Gridded maps of building patterns throughout sub-Saharan Africa, version 1.1*. University of Southampton. <https://doi.org/10.5258/SOTON/WP00677>
- Draper, N. R., & Smith, H. (2014). Applied regression analysis. In *Applied Regression Analysis*. Wiley. <https://doi.org/10.1002/9781118625590>
- Duque, J. C., Patino, J. E., Ruiz, L. A., & Pardo-Pascual, J. E. (2015). Measuring intra-urban poverty using land cover and texture metrics derived from remote sensing data. *Landscape and Urban Planning*, 135, 11–21. <https://doi.org/10.1016/j.landurbplan.2014.11.009>
- Eck, K. (2012). In data we trust? A comparison of UCDP GED and ACLED conflict events datasets. *Cooperation and Conflict*, 47(1), 124–141. <https://doi.org/10.1177/0010836711434463>
- Eksner, H. J. (2013). Revisiting the “ghetto” in the New Berlin Republic: Immigrant youths, territorial stigmatisation and the devaluation of local educational capital, 1999–2010. *Social Anthropology*, 21(3), 336–355. <https://doi.org/10.1111/1469-8676.12032>
- Engstrom, R., Sandborn, A., Yu, Q., Burgdorfer, J., Stow, D., Weeks, J., & Graesser, J. (2015, June 9). Mapping slums using spatial features in Accra, Ghana. *2015 Joint Urban Remote Sensing Event, JURSE 2015*. <https://doi.org/10.1109/JURSE.2015.7120494>
- Fekete, A. (2009). Validation of a social vulnerability index in context to river-floods in Germany. *Natural*

- Hazards and Earth System Science*, 9(2), 393–403. <https://doi.org/10.5194/nhess-9-393-2009>
- Fink, G., Arku, R., & Montana, L. (2012). The health of the poor: women living in informal settlements. *Ghana Medical Journal*, 46(2), 104–112.
- Frigerio, I., Ventura, S., Strigaro, D., Mattavelli, M., De Amicis, M., Mugnano, S., & Boffi, M. (2016). A GIS-based approach to identify the spatial variability of social vulnerability to seismic hazard in Italy. In *Applied Geography* (Vol. 74, pp. 12–22). Elsevier Ltd. <https://doi.org/10.1016/j.apgeog.2016.06.014>
- Georganos, S., Gadiaga, A. N., Linard, C., Grippa, T., Vanhuyse, S., Mboga, N., Wolff, E., Dujardin, S., & Lennert, M. (2019). Modelling the Wealth Index of Demographic and Health Surveys within Cities Using Very High-Resolution Remotely Sensed Information. *Remote Sensing*, 11(21), 2543. <https://doi.org/10.3390/rs11212543>
- Gilbert, A. (2007). The return of the slum: Does language matter? *International Journal of Urban and Regional Research*, 31(4), 697–713. <https://doi.org/10.1111/j.1468-2427.2007.00754.x>
- Gill, B. (2015). The English Indices of Deprivation 2015. In *Statistical Release*. <https://doi.org/10.1016/j.ajem.2017.08.036>
- Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. *Journal of Machine Learning Research*, 15, 315–323. <http://proceedings.mlr.press/v15/glorot11a.html>
- Gruebner, O., Sachs, J., Nockert, A., Frings, M., Khan, M. M. H., Lakes, T., & Hostert, P. (2014). Mapping the Slums of Dhaka from 2006 to 2010. *Dataset Papers in Science*, 2014, 1–7. <https://doi.org/10.1155/2014/172182>
- Hair, J. (2006). Multivariate Data Analysis (6th ed.). In *Pearson Prentice Hall* (6th ed.). <https://digitalcommons.kennesaw.edu/facpubs/2925>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *Unsupervised Learning* (pp. 485–585). Springer, New York, NY. https://doi.org/10.1007/978-0-387-84858-7_14
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-December*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Hecht-Nielsen, R. (1989). *Theory of the backpropagation neural network*. 593–605. <https://doi.org/10.1109/ijcnn.1989.118638>
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417–441. <https://doi.org/10.1037/h0071325>
- HOTOSM. (2020). *Humanitarian OpenStreetMap Team - HOT*. <https://data.humdata.org/organization/hot>
- Huntington, J. L., Hegewisch, K. C., Daudert, B., Morton, C. G., Abatzoglou, J. T., McEvoy, D. J., & Erickson, T. (2017). Climate engine: Cloud computing and visualization of climate and remote sensing data for advanced natural resource monitoring and process understanding. *Bulletin of the American Meteorological Society*, 98(11), 2397–2409. <https://doi.org/10.1175/BAMS-D-15-00324.1>
- Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Sage.
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *32nd International Conference on Machine Learning, ICML 2015, 1*, 448–456. <http://proceedings.mlr.press/v37/ioffe15.html>
- Ipsum, L., Sit, D., Rippin, N., Alkire, S., Foster, J. E., Seth, S., Santos, M. E., Roche, J. M. J. M. J. M., Ballon, P., Alkire, S., Foster, J. E., Seth, S., Santos, M. E., Roche, J. M. J. M. J. M., Ballón, P., Emma, M., Roche, J. M. J. M. J. M., Ballon, P., Gianni, B., ... Banks, N. (2015). Overview of Methods for Multidimensional Poverty Assessment years. In *Multidimensional Poverty Measurement and Analysis* (Vol. 45, Issues 2–3). <http://www.ophi.org.uk/measuring-multidimensional-poverty-in-latin-america-previous-experience-and-the-way-forward/%5Cnhttps://ideas.repec.org/p/qeh/ophiwp/ophiwp040.html%5Cnhttp://link.springer.com/article/10.1007/s10888-011-9210-3%5Cnhttp://link.springer>
- Jankowska, M. M., Weeks, J. R., & Engstrom, R. (2011). Do the most vulnerable people live in the worst slums? A spatial analysis of Accra, Ghana. *Annals of GIS*, 17(4), 221–235. <https://doi.org/10.1080/19475683.2011.625976>
- Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301), 790–794. <https://doi.org/10.1126/science.aaf7894>
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. In *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (Vol. 374,

- Issue 2065). Royal Society of London. <https://doi.org/10.1098/rsta.2015.0202>
- Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>
- Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A convolutional neural network for modelling sentences. *52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014 - Proceedings of the Conference*, 1, 655–665. <https://doi.org/10.3115/v1/p14-1062>
- Karanja, I. (2010). An enumeration and mapping of informal settlements in Kisumu, Kenya, implemented by their inhabitants. *Environment and Urbanization*, 22(1), 217–239. <https://doi.org/10.1177/0956247809362642>
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15. <https://arxiv.org/pdf/1412.6980.pdf> %22 entire document
- Kit, O., Lüdeke, M., & Reckien, D. (2012). Texture-based identification of urban slums in Hyderabad, India using remote sensing data. *Applied Geography*, 32(2), 660–667. <https://doi.org/10.1016/j.apgeog.2011.07.016>
- Kohli, D., Sliuzas, R., Kerle, N., & Stein, A. (2012). An ontology of slums for image-based classification. *Computers, Environment and Urban Systems*, 36(2), 154–163. <https://doi.org/10.1016/j.compenvurbsys.2011.11.001>
- Kohli, D., Sliuzas, R., & Stein, A. (2016). Urban slum detection using texture and spatial metrics derived from satellite imagery. In *Journal of Spatial Science* (Vol. 61, Issue 2, pp. 405–426). Mapping Sciences Institute Australia. <https://doi.org/10.1080/14498596.2016.1138247>
- Kraff, Nicolas J., Taubenbock, H., & Wurm, M. (2019, May 1). How dynamic are slums? EO-based assessment of Kibera's morphologic transformation. *2019 Joint Urban Remote Sensing Event, JURSE 2019*. <https://doi.org/10.1109/JURSE.2019.8808978>
- Kraff, Nicolas Johannes, Wurm, M., & Taubenbock, H. (2020). Uncertainties of human perception in visual image interpretation in complex urban environments. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 4229–4241. <https://doi.org/10.1109/JSTARS.2020.3011543>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 2, 1097–1105.
- Kuffer, M., Pfeffer, K., & Sliuzas, R. (2016). Slums from space-15 years of slum mapping using remote sensing. *Remote Sensing*, 8(6). <https://doi.org/10.3390/rs8060455>
- Kuffer, M., Pfeffer, K., Sliuzas, R., & Baud, I. (2016). Extraction of Slum Areas From VHR Imagery Using GLCM Variance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5), 1830–1840. <https://doi.org/10.1109/JSTARS.2016.2538563>
- Kuffer, M., Pfeffer, K., Sliuzas, R., Baud, I., & van Maarseveen, M. (2017). Capturing the diversity of deprived areas with image-based features: The case of Mumbai. *Remote Sensing*, 9(4). <https://doi.org/10.3390/rs9040384>
- Kuffer, M., Pfeffer, K., Sliuzas, R., Taubenbock, H., Baud, I., & Van Maarseveen, M. (2018). Capturing the Urban Divide in Nighttime Light Images from the International Space Station. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(8), 2578–2586. <https://doi.org/10.1109/JSTARS.2018.2828340>
- Kuffer, M., Thomson, D. R., Boo, G., Mahabir, R., Grippa, T., Vanhuyse, S., Engstrom, R., Ndugwa, R., Makau, J., Darin, E., de Albuquerque, J. P., & Kabaria, C. (2020). The Role of Earth Observation in an Integrated Deprived Area Mapping “System” for Low-to-Middle Income Countries. *Remote Sensing*, 12(6), 982. <https://doi.org/10.3390/rs12060982>
- Kuffer, M., Wang, J., Nagenborg, M., Pfeffer, K., Kohli, D., Sliuzas, R., & Persello, C. (2018). The scope of earth-observation to improve the consistency of the SDG slum indicator. In *ISPRS International Journal of Geo-Information* (Vol. 7, Issue 11, p. 428). MDPI AG. <https://doi.org/10.3390/ijgi7110428>
- Larsson, G., Maire, M., & Shakhnarovich, G. (2016). *FractalNet: Ultra-Deep Neural Networks without Residuals*. <http://arxiv.org/abs/1605.07648>
- Lathuiliere, S., Mesejo, P., Alameda-Pineda, X., & Horaud, R. (2020). A Comprehensive Analysis of Deep Regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(9), 2065–2081. <https://doi.org/10.1109/TPAMI.2019.2910523>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. <https://doi.org/10.1038/nature14539>
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2323. <https://doi.org/10.1109/5.726791>

- Leonita, G., Kuffer, M., Sliuzas, R., & Persello, C. (2018). Machine learning-based slum mapping in support of slum upgrading programs: The case of Bandung City, Indonesia. *Remote Sensing*, 10(10). <https://doi.org/10.3390/rs10101522>
- Li, L., Han, L., Ding, M., Liu, Z., & Cao, H. (2020). Remote Sensing Image Registration Based on Deep Learning Regression Model. *IEEE Geoscience and Remote Sensing Letters*. <https://doi.org/10.1109/LGRS.2020.3032439>
- Lilford, R., Kyobutungi, C., Ndugwa, R., Sartori, J., Watson, S. I., Sliuzas, R., Kuffer, M., Hofer, T., Porto De Albuquerque, J., & Ezech, A. (2019). Because space matters: Conceptual framework to help distinguish slum from non-slum urban areas. *BMJ Global Health*, 4(2), 1267. <https://doi.org/10.1136/bmjgh-2018-001267>
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 318–327. <http://arxiv.org/abs/1708.02002>
- Liu, R., Kuffer, M., & Persello, C. (2019). The temporal dynamics of slums employing a CNN-based change detection approach. *Remote Sensing*, 11(23). <https://doi.org/10.3390/rs11232844>
- Lloyd, C. T., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F. R., Gaughan, A. E., Nieves, J. J., Hornby, G., MacManus, K., Sinha, P., Bondarenko, M., Sorichetta, A., & Tatem, A. J. (2019). Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data*, 3(2), 108–139. <https://doi.org/10.1080/20964471.2019.1625151>
- Lucci, P., Bhatkal, T., & Khan, A. (2018). Are we underestimating urban poverty? *World Development*, 103, 297–310. <https://doi.org/10.1016/j.worlddev.2017.10.022>
- Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017). A review of supervised object-based land-cover image classification. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 130, pp. 277–293). Elsevier B.V. <https://doi.org/10.1016/j.isprsjprs.2017.06.001>
- Mahabir, R., Agouris, P., Stefanidis, A., Croitoru, A., & Crooks, A. T. (2020). Detecting and mapping slums using open data: a case study in Kenya. *International Journal of Digital Earth*, 13(6), 683–707. <https://doi.org/10.1080/17538947.2018.1554010>
- Mahabir, R., Croitoru, A., Crooks, A., Agouris, P., & Stefanidis, A. (2018). A Critical Review of High and Very High-Resolution Remote Sensing Approaches for Detecting and Mapping Slums: Trends, Challenges and Emerging Opportunities. *Urban Science*, 2(1), 8. <https://doi.org/10.3390/urbansci2010008>
- Mahabir, R., Crooks, A., Croitoru, A., & Agouris, P. (2016). The study of slums as social and physical constructs: challenges and emerging research opportunities. *Regional Studies, Regional Science*, 3(1), 399–419. <https://doi.org/10.1080/21681376.2016.1229130>
- Makau, J., Dobson, S., & Samia, E. (2012). The five-city enumeration: The role of participatory enumerations in developing community capacity and partnerships with government in Uganda. *Environment and Urbanization*, 24(1), 31–46. <https://doi.org/10.1177/0956247812438368>
- Mapillary. (2021). *Mapillary - Street-level imagery, powered by collaboration and computer vision*. <https://www.mapillary.com/app/?lat=20&lng=0&z=1.5>
- Martínez, J., Pfeffer, K., & Baud, I. (2016). Factors shaping cartographic representations of inequalities. Maps as products and processes. *Habitat International*, 51, 90–102. <https://doi.org/10.1016/j.habitatint.2015.10.010>
- Maxar. (2021). *Ecopia Building Footprints powered by Maxar*. <https://www.maxar.com/products/building-footprints>
- Mboga, N., Persello, C., Bergado, J. R., & Stein, A. (2017). Detection of informal settlements from VHR images using convolutional neural networks. *Remote Sensing*, 9(11), 1106. <https://doi.org/10.3390/rs9111106>
- Mishra, S. V. (2018). Urban deprivation in a global south city-a neighborhood scale study of Kolkata, India. *Habitat International*, 80, 1–10. <https://doi.org/10.1016/j.habitatint.2018.08.006>
- Mulei, S. (2012). Water Quality Degradation Trends in Kenya over the Last Decade. In *Water Quality Monitoring and Assessment*. InTech. <https://doi.org/10.5772/32010>
- Nelson, J. K., & Brewer, C. A. (2017). Evaluating data stability in aggregation structures across spatial scales: revisiting the modifiable areal unit problem. *Cartography and Geographic Information Science*, 44(1), 35–50. <https://doi.org/10.1080/15230406.2015.1093431>
- Nolan, L. B. (2015). Slum Definitions in Urban India: Implications for the Measurement of Health Inequalities. *Population and Development Review*, 41(1), 59–84. <https://doi.org/10.1111/j.1728->

4457.2015.00026.x

- O'Shea, K., & Nash, R. (2015). *An Introduction to Convolutional Neural Networks*.
<http://arxiv.org/abs/1511.08458>
- Ogut, F., Kimata, D., & Kweyu, R. (2019). *A Spatial Analysis of Unplanned and Mushrooming Dumpsites and Environmental Governance in Nairobi County*. 9(10). <https://doi.org/10.7176/JEES/9-10-17>
- OpenStreetMap. (2020). <https://www.openstreetmap.org/>
- Patel, A., Koizumi, N., & Crooks, A. (2014). Measuring slum severity in Mumbai and Kolkata: A household-based approach. *Habitat International*, 41, 300–306.
<https://doi.org/10.1016/j.habitatint.2013.09.002>
- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), 559–572.
<https://doi.org/10.1080/14786440109462720>
- 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>
- Prabhu, R., Parvathavarthini, B., & Alaguraja, A. R. (2021). Integration of deep convolutional neural networks and mathematical morphology-based postclassification framework for urban slum mapping. *Journal of Applied Remote Sensing*, 15(01). <https://doi.org/10.1117/1.jrs.15.014515>
- Pratomo, J., Kuffer, M., Martinez, J., & Kohli, D. (2017). Coupling uncertainties with accuracy assessment in object-based slum detections, case study: Jakarta, Indonesia. *Remote Sensing*, 9(11), 1164.
<https://doi.org/10.3390/rs9111164>
- Pyo, J. C., Duan, H., Baek, S., Kim, M. S., Jeon, T., Kwon, Y. S., Lee, H., & Cho, K. H. (2019). A convolutional neural network regression for quantifying cyanobacteria using hyperspectral imagery. *Remote Sensing of Environment*, 233, 111350. <https://doi.org/10.1016/j.rse.2019.111350>
- Ren, X., Li, T., Yang, X., Wang, S., Ahmad, S., Xiang, L., Stone, S. R., Li, L., Zhan, Y., Shen, Di., & Wang, Q. (2019). Regression Convolutional Neural Network for Automated Pediatric Bone Age Assessment from Hand Radiograph. *IEEE Journal of Biomedical and Health Informatics*, 23(5), 2030–2038. <https://doi.org/10.1109/JBHI.2018.2876916>
- Roy, D., Bernal, D., & Lees, M. (2020). An exploratory factor analysis model for slum severity index in Mexico City. *Urban Studies*, 57(4), 789–805. <https://doi.org/10.1177/0042098019869769>
- Ruktanonchai, C. W., Ruktanonchai, N. W., Nove, A., Lopes, S., Pezzulo, C., Bosco, C., Alegana, V. A., Burgert, C. R., Ayiko, R., Charles, A. S., Lambert, N., Msechu, E., Kathini, E., Matthews, Z., & Tatem, A. J. (2016). Equality in Maternal and Newborn Health: Modelling Geographic Disparities in Utilisation of Care in Five East African Countries. *PLOS ONE*, 11(8), e0162006.
<https://doi.org/10.1371/journal.pone.0162006>
- Schmidt, M. C., Deutsch, R. C., Piegorsch, W. W., & Cutter, S. L. (2008). A Sensitivity Analysis of the Social Vulnerability Index. *Risk Analysis*, 28(4), 1099–1114. <https://doi.org/10.1111/j.1539-6924.2008.01072.x>
- Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks For Large-scale Image Recognition*.
<http://www.robots.ox.ac.uk/>
- SLUMAP. (2020). *SLUMAP research project - Mapping slums with remote sensing*. <https://slumap.ulb.be/>
- Song, J., Gao, S., Zhu, Y., & Ma, C. (2019). A survey of remote sensing image classification based on CNNs. *Big Earth Data*, 3(3), 232–254. <https://doi.org/10.1080/20964471.2019.1657720>
- Stoler, J., Daniels, D., Weeks, J., Stow, D., Coulter, L., & Finch, B. (2012). Assessing the utility of satellite imagery with differing spatial resolutions for deriving proxy measures of slum presence in Accra, Ghana. *GIScience and Remote Sensing*, 49(1), 31–52. <https://doi.org/10.2747/1548-1603.49.1.31>
- Subbaraman, R., O'Brien, J., Shitole, T., Shitole, S., Sawant, K., Bloom, D. E., & Patil-Deshmukh, A. (2012). Off the map: the health and social implications of being a non-notified slum in India. *Environment and Urbanization*, 24(2), 643–663. <https://doi.org/10.1177/0956247812456356>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 07-12-June-2015, 1–9.
<https://doi.org/10.1109/CVPR.2015.7298594>
- Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics, 5th ed. In *Using multivariate statistics*, 5th ed. Allyn & Bacon/Pearson Education.
- Tatem, A., Peter, G., Bhatt, S., Dan, W., & Carla, P. (2013). *WorldPop :: Development and Health Indicators*.
<https://doi.org/10.5258/SOTON/WP00127>

- Taubenböck, H., & Kraff, N. J. (2014). The physical face of slums: A structural comparison of slums in Mumbai, India, based on remotely sensed data. *Journal of Housing and the Built Environment*, 29(1), 15–38. <https://doi.org/10.1007/s10901-013-9333-x>
- Taubenböck, H., Kraff, N. J., & Wurm, M. (2018). The morphology of the Arrival City - A global categorization based on literature surveys and remotely sensed data. *Applied Geography*, 92, 150–167. <https://doi.org/10.1016/j.apgeog.2018.02.002>
- Taubenböck, H., Wurm, M., Setiadi, N., Gebert, N., Roth, A., Strunz, G., Birkmann, J., & Dech, S. (2009). Integrating remote sensing and social science: The correlation of urban morphology with socioeconomic parameters. *2009 Joint Urban Remote Sensing Event*. <https://doi.org/10.1109/URS.2009.5137506>
- Thomson, D. R., Kuffer, M., Boo, G., Hati, B., Grippa, T., Elsey, H., Linard, C., Mahabir, R., Kyobutungi, C., Maviti, J., Mwaniki, D., Ndugwa, R., Makau, J., Sliuzas, R., Cheruiyot, S., Nyambuga, K., Mboga, N., Kimani, N. W., de Albuquerque, J. P., & Kabaria, C. (2020). Need for an integrated deprived area “slum” mapping system (IDEAMAPS) in low-and middle-income countries (LMICS). *Social Sciences*, 9(5), 80. <https://doi.org/10.3390/SOCSCI9050080>
- Tigre, G. (2018). *Multidimensional Poverty and Its Dynamics in Ethiopia* (pp. 161–195). Springer, Singapore. https://doi.org/10.1007/978-981-10-8126-2_8
- Toshev, A., & Szegedy, C. (2014). DeepPose: Human pose estimation via deep neural networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1653–1660. <https://doi.org/10.1109/CVPR.2014.214>
- Tusting, L. S., Bisanzio, D., Alabaster, G., Cameron, E., Cibulskis, R., Davies, M., Flaxman, S., Gibson, H. S., Knudsen, J., Mbogo, C., Okumu, F. O., von Seidlein, L., Weiss, D. J., Lindsay, S. W., Gething, P. W., & Bhatt, S. (2019). Mapping changes in housing in sub-Saharan Africa from 2000 to 2015. *Nature*, 568(7752), 391–394. <https://doi.org/10.1038/s41586-019-1050-5>
- UN-Habitat. (2003). *The Challenge of Slums—Global Report on Human Settlements*. [https://www.un.org/ruleoflaw/files/Challenge of Slums.pdf](https://www.un.org/ruleoflaw/files/Challenge%20of%20Slums.pdf)
- UN-Habitat. (2011). *UN-habitat and the Kibera slum upgrading initiative 2011*. <https://unhabitat.org/books/un-habitat-and-kenya-slum-upgrading-programme-kensup/>
- UN-Habitat. (2016). Slum almanac 2015–2016: Tracking improvement in the lives of slum dwellers. *Participatory Slum Upgrading Programme*.
- United Cities and Local Governments (UCLG). (2018). *Towards the localization of the SDGs*. https://www.uclg.org/sites/default/files/towards_the_localization_of_the_sdgs.pdf
- United Nations. (2015). *The Millennium Development Goals Report 2015* (Issue June). <https://www.undp.org/content/undp/en/home/librarypage/mdg/the-millennium-development-goals-report-2015.html>
- United Nations. (2018). *The Sustainable Development Goals Report 2018*. <https://unstats.un.org/sdgs/files/report/2018/TheSustainableDevelopmentGoalsReport2018-EN.pdf>
- United Nations. (2019). World Urbanization Prospects: The 2018 Revision. In *World Urbanization Prospects: The 2018 Revision*. <https://doi.org/10.18356/b9e995fe-en>
- Van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., Lyapustin, A., Sayer, A. M., & Winker, D. M. (2016). Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. *Environmental Science and Technology*, 50(7), 3762–3772. <https://doi.org/10.1021/acs.est.5b05833>
- Vanhuyse, S., Georganos, S., Kuffer, M., Grippa, T., & Wolff, E. (2021). *Gridded urban deprivation probability from open optical imagery and dual-pol SAR data*. 1, 0–3. https://igarss2021.com/view_paper.php?PaperNum=3754
- Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: How to use principal components analysis. *Health Policy and Planning*, 21(6), 459–468. <https://doi.org/10.1093/heapol/czl029>
- Walach, E., & Wolf, L. (2016). Learning to count with CNN boosting. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9906 LNCS, 660–676. https://doi.org/10.1007/978-3-319-46475-6_41
- Wang, J., Kuffer, M., Roy, D., & Pfeffer, K. (2019). Deprivation pockets through the lens of convolutional neural networks. *Remote Sensing of Environment*, 234, 111448. <https://doi.org/10.1016/j.rse.2019.111448>
- Whelan, C. T., & Maître, B. (2005). Vulnerability and multiple deprivation perspectives on economic

- exclusion in Europe: A latent class analysis. *European Societies*, 7(3), 423–450. <https://doi.org/10.1080/14616690500194050>
- Williams, T. K. A., Wei, T., & Zhu, X. (2020). Mapping Urban Slum Settlements Using Very High-Resolution Imagery and Land Boundary Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 166–177. <https://doi.org/10.1109/JSTARS.2019.2954407>
- Wurm, M., & Taubenböck, H. (2018). Detecting social groups from space – Assessment of remote sensing-based mapped morphological slums using income data. *Remote Sensing Letters*, 9(1), 41–50. <https://doi.org/10.1080/2150704X.2017.1384586>
- Wurm, Michael, Stark, T., Zhu, X. X., Weigand, M., & Taubenböck, H. (2019). Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 150, 59–69. <https://doi.org/10.1016/j.isprsjprs.2019.02.006>
- Xu, Y., & Goodacre, R. (2018). On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. *Journal of Analysis and Testing*, 2(3), 249–262. <https://doi.org/10.1007/s41664-018-0068-2>
- Yuan, Y., & Wu, F. (2014). The development of the index of multiple deprivations from small-area population census in the city of Guangzhou, PRC. *Habitat International*, 41, 142–149. <https://doi.org/10.1016/j.habitatint.2013.07.010>
- Zhang, X. Q. (2016). The trends, promises and challenges of urbanisation in the world. *Habitat International*, 54, 241–252. <https://doi.org/10.1016/j.habitatint.2015.11.018>

APPENDIX

Annex 1: The records of all conducted experiments for hyper-parameter tuning.

Order	Hyper-parameter	Value	MAE		RMSE		R ²	
			Training	Validation	Training	Validation	Training	Validation
1	Batch size	32	0.0375	0.0420	0.0556	0.0615	0.6668	0.6282
		64	0.0363	0.0427	0.0529	0.0612	0.6979	0.6315
		128	0.0324	0.0412	0.0479	0.0594	0.7525	0.6526
		256	0.0069	0.0419	0.0129	0.0585	0.9821	0.6632
		512	0.0262	0.0425	0.0395	0.0599	0.8319	0.6469
2	Learning rate	0.005	0.0341	0.0413	0.0495	0.0590	0.7357	0.6574
		0.001	0.0069	0.0419	0.0129	0.0585	0.9821	0.6632
		0.0001	0.0048	0.0428	0.0090	0.0599	0.9913	0.6472
		0.0005	0.0070	0.0454	0.0108	0.0647	0.9875	0.5876
3	Loss function	MAE	0.0069	0.0419	0.0129	0.0585	0.9821	0.6632
		MSE	0.0373	0.0428	0.0486	0.0590	0.7447	0.6574
		Log-cosh	0.0089	0.0427	0.0105	0.0602	0.9880	0.6429
		Attention loss (1.2)	0.0238	0.0418	0.0332	0.0597	0.8809	0.6487
		Attention loss (1.15)	0.0034	0.0419	0.0054	0.0588	0.9969	0.6599
		Attention loss (1.1)	0.0065	0.0422	0.0105	0.0585	0.9881	0.6637