

**A machine learning-guided data integration framework to measure
multidimensional poverty**

by

TAPERA MUSUNGWA

submitted in accordance with the requirements for
the degree of

MAGISTER TECHNOLOGIAE

In the subject of

INFORMATION TECHNOLOGY

at the

University of South Africa

Supervisor: Professor Ernest Mnkandla

March 2021

Declaration

Name: TAPERA MUSUNGWA

Student number: 44814526

Degree: MAGISTER TECHNOLOGIAE IN INFORMATION TECHNOLOGY

Exact wording of the title of the dissertation as appearing on the copies submitted for examination:

A MACHINE LEARNING-GUIDED DATA INTEGRATION FRAMEWORK TO MEASURE

MULTIDIMENSIONAL POVERTY

The author declares that the above dissertation is my own work and that all sources that I have used or quoted have been indicated and acknowledged by means of complete references.



SIGNATURE

25/03/2021

DATE

Abstract

As developing nations like South Africa chart a path of socio-economic development, the spatialisation of progress, opportunity, and neglect is a critical antecedent to policy-making and regional interventionism. Efforts to capture meaningful data using household surveys and censuses face a diluted accuracy due to sampling, surveying, and quantification errors. The reliability and regularity of these traditional methods is also constrained since the processes are costly and time-consuming. Recent investigations in the field of machine learning and satellite imaging have presented a viable proof-of-concept technique to exploit specific economic indicators to demonstrate economic development patterns across regional areas. The current study adopts several interrelated approaches encompassed within the field of remote sensing in order to evaluate and model poverty in the South African landscape. By adopting publicly accessible information for classification to indicate the intensity of poverty, this study proposed an inexpensive solution to poverty estimation. Concretely, the solution combined satellite imagery and geospatial data with regional poverty data exploiting an ensemble approach to poverty diagnosis. The solution is based upon multidimensional indicators and multi-layered insights that can be extrapolated from overlapping models to bolster them and help with socio-economic well-being estimations.

Through machine learning techniques and object-oriented training of a convolutional neural network, this study revealed that a naïve combination of distinct data sources shows patterns of socio-economic well-being in South Africa by achieving an R^2 of 0.56 wealth estimation compared to 0.54 from satellite imagery. This outlined variability and incongruity within landscapes that not only reflect the persistent subdivisions of apartheid-era enclavisation, but indicate critical gaps in domestic social services, infrastructure, and developmental pathways. This study is applicable to policy makers in low- and middle-income countries that lack accurate and timely data on economic development as an important precursor to public support, policy making, and planned expenditures.

Keywords: Poverty prediction, Machine learning algorithm, Satellite imagery, Deep learning, Convolutional neural network, Computer vision, Geospatial information, Night-time lights, Data for development.

Acknowledgements

First of all, I want to sincerely acknowledge my supervisor, Prof. Ernest Mnkandla, to whom I owe gratitude for his guidance, expertise, feedback, and encouragement throughout my studies. You are an awesome mentor! Thank you for believing that this work would be successful. I would not have achieved this much without your support.

I dedicate this dissertation to my parents and family, whose love and guidance made me the person I am today. To my mother who is receiving chemotherapy for colorectal cancer, I pray for your healing. This could not come in a worse year, when I cannot travel to be with you. To my friends and colleagues, thank you for your never-ending support throughout the way. My journey is a dream full of blessings. It could not have been better: thanks be to God. Lastly, I would like to thank the University of South Africa for granting me the opportunity to conduct my studies. The greatest outcome from this research was the discovery of myself.

Table of Contents

Declaration	ii
Abstract	iii
Acknowledgements	v
Table of Contents	vi
List of Figures	ix
List of Tables	xi
List of Acronyms or Abbreviations	xii
Citation management and reference method	xiv
Chapter 1. Purpose and rationale of the study	1
1.1 Introduction _____	1
1.2 Research background _____	1
1.3 Research rationale _____	3
1.4 Problem statement _____	4
1.5 Research question _____	6
1.6 Research objectives _____	7
1.7 Research methodology _____	7
1.8 Limitations and delimitations of the study _____	9
1.9 Dissertation outline _____	10
1.10 Chapter summary _____	11
Chapter 2. Review of the literature	13
2.1 Introduction _____	13
2.2 Measurement of poverty _____	13
2.3 Geospatial information systems and remote sensing _____	15
2.4 Detection of poverty using machine learning _____	17
2.5 Nightlight analysis and spatial imaging _____	20
2.6 Other remote sensing solutions _____	26
2.7 A background overview of machine learning _____	30
2.7.1 Deep learning	31
2.7.2 Convolutional Neural Network (CNN)	32
2.7.3 Transfer learning.....	35
2.8 Conceptual framework _____	36
2.9 Chapter summary _____	38
Chapter 3. Research methodology	40
3.1 Introduction _____	40
3.2 Research paradigm _____	41

3.3 Research approach	41
3.4 Research strategy	42
3.4.1 Design science research	44
3.5 Target country	48
3.6 Data sources	50
3.6.1 Demographic and health survey	51
3.6.2 Night-time luminosity	52
3.6.3 Daytime satellite imagery	54
3.6.4 OpenStreetMap data	55
3.7 Research experiment	56
3.7.1 Satellite imagery approach	58
3.7.2 Crowd-sourced information approach	59
3.7.3 Classification and regression models	61
3.8 Metrics and evaluation	63
3.8.1 R-squared value	63
3.8.2 Root mean squared error	64
3.8.3 Cross-validation	65
3.8.4 Data analysis and assessment	66
3.8.5 Descriptive statistics	67
3.8.6 Poverty indicators and calculations	67
3.8.7 Reliability and validity	68
3.9 Ethical considerations	69
3.10 Chapter summary	70
<i>Chapter 4. Implementation and results</i>	71
4.1 Introduction	71
4.2 Data collection	71
4.3 Pre-processing	72
4.3.1 Demographic and health survey	72
4.3.2 Night-time lights	75
4.3.3 Daytime satellite imagery	76
4.3.4 OpenStreetMap	77
4.4 Experimentation	80
4.4.1 Experiment setup	81
4.4.2 Satellite imagery approach	83
4.4.3 Crowd-sourced information approach	88
4.5 Feature extraction	90
4.6 Chapter summary	93
<i>Chapter 5. Discussion and analysis</i>	94
5.1 Introduction	94
5.2 Satellite imagery approach	94
5.3 Crowd-sourced information approach	98
5.4 Features and further evaluation	101

5.4.1 R-squared	101
5.4.2 Root Mean Square Error	102
5.4.3 Descriptive statistics	102
5.5 Poverty mapping _____	107
5.6 Reliability, validity, and generalisability _____	109
5.7 Chapter summary _____	111
<i>Chapter 6. Conclusions and recommendations</i>	<i>113</i>
6.1 Introduction _____	113
6.2 Overview of the study _____	113
6.3 Revisiting the problem statement and research questions _____	114
6.4 Summary of contributions _____	116
6.5 Concluding remarks _____	117
6.6 Recommendations and future research _____	120
6.7 Limitations _____	122
6.8 Chapter summary _____	122
<i>References</i>	<i>124</i>
<i>Appendices</i>	<i>134</i>
Appendix A: Ethical Clearance Certificate _____	134
Appendix B: Language-editing confirmation _____	137
Appendix C: Turnitin Report _____	138

List of Figures

Figure 1.1: Design science research activities, according to Peffers et al. (2007, p. 58, Figure 1).....	8
Figure 2.1: Three-model comparison of daylight (top), nightlight (middle) and filtered (bottom) assets across (left to right) urban, nonurban, water and roads (Jean et al. 2016, p. 792, Figure 2).....	21
Figure 2.2: Ethnic group nationality partitioned luminosity measures (Michalopoulos & Papaioannou 2014, p. 154, Figure IIB).....	24
Figure 2.3: Light-recorded long-term growth on Korean Peninsula (Henderson et al. 2012, p. 1002, Figure 2).....	26
Figure 2.4: Correlation of typeface with amenity type (Ma et al. 2019, p. 3, Figure 2).....	29
Figure 2.5: A 3-layer neural network with three inputs, two hidden layers of four neurons each, and one output layer (Karpathy 2020, p. 1). The connections (synapses) between neurons are across layers but not within a layer.....	32
Figure 2.6: A 16-layer VGG16 pipeline model architecture for $224 \times 224 \times 3$ image sizes, with 1000 classes in the softmax layer (rightmost). It shows all the stages in the layer, including the convolutional, pooling, and a fully connected final layer (ul Hassan 2018).....	34
Figure 2.7: Transfer learning methods (Yamashita et al. 2018, p. 621, Figure 10)....	35
Figure 2.8: Conceptual framework	38
Figure 3.1: Design science research cycles (Hevner 2007, p. 88, Figure 1).....	43
Figure 3.4: Design science research activities, according to Peffers et al. (2007, p. 58).....	45
Figure 3.2: Visual distinction between upper middle class and poor residential areas in South Africa (The Guardian 2016)	50
Figure 3.3: Night-time imagery for Cape Town, South Africa (NOAA 2020)	54
Figure 3.5: Research organization	57
Figure 3.6: Satellite-based transfer learning model.....	59
Figure 3.7: OSM model for the study	60
Figure 3.8: 5-fold cross-validation data split	65
Figure 4.1: Python code for DHS dataset objects count.....	73
Figure 4.2: South African DHS clusters.....	74

Figure 4.3: South African night-time luminosity data (left) and night-time data merged with DHS coordinates for Gauteng province (right)	76
Figure 4.4: Sample Google satellite image for South Africa downloaded from (Google 2020)	77
Figure 4.5: Visualised OSM data	78
Figure 4.6: QGIS OSM datapoints showing roads (blue), POI (green), buildings (orange), and DHS clusters (red dots)	80
Figure 4.7: Satellite imagery approach steps	83
Figure 4.8: Examples of daytime satellite images corresponding to light intensities	85
Figure 4.9: Data split between training and validation	86
Figure 4.10: Learning plot	87
Figure 4.11: Night-time classification saliency maps generated from satellite imagery	87
Figure 4.12: Sixteen-layer VGG16 model with parameters for this study	91
Figure 4.13: Model of the CNN architecture	93
Figure 5.1: South Africa ground-truth wealth index and the model poverty prediction results for the study	97
Figure 5.2: Results of the South African ground truth wealth index and poverty prediction model	100
Figure 5.3: Night-time lights summary data	103
Figure 5.4: Average night-time light distribution per cluster	104
Figure 5.5: Distribution of the light intensity at night per cluster	104
Figure 5.6: Relationship between Night-Time Light Luminosity and DHS Indicators	105
Figure 5.7: t-SNE plot for night-time lights	106
Figure 5.8: Spearman and Pearson’s correlation for important features	107
Figure 5.9: District-level ground truth and predicted wealth indices	108
Figure 5.10: Regional localisation of domestic poverty (South Africa Gateway 2019, p. 1)	109

List of Tables

Table 4.1: Training data setup	72
Table 4.2: Descriptive statistics of the DHS clusters.....	74
Table 4.3: OSM feature categories	79
Table 4.4: AWS instance specifications	81
Table 4.5: Software and library versions	82
Table 4.6: Satellite imagery and distribution between training and validation sets	84
Table 5.1: Satellite imagery regression results comparison table.....	96
Table 5.2: Comparison table of OSM and VIIR data regression results	98
Table 5.3: R2 output comparison	101
Table 5.4: Gaussian mix model image bin assignment	103

List of Acronyms or Abbreviations

AI	Artificial Intelligence
ANOVA	Analysis of Variance
API	Application Programming Interface
AWS	Amazon Web Services
CNN	Convolutional Neural Network
CSV	Comma Separated-Values
DHS	Demographic and Health Survey
EPSG	European Petroleum Survey Group
GBDT	Gradient Boosting Decision Tree
GBR	Gradient Boosting Regressor
GCP	Google Cloud Platform
GDP	Gross Domestic Product
GIS	Geographic Information System
GPS	Global Positioning System
GPU	Graphical Processing Unit
GSV	Google Street View
GUI	Graphical User Interface
IPUMS	Integrated Public Use Microdata Series
KDD	Knowledge Discovery in Databases
LED	Living Environment Deprivation
ML	Machine Learning
MPI	Multidimensional Poverty Index
NCEI	National Centre for Environmental Information
NOAA	National Oceanic and Atmospheric Administration
OBIA	Object-Based Image Analysis
OSM	OpenStreetMap
POI	Point-of-Interest
QGIS	Quantum Geographic Information Systems
ReLU	Rectified Linear Unit
RGB	Red, Green, Blue
RMSE	Root Mean Squared Error
S-NPP	Suomi National Polar-orbiting Partnership
SAR	Spatial Auto-Regressive

t-SNE	t-distributed Stochastic Neighbour Embedding
TIFF	Tag Image File Format
UK	United Kingdom
UNDP	United Nations Development Program
UNISA	University of South Africa
UNSDG	United Nations Sustainable Development Goals
USAID	United States Agency for International Development
USD	United States Dollar
VGG	Visual Geometry Group
VGI	Volunteered Geographic Information
VIIRS DNB	Visible Infrared Imaging Radiometry Suite Day/Night Band
ZAR	South African Rand

Citation management and reference method

References were managed electronically using Mendeley. The Monash University Harvard citation style was used for referencing throughout this dissertation. Each contribution and quotation in this dissertation derived from the work of others is attributed, cited, and referenced.

Chapter 1. Purpose and rationale of the study

1.1 Introduction

This chapter delineates the scope of the study and its intended contribution to the discourse. Central to the study is an endeavour to assess a variety of methods of measuring the poverty data gap using remote sensing, crowd-sourced information, night-time lights, and daytime satellite imagery in addition to socio-economic survey datasets.

1.2 Research background

As the global community continues to strive towards a more equitable standard of poverty eradication, the ability to assess, monitor, and manage welfare concerns is inherently linked to the ability to accurately and consistently measure the social effects of domestic poverty. From an intervention perspective, Jean et al. (2016, p. 790) remind us that measuring the economic characteristics of residents accurately influences research and policy in a critical way, suggesting that gaps and limitations in data collection and reliability can lead to inconsistencies in support and services for the most vulnerable global populations. While key poverty indicators such as the United Nations Sustainable Development Goals (UNSDG) provide a tangible basis for the measurement and benchmarking of related factors, Ayush et al. (2020, p. 1) report that most assessments are based upon costly, time-consuming processes such as household surveys. Due to such restrictions, the regularity and reliability of these surveys has been constrained, with even the most surveyed nations in Africa, like Uganda, conducting household surveys only every few years and only for a small representative population of villages or regional populations (Ayush et al. 2020, p. 1). As governments commit a broader range of resources to the diagnoses and mitigation of poverty, the breadth of regional surveying and analysis must scale upwards, encompassing a much more inclusive standard of practice in order to ensure that administrative strategies and investments are appropriate and effective.

Whilst the ubiquity of big data and spatial imaging datasets has provided researchers with unprecedented access to on-demand analysis, Xie et al. (2016, p. 2) argue that the concentration of these indicators amongst developed, wealthier nations has left developed countries lagging behind in a new digital divide. Accordingly, Zhao et al. (2019, p. 14) and Tingzon et al. (2019, p. 425) reflect that technological advances in geospatial information, computer-mediated search capabilities, and machine learning have enabled innovative approaches for estimating socio-economic indicators. By combining machine learning and geospatial information that is often retrievable from open source datasets, it becomes possible to chart poverty indicators in real time and across longitudinal models (Ayush et al. 2020, p. 1; Tingzon et al. 2019, p. 425; Zhao & Kusumaputri 2016, p. 1; Zhao et al. 2019, p. 1). It is through a low-cost, high-fidelity inspection of spatial patterns via remote sensing as well as the accuracy afforded by machine learning insights and capabilities that poverty datasets are becoming increasingly accessible to a broader spectrum of both developed and developing nations.

Central to the underlying motivations surrounding remote sensing and spatio-structural analysis of systemic inequalities are the cost-related hurdles of social surveys and regionalised investigations, many of which take years to plan and, in some cases, such as South Africa, decades to implement (Gebru et al. 2017, p. 1; South Africa Gateway 2019, p. 2; Statistics South Africa 2019b, p. 2). In a variation study of satellite imagery and the classification potential of machine learning for spectral, textural, and structural differentiation, Engstrom, Hersh, and Newhouse (2017, p. 2) confirmed that satellite observations are 'highly predictive of economic well-being', drawing correlations with welfare and key regional features such as population density and construction materials (e.g., roofing). Due to the critical role of evidential insights and multi-factorial data in the diagnosis and interpretation of regional development goals, poverty, and social inequality, Holloway and Mengersen (2018, p. 10) suggest that advances in machine learning and the adoption of 'ensemble approaches' to structural diagnosis offer discrete advantages that can be used to accelerate research and improve evidential fidelity. The current study has adopted a similar paradigm, exploring the relationship between remote sensing capabilities, machine learning, and poverty diagnosis in a nation where regional

inequalities are spatially complex and oftentimes visually indistinguishable: South Africa.

In spite of government efforts to capture meaningful data from household surveys, Blumenstock, Cadamuro, and On (2015, p. 1073) report that the precision of these infrequent research instruments is diluted by up to 50% due to a variety of sampling, surveying, and quantification errors. Given the inadequacy of traditional poverty assessment methods, a range of innovative technological solutions have recently emerged as advances in remote sensing, digital imaging, and data mining have created a new range of opportunities for innovative researchers. For example: Michalopoulos and Papaioannou (2014, p. 154) devised a luminosity-based cross-border poverty assessment resource to evaluate institutional efficiency in African nations; Ma et al. (2019, p. 4) introduced a typeface-based remote sensing model that correlated the relationship between signage fonts and regional amenity density in modern urban centres such as London; and Blumenstock et al. (2015, p. 1073) utilised mobile data usage statistics to model technological penetration and changing patterns in urban prosperity across emerging markets. Each of these techniques offers a different but comparable proposition: to utilise publicly accessible data resources from remote repositories to quantitatively model the relationship between regional characteristics and socio-economic conditions.

1.3 Research rationale

The range of emerging analytical techniques being proposed for remote sensing and long-distance surveying of socio-economic conditions is broad and innovative, highlighting the capacity-boosting influences of new technology-supported solutions. In South Africa, poverty remains an obvious but under-defined problem that often neglects the heterogeneous nature of social inequality and spatial stratification. In a recent nightlight-sensing report presented by Pfeifer, Wahl, and Marczak (2018, pp. 31–2), it was concluded that regionalised developments in the form of infrastructure, enhanced transportation, and the localisation of government services had distinct impacts on luminosity, accelerating ‘*catch-up processes*’ — particularly in

urban spaces. At the same time, Mveyange (2016, p. 20) has revealed that the conceptual link between remote sensing of night-time and Geographic Information System (GIS) images and domestic poverty is often obscured by the inability to align more tangible indicators of exclusion (e.g., education, healthcare, transportation, infrastructure) with regional developmental patterns (e.g., density, luminosity). In fact, Henderson, Storeygard, and Weil (2012, p. 1015) as well as Pinkovskiy and Sala-i-Martin (2014, p. 8) have demonstrated that remote sensing models using nightlight assessment of luminosity over time are positively correlated with datasets based on national Gross Domestic Product (GDP), indicating a high degree of complementarity. For this reason, additional research on South African exclusion and the structural relationship between inequality and poverty was justified and required by the central gaps in the existing academic register.

Other researchers including Blumenstock et al. (2015, p. 1073), Ma et al. (2019, p. 4), Suel et al. (2019, p. 7), and Zhao et al. (2019, p. 1) have employed remote sensing capabilities to compare satellite or street-level imagery with economic indicators, night-time lights, or crowd-sourced information in order to find correlations between key features and socio-economic traits. The problem with these empirical solutions is that they are not reliable measures of socio-economic disadvantage: they have a broad range of inconsistencies, conflicting definitions, and unreliable techniques, leading to ineffective measures that do not add sufficient value to the long-term assessment of population welfare (Jean et al. 2016, p. 790). It is important to realise that, as Blumenstock et al. (2015) pointed out, call detail records such as telephone calls or mobile phone usage records used in some research efforts are not publicly available and consequently are hard to rely on. With evidence limited to generalisable conditions, the relative effectiveness of innovative models and exploratory technologies is constrained by their practical applications in relation to large-scale, replicable, and revelatory remote sensing capabilities.

1.4 Problem statement

Problem with the model developed by Jean et al. (2016), as well as the more recent model by Lam et al. (2018, pp. 7–8) is that it lacks the quantitative granularity

to distinguish between poverty characteristics: it generalises clustered housing groups and population sets (e.g., rural, urban) into general outputs rather than assessing quantifiable traits, rendering the model incapable of reconciling variations between municipalities, different socio-economic patterns, industrial versus residential areas, and a variety of other predictive indicators. In contrast, the model developed by Zhao and Kusumaputri (2016, p. 2) does not capture the qualitative differences in the geospatial features that are quantified from the satellite images. The current study was carried out as a proof of work investigation to apply machine learning techniques using only publicly available data that are readily available for processing. This work demonstrates the close relationship of these economic activities in the area and thus shows features that can elucidate poverty levels. This study has investigated the following core problem statement:

Despite accelerated and sustained domestic economic development, South Africa is characterised by robust socio-economic inequality, with urban exclusion represented by key physical indications, including housing quality, night-time luminosity, and gaps in social services and domestic infrastructure. Due to the high cost of social surveys and the challenges of physical in situ data collection, there is an immediate need for alternative remote sensing solutions capable of diagnosing, modelling, and interpreting systemic exclusion conditions to improve policymaking and target future interventions.

Whilst it is contended that the existing approaches to using non-traditional data are promising given that they are cheaper and inherently scalable, their broader adoption by policymakers is still lagging behind traditional data usages. Providing a basis for this investigation, researchers, including Jean et al. (2016, p. 794) and Zhao and Kusumaputri (2016, p. 6) have confirmed the viability of machine learning applications to evaluate household poverty indicators, enabled by geospatial feature assessment capabilities. By training software to make autonomous estimations about the spatial distribution of socio-economic well-being, and then comparing those outputs with regional or domestic household surveys, it is possible to confirm the degree of accuracy and consistency of the analytical tool (Jean et al. 2016, p. 794). This investigation has focused on patterns of regional socio-economic inequality that are oftentimes obscured by the tenuous relationship between urbanisation and social

migration. As a result, in nations like South Africa, where regional density patterns mirror the post-colonial heritage of exclusion and enclavisation reflected during apartheid, demonstrating the structural relationship between economic opportunity and regionalisation can be critical to understanding systemic gaps and inefficiencies.

1.5 Research question

Central to the pursuit of a low-cost, high-fidelity remote sensing solution for modelling poverty in developing nations, the following core research question has been answered during this investigation:

Given the heterogeneity of economic characteristics in developing nations like South Africa, which cost-efficient machine learning methods can be used to determine the relationship between special features and socio-economic variations?

The sub-research questions that were addressed to support the main research question of this study are as follows:

- 1) Which technique is suitable for conducting satellite imagery remote sensing and how can this approach be adapted to analyse poverty in South Africa? This question seeks to establish the strengths and weaknesses of remote-sensing grounded approaches and satellite imagery classification tools.
- 2) What are the challenges that affect machine learning capabilities and their application to satellite-based imagery of varied regional cityscapes? This question determines whether the characteristics of regional poverty in South African regions can be determined using cost-effective remote sensing geospatial techniques.
- 3) What best practices and techniques can be applied to architect predictable remote sensing capabilities in future poverty assessment studies? This question seeks to establish whether the practices and techniques discovered during the study can be generalised and recommend suggestions for future studies.

1.6 Research objectives

Due to the high level of complexity, investment, and time commitments required for ground-level field-based poverty surveys in African nations, the primary objective of this investigation was to critically assess the viability of cost-efficient data sources and machine learning methods for poverty estimation in order to confirm the proof-of-concept and provide a viable alternative for poverty assessment in emerging but developed economies such as South Africa. By deploying an integrative approach to combine publicly available geospatial data, as well as using machine learning capabilities and analysis of relationships between special features and socio-economic variations, this study achieved the following central research objectives:

- To establish the strengths and limitations of recent research regarding remote sensing capability approaches and satellite imagery classification tools;
- To determine whether the regional poverty characteristics in South African regions can be determined using cost-efficient remote sensing geospatial techniques; and
- To establish whether the practices and techniques discovered during the study can be generalised and recommend suggestions for future studies.

1.7 Research methodology

To achieve the objectives of the study, the related literature and studies were reviewed in order to draw insights and lessons that would help formulate a computational framework as an artefact. Subsequently, a working model capable of predictably and reliably estimating socio-economic well-being in South Africa was developed. This study followed a positivist paradigm and design science approach as presented by Hevner, March, Park, and Ram (2004, p. 82) to the structure and application of cost-efficient and readily available data sources in the design of wealth estimation models. Hevner et al. (2004, p. 78) describe design science as a series of

related expert activities that are done to come up with an artefact, which is then evaluated in order to improve both the artefact's quality and the process required to design the artefacts. Figure 1.1 shows the adopted approach that follows Peffers, Tuunanen, Rothenberger, and Chatterjee (2007, p. 58) modelled for this study. The approach follows six steps: problem identification and motivation, definition of solutions objectives, design and development, demonstration, evaluation, and communication.

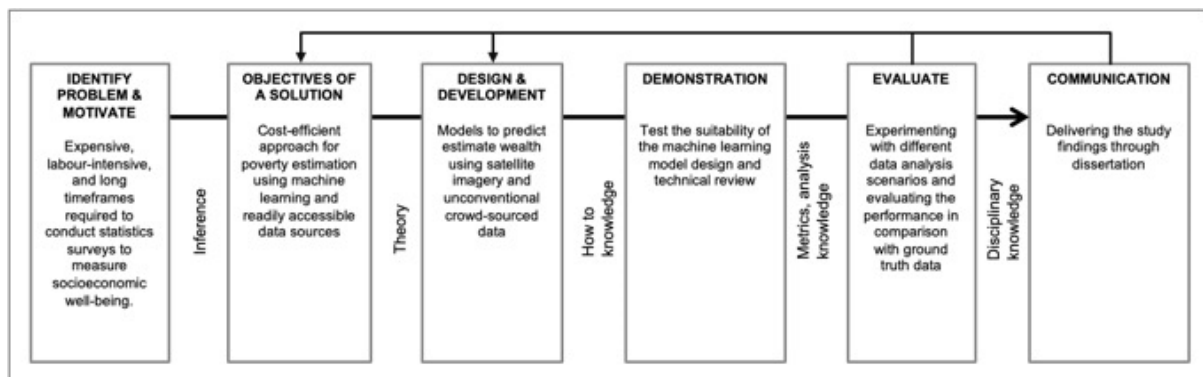


Figure 1.1: Design science research activities, according to Peffers et al. (2007, p. 58, Figure 1)

The individual steps serve to structure the remaining study, and, at first, the research layout will start with a literature review as motivation in Chapter 2. The objectives of the study, which are to establish the strengths and limitations of recent research on remote sensing capabilities approaches and satellite imagery classification tools, as well as to determine whether regional poverty characteristics in South African regions can be determined using cost-efficient remote sensing geospatial techniques, are defined in the current chapter. In the design stage, the research extends the precedents established by Jean et al. (2016, p. 791), Zhao and Kusumaputri (2016, p. 4), and Head et al. (2017, p. 2) to design a predictive model capable of analysing, classifying, and summarising the indicative prevalence of poverty in South Africa detailed in Chapter 3 and has the following two approaches: (i) predict night-time lights as a proxy task of satellite imagery and then use them to compute the average features embedded per cluster to estimate wealth and (ii) include unconventional crowd-sourced datasets with night-time lights to infer wealth for cases where frequent surveys are not possible. The results will be demonstrated through experiments of the machine learning model grounded in the proposed artefacts (Chapter 4) and the performance of the models is evaluated using statistical methods

and ground truth data to verify if the designed artefact meets the pre-set requirements (Chapter 5). This is performed by iteratively improving the parameters to expand the performants and compare the results before selecting the best findings.

The categorical dimensions for the models were extracted by incorporating a model to examine the extent to which merging geospatial remote-sensed and crowd-sourced data can aid cost-effective and scalable means to estimate welfare in South Africa. The results were based on a critical, quantitative comparison of the classification results and socio-economic reports published by the South African government, thus confirming the transferability and consistency of the proposed model. The research is data-centric and multi-disciplinary, attempting to create knowledge that connects and spans the disciplines involved.

1.8 Limitations and delimitations of the study

The method of predicting human development with satellite data, geospatial data, and deep learning requires ground-truth data for training and evaluating the models. The first and most used approach is to use official socio-economic statistics as ground truth (Jean et al. 2016, p. 794). For this research, the Demographic and Health Survey (DHS) produced by the United States Agency for International Development (USAID) was used as the ground truth dataset (i.e., maximally objective data). DHS survey data have some caveats on spatial and temporal precision regardless of being reliably sampled. To preserve the privacy of the survey respondents' data, noise is intentionally added by randomly compensating the Global Positioning System (GPS) coordinates of each household cluster by up to 2 km in urban areas and up to 5 km for rural areas. Noise offset will likely introduce bias in spatial analysis, and therefore buffer zones will be introduced as an attempt to mitigate the effects. Second, ground truth data was collected in 2016 and therefore features and wealth predictions were also adjusted for 2016. Recent values can only be rolled out when the next surveys are conducted.

In terms of scope, the models were trained using satellite imagery data from South Africa. No comparisons were made with other countries within the region or that

have an almost similar GDP. Furthermore, while the accuracy and performance of the neural network is critical for the accurate predictions and determinations of the model during the assessment process (Ilic, Sawada, & Zarzelli 2019, p. 3; Suel et al. 2019, p. 7), the scope of the research involves a single model architecture for satellite imagery, a VGG16 architecture from Visual Geometry Group pre-trained on the ImageNet data set.

Lastly, there may be concerns about the choice of the dataset used as ground truth given the Anglo-American military and economic interventions in Africa. This history is an extraneous detail when considering the socio-economic data for this study. It is assumed that the international institutional arrangements (including the dynamics of colonisation) do not have any epistemic impact on the quality of the data.

1.9 Dissertation outline

This chapter has provided the underlying motives and practical basis for conducting this exploratory study of data integration measures to support poverty identification in emerging economies. The research background section of this chapter showed the need for an additional in-depth understanding of the poverty phenomenon as an important step to inform research studies and guide policymakers in the development of timely and accurate policies for poverty intervention and subsequent impact monitoring. The study justification and problem statement, after which the research objectives and questions were detailed. The following paragraphs present a brief overview of each of the subsequent chapters that comprise a progressively technical exploration of targeted datasets, applying techniques and technologies that were designed to explore the relationships accurately and systematically between regional socio-economic indicators.

Chapter 2: Literature review. Drawing upon a growing body of research regarding remote sensing, GIS imaging, and distance-based socio-economic analysis, this chapter outlines a variety of prior research in this field and identifies both the strengths and weaknesses of prior techniques, focusing on machine learning

capabilities and image-based analysis. It provides the conceptual and theoretical basis for the experimental design and empirical study.

Chapter 3: Research methodology. This chapter outlines the range of methods adopted and deployed to capture GIS-based evidence from South Africa and its socio-economically diverse urban spaces. The central sources of evidence and the underlying techniques and technologies employed during this study are introduced and discussed. It formalised the design and conceptual basis for the key classification methods and identifiers that were adopted for the assessment of poverty in South Africa.

Chapter 4: Implementation and results. This chapter presents the results of the remote sensing study, highlighting the overlapping similarities between day- and night-time-based indicators of poverty and various socio-economic determinations. It outlines the stages of data collection, organisation, and analysis, introducing the reliability and accuracy of data to confirm the effectiveness of the remote sensing capabilities of this poverty detection model.

Chapter 5: Discussion and analysis. To triangulate these findings and confirm the validity and transferability of the techniques employed during this study, this chapter discusses evidence and models in relation to previous research in the field of applying machine learning to computational social sciences.

Chapter 6: Conclusions and Recommendations. This final chapter draws summative conclusions from the entire study, highlighting the effectiveness and limitations of the methods used for this research, and offers recommendations to improve future results in additional remote sensing studies and experiments in ongoing poverty analysis.

1.10 Chapter summary

This chapter provided a synopsis of the study, including the underlying core problem under study, together with the objectives of the work, relevant research

questions, methodology, and the limitations of the investigation in relation to South Africa. Furthermore, the paradigm underpinning the study and the approach followed to solve the identified problems were also discussed. The following chapter is the outcome of an extensive review of the literature conducted to construct the main argument of this study.

Chapter 2. Review of the literature

2.1 Introduction

The field of remote sensing and purpose-constructed spatial imaging analysis has evolved towards increasingly advanced and innovative solutions as researchers have combined models and methods into discrete solutions. While daytime satellite imagery and structural recognition were once the standard for GIS-based analysis, modern innovations in machine learning and system adaptation have created opportunities for novel solutions such as night-time and geospatial analysis. The field of study surrounding data integration is increasingly broad, encompassing a variety of conceptual foundations related to machine learning, remote sensing, and probabilistic analysis. In order to narrow down the appropriateness and effectiveness of these techniques, this chapter compares multiple experimental and theoretical studies, outlining the potential advantages of these varied approaches for capturing data specifically related to regional poverty indicators. The following sections explore this range of techniques, outlining the advantages of specific, targeted approaches, and the potential challenges encountered in mapping various phenomena across diversified geographic and technological networks. This establishes a background for the research problem and identifies the gaps within the research questions posed. Lastly, the chapter briefly introduces machine learning, as it formed the backbone of the study at hand.

2.2 Measurement of poverty

Poverty has traditionally been conceptualised as one-dimensional: usually based on income, sometimes interchangeably referred to as consumption, which is also called income poverty (Alkire & Santos 2013, p. 5). Over time, combining various approaches has been proposed and doing so has led to the inclusion of economic well-being, capability, and social inclusion (Sen 1993, p. 31; Wagle 2008, p. 15) in these concepts. A more recent international measure is the Multidimensional Poverty Index (MPI), which recognises poverty as a multidimensional phenomenon and its

usage has increased rapidly over the last decade (Wagle 2008). MPI was initially developed by Alkire and Santos (2013) and later adopted by UNDP for its Human Development Reports (Oxford Poverty and Human Development Initiative 2015). The MPI offers insights into the extent and intensity of multidimensional poverty by combining ten indicators intended to reflect experiences of poverty-related deprivation from critical dimensions of health, education, and living conditions. The most reliable way to measure and estimate poverty is through intensive socio-economic household surveys and census data collection. However, these approaches are costly and time consuming. Studies have been conducted to highlight the relationship between poverty and spatial factors that are measurable from orbital space (e.g., satellites) such as buildings, agricultural activities, proximity to essential services, etc. (Lillesand, Kiefer, & Chipman 2015).

Underscoring the central principles of this poverty modelling is what Bertram-Hümmer and Baliki (2015, p. 770) describe as '*visible wealth*'. Although it is an ideal socio-economic differentiation mechanism, the researchers note that 'visibility' by definition is unclear—a problem which can only be reconciled if some tangible dimension, category, or indicator can be identified relative to the underlying condition evaluated: wealth (Bertram-Hümmer & Baliki 2015, pp. 770–3). To distil the range of variables (e.g., dwelling size, outbuildings, vehicles, possessions) into a quantitative indicator of visibility, Bertram-Hümmer and Baliki (2015, p. 781) tested the theory of '*relative deprivation*' by comparing individual interpretations of '*visible wealth*' by weighing the relative importance of multiple indicators. The findings confirmed that tangible indicators of visible wealth such as being domiciled in a house or having possessions had a direct impact on the perceptions of deprivation held by the participants, confirming that when applied to a residential environment on a comparative basis, perceptions of poverty (or lack thereof) are likely to be contingent upon comparative interpretations of visible regional wealth (Bertram-Hümmer & Baliki 2015, p. 781). Although Bertram-Hümmer and Baliki's (2015, p. 781) study does not apply remote sensing techniques to the assessment of poverty, it does highlight the weight ascribed to visible determinants of wealth or deprivation in a given society, justifying the assessment of visible indicators as proxy measures for differences in socio-economic status. Consequently, remote sensing and machine learning

algorithms can be developed to identify poverty traps and deprivations from features that are visible in geospatial data as an aid for policymakers to develop interventions.

2.3 Geospatial information systems and remote sensing

As satellite imagery, Geospatial Information Systems (GIS), and crowd-sourcing technologies increase the visibility and transparency of global landscape models, the potential for mapping both urban and rural development patterns continues to motivate a diversified range of exploratory studies (Hillson et al. 2014, p. 2). Combining data from Global Positioning Systems (GPS) via ground-based mapping of structural features and characteristics and GIS imagery to digitise rooftop features, Hillson et al. (2014, p. 11) revealed significant inaccuracies in top-down assumptions based on rooftop surface area calculations and Object-Based Image Analysis (OBIA). For this reason, the combination of physical, crowd-sourced, and ground-based assessments with remote sensing observations of the same regional characteristics created an opportunity to diagnose the likelihood of error and bias in the mapping outputs, confirming the potential for residential density bias in regions where both industrial and residential spaces are frequently interspersed (Hillson et al. 2014, p. 11).

Central to the challenges of remote sensing is the significant scale of classification data that is needed in order to draw meaningful interpretations from GIS models. Challenges such as regional granularity, texture, and feature-based recognition create hurdles for training and classifying datasets (Lam et al. 2018, p. 2). Tingzon et al. (2019, p. 425), for example, focused their research on determining whether a range of satellite-based methods could be applied to poverty prediction in varying urban landscapes, assessing the intersection between publicly accessible geospatial information and emergent GIS-based imagery in terms of regional coverage and accuracy. While Xie et al. (2016, p. 3935) have concluded that deep learning advances and the architecture of data-rich identification proxies indeed provide a solution for transfer learning and efficient GIS modelling, these techniques require the ability to train systems to identify complex features and terrains. Jean et al. (2016, p. 794) propose that GIS-based models are capable of outperforming other passive data

collection instruments such as cell phone data modelling, but cluster effects and regional incongruities can create false positives or inaccuracies that may lead to a lack of fidelity in the model output.

Crowd-sourced geographical information is another emerging data source for assessing urban density, socio-economic development and inequalities to identify land use, land expansion, proximity of services, building polygons, develop urban spatial structure and road networks (Gervasoni, Fenet, Perrier, & Sturm 2019, pp. 595–596; Mahabir et al. 2018, p. 5). Volunteered Geographic Information (VGI) contains billions of entries that are maintained by a massive community of global mappers whose common objective is to curate precise and complete geospatial data. As computational resources increase and the ability to use computer vision on satellite imagery evolves, research has increasingly used geospatial datasets as well (Yan et al. 2020, p. 1781). Yue et al. (2015, p. 464) acknowledge that there is an increased availability of GIS data, which has enabled the development of advanced machine learning algorithms for geospatial solutions.

Building upon the early models that relied on VGI, there is a distinct and significant opportunity for the development of effective visual sensing capabilities that can accurately compare economic welfare analyses. Poverty characteristics are complex and, therefore, it is difficult to effectively evaluate poverty with a single type of data in both theory and practice. Leveraging the approach to night-based luminosity and satellite imagery, Zhao et al. (2019) conducted studies in an attempt to identify and estimate poverty in Bangladesh using data from multiple sources, including road maps and land cover data. Studies by Jean et al. (2016) and Zhao et al. (2019) also acknowledge that differences between wealthy, sparsely populated areas and poor, densely populated areas within night-time luminosity data would likely produce similar results, making it difficult to tell the difference. Combining data from multiple sources provides a comprehensive evaluation of poverty from multiple dimensions. A similar study conducted by Pan and Hu (2018, p. 1104) on multidimensional poverty measurement using night-time luminosity data recognises that the availability of education, health, transport, and financial services related to poverty have become important factors impacting poverty. In order for related analyses to be highly effective,

however, it is critical that they use data from multiple sources to accurately predict poverty measures within an area.

Efforts to crowd-source geospatial data are increasing. A good example is OpenStreetMap (OSM), the largest free collaborative platform that produces a crowd-sourced geographic database, and one of the major platforms that promotes the use of geospatial data in humanitarian and community development fields. Big corporations such as Microsoft, Uber, Grab, Facebook, Amazon, Baidu, and Weibo have been making substantial contributions to community mappers from around the world who are curating accurate and representative geospatial data on sites such as OSM (Anderson, Sarkar, & Palen 2019, pp. 6–7). The OSM open map dataset contains geospatial data which provides critical infrastructure that shows the presence of roads, rivers, built-up areas, and points of interest (POI), enabling researchers to investigate the association among present characteristics and socio-economic conditions. It helps us to understand our planet and the functioning of the society. More than 100 million changesets have been uploaded to OSM, representing a collective global contribution of nearly 1 billion features in the last 16 years. Hu, Yang, Li, and Gong (2016, p. 4) and Ye et al. (2019, p. 938) conducted studies to demonstrate that crowd-sourced data from Baidu and OSM can provide details of the spatial distribution of economic activities.

2.4 Detection of poverty using machine learning

The ubiquity of artificial intelligence and machine learning capabilities in modern social studies is enhanced by the perceived effectiveness of these overlapping models in the diagnosis and interpretation of socio-economic inequality and regionalised variations in economic growth and development (Holloway & Mengersen 2018, p. 2). OBIA involves analysing pixels grouped together into matching sections or objects rather than individual pixels (Holloway & Mengersen 2018, p. 6). From variations in structural characteristics in Ghana to the impacts of a recent cyclone in the Philippines, Holloway and Mengersen (2018, p. 6) reveal that the separation and classification of structural characteristics through remote sensing (e.g., point-based, edge-based, region-based) initiatives allows researchers to diagnose the impacts and

consequences of both isolated and persistent socio-economic events. It is this practical applicability of machine learning to complex visual relationships that elevates this solution to a high-interest domain in the field of poverty and welfare analysis.

By applying indicators such as the Living Environment Deprivation (LED) index to remote sensing of regional poverty and public health, Arribas-Bel, Patino, and Duque (2017, p. 4) predict that researchers can chart regional characteristics and features over time to measure the effects of deprivation. By automating this process of feature detection through machine learning capabilities, it is possible to measure textural variations (e.g., greenbelts versus urbanisation) across pixel subsets and chart the intersection between city features and poverty indicators (Arribas-Bel et al. 2017, pp. 8–9). To automate this process, Arribas-Bel et al. (2017, p. 10) adopted a random forest algorithm that leverages other machine learning techniques (e.g., bagging) to aggregate a series of models into a single prediction. This decision tree model was based upon a full dataset and its multi-layered subsets, which could be sampled randomly in order to de-correlate identifiers that might otherwise lead to model inaccuracies (e.g., single versus strand trees) (Arribas-Bel et al. 2017, p. 11). A Gradient Boosting Regressor (GBR) algorithm was used to ‘*boost*’ the output of other models with the result being only one prediction variable. Through the application of these techniques to a specific urban landscape (e.g., Liverpool, UK), Arribas-Bel et al. (2017, p. 18) confirmed that textural variations with particular emphasis on urban green spaces could be effective measures of LED characteristics, with lower land cover variability indicating higher LED.

The central advantage of a remote sensing model is predicated upon its transferability to other geographical perspectives as researchers seek to develop general, data-driven solutions that can be replicated with a high level of accuracy (Suel et al. 2019, p. 7). Through a restrained training methodology, Suel et al. (2019, p. 8) have demonstrated how the improvement of Convolutional Neural Network (CNN) analysis can be coordinated across multiple datasets by implementing an image-outcome pairing standard that allows for intra-regional confirmation. This contextual solution considers that within any given image set, training can be executed across a majority percentage of the images, with minority imagery from the same dataset being withheld for assessment purposes (e.g., 80/20) (Suel et al. 2019, p. 8). When applied

to other municipalities or visual reference groups, it becomes possible to use this segmented training methodology to validate regionally accurate visual cues (e.g., building traits in a given municipality) and subsequently confirm whether the training sets retain their accuracy when transferred to other municipal areas (e.g., London versus Manchester) (Suel et al. 2019, pp. 8–9).

As increasingly innovative deep learning technologies are developed, the accuracy of visual mapping and street-view analysis is a critical antecedent to the value and transferability of remote sensing analysis (Ilic et al. 2019, p. 1). Deep mapping techniques involve developing multiple representations of similar classifications (e.g., roofing materials, housing characteristics) that can be applied across multiple urban landscapes in order to develop a meaningful outcome, insight, or interpretation (Ilic et al. 2019, p. 3). For example, by developing a CNN-based model for Google Street View (GSV) mapping, Ilic et al. (2019, p. 3) assessed improvements in the quality of the frontage of residential properties in Ottawa, Canada, to determine the level of gentrification in various urban regions. Although the model was vulnerable to false positives and negatives due to the complex and varied definition of ‘frontage quality’, evidence suggested that this technique could be applied to urban analysis in order to assess changes in landscape characteristics over time as planning permits are approved or withheld in environments of varying socio-economic conditions (Ilic et al. 2019, p. 6).

It is the practical applicability of remote sensing CNN models to real-world problems that presents the greatest opportunities for policy-making and urban development. Nguyen et al. (2019, p. 7), for example, revealed how high-level satellite images could be used to classify the built environment and its effects on regional amenities, targeting health-improving outcomes that are defined in relation to hospital and service density. When considering the disparities between quality of health across urban and rural areas, this model allowed regional data on health pathology to be compared with visual data on built services and social support over time (Nguyen et al. 2019, p. 7). Gebru et al. (2017, p. 13108) offered a different classification model based upon automobile characteristics that demonstrated the applicability of GSV images to a CNN-based assessment of regional demographics. The findings revealed a direct correlation between inferred classifications (e.g., income, race, education,

voting patterns) and the actual demographic indicators that were captured for a given neighbourhood or urban region (Gebru et al. 2017, p. 13108). The implications of such findings are linked to future voting profiling and social services programmes, suggesting that the implications of visual mapping can be extended beyond the experimental basis for these targeted models (Gebru et al. 2017, p. 13113).

2.5 Nightlight analysis and spatial imaging

One emergent solution for assessing urban density and poverty-related patterns in discrete geographic models is to evaluate luminosity at night in order to predict changes in productivity and economic activity over time (Chen & Nordhaus 2011, p. 8589; Ghosh, Anderson, Elvidge, & Sutton 2013; Henderson et al. 2012, p. 995; Jean et al. 2016, p. 791; Michalopoulos & Papaioannou 2014, p. 152; Mveyange 2016, p. 2; Pfeifer et al. 2018, p. 2; Pinkovskiy & Sala-i-Martin 2014, p. 3). Jean et al. (2016, p. 791) recognise that night-time lights were universally reliable and existing proxy for economic activity regardless of them being noisy. Through machine learning training of a CNN, several critical stages are required for developing a representation of economic activity, as both daytime variations and night-time light features can be used to increase the accuracy of the dataset (Jean et al. 2016, p. 791). It is the recognition of night-time light variations at higher income levels that Jean et al. (2016, p. 791) proposed can be applied to the detection of variations amongst lower-income, more densely populated urban areas; however, inconsistencies in the poverty assessment can lead to false positives or assumptions of higher income regions. By training the model to recognize some *'livelihood-related characteristics of the landscape'* in the daytime model, the model can then discern specific features that are correlated with night-time light indicators, resulting in an automated and high-accuracy representation of household consumption expenditure and wealth (Jean et al. 2016, p. 791). Figure 2.1 provides a visual representation of this multi-modal assessment, utilising variable landscapes to not only train the daytime assessment, but also the nightlight interpretation of luminosity and regional density (Jean et al. 2016, p. 791).



Figure 2.1: Three-model comparison of daylight (top), nightlight (middle) and filtered (bottom) assets across (left to right) urban, nonurban, water and roads (Jean et al. 2016, p. 792, Figure 2)

Similar visual comparisons were recently employed by Engstrom et al. (2017, p. 6) to draw verifiable comparisons between night-time light density, the regional poverty rate, and the mean population density. Their comparison revealed a higher level of correlation between night-time light areas and population density than welfare indications, suggesting a potentially incomplete representation of welfare and poverty-specific representations. To further classify and distinguish between visual features via GIS images, Engstrom et al. (2017, p. 7) applied deep learning object-based classification techniques to develop object predictions based on specific traits such as roof type, road characteristics, structural density, automobile density, and landscape usage (e.g., agriculture, urban). Evidence revealed a negative correlation between regional poverty surveys and vehicle density (e.g., lower density equals higher poverty), a negative correlation between road density (e.g., higher density equals lower poverty), and a positive correlation between regional poverty and agricultural land (e.g., higher agricultural land equals higher poverty) (Engstrom et al. 2017, pp. 14–15). Specific to the role of night-time lights in explaining poverty variances observed in the models, Engstrom et al. (2017, p. 19) confirmed a 7-12% explanation in the variance of per capita consumption, suggesting a potentially inaccurate, and only marginally supportive solution for GIS poverty modelling.

A comparison of recent GIS-based machine learning studies, including those by Jean et al. (2016, p. 794), Xie et al. (2016, p. 3934), and Babenko et al. (2017, p. 3), revealed that the efficacy and accuracy of the analysis are predicated upon the effectiveness of the training programme devised for convolutional neural networks (CNNs) and the correlated accuracy of feature-based observations with their capacity to predict poverty-related factors (e.g., urban density, access to electricity). In a broad, global assessment of the night-time light assessment of poverty, Ghosh et al. (2013, p. 4712) revealed that when applied across multiple urban landscapes and compared with correlated data from annual poverty reports such as the World Bank, it becomes possible to track patterns of socio-economic deprivation and development over time by using comparative models. Alternatively, Mveyange (2016, p. 15) used an Integrated Public Use Microdata Series (IPUMS) dataset from ten African countries to compare night-time light data with regional employment, education, and urban density statistics, generating a comparative indicator of poverty that was representative of a broad, multi-dimensional scale. Pfeifer et al. (2018, pp. 20–30) localised this approach to a regional study of South Africa, drawing evidence from multiple night-time light datasets that revealed the affective influence of infrastructure improvements made in preparation of hosting the 2010 FIFA World Cup on regional poverty as measured by improved structural solutions and socio-economic welfare. By specifically identifying features (e.g., train stations, bus depot) that could be correlated with the 2010 FIFA World Cup investments, post-event night-time indicators revealed regional urban luminosity indicators that suggested positive economic improvements (Pfeifer et al. 2018, p. 29).

Systematising the approach to night-based luminosity analysis, Pinkovskiy and Sala-i-Martin (2014, p. 7) developed a pixel-based assessment of luminosity that assigned a score to each spatial indicator (one square kilometre) in order to model the relative luminosity of a given region or geographic area. The premise for the nightlight-based assessment is based upon the positive correlation between luminosity and national accounts (e.g., GDP) which has been demonstrated over a range of critical socio-economic events (e.g., decline in luminosity during the Asian Financial Crisis) (Pinkovskiy & Sala-i-Martin 2014, p. 8). A similar study conducted by Chen and Nordhaus (2011, p. 8589) acknowledges that when comparing light saturation with the

estimated urban population, a log-log linear relationship between urban scale and population can be calculated. However, in order for such analyses to be effective, it is critical for these models to be able to resolve inconsistencies such as scanning errors, water vapour, luminous noise (e.g., fires), and other irregularities that are likely to negatively influence the accuracy of the datasets (Chen & Nordhaus 2011, p. 8589).

Building upon these early models, there is a distinct and significant opportunity to develop effective visual sensing tools that are capable of accurate and comparative analysis of economic welfare via satellite imagery. Research by Pinkovskiy and Sala-i-Martin (2014, p. 8) in Angola, for example, demonstrated a doubling of per capita income between 2000 and 2009 that was positively correlated with a significant increase in night-time lighting recorded from satellite images taken during this same nine-year period. Similarly, in India, where per capita income more than doubled between 1994 and 2010, satellite imagery demonstrated that lighting intensity both within urban areas and across previously unlit areas of the nation increased significantly during a period of accelerated economic growth (Pinkovskiy & Sala-i-Martin 2014, p. 8). When luminosity was compared across ethnic groups in Africa separated by national boundaries (e.g., Ambo Group in Angola and Namibia), the differentiation between luminosity within a single group had direct implications for assessing institutional performance between countries, as visualised in Figure 2.2 (Michalopoulos & Papaioannou 2014, p. 154). These findings revealed an important visual differentiation between the economic welfare of members of a single ethnic group when subdivided across international borders (Michalopoulos & Papaioannou 2014, p. 154).

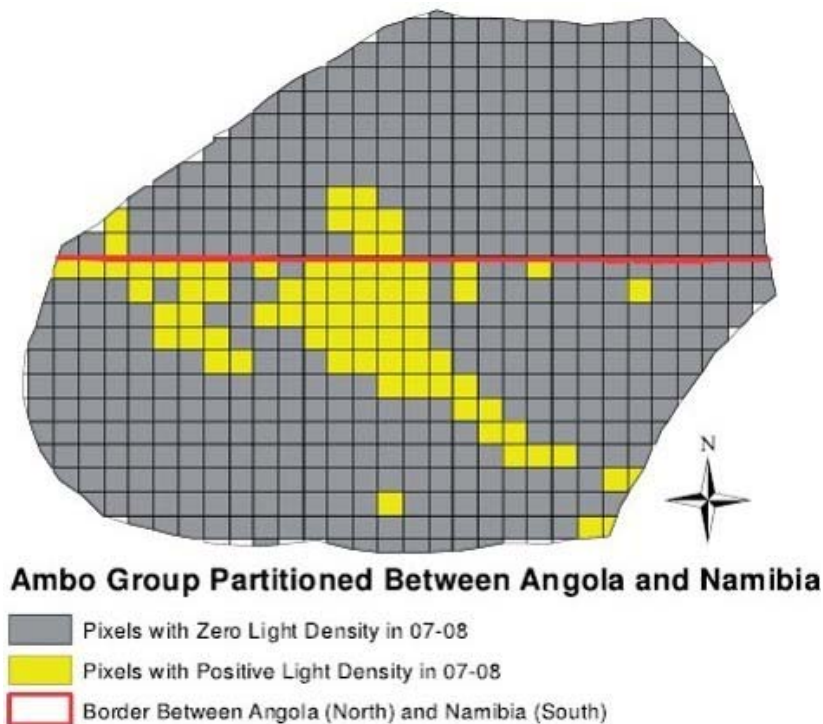


Figure 2.2: Ethnic group nationality partitioned luminosity measures (Michalopoulos & Papaioannou 2014, p. 154, Figure IIB)

A critical problem with the pixel-based luminosity equation is that the digital number output (luminosity score) at the bottom or top end of a model is likely to reflect either artificial dimming (e.g., lack of light continuity) or blooming (e.g., luminosity carryover from pixel to pixel) (Pinkovskiy & Sala-i-Martin 2014, p. 8). Furthermore, although the efforts of Michalopoulos and Papaioannou (2014, p. 159) were originally constructed to highlight the significant gaps in national institutions across international borders, the findings revealed regionalised heterogeneity in cross-border luminosity which could not be explained by institutional variations. Instead, other contextual factors within the region, such as distance from the urban centre, were identified as primary impact factors that played a direct role in determining the degree of institutional penetration across regional populations (Michalopoulos & Papaioannou 2014, p. 159). This finding is problematic, particularly when considering luminosity as a representative indicator of poverty, as remote sensing experiments would be unable to reconcile variations between municipalities, different socio-economic patterns, industrial versus residential areas, and a variety of other predictive indicators.

In order for nightlight mapping to provide an effective proxy for domestic welfare, it is important not only to observe a single representation of lighted urban or

residential spaces, but also to compare the characteristics of night-time lighting over a time-series sample (Henderson et al. 2012, p. 999). Employing data aggregated by the National Oceanic and Atmospheric Administration (NOAA) from US defence satellites, Henderson et al. (2012, p. 999) determined a reliable source of nightlight imaging that could be compared across regions and time periods to assess luminosity on a scale from zero (no lights) to a high score of 63 (brightest saturation of lights). Figure 2.3 visualises two long-term comparative images of the Korean Peninsula and the change in luminosity from 1992 until 2008 as urban landscapes have expanded and rural areas have gained access to electricity (Henderson et al. 2012, p. 1002). By applying a pixel-based classification technique, Henderson et al. (2012, p. 1015) propose calculating economic growth over time by comparing luminosity across time-series visual models and comparing positive and negative lighting effects within the datasets. Similar to the findings presented by Pinkovskiy and Sala-i-Martin (2014, p. 8) with respect to the Southeast Asian financial crisis, Henderson et al. (2012, p. 1015) suggest that economic recessions can be observed via regressive patterns in regional luminosity, while expanding GDP can be observed via an increased regional level of consistent and efficient brightness. However, despite the advantages associated with a general assessment of luminosity for urban analysis and activity mapping, Jean et al. (2016, p. 790) remind the research community that in areas below the poverty line (less than USD \$1.90 per capita per day), the lack of electric capability is likely to mean limited luminosity and, therefore, limited variations in levels over time.



Figure 2.3: Light-recorded long-term growth on Korean Peninsula (Henderson et al. 2012, p. 1002, Figure 2)

2.6 Other remote sensing solutions

As researchers struggle to apply methods such as night-time light analysis to the assessment of economic growth and regional welfare, other social models and investigations are providing a comparative basis for validation and confirmation. One important revelation presented by Pinkovskiy and Sala-i-Martin (2014, p. 8) in relation to the results captured from Angolan and Indian remote sensing experiments is related to the gap in domestic surveys (e.g., reported income increase) and national accounts (e.g., World Bank Per Capita Income Data). Specifically, there is a stronger correlation between luminosity and per capita income data than with the household surveys and social reports, and the researchers attribute the weakness in the latter category to reliability and consistency issues associated with these field-level investigations (Pinkovskiy & Sala-i-Martin 2014, p. 8). Yet Blumenstock et al. (2015, p. 1073) recognise that many nations are restricted to fieldwork surveys and ground-level assessments of poverty data which have been proven to vary in terms of accuracy by as much as 50% from the ground-truth measures of regional wealth and economic productivity. For this reason, researchers are seeking quantitative correlations which they can use not only to model poverty characteristics over time, but also to improve the reliability of the relationship between a given metric (e.g., luminosity, housing

density) and its predictive effects for poverty modelling (Blumenstock et al. 2015, p. 1073).

As an alternative to night-time light sensing techniques, Jean et al. (2016, p. 790) describe a new machine learning method for the extraction of socio-economic data from daytime satellite imagery in high resolution, applying the technique to a five-country case study in order to compare readouts and analysis against local, domestic georeferenced datasets regarding economic traits. Due to the lack of trained sets and object-based recognition patterns, Jean et al. (2016, p. 790) adopted a '*transfer learning approach*' to remote sensing of poverty levels, adopting a '*noisy but easily obtained proxy for poverty*' that could be used to train a deep learning system. Specifically, this approach used a night-time light training model to help a CNN learn classification strata for a range of urban characteristics (e.g., urban density, roofing materials, agricultural areas) (Jean et al. 2016, p. 791). When applied to daytime classification models, distinctions between key features such as roofing material and distance to an urban area were accurately filtered, and the researchers applied the trained sets to the asset-based differentiation between key image assets (Jean et al. 2016, p. 792). By comparing the accuracy of the transfer model with other remote sensing exercises, such as nightlight calculations or cell phone data, Jean et al. (2016, pp. 792–3) demonstrated an approach with increased efficiency that is effective for poverty-oriented cluster assessment. When efforts to assess differences within clusters are employed, the model becomes less effective, as granularity is constrained by both the capabilities of the CNN-derived coordinates and the resolution of the geographical imagery (Jean et al. 2016, p. 793).

As technological models evolve, the intersection between physical and digital assessment is diversifying, with researchers such as Blumenstock et al. (2015, p. 1074) proposing that other indicators such as mobile phone use can serve as proxies for wealth in developing nations. The rationale behind this approach is based upon technological penetration—a representation of individual wealth and economic means that is characterised by an increase in mobile data consumption that can be merged with personal economic data in order to predict wealth (Blumenstock et al. 2015, p. 1073). Through a paired survey and remote sensing experiment, Blumenstock et al. (2015, p. 1076) present a regional snapshot model which suggests that in emerging

economies like Rwanda, the pattern of mobile phone activity over time is positively correlated with regional wealth, allowing researchers to model patterns of growth reflecting urbanisation processes and district services.

In another study, Ma et al. (2019, p. 2) adopted a novel approach to spatial economic analysis, predicting that there may be a correlation between typeface usage in building signage and regional economic wealth, classifying nine class-specific typefaces according to their relationship with modern amenities. This novel approach involved two central stages: (1) typeface classification based upon urban audits to determine amenity type/frequency (see Figure 2.4), and (2) regional classification and comparison (Ma et al. 2019, p. 3). By applying this technique to a geographical area with robust socio-economic data (London), Ma et al. (2019, p. 4) were able to successfully map typeface characteristics using remote sensing capabilities, Google Street View, and a deep CNN capable of learning and automatically evaluating typefaces and their associated characteristics relative to this classification matrix. Similar to the scalar limitations reflected in Blumenstock et al. (2015, p. 1074), the generalisation of typeface characteristics presented by Ma et al. (2019, p. 5) is vulnerable to a range of heterogeneous influences including urban rehabilitation, regional preferences, socio-cultural biases, and other interference categories, which make the narrow visualisation of typeface an unreliable solution for socio-economic modelling.

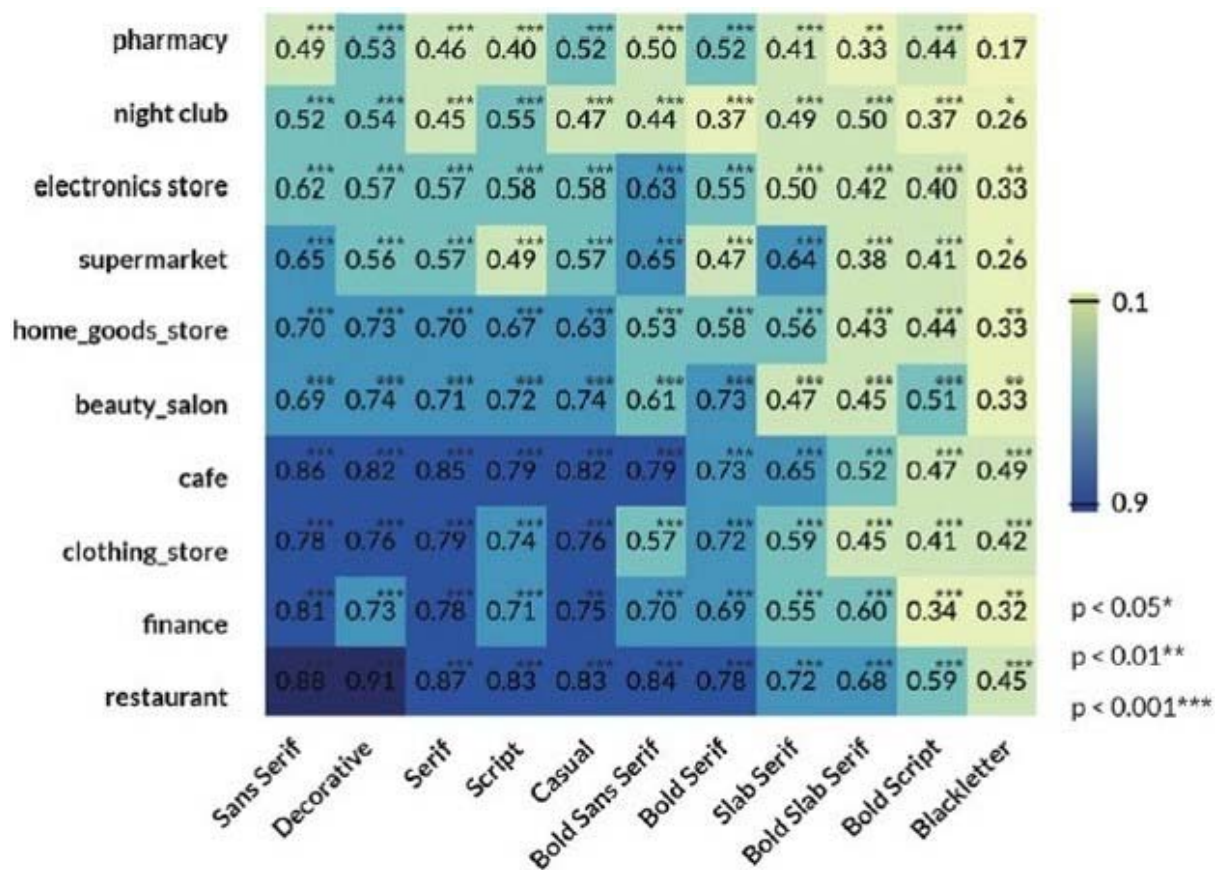


Figure 2.4: Correlation of typeface with amenity type (Ma et al. 2019, p. 3, Figure 2)

Models proposed by Blumenstock et al. (2015, p. 1074) and Ma et al. (2019, p. 5) uses different contextual and semantic triggers to visually track poverty according to a discrete set of cues or characteristics. One of the challenges is what Law, Paige, and Russell (2019, p. 3) refer to as the challenge of 'heterogeneous good', a value determination that is based on the weighted evaluation or interpretation of a quality relative to the assessment metric. The typeface correlation model developed by Ma et al. (2019, p. 5) demonstrated a classification problem that was resolved by repeated comparisons across a regional landscape in order to identify weighted similarities. In relation to housing, Law et al. (2019, p. 3) describe a '*hedonic price approach*' which suggests that housing attributes can be used to predict real estate values, where variations (e.g., one bedroom versus two bedrooms) can be used to classify the value relative to competing traits. To realise a maximal '*predictive accuracy*' of the hedonic price model, Law et al. (2019, pp. 10–11) compared CNN perceptual calculations based on attributional features with real housing prices, calculating correlational relationships between image orientations (e.g., street, aerial) and determining the degree of error for various street orientations and visual characteristics (e.g.,

obfuscation). Although the approach revealed encouraging results, the lack of continuity in housing facades, designs, and attributes resulted in a high degree of variation, which was magnified when the training set was transferred to another region or neighbourhood where structural traits were regionally indicative or highly heterogeneous (Law et al. 2019, p. 16).

For emerging economies, the accurate representation of regional poverty through remote sensing is a critical problem that is motivating innovative tactics in CNN and object-based classification according to feature characteristics (Ayush et al. 2020, p. 3). The following section evaluates remote sensing solutions and discusses the challenges of reliability and consistency associated with field-level investigations, call data records from mobile operators not being readily accessible, the possibility of granularly assessing within lowest levels of clusters, heterogeneous influences, and transferability to other regions. To overcome these challenges, it is imperative that further research be required to use cost-effective algorithms that not only use readily accessible data, but also enable granular poverty assessments supported by artificial intelligence (AI) and can be transferred to other less developed countries for socio-economic policy development and interventions.

2.7 A background overview of machine learning

Kovalerchuk (2018, p. vii) acknowledges that in modern problem solving, high-dimensional data shapes information analysis, management, and strategy-making as researchers visually compare hundreds of complex, multidimensional relationships across a discrete dataset. Knowledge discovery is characterised by Kodratoff (1999, p. 1) as a data mining process that is oversimplified, a field of research that relies heavily upon inductive reasoning in order to draw meaningful conclusions from otherwise unrelated datasets. A subset of AI, machine learning is implemented by leveraging mathematical models that use training data to make predictions, yet have not been programmed to make the said decisions explicitly (Zhang 2020, p. 223). As machines are trained and focused, they imitate the human learning process and their capacity for rule-based identification of targets or critical data allows researchers to apply replicable comparative techniques to the assessment of varied phenomena

(e.g., the relationship between nightlights and poverty) (Jean et al. 2016, p. 792). In fact, significant advances have been made in the use of nightlights, satellite imagery, and crowd-sourced map data in estimating socio-economic equality in recent years, and therefore these indicators can act as reliable predictors of wealth levels.

Machine learning is mostly categorised as either supervised or unsupervised. Supervised learning involves learning from examples of labelled training data whose outcome is known, while unsupervised learning involves using an unlabelled dataset and then finding patterns or knowledge as a solution (Zhang 2020, p. 224). In supervised learning, the known data is used for the model training process continuously until the required accuracy level is reached, and then, through knowledge transfer, will generalise on unknown data. With unsupervised learning, the model uses clustering or association to reduce redundancy or organise by similarity to extract rules. The following sections discuss the main machine learning concepts that were applied to this study.

2.7.1 Deep learning

Emulating the brain, deep learning is one of the ML methods that uses representation (feature) learning based on a neural network to process external input features to an output (Goodfellow, Bengio, & Courville 2016, pp. 5–8). Neural networks consist of neurons that are connected to each other and form deep neural networks, similar to neurons in the nervous system. The neural network consists of three main different model layers: the input layer, hidden layers, and an output layer (Figure 2.5).

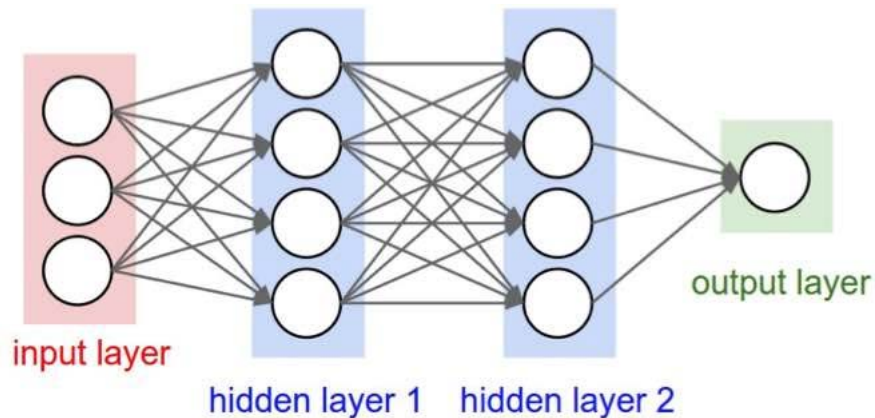


Figure 2.5: A 3-layer neural network with three inputs, two hidden layers of four neurons each, and one output layer (Karpathy 2020, p. 1). The connections (synapses) between neurons are across layers but not within a layer.

The input layer is also called the visible layer as it contains variables that are observable, takes in raw data such as the image, and stores each pixel value in a neuron. A series of concealed layers extract the abstract features from the image, where a neuron will serve as a sifter that is triggered each time a precise feature is detected. It is called '*hidden*' because the input is not given as a value but visualisations of relationships in the observed data for each feature represented (Goodfellow et al. 2016, p. 6). The output layer recognises the objects and categories that are present in the image. The success of how an algorithm performs vision-based tasks is determined by how the hidden layers perform. There are several variants that can be implemented as hidden layers to control their behaviour, and among them is CNN, which is implemented to analyse satellite imagery for this study. The algorithm is trained to discern human eye-recognizable features, such as points of interest, buildings, or roads which can be co-related with night-time light intensity, which will then be transferrable to predict poverty (Jean et al. 2016, p. 790).

2.7.2 Convolutional Neural Network (CNN)

CNNs have been applied to deep learning solutions in order to learn the important features of input images to avoid an attempt to come up with a method to extract the features (Long, Gong, Xiao, & Liu 2017, p. 2486). In a CNN, the convolutional layer is responsible for extracting features by learning from high-dimensional data when it slides or '*convolves*' over input images. Down-sampling is

required to address sensitivity issues where the location of specific features impacts classification from shifts and distortions (Yamashita, Nishio, Do, & Togashi 2018, p. 615). In order for the CNN to generate optimal output, it must be trained based on a training set that includes labelled remote sensing images and the corresponding target outputs that confirm or reject the identification of the feature (Yu, Wu, Luo, & Ren 2017, p. 6). The VGG16 CNN model architecture (Simonyan & Zisserman 2015, p. 8) visualised in Figure 2.6 includes convolution layers, pooling layers, and full-connection layers (Long et al. 2017, p. 2488) and is now a significantly popular model that is preferred nowadays for classification tasks in computer vision. The layer's functions are described below.

- Convolution Layer: Utilises several filters on local receptive fields to generate different feature maps using the maps of the previous layer or input (Long et al. 2017, p. 2488).
- Pooling Layer: Additionally, it uses filters as well to generalise the brief representation of the convolution layer to reduce the number of parameters (Long et al. 2017, p. 2488) in order to reduce the spatial information of features (e.g., adjacency and rotation). It can include maximum or average pooling. It is also a way to down-sample an input image and introduce transitional invariance to the network in order to identify objects in images regardless of the appearance.
- Full-Connection Layer: Extracts the output of the convolution or pooling exercises and predicts the best label to describe the associated image.

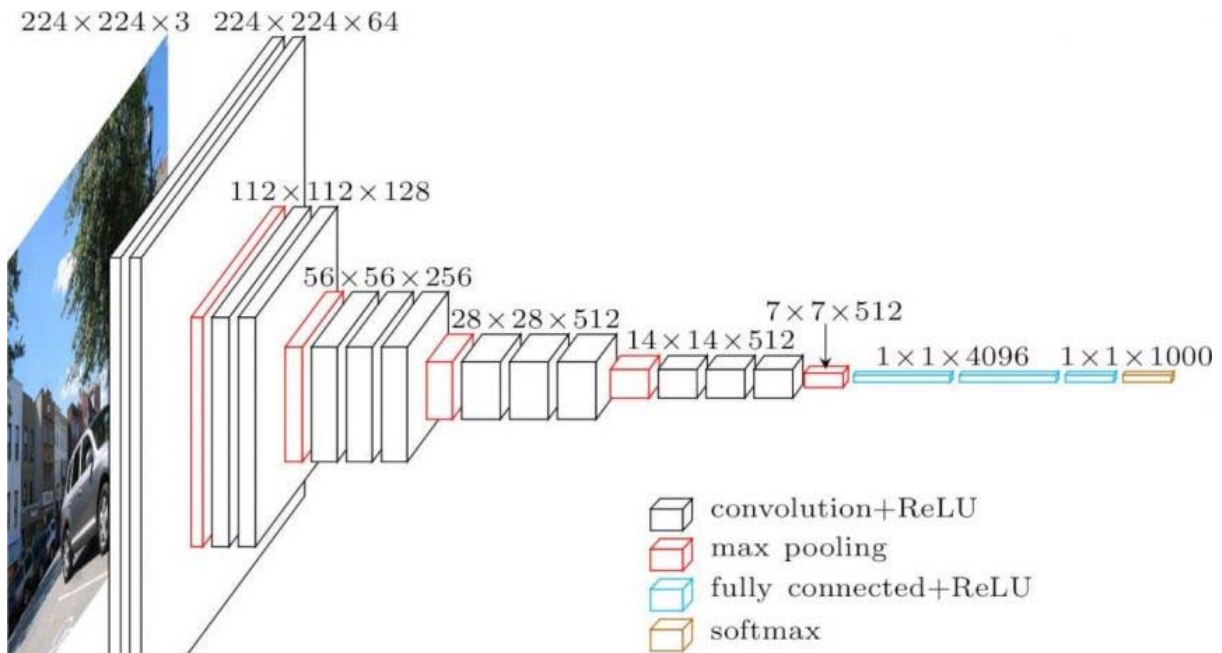


Figure 2.6: A 16-layer VGG16 pipeline model architecture for $224 \times 224 \times 3$ image sizes, with 1000 classes in the softmax layer (rightmost). It shows all the stages in the layer, including the convolutional, pooling, and a fully connected final layer (ul Hassan 2018)

The major component in extracting features from satellite images is a CNN model architecture. A CNN with VGG16 architecture is a convolutional neural network that consists of 16 blocks, also known as trainable layers (Simonyan & Zisserman 2015, p. 2). The authors have designed the VGG to have 16 layers that are composed of stakes of convolutional base layers which implement neurons with 3×3 receptive fields and max-pooling layers for spatial pooling (red-coloured in Figure 2.6) to aid in down-sampling. The convolutional base is followed by a Rectified Linear Unit (ReLU), an activation non-linearity function. In a network layer, activation functions are required to fire exact neurons (Krizhevsky, Sutskever, & Hinton 2017, pp. 85–86). The last three layers are called fully connected layers, which have 4096 neurons and whose output are feature maps that are fed into a 1000-way softmax one-dimensional array of vectors that produces a distribution over the 1000 class labels (Krizhevsky et al. 2017, p. 87). The input holds a raw pixel value of an image, and $224 \times 224 \times 3$ refers to width 224, height 224, and with three colour channels red, blue, green (RGB). When comparing the performance of neural networks, the advantage of using CNN architecture presented by Yamashita et al. (2018, p. 612) is related to its computational efficiency in filtering important features in image processing automatically without the need for human supervision. CNN models can be pre-trained to recognise 1,000

classes of a dataset such as ImageNet and fine-tuned for research that uses satellite imagery datasets (Simonyan, Vedaldi, & Zisserman 2014, p. 1).

2.7.3 Transfer learning

As a critical innovation in machine learning and artificial intelligence (AI) deployment, transfer learning involves the use of both the 'data in the target task domain' and any of the learning processes in the source domain, including the training data, models, and task description' (Yang, Zhang, Dai, & Pan 2020, p. 8). Due to their dependence upon massive training datasets, deep learning models require large banks of data in order to promote an understanding of the latent patterns and characteristics that improve predictability and replication of the identification or modelling output (Tan et al. 2018, p. 270). By rejecting the 'hypothesis that the training data must be independent and identically distributed with the test data' (Tan et al. 2018, p. 271), deep transfer learning uses independent tasks and a CNN model to extract features and classification measures from high-level representations (e.g., satellite classification data) (Xie et al. 2016, p. 3929). Figure 2.7 shows the two basic ways pre-training can be applied for transfer learning, which are fixed feature extraction and fine-tuning (Yamashita et al. 2018, p. 620).

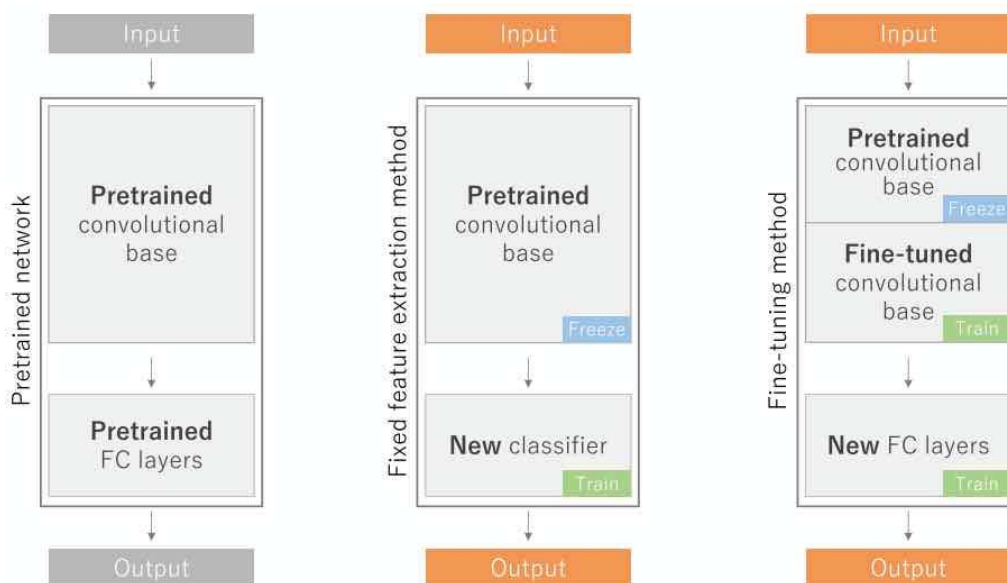


Figure 2.7: Transfer learning methods (Yamashita et al. 2018, p. 621, Figure 10)

Transfer learning enables a model to recognise features of an image (as the network would have learnt them from a small dataset) where the network is pre-training on an extremely large dataset from across many domains, such as ImageNet, which yields better performance and faster training times when applied to a specific task of interest (Jean et al. 2016, p. 790). ImageNet contains over 15 billion high-resolution images of over 20,000 categories, which were labelled through crowd-sourcing (Krizhevsky et al. 2017, p. 85). For socio-economic studies that use geospatial data, the approach involves training an algorithm to predict the intensity of night-time lights first in order to obtain more detailed poverty data. This follows an assumption that brighter lights at night signify a more economically developed place, compared to less well-lit areas, thereby using night-time lights as a substitution for socio-economic activities.

2.8 Conceptual framework

Although the central foundations of these studies are based on a variety of overlapping and varied remote sensing techniques, night-time light data collection, and crowd-sourced map, there are several unifying considerations that predict the efficacy of any model in future investigations. Through a review of the literature in this field, it is evident that there are several tenuous and varied relationships that need to be tested in order to resolve the array of procedural and systemic gaps related to multimodal remote sensing techniques:

- **Core Concept 1:** Machine learning and object-oriented modelling allow researchers to distinguish between geospatial and satellite map features that present direct indications of poverty profiles (Arribas-Bel et al. 2017, p. 4; Holloway & Mengersen 2018, p. 2).
- **Core Concept 2:** There is a strong and testable relationship between geospatial data sources and poverty that can be compared with key social indicators (e.g., education levels, access to electricity, access to water, durable household goods, and road network) in order to determine changing patterns

of regional deprivation (Fatehkia et al. 2020, p. 1; Jean et al. 2016, p. 791; Xie et al. 2016, p. 3934; Zhao and Kusumaputri 2016, p. 1).

- **Core Concept 3:** A multimodal, multi-banded GIS-based solution that combines various datasets will likely reveal greater accuracy in predicting poverty than a single band day or night-time model alone (Engstrom et al. 2017, p. 7).
- **Core Concept 4:** Post-processing benchmarking is critical for confirming the accuracy of the model and predicting whether the observations achieved during the assessment are comparable to government reports or regional publications (Jean et al. 2016, pp. 792–3; Pinkovskiy & Sala-i-Martin 2014, p. 8).

Figure 2.8 provides a conceptual framework that aligns the findings of the literature to prescribe the structural and evidence-related requirements to design and apply a crowd-sourced and remote sensing strategy to the assessment of poverty in South Africa.



Figure 2.8: Conceptual framework

2.9 Chapter summary

This chapter has provided an in-depth overview of the range of remote sensing techniques that have been explored as researchers strive to extrapolate meaningful evidence from a growing spectrum of GIS databases. Advanced imaging capabilities being developed in a variety of recent experimental studies are indicative of the power of remote sensing and data integration exercises. By applying systemised analytical techniques, machine learning, and extensive model training, it becomes possible for researchers to exploit high-quality, high-fidelity images to illuminate meaningful

patterns and phenomena. For nations in which the cost of large-scale population studies is likely to improve investment, remote sensing capabilities offer an economical solution that can be administered across geographic channels by specialists in particular socio-economic fields. The models reviewed during this chapter have highlighted several targeted solutions for designing, administering, and analysing an experimental platform that can be deployed across a broad range of locations and indicators. The following chapter will outline the methodological solutions that were used to craft and administer remote sensing to a discrete problem of poverty in South Africa, a nation of extreme socio-economic inequality.

Chapter 3. Research methodology

3.1 Introduction

The spectrum of methodological techniques and poverty assessment indicators introduced in the preceding chapter is broad and multi-faceted. As researchers struggle to parse the noise and complexity of poverty indicators, we are faced with a variety of hurdles, ranging from data availability, consistency, to measurement accuracy. Methodological fidelity and accuracy are critical to the transferability of empirical findings and the applicability of mapping and modelling techniques to more complex, systematic problems. Consequently, the design and implementation of an experimental resource require selecting the techniques that are most appropriate and then applying them to a dataset that is verifiable, replicable, and inherently comparable to the existing overlapping discourse. This helps researchers plan and execute a study to answer the research questions (Saunders, Lewis, & Thornhill 2015, p. 163). By comparing studies that have focused on the challenges of night-time light assessment and classification, this chapter outlines specific sources of evidence and analytical techniques applied to the methodological approach. The following sections will outline the methodological selection process employed over the course of this study and will discuss both the origins of the data collected for this experiment and the techniques applied to facilitate this analytical solution applied to the machine learning problem.

The presentation of this chapter broadly follows the scheme proposed by Saunders et al. (2015, p. 124) where each onion layer is a detailed description of the research process stages followed. The research paradigm in Section 3.2 explains the philosophical considerations of the study, South Africa as the study area is discussed in Section 3.3, and the research approach established to deduce the truth is discussed in Section 3.4. Furthermore, this chapter will delve into the details of the data sources and uses of these data in Section 3.5. At the end of the chapter, ethical concerns are outlined. The chapter not only discusses the selected options, but also explains the logic behind the selected choice and techniques to answer research questions.

3.2 Research paradigm

Different approaches are available to do research. The method chosen depends on the research questions to be answered related to the research topic. The term 'paradigm' originated with Kuhn (1970) to define how members of scientific communities share beliefs, assumptions, and world views. A paradigm is a system of thinking or belief which is used to organise observations and reasoning (Bhattacharjee 2012, p. 17). As a basis for research design and methodological development, Bryman (2012, p. 24) acknowledges that the research paradigm serves as the epistemic foundation for selecting and applying techniques to a given social problem. Whilst many poverty studies and poverty models depend upon complex, multi-dimensional social surveys, the primary goal of these approaches is the extrapolation of transferrable insights and interpretations that can be used to inform and support improved domestic policies.

The extrapolation of evidence from empirical or observational sources through analytical and deductive techniques involves a series of structured steps and approaches that originated in the natural sciences and represents a positivist paradigm (Bryman 2012, p. 24). Indicative of patterns and predictable, measurable representations of evidence related to a particular problem (e.g., the propagation of poverty in urban versus rural spaces), a central advantage of the positivist paradigm is its ability to illuminate likelihood and statistical relationships, allowing for repetition and hypothesis testing (Jonker & Pennink 2009, p. 30). For the current study, the positivist paradigm, which is based on the empiricist view of epistemology, formed the basis for the deductive evidential techniques that were applied to overlapping datasets from South Africa.

3.3 Research approach

The research approach must be established before an investigation can begin. To discover the truth, there are two fundamental approaches: the deductive and inductive approaches. A deductive approach begins by defining basic statements about confirming how the world works based on research questions and the

hypothesis developed. This differs from the inductive approach, which requires defining truths that explain the phenomenon as the focus of the investigation by collecting and analysing qualitative data to establish links with research objectives before drawing new conclusions. The positivist empiricist view of epistemology, which requires obtaining knowledge through experimentation and observation, synchronises with the deductive nature of this study whose objective is to compare the performance of predictive techniques that crowd-sourced geospatial data over satellite imagery for South African poverty estimation. Therefore, the study follows a deductive approach which involves a series of steps taken to design and experiment with several models to choose the best performing.

3.4 Research strategy

The research strategy is a general approach used to respond to the research questions (Oates 2006, p. 25). There are several strategies available, including survey, case study, ethnography, experiment, and design and creation. This research can be classified as a knowledge discovery process that involves both elements of data science and engineering, as it involves building a model that uses data to test and confirm whether it works as hypothesised. It combines data science and knowledge discovery processes to adopt the design science research method. Originating in the field of natural sciences, design science research was proposed by Simon (1996, p. 198) who suggested that researchers “devise artefacts to attain goals” in order to structure analytical instruments according to the ‘*character of the artefact and the environment in which the artefact performs*’. Artefacts are further explicated by Dresch, Lacerda, and Antunes (2015, p. 56) as “meeting points between the inner environment’ and the ‘surroundings in which they operate””; therefore, the purpose of design science is to create new knowledge by developing “new artefacts that contribute to better human performance”. Although it is idealistic in such a broad representation of its affective capacity, by establishing problem solving as the basis for the design of the artefact, design science makes the relationship between structure and optimisation intrinsic to the focus and purpose of the design process itself.

For the current study, design science informed several considerations that shaped this software-based, but sociologically and data-science-oriented, problem. Hevner's (2007, p. 88) design science research cycle is linked with the Knowledge Discovery in Databases (KDD) process of Fayyad, Piatetsky-Shapiro, and Smyth (1996, p. 29) to come up with an analytical solution that works as a research intervention instrument. Design science follows a structured process in its effort to support the design of artefacts, which is comprised of the relevance cycle, design cycle, and rigour cycle (Hevner 2007, p. 88), as shown in Figure 3.1.

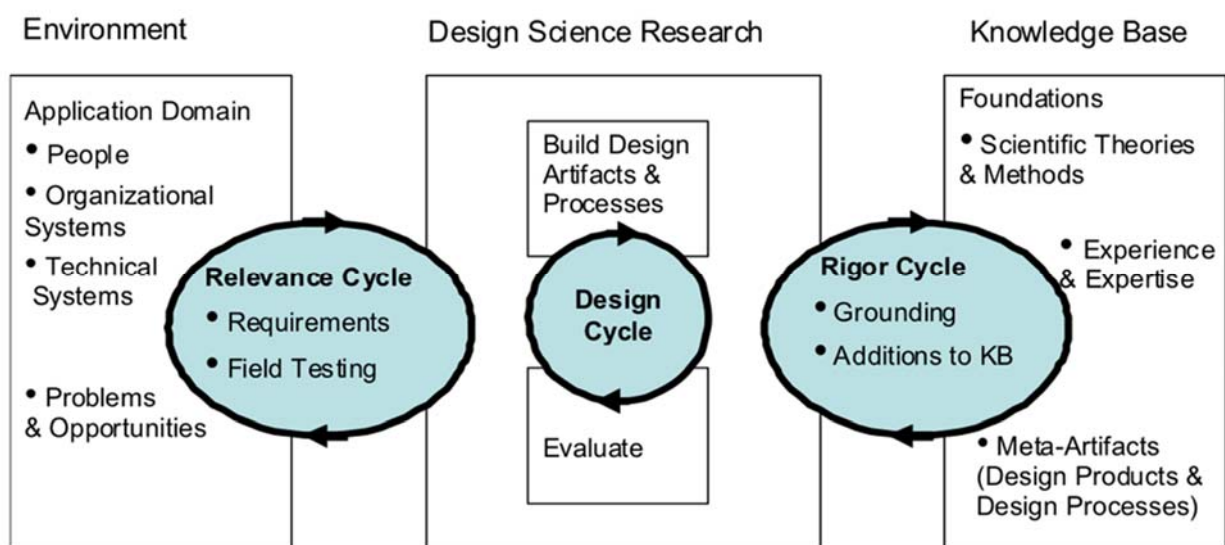


Figure 3.1: Design science research cycles (Hevner 2007, p. 88, Figure 1)

The design science research methodology has iterative expert activity steps to design and evaluate an artefact which, when followed, will result in the output being fed into future iterations, thereby enhancing the overall understanding of the phenomenon under study and refining the characteristics of the resulting artefact (Hevner et al. 2004, p. 85). The relevance cycle aims to introduce artefacts in a domain consisting of people and the organisational and technical system that interacts to improve the current environment (Hevner 2007, pp. 88–89). In Chapter 2, existing research by experts was reviewed to propose the requirements for developing a cost-efficient machine learning model to estimate wealth indices from readily accessible datasets. Essentially, the design cycle is where the artefact will be developed and evaluated iteratively until a satisfactory design is achieved following the work of Peffers et al. (2007, p. 54) design science research methodology adopted for this study (cf.

Chapters 3 and 4). Lastly, the rigor cycle ensures that the artefact is innovative and meticulously developed. The artefacts were tested using experiments, evaluated, and discussed (cf. Chapters 5 and 6).

For prior solutions, Blumenstock et al. (2015, p. 1074) and Ma et al. (2019, p. 2) have proposed novel measures of classification and structural identification without defining any significant or purposive objective. Classifying building font choices, for example, in Ma et al. (2019, p. 2) is a novel representation of deep learning capabilities, but it neglects the broader concept of gentrification upon which the research originated. By adopting a design science basis for this study, the problem, poverty in South Africa, informed the methods, categorisation, and optimisation of the analytical tools designed for this research. Accordingly, this process has revolved not only around the expectation of a novel software-based proof of work output, in addition to a positive contribution of this artefact to the field of research surrounding poverty, regional inequality, and the lagging process of socio-economic development in under-resourced urban spaces.

3.4.1 Design science research

Design research originated from engineering and the sciences of the artificial (Hevner et al. 2004, p. 76; Simon 1996, p. 216). Its importance is highly acknowledged in the information systems literature and has received increased attention in the last decade from various authors such as (Gregor 2002), (Hevner et al. 2004), (Peffer et al. 2007), and (Simon 1996) among others. Design science research concerns itself with the knowledge and understanding of the domain of a problem and then gains new knowledge by developing a novel artefact and analysis of the artefact in use with abstraction (Hevner et al. 2004). As a result of following specific guidelines for evaluation and iteration, Hevner et al. (2004, p. 85) says that the end product of a design research should help to close the gap between the current state and an ideal goal. Although its application is widespread among research fields, it is most notable in engineering and computer science disciplines.

Peppers et al. (2007) formalised the design science process and presented a methodology that acts as a model for carrying out and evaluating research. Unlike Hevner et al. (2004), the model by Peppers et al. (2007, p. 56) which is adopted for this study provides an explicit process description. The methodology guided the research process, in addition to legitimising it by providing overarching elements on which an ideal research is formed.

The adopted model proposed by Peppers et al. (2007) is depicted in Figure 3.4. The Peppers et al. (2007) methodology starts with addressing a problem and the objective motivation, and then iteratively moves to the rest of the stages which involve building and evaluating the artefact (rigour cycle) until its application in the relevance cycle (Hevner 2007, pp. 1–2; Hevner et al. 2004, pp. 87–88). The purpose of design science goes hand in hand with the goal of this research, which is to solve insider problems by developing a model for integrating remote-sensing and crowd-sourced data.

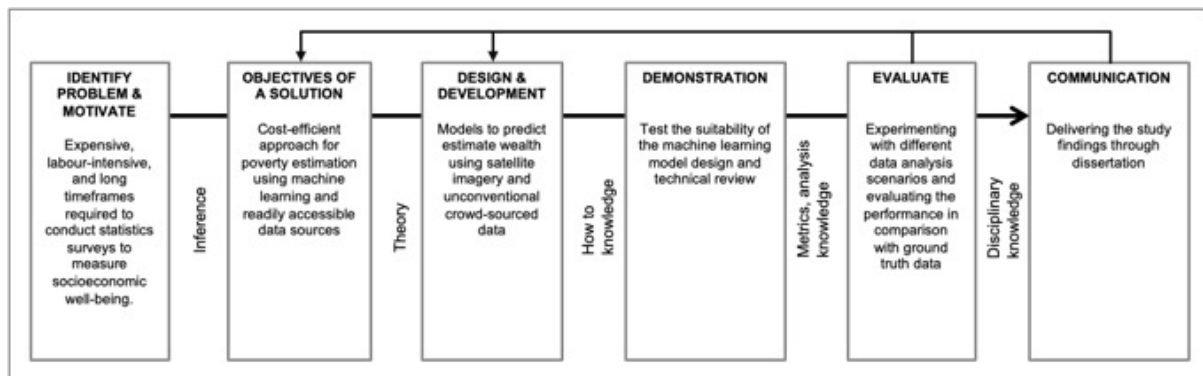


Figure 3.2: Design science research activities, according to Peppers et al. (2007, p. 58)

The Peppers' et al. (2007) process methodology consists of six elements, which will now be discussed in detail regarding how they are applied in this research study.

3.4.1.1 Identify the problem and motivate

Estimating poverty indicators using accurate, granular, and up-to-date data is essential for a stable economy and political environment, as well as for the identification of vulnerable regions for humanitarian assistance efforts. Traditional data collection methods for poverty statistics are conducted through on-the-ground

household surveys, which requires heavy organisation, is time-consuming, is updated after long time frames, is sometimes limited to accessible areas, and is sometimes considered subjective, making it difficult to rely on one for up-to-date socio-economic data (Engstrom et al. 2017, p. 2; Jerven 2014, p. 6). The main research question for this study focused on cost-effective approaches that can be adopted to estimate poverty. The need for an artefact (in this case, a machine learning model) for the estimation of poverty on a large scale using readily accessible data in a faster and easier way.

By evaluating recent machine learning and satellite imagery approaches, the motivation of the study is for the adoption of open-source and freely available data to mitigate the challenges of sparse data, shortcomings and weakness can be identified and improved upon leading to cost-efficient machine learning models to predict the poverty indicators that will benefit resource-constrained agencies. The main research question as stated in Chapter 1, Section 1.5 is: *Given the heterogeneity of economic characteristics in developing nations like South Africa, which cost-efficient machine learning methods can be used to determine the relationship between special features and socio-economic variations?*

3.4.1.2 Define the objectives of the solution

The activity of defining the objectives of the machine learning model is the second step in the design science research methodology suggested by Peffers et al. (2007). In line with the main and sub-research questions proposed for the study, the required information and data were gathered through various means and sources. The databases of academic literature and the internet were used to review the existing literature on machine learning, remote sensing, and poverty measurement and their application in South Africa from Chapter 2 was combined with the problem statement from Chapter 1 (Section 1.4). The output of the activity provided a background for understanding the research problem; motivate its relevance and lay the foundation for designing a machine learning model (artefact).

3.4.1.3 Design and development

The aim and function of the third activity in the methodology is to address the design specifications and development requirements of an IT artefact through specification, which forms the output of the study. The artefact can be in the form of a construct, model, method, or an instantiation (Hevner et al. 2004, p. 78) and is a socio-technological innovation, which according to Hevner et al. (2004, pp. 81–83), can be grounded in theory of a problem to define ideas, practices, or products. In this study, the design requirements are converted into machine learning models that are developed in the following phases (cf. Chapters 3 and 4) as the artefacts:

- i. Predict wealth using satellite imagery: This model replicates and improves on a study by Jean et al. (2016) where night-time lights intensity is used as a proxy task, which then are used to compute average features embeddings per cluster to estimate wealth using regression models.
- ii. Predict wealth using cost-efficient crowd-sourced data: Use exploratory data analysis of night-time lights and OSM data to perform regression analysis of predictive models.

Through experiments, the performance of the developed models is evaluated. The first model which uses satellite imagery is evaluated using South African DHS data as ground truth by measuring co-relation between night-time luminosity and the average household wealth index. The second experiment is conducted on night-time luminosity data alone, and then together with the OSM data using several regression models to compare if the crowd-sourced readily accessible data performs better than proprietary or licensed data.

3.4.1.4 Demonstration

In this phase, ground truth data were applied to wealth prediction models to measure their performance and demonstrate the feasibility of the concept as experiments. The first model used night-time luminosity proxy from satellite imagery using CNN, deep learning, and transfer learning. The second model applied random forest and ridge regression classification algorithms to crowd-sourced data and night-time data separately and combined. The classification task was implemented using

Python to rapidly prototype the patterns using Google's Colab Notebook and machine learning libraries.

3.4.1.5 Evaluation

The evaluation phase involves evaluating the artefact implemented against the objectives formulated in Phase 2. The results of the demonstration phase are observed and compared with the objectives to formulate conclusions on the performance of the models and how they meet the objectives. Evaluation and testing of the models must be rigorously demonstrated to ensure utility, quality, and efficacy, which are necessary for conducting research in order to produce some poverty indicators (Hevner et al. 2004, p. 85). Improvements can be made to the artefact iteratively until a decision is made that the model performance is comparable and good enough before proceeding to the next phase.

3.4.1.6 Communication

The last element of the methodology involves synthesising and concluding the findings of the research phases, outcomes, and processes. Additionally, it includes guidelines for using the model and generalisation possibilities as well. Research questions are answered, reflections on the research are formed, conclusions are formed, and recommendations for future research are formulated from the conclusions. Moreover, the main part of this phase is the presentation of the results and research in a scholarly or scientific format, as documented in Chapter 6. The next section discusses the adopted research strategy and experimentation utilised to build and test the implementation models.

3.5 Target country

The current study focuses on South Africa, a nation characterised by widespread, regionalised poverty, with approximately 40% of the population reportedly living below the upper-bound poverty line (Statistics South Africa 2019a). South Africa

reflects a dualistic economic system with extremely wide gaps in wage equality and a Gini coefficient of 0.63 in 2015, an indicator that had actually increased since the appointment of the modern regime in 1994 (World Bank 2019, p. 1). Despite an expanding domestic economy, more than 30.4 million South Africans live on less than \$75 USD per month; With more than 50% of the population below the age of 35 currently unemployed, the domestic disparities between the rich and poor populations offer a glaring indication of pro-business, anti-poor domestic policies (Worstall 2017).

Currently, approximately 10% of the upper-class population holds around 71% of the net wealth, whilst low levels of intra-generational mobility result in pass-through poverty and exhibit persistent inequality across the population demographics (World Bank 2019, p. 1). With more than 18.8% of the population living on less than \$1.90 USD per day, the South Africa Gateway (South Africa Gateway 2019, p. 1) reports that domestic poverty is systemic and geographically concentrated, resulting in the regionalisation of poverty. Characterised as multi-dimensionally poor, poverty in these critical regions is measured according to four core measures: health, education, living standards, and economic activity (South Africa Gateway 2019, p. 1). Although more than 22 years have passed since the collapse of apartheid, Figure 3.2 presents a stark juxtaposition of wealth and poverty in South Africa, with major roads, rivers, and industrial zones providing divisionary boundaries for class-based separation (*The Guardian* 2016). As of 2018, Statista reports that more than 66% of the South African population resides in urban areas and more than 70% of that population is employed in white-collar industries such as the service sector.



Figure 3.3: Visual distinction between upper middle class and poor residential areas in South Africa (The Guardian 2016)

South Africa is made up of nine administrative units called provinces, divided into 52 administrative units called districts (eight of which are metropolitan cities) with 205 local municipalities. At the finest level of administrative units, instances are called wards. There are 4,392 wards. South Africa used to conduct a census every five years and has since extended the interval to ten years due to a lack of capacity within Statistics South Africa (Statistics South Africa 2020).

3.6 Data sources

The compositional approach adopted by Jean et al. (2016, p. 791) provided a viable, multi-layered solution to the problem of data fidelity, poverty analysis, and density errors in remote sensing of both urban and rural areas. As demonstrated by Pinkovskiy and Sala-i-Martin (2014, p. 8), poverty visualisation must be confirmed to ensure that the increase in lights being seen is closer to what is suggested by surveys before assumptions about the effectiveness of the machine learning model can be verified, as additional economic activity could be benefiting those who are already well-off, or images are recorded by aged satellites which are almost to be retired. Similarly, Zhao et al. (2019, p. 3), Zhao and Kusumaputri (2016, p. 1), and Head et al. (2017, p. 2) outlined specific database sources that could be used to systematically and comparatively analyse the relationships between poverty indicators and spatial

imaging analysis. Consequently, this study used overlapping visual data, crowd-sourced maps, and statistical datasets for comparative analysis, allowing regional indicators of poverty to be identified alongside the weighted output of the image and feature classification model. For example, as maximally objective data, the DHS USAID (2020) provided baseline indicators of regional poverty and socio-economic welfare. For night-time imagery dataset, the NOAA (2020, p. 1) imagery was downloaded. Regional satellite images were downloaded from Google (2020), whilst the volunteered geospatial data was downloaded from Geofabrik GmbH (2020). The data used for this analysis were captured from different critical resources as enumerated below.

3.6.1 Demographic and health survey

The publicly available Demographic and Health Survey (DHS) household survey data for South Africa were used as the ground truth, namely, to serve as a reference point to verify the validity of the analysis. It provides baseline indicators of regional poverty and socio-economic welfare outcomes for validating the machine learning models by comparing them with the wealth predictions. The most recent survey included in the analysis was conducted in 2016. DHS contains nationally representative household survey information on a wide variety of demographic, social, economic, and health-related indicators by sampling villages proportional to their size and is funded by USAID (2020). The South African DHS data contained 750 clusters. Key indicators identified within this model include education levels, access to electricity, access to water, and wealth index.

DHS provides geocoded data where the approximate location of each household is re-coded using the latitude and longitude coordinates and assigned a cluster or enumeration area surveyed in addition to the asset data. For each cluster, DHS also collects and provides the centroid's geographic coordinates. Each cluster is roughly equivalent to a village in rural areas and a neighbourhood in urban areas. DHS protects the privacy of the surveyed households by anonymising: the GPS coordinates are jittered by up to 2km for urban clusters and 5km for rural clusters (Burgert, Colston, Roy, & Zachary 2013, p. vii). A further 1% of rural clusters are jittered by up to 10km.

Displacements are a source of noise in training data and can add considerable measurement errors to regional wealth estimates. Although the DHS instrument does not collect consumption or income data, its measures of asset ownership, health, and education are commonly used in calculating MPI. The study included four responses from the DHS survey for socio-economic indicators (Tingzon et al. 2019, p. 426).

- **Wealth Index.** The DHS has a factor called 'wealth index', which is a continuous scale to measure socio-economic well-being. It is calculated as the first principal component of survey responses related to a household's wealth (Sahn & Stifel 2003), for example, roof material, television ownership, water supply. No further transformations were done on this column.
- **Education Completed in Years.** This is the most reliable measure of education completed in years for all respondents to the household survey who are at least six years old. In most cases, the respondent would have been the head of the household. The aggregate information required for the study is computed by calculating the average years of education completed by all households present in a cluster.
- **Access to Electricity.** This is an aggregate of household survey respondents who affirmatively confirmed that they have access to electricity within a DHS cluster.
- **Access to Water.** The survey has responses for the time it takes for each household to reach a drinking water source, measured in minutes. The time is set to zero in cases where the source of water is in situ (i.e., tap water, water well or bottled water). Responses are aggregated to compute the mean for all household respondents in a cluster.

3.6.2 Night-time luminosity

These datasets provide records of artificial light on the Earth's surface collected during the night using average radiance composite images using night-time data from the Visible Infrared Imaging Radiometry Suite Day/Night Band (VIIRS DNB), a remote

sensing instrument on the Suomi National Polar-Orbiting Partnership (S-NPP) satellite. The data were pre-processed by the National Oceanic and Atmospheric Administration (NOAA) National Centre for Environmental Information (NCEI) to eliminate stray light whilst emphasising light from towns. The dataset comprises cloud-free composite maps of all calendar years which are developed from all the available archived Operational Linescan System smooth resolution data. The NOAA NCEI suite is a version 1 product; version 1 products span the globe with light intensity data. The data is provided as inter-calibrated tiled raster layers.

The imagery recognises wavelengths from green to near-infrared and is distributed in the form of monthly composite cloud-free maps to maximise the amount of light received from the Earth's surface. Satellite imagery data was downloaded from intra-period indicators in annual bands at 15 arc-second geographic grids at a resolution of 0.742km × 0.742km grid cells (see Figure 3.3). The data sets are made available for free in GeoTIFF format as a set of six tiles on the NOAA/NCEI (2020) website. The continuous luminosity level was modulated from level 0 to 122 in South Africa and grouped into five luminosity intensity classes (1-5): low intensity (0), moderately low intensity (0.05-2), medium intensity (2-15), moderately high intensity (15-30), and high intensity (30-122) (Tingzon et al. 2019, pp. 426–427). The light luminosity intensity data is measured in units of $nW/cm^2 - sr$ (nanoWatts/cm² - steradian), and 2016 data was downloaded for this study.



Figure 3.4: Night-time imagery for Cape Town, South Africa (NOAA 2020)

Each tile is a set of images that are cut at the equator containing average radiance values and spans 120 degrees of longitude (NOAA/NCEI 2020, p. 1). The images were extracted on the basis of the latitude and longitude of the DHS clusters. Specific tiles with night-time luminosity data for South Africa were downloaded with Tag Image File Format (TIFF) data.

3.6.3 Daytime satellite imagery

The primary strategy for the daytime imagery model was to obtain high-quality satellite image samples to use in the training of the convolutional architecture for wealth prediction. The requirement was to achieve this merely from freely accessible sources of data or from those that keep their costs close to zero. For this reason, the Google Static Maps Application Programming Interface (API) was chosen as the sole source to download daytime satellite imagery. Daily Google API requests are restricted to 25,000 free of charge, after which billing is required or the user will get errors. Although the night-time satellite imagery allowed for classification of luminosity relative

to each of the regional sectors and density bands, it was critical to differentiate between the contributing factors. The banding approach employed by Jean et al. (2016, p. 792) was similarly adopted by Tingzon et al. (2019, p. 426) and involved obtaining daytime satellite imagery data for the target regions. For this research, the daytime satellite imagery data was collected corresponding to the DHS cluster centroids. Since DHS introduces noise to preserve privacy for the mean latitude and longitude of each household, large neighbourhoods will be considered to address noise in the cluster coordinates. Each DHS cluster comes already tagged as a rural or urban location. The geolocation information coordinates are also necessary in order to get images consistent with the DHS cluster coordinates.

After completing the Google Cloud registration process, an API and a Secret Key were provided to enable the extraction of regional images using the DHS geolocation and the required zoom level. The zoom level of Google Maps ranges from 0 to 19 describing the scale of the map. Google Maps is built on a 256×256 pixel tile system where the zoom level is 0 for a 256×256 pixel image of the whole earth. A 256×256 tile for zoom level 1 enlarges a 128×128 pixel region from zoom level 0. The download was carried out with a zoom level = 17, scale = 1, and image size = 400×400 pixels for 2016. A zoom level of 17 (1 pixel = 2.387m) means that each image covered ~1km in width and height. The resolution of the images was set at 400×400 pixels to match 0.25km^2 which is equivalent to a single pixel of the night-time lights data, which it typically covers. This will accelerate model training and extraction of geographical features during the unsupervised training process.

3.6.4 OpenStreetMap data

Despite the ease of access, Google Static Map imagery is proprietary and may incur costs if you request to download more than 25,000 static images per day, as this causes delays in downloading the required number of images for conducting studies. This research proposes a methodology that uses an open-source repository of geographic information contributed by volunteers. OpenStreetMap (OSM), an online publicly accessible database, is amongst the most established geographic mapping and analysis crowd-sourcing platforms. It is an open-source, large-scale collaborative

geospatial project that contains records of volunteered geographic information. OSM has grown to be the most popular geospatial data crowd-sourcing platform. The platform allows users, a community of global mappers, to accurately depict any interesting features they encounter that producers of commercial maps may have never represented. The completeness of projects that involve mass collaboration is an ongoing challenge (OSM 2020). In addition to the advantages of being open source, OSM is also used by humanitarian teams in health outreach projects, as it has the most complete coverage even when compared to commercial map engines such as Google Static Maps (Humanitarian OpenStreetMap Team 2020). The main focus for OSM in 2004 when it was initially launched was on roads and streets; however, it was broadened later on to include mapping of buildings, land uses, points of interest, and other notable geographic features (Barrington-Leigh & Millard-Ball 2017, p. 2). It uses nodes where each node has an ID, coordinates, and metadata to describe it. In other words, for each road, the metadata includes its capacity, category, whether it is paved, etc.

For this study, OSM data for South Africa were acquired from the Geofabrik GmbH (2020) website, an online repository containing OSM data as ESRI shapefiles. Data were used to train a model according to buildings, roads, points of interest, railways, water or waterways, transport, and land use related to areas of interest, increasing the granularity of the training set and supporting a more robust interpretation of the output. Three steps were followed to characterise the OSM data: (1) downloaded the South African OSM data; (2) features were extracted and classified according to roads, buildings, and POIs using QGIS algorithms, an open-source GIS software; (3) using shapefiles from the DHS GPS dataset, the survey clusters and OSM feature locations files were merged.

3.7 Research experiment

In order to design the experimental features of this study, it was important to review a variety of prior techniques before developing a solution that would have an adequate degree of specificity and predictive capacity to achieve the overarching goal of the research. This section introduces and discusses the applied methods used to

predict socio-economic well-being. The design of this experimental prediction model and the methods implemented in this study derived structure and instrumentality from Jean et al. (2016, p. 790), Tingzon et al. (2019, p. 425) and Head et al. (2017, pp. 1–2) and the precipitating night-time model introduced by Xie et al. (2016, p. 3931). Two computational frameworks are proposed to estimate wealth using satellite imagery and crowd-sourced geospatial information. The process of predicting regional socio-economic indicators using night-time light intensity as a proxy task for satellite imagery data to compute the average feature embeddings per cluster and use regression models on the cluster-level feature vectors. The second approach involved using OSM data and night-time lights data, first separately and then combined, with the hypothesis that integrating data from multiple sources will lead to performance improvements of the regression model compared to using data from a single source.

The analysis for this study will use design experiments as a means to evaluate design alternatives with respect to their impact on reality. Experiments will be conducted to gain general knowledge about the artefact to ensure verifiability and to find ways to improve the product design. Firstly, the datasets that were used in the development and testing of the computational algorithm are discussed in Section 3.6. Following that, the two different approaches for the estimation of socio-economic well-being are discussed. Figure 3.5 shows the research stages that were followed to carry out the study.

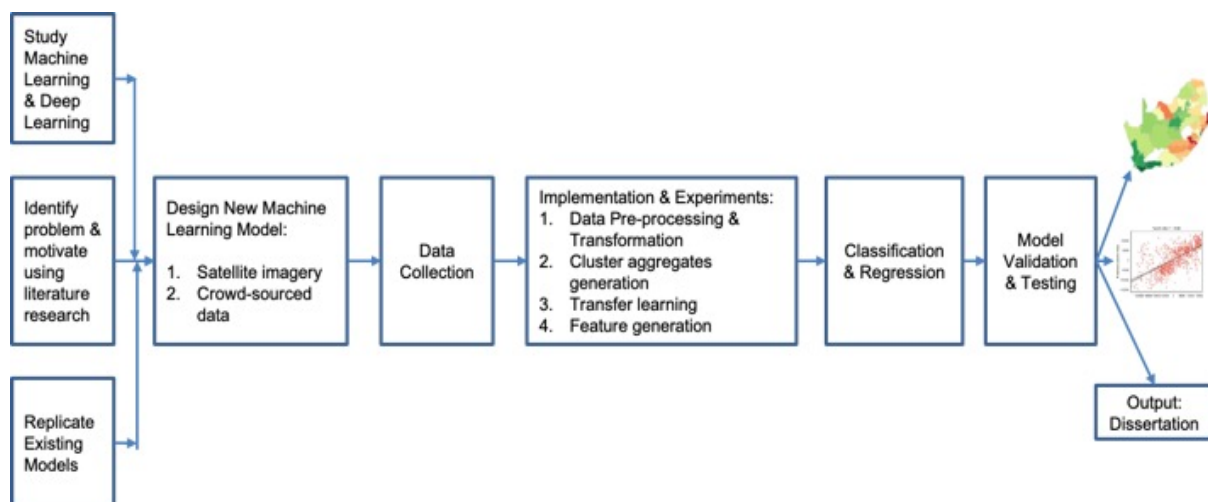


Figure 3.5: Research organization

This organisation follows the design science research process adopted for the study. The satellite imagery approach based on CNN and the transfer learning architecture was also developed as discussed in Section 3.8.1. The crowd-sourced information model was discussed in Section 3.8.2. Regression models were implemented to predict socio-economic well-being and simulations were performed to validate and measure performance as discussed in Chapter 5. Finally, in Section 3.8.3, the machine learning classification models used in the study are discussed.

3.7.1 Satellite imagery approach

The first method applied uses the transfer learning strategy, which means that an existing CNN architecture is modified to learn and classify images for the current study. The VGG-16 is an existing architecture that was developed by Simonyan and Zisserman (2015, p. 12) of the Visual Geometry Group at the University of Oxford. This is a very large-scale network that has 16 layers with trainable weights (13 convolutional of increasing depths and 3 fully connected). A singular critical assumption in designing this study was based upon the hypothesis that night-time light modelling would yield a positive proxy for economic activity (Jean et al. 2016, p. 792; Xie et al. 2016, p. 3930). Training a network of this magnitude requires tremendous computational power, which would not be feasible in the context of this project. However, since the network has been extensively pre-trained on the ImageNet database (Deng et al. 2009, p. 251), it contains an enormous variety of images belonging to many class varieties and has weights that are publicly available. The satellite images used for this study will have extractable features similar to the image features that were used to train the convolutional network. The model will classify the extracted image features based on night-time luminosity intensity classes at a DHS cluster-level corresponding to the daytime satellite images. The hypothesis for this experiment is that night-time lights have features that can be used as a proxy for economic development in the sense that wealthier places will have brighter lights at night compared to less affluent places. The following steps are followed for the experiment:

- Download the required datasets from sources and pre-process the data before it can be used in the models to ensure that it was in a format that can be understood by the model in addition to data cleaning, transformation, and reduction.
- Generate cluster aggregates for information.
- Predict night-time light intensity as a proxy task using the transfer learning approach to train a VGG architecture on satellite images for predicting the night-time lights bins and average them across a cluster.
- Compute the average of 4096 size feature embeddings per cluster to estimate wealth using ridge and random forest regression models.

After the extraction of vectors of learnt representations, they are averaged at cluster level and used as input to a regressor for the prediction of socio-economic indicators as shown in Figure 3.6.

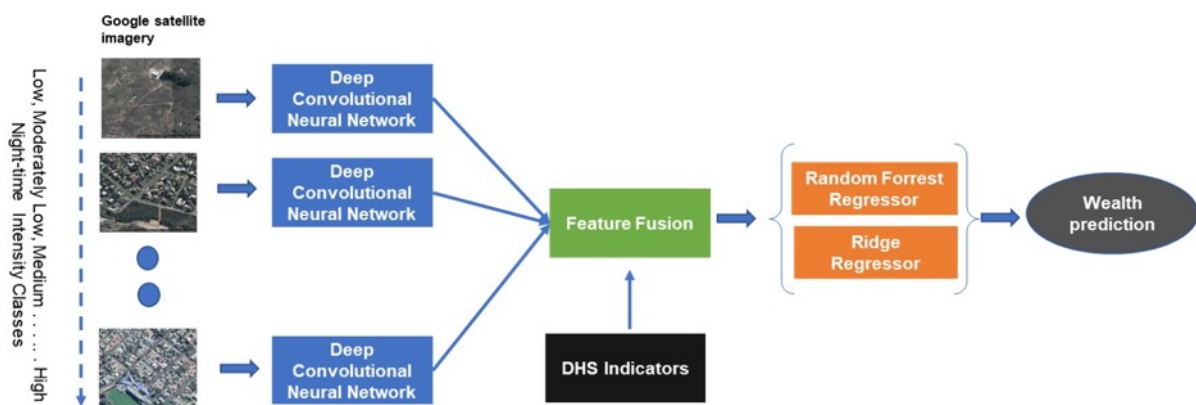


Figure 3.6: Satellite-based transfer learning model

3.7.2 Crowd-sourced information approach

The OSM model compared regressor performance on OSM features, both separately and combined, to predict socio-economic well-being. Furthermore, experiments are set to test the performance using features from both night-time and OSM datasets. Another critical assumption in designing this model was that using data from multiple cost-effective sources to train models will enhance the performance of the model. The model design is shown in Figure 3.7. Night-time features consist of

summary statistics and histograms extracted from nightlight pixels within each cluster using a script and saved as a .csv file.

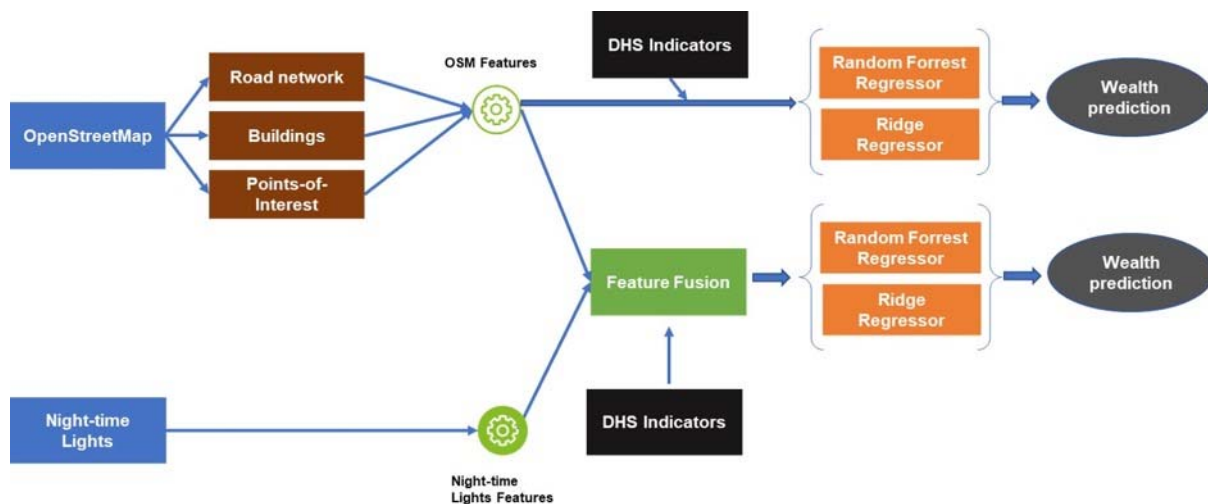


Figure 3.7: OSM model for the study

The model results will be compared as part of the model performance evaluation. Although this is a valuable starting point for this experiment, it was important to focus the training set on specific contextual, geographic, and structural characteristics associated with South African urban spaces. There were several stages to this modelling methodology in order to evaluate the merging of datasets for poverty estimation which are outlined as follows:

By selecting regionally specific images that were tailored to South African spaces with varied levels of daytime poverty indicators identified using the OpenStreetMap (OSM) model, any indications of poverty identified by the night-time model could be verified and assessed relative to their daytime counterparts. This secondary assessment model employed a multi-regional, multiple timestamp series GIS solution adopted from (Jean et al. 2016, p. 792) and trained to address the economic poverty indicators identified during the initial training sessions. In order for the OSM model to yield productive results, it was critical to first identify and classify the various building types (e.g., housing, mixed land use, storage, mercantile, medical institution), allowing cluster areas to be mapped relative to structural densities within each particular class. The critical analysis of this output thereby employed a comparison of several regression models trained on various OSM features, establishing the justified basis for socio-economic welfare.

3.7.3 Classification and regression models

In order to realise meaningful data-driven comparisons of the model performance in learning features, statistical techniques are applied to the resolution and analytical granularity of the knowledge base comprising the data that will be analysed. The study problems were treated as classification tasks. Classification techniques in machine learning include both supervised and unsupervised categories, as discussed earlier. The deductive approach to machine learning involves applying a set of reasoning rules from a training set or a range of premises to create a classifying output that can be applied to new instances (Kotsiantis 2007, p. 4). The unsupervised approach (e.g., cluster analysis) requires only input data without an associated response to find relationships or groups in the data. In contrast, supervised learning uses data that are already labelled with the correct values to train the algorithm. Most statistical learning methods are used for supervised learning tasks where the learning problems are either classification or regression. To improve classification and regression for prediction and avoid overfitting, James et al. (2013, p. 204) proposes regularisation and bagging among other options. To improve the performance of the models to derive features, the Ridge and Random Forest regressors were implemented as follows.

- **Ridge Regression:** A simple linear regression is an approach for estimating a quantitative response Y based on a single predictor variable X , assuming there is an approximately linear relationship between X (the input variables) and Y (the target variables) (James et al. 2013, p. 61). The primary goal of linear regression is to fit a function so that the residual sum of squares (loss function) is minimized. The model coefficients are defined through an optimisation method that seeks to minimise the sum-squared error between the predictions and the expected target values. Ridge regression (also known as Tikhonov) is a regularisation method that shrinks the regression coefficients by imposing a penalty term on them, which helps to reduce variance for better prediction where linear regression could become unstable due to input sensitivity. Since some of the features will be more important than others in predicting the wealth

level of a cluster, a substantial number of parameters may present challenges of low bias and high variance resulting in overfitting and less generalisability. The values of the ridge regression coefficient parameter minimise

$$\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j^2 = RSS + \sum_{j=1}^p \beta_j^2 \quad (3 - 1)$$

over β for a given λ , where λ is a tuning parameter ($\lambda \geq 0$) which is used to seek coefficient estimates that fit the data well by making the residual sum of squares (RSS) small, p is the number of variables available for use during prediction, n is the number of distinct data points or observations, i is used to index the samples/observations (from 1 to n) and j will be used to index the variables (from 1 to p). x_{ij} represent the value of the j^{th} variable for the i^{th} observation, where $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, p$ (James et al. 2013, p. 10). The shrinkage penalty, $\sum_{j=1}^p \beta_j^2$ has the effect of shrinking the estimates towards zero.

- **Random Forest:** A supervised classification mechanism that uses decision trees to allow training sets to be mapped or split into individual nodes that operate as an ensemble committee and are then combined to have one strong predictor. A central advantage is its simplicity; however, it is likely to require multiple training sessions and may not withstand more complex diagonal partitioning or replication challenges. An RSS splitting criterion is used as follows:

$$RSS = \sum_{left} (Y_i - Y_L^*)^2 + \sum_{right} (Y_i - Y_R^*)^2 \quad (3 - 2)$$

where Y_L^* = mean y-value for the left node, Y_R^* mean y-value for the right node, and y_i is the actual observed value.

3.8 Metrics and evaluation

A major liability that results from machine learning models is the phenomenon called overfitting. When a model is trained, it learns both meaningful signals and random noise such that the model will fit its parameters to both the noise and the signal. It will appear as an excellent fit for the input data, but will perform poorly when applied to previously unobserved data. There are different methods that can be employed to evaluate the performance of machine learning models depending on the phenomenon under study. A technique called cross-validation was used to protect against overfitting. To measure performance, the value of the models for this study, normally an estimate of the wealth index which is continuous and ranges between - one and +one was used. A value of ± 1 indicates a perfect degree of association and a correlation coefficient closer to zero represents a weaker relationship. To assess the performance of the models, one has to look at the correlation, which measures how well the model fits the data, which is the association between the independent and response variables, and how well the model estimates the response variable (direction). Models were evaluated using ridge and random forest regression machine learning techniques. The scoring metrics for the models is done using R-squared (R^2), correlation between real and predicted values, and Root Mean Squared Error (RMSE) to compare the performance of poverty prediction. Descriptive statistics were used to explain the characteristics of the data for easier understanding. The methods and techniques are described in detail below.

3.8.1 R-squared value

The accuracy of a machine learning model is measured using R-squared, represented as R^2 , or also known as coefficient of determination. R^2 is a statistical measure of fit metric that is based on a linear relationship representing the proportion of the variance between the response variable (Y) and an independent variable (X) in a regression model. It takes values between zero and one ($0 \leq R^2 \leq 1$) and is independent of the scale of the response variable. It is computed by:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} = 1 - \frac{RSS}{TSS} \quad (3 - 3)$$

where $\sum_{i=1}^N (y_i - \bar{y}_i)^2$ is also known as the “*Total Sum of Squares*” (TSS) which measures the total variance (summation) inherent to the actual response variable output and average value, whilst $\sum_{i=1}^N (y_i - \hat{y}_i)^2$ is the “*Residual Sum of Squares*” (RSS) measuring the total variation of the discrepancies between the actual values of Y and those predicted by the regression equation (James et al. 2013, pp. 69–71). An R^2 value which is close to one indicates a high proportion of response variability between the response and independent variables that has been explained and captured in the regression model. Whereas an R^2 close to zero indicates that the regression model is unable to sufficiently capture and explain the variability of the response variable. In this research, R^2 is used to measure the proportion (%) of variation of the wealth index survey data and the predictions generated by OSM, night-time luminosity, and satellite imagery during cross-validation.

3.8.2 Root mean squared error

Root Mean Squared Error (RMSE) is useful to measure and compare the prediction performance of a regression model using the error rate, where errors are measured at similar rates. It does this by measuring the difference between predicted values and the real values and, therefore, is the sample standard deviation of that difference. To compute it, you need to determine the residuals first denoted as $\hat{y}_i - y_i$ and it is based on the cross-validated predictions which are computed by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N}} \quad (3 - 4)$$

where i is the variable i , N is the number of missing data points, \hat{y}_i is the predicted value, and y_i is the actual observed value for the i^{th} observation. It is not always straight forward if one has obtained a good RMSE or R^2 value, which is why it is important to evaluate the performance of the model using multiple metrics. In this

case, R^2 provides additional insight especially since R^2 has an interpretability advantage over RMSE.

3.8.3 Cross-validation

Cross-validation is used to protect the models from overfitting. The k-fold cross-validation, sometimes known as rotation estimation, is similar to random sampling but uses subsets for testing and training instead. The sparsity of poverty indicators usually causes differences between the distribution of training and testing data in subsets, leading to results that can be quite different between splits. To avoid these challenges, the performance of the baseline model was improved using a k-fold validation with the data used by the training model split into five parts, and each part is a fold. The model is trained with all k-1 folds and then tested with the remaining fold. This will be repeated k number of times, as shown in Figure 3.8. To overcome overfitting problems, we set the shuffle-true parameters in addition to the k-fold parameter. Cross-validation allows researchers to tune hyperparameters using their original training subset, enabling them to keep the test as a truly unseen dataset for selecting the final model. This allows the diagnostic values from the testing set to provide a more accurate reflection of how the model would perform in a realistic out-of-sample scenario.

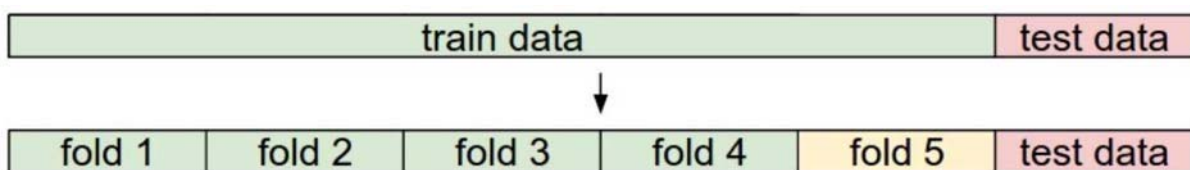


Figure 3.8: 5-fold cross-validation data split

The performance of the baseline wealth index prediction model based on the DHS survey wealth index was implemented using five-fold cross-validation where, across five iterations, the training data are split into five equal folds, and the model is trained on four folds of the data, then evaluated with the remaining fold. The five-fold nested cross-validation scheme was used to validate all the models in the training set to select the 'optimal' model. After the data was merged, Spearman and Pearson correlation was investigated to check the importance of the data features on the

predicted values. Pearson benchmarks linear relationships, while Spearman benchmarks monotonic relationships. Regression plots were used to visualise the relationship between independent and predicted indicators. A t-distributed stochastic neighbour embedding plot (t-SNE) was used to visualise the mean embeddings obtained from the deep learning model.

3.8.4 Data analysis and assessment

One of the challenges involved in using Google Earth and satellite images is that the underlying records are the result of multiple satellite patches and stitched images. For this reason, the outputs may be inconsistent or incongruent, with cross-locational images resulting in inaccuracies or a lack of visual conformity. Therefore, the data analysis for this study required the selection of image sets that were consistent and representative of a singular, regionally conforming output. Statistical correlations and an Analysis of Variance (ANOVA) comparison of means tests were used to confirm the statistically significant relationships between variables within the dataset (Singh 2007, p. 238). However, for visual comparison, poverty predictions were based on regional forecasts (e.g., percentage of impoverished households) and subdivided regional GDP datasets that could predict the likely bounds of poverty within a given geographic area (e.g., percent of the population in Cape Town earning less than \$1.50 USD per day). As the primary research objective was to build a proof of concept model for the solution, the primary analytical output was a correlated comparison between the poverty predictions of the neural network analysis (e.g., percentage of impoverished members of the population for the sample region) and the quantitative data captured via regional household surveys.

One of the challenges in developing an accurate visualisation instrument was that many of the housing characteristics within the targeted regions are similar or structurally indiscriminate. For example, making a distinction between residential structures and multi-family housing is often impossible given the top-down view offered by the satellite model. This issue also arises when considering night-time light data and in order to estimate the degree of error in night-time assessment models, Chen and Nordhaus (2011, p. 8591) averaged the luminosity scores recorded for multiple

satellite image records, resolution levels, and scalar records. By acknowledging particular factors that might interfere with the accuracy of the dataset, such as image resolution and artificial influences (e.g., fires, factories), predictions can be made about errors before the analysis is completed (Chen & Nordhaus 2011, p. 8592). For this study, regional generalisations were made based upon the variations between ground truth data and the calculations made by the model.

3.8.5 Descriptive statistics

When data are gathered, it is initially read and understood by applying statistical techniques before applying algorithms and making predictions. Descriptive statistics enable the summarising and organising of large quantities of data using limited numerical values which highlight the important data features so that they can easily be understood, that is, to describe variables numerically (Saunders et al. 2015, p. 444). The DHS dataset will be summarised using frequency tables and statistics including measuring central tendency (e.g., mean, median, and mode) and measures of spread/dispersion (e.g., standard deviation, mean deviation, percentiles, variance, quartiles). During the exploratory analysis of night-time luminosity data, characteristics of the distribution consisting of summary statistics and histogram-based features will be produced for each DHS cluster. These include the following additional statistics: Mean, Maximum, Minimum, Median, Coefficient of Variation, Standard Deviation, Skewness, and Kurtosis. Skewness is the measure of symmetry or lack thereof, whilst kurtosis measures the pointiness of the data. The correlation, which considers whether there is a relationship between variables in the data, apart from characteristics, is already discussed in Section 3.9.1.

3.8.6 Poverty indicators and calculations

The validity of visible wealth as a proxy for poverty has been empirically confirmed through a survey assessment by Bertram-Hümmer and Baliki (2015, p. 781) which suggests that individuals will perceive their own personal status and relative deprivation differently according to regional characteristics. Therefore, differentiation

between owned assets, proximity to services, building footprints, road network, and remotely sensed data on night-time light data were identified to serve as proxies for poverty characteristics that could be visually correlated with household surveys. In order for this experimental model to be sufficiently valuable, it needed to be extendable and generalisable across multiple geographic regions. For example, Jean et al. (2016, p. 790) employed household expenditures or average household wealth to calculate poverty characteristics at the cluster level (villages in rural areas or wards in urban areas). Extending this approach, Bency et al. (2017, p. 8) highlighted how building type, rent value, roads occupying more space, and neighbourhood features can be used to distinguish key patterns such as gentrification and poverty. Therefore, this study used multiple predictors of regional poverty not only to assess the precision of the trained solution, but also to confirm the transferability of this model to other regions in South Africa and other African nations based on contextual or categorical indicators. Therefore, the central dimensions used to assess poverty levels included the following:

- Building type
- Distance to nearest essential services
- Education level attained
- Access to drinking water
- Road network
- Amenities or services (frequency of parks, amenities such as community centres, government buildings)

Once the visual classification models had been compiled, statistical comparisons between the visual findings and the quantitative data sets needed to be performed. Quantitative comparative measures developed by Pinkovskiy and Sala-i-Martin (2014, p. 9) modelled the GDP data from the World Bank over time to calculate an indicator that compared luminosity with per capita income in order to confirm the degree of correlation between visible and economic indicators over time. This approach was inappropriate for the degree of granularity used in the current assessment but predicted that regional statistical data could be used to compare visual cues and outputs in order to calculate poverty rates and characteristics.

3.8.7 Reliability and validity

Once a machine learning artefact is developed, trained, and tested, reliability and validity are two primary factors that require attention. Reliability refers to the degree of trust in the variables produced by the model. This is done to ensure that they are consistent, replicable, and free from measurement errors. On the other hand, validity refers to the accuracy of the model results that are produced on the test data.

In order to test the validity of the wealth estimation for the purposes of this study, the findings were compared with the average wealth index of the ground truth data, which is based on the South African DHS survey for 2016. Ground truth data is obtained by direct observation or survey-based techniques rather than by modelling or inference. Validation challenges exist for big data (and remote sensing research in general) as the data are not specifically collected for socio-economic analysis. Reliability concerns the more mechanical aspects of a procedure by measuring the degree to which there is consistency rather than error in the measurements.

3.9 Ethical considerations

This study focuses on the assessment of a controversial social indicator: poverty. Although the evidence captured and analysed during the course of this research is based on publicly accessible secondary datasets, the vulnerability of the populations that this information targets remains a significant concern. Central to the responsible execution of any exploratory, data-oriented research, Bryman (2012, p. 134) proposes that in any social research, ethical integrity is of paramount importance and through targeted controls, efforts must be made to mitigate harm and realise objective, broadly valuable outcomes. Similarly, Punch (2013, p. 49) suggests that by triangulating empirical evidence against underlying theoretical and conceptual foundations, interpretation and analysis can consider the broader applications and consequences of the empirical findings. The GIS imaging employed is generally anonymised and therefore protects participants from exposure. For this study, the controls employed in capturing and analysing the evidence for this data integration experiment were clearly defined before the visual data were downloaded. Geographic location data was scrubbed from visual representations of evidence to protect the

anonymity of residents and mitigate any potential association between poverty judgments and particular groups or individuals. Although regional specificity still played a critical role in this study, the generalisability applied to the machine learning techniques used will likely prove beneficial in future studies where these stages of research can be replicated and applied to different datasets (Carrig & Hoyle 2011, p. 129). All findings were analysed by applying objective, quantitative comparisons, and visual grouping measures, thus reducing the likelihood of subjectivity and maintaining the objective output of the regional insights (Bryman 2012, p. 134). The researcher obtained an ethics clearance from the University of South Africa (UNISA) before data collection and followed the UNISA research ethics policy throughout the study. The research ethical clearance certificate is attached in the Appendices section.

3.10 Chapter summary

This chapter presented detailed methodological solutions that were used to craft and administer remote sensing to a discrete problem of poverty in South Africa, a nation of extreme socio-economic inequality. The chapter showed that by adopting a range of previously validated remote sensing techniques and applying regionally specific datasets to the systematic analysis of both daytime and night-time images, the methods developed for this study have illuminated a range of socio-economic patterns that are observable even in high-density urban areas. It was shown that for South Africa, this regionalisation and structural isolation results in several critical implications, including mapping and analysing the effectiveness of government programs, assessing network gaps and poverty enclaves, and predicting factor-specific links between energy density and regional poverty.

The following chapter will present these findings in their entirety, drawing upon specific South African sources of evidence to draw conclusions about the current state of regional poverty in this developing but high-inequality nation.

Chapter 4. Implementation and results

4.1 Introduction

This chapter discusses the experiments which were carried out for this study and aims to delve into the details of the data collection, pre-processing, and analysis thereof. The chapter begins with an introduction of the processing that was done to prepare the data for the practical implementation of both the feature-based and computer vision models. Several tools and scripts were used to prepare the data and generate .csv files. The chapter subsequently discusses the statistical models, the results of the implemented models, and a thorough evaluation of the performance of the models. In addition, this chapter discusses the setup of the environment, data preparation, and the corresponding evaluation metrics.

4.2 Data collection

Data collection is the method of acquiring data from relevant sources that enables the researcher to answer the research questions and evaluate the results. Dataset preparation is a time-consuming effort and requires a significant volume of high-quality images that are gathered from areas within the clusters targeted for the study. The data for this research are notoriously unstructured and noisy with varying degrees of difficulty in handling. This was a major challenge that took a lot of time and required patience. Google limits the maximum number of images that can be downloaded per day to 25,000 and will return blank messages instead of an error if you attempt to download more than the allowed daily limit. This led to delays in model development and training.

The main requirement for selecting a data source was to enable the models to use readily accessible data sources to allow cost-efficient approaches to poverty estimation using machine learning. It is costly to purchase high-resolution images, and additionally, implementing deep learning models requires powerful computers with high performing GPUs, which can act as adoption barriers for the tools to the majority

of developmental organisations with limited resources (Ayush et al. 2020, p. 1; Jean et al. 2016, p. 794). Table 4.1 provides information on the data used to train each model for the experiments. The data sets described in the previous chapter will be used. Models were built using different types of training data set, depending on the objective and methodology of the research.

Table 4.1: Training data setup

Model	Approach	Training Dataset
1	Satellite image-based	Daytime satellite, night-time lights, and DHS
2	OSM-based	OSM and DHS
3	Night-time and OSM	OSM data, night-time lights, and DHS

4.3 Pre-processing

Some pre-processing techniques must be applied to make the obtained data useful for analysis in this study before the actual image analysis and classification can be conducted. The pre-processing includes atmospheric correction of the image, bit-depth conversion, georeferencing, and co-registration of the images.

4.3.1 Demographic and health survey

The pre-processing tasks involved processing a shapefile that was downloaded from the DHS website to get a layer in QGIS with buffers of different sizes for urban or rural DHS clusters. These were different sizes as shown in Figure 4.1. The process was carried out using a buffer algorithm along with expressions in the specific column. The researcher also changed the initial shapefile from EPSG Geodetic Parameter Dataset which was initially created by European Petroleum Survey Group (EPSG) to that for South Africa in order to improve accuracy. A Python script was developed to process the data. Two data files, Zahr71FL.DTA and Zahr71FL.DO, the file formulated with the household strata and the keys to the questionnaires, respectively, were downloaded and unzipped from the data file. The files were merged to generate a .csv file with 11,083 rows and 794 columns as shown in the code in Figure 4.1.

```
▶ dhs = pd.read_stata(dhs_file, convert_categoricals=False)
dhs_dict = get_dhs_dict(dhs_dict_file)
dhs = dhs.rename(columns=dhs_dict).dropna(axis=1)
print('Data Dimensions: {}'.format(dhs.shape))
```

```
Data Dimensions: (11083, 794)
```

Figure 4.1: Python code for DHS dataset objects count

The cluster column specified the exact DHS cluster where the survey was conducted. In order to obtain a meaningful prediction target, the mean or median values were calculated for the required columns. The mean was calculated for the columns of the cluster “*Has electricity*”, “*Education completed in years*”, and “*Wealth index factor score combined*”. The median value was calculated for the factor “*Time to get to water source (minutes)*”. A .csv file was generated with 746 columns and 4 rows for each column grouped by cluster number. Figure 4.2 shows the spatial distribution of the clusters obtained by simulation of a DHS shapefile in QGIS where green represents urban clusters and red represents rural clusters.

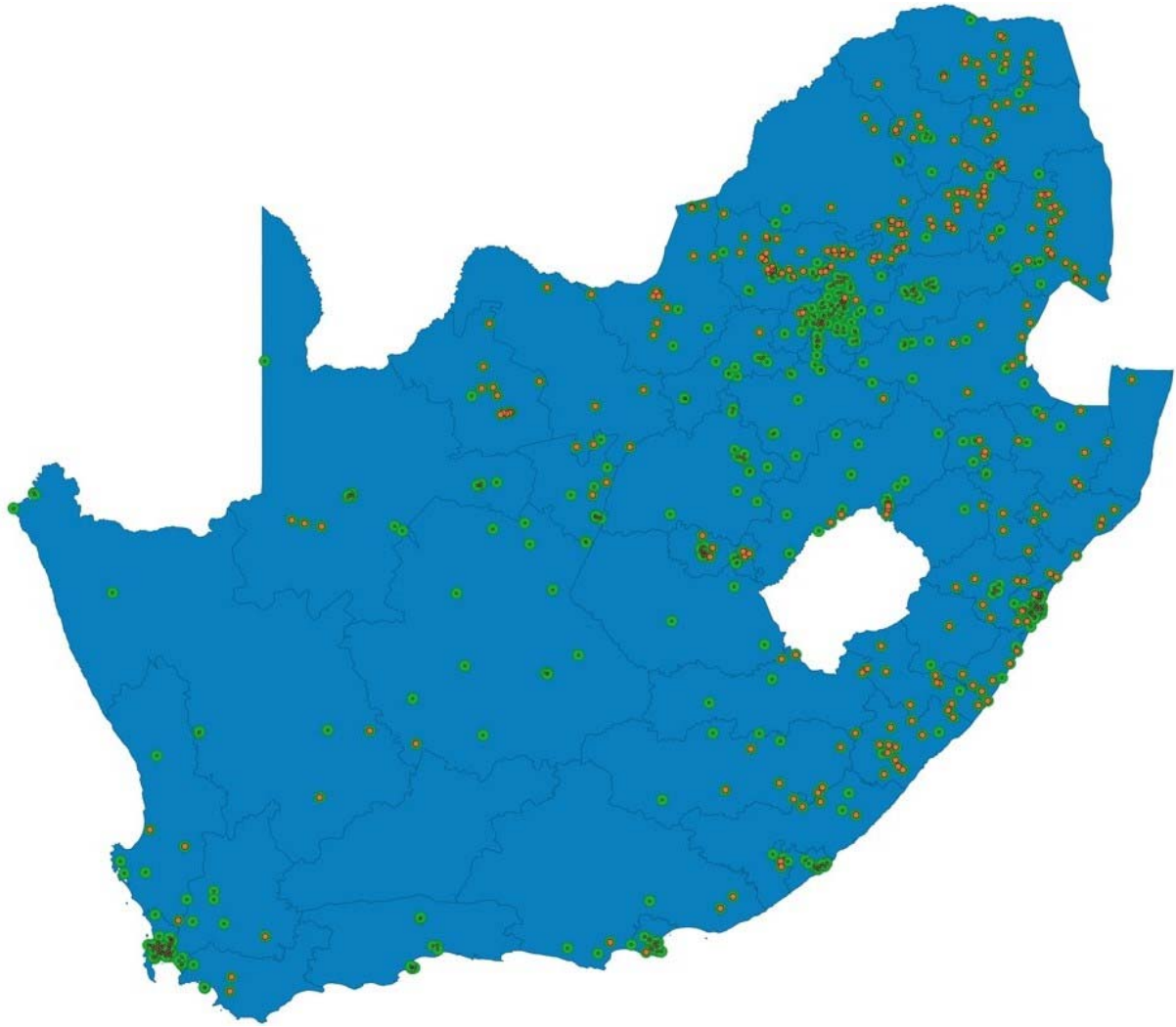


Figure 4.2: South African DHS clusters

Table 4.2 shows the summary data describing the statistics of the DHS clusters that were downloaded and pre-processed using a Python script.

Table 4.2: Descriptive statistics of the DHS clusters

	Wealth Index	Education Completed (years)	Access to Electricity	Access to Water (minutes)
Count	746	746	746	746
Mean	10,847.23	9.99	0.90	4.01
Std Dev	91,384.04	3.89	0.24	10.62
Min	-246,990.20	0.77	0.00	0.00
25%	-53,220.37	7.27	0.94	0.00
50%	12,694.50	9.51	1.00	0.00
75%	83,642.22	12.43	1.00	0.00

Max	167,679.00	37.33	1.00	105.00
-----	------------	-------	------	--------

4.3.2 Night-time lights

The night-time lights images were downloaded in GeoTIFF format. The TIFF files were used to generate .csv files with night-time luminosity data for the coordinates around the DHS clusters. In order to generate the file, first of all the point sampling tool plugin for QGIS was installed and used to open the VIIRS data. By default, the VIIRS data cover an exceptionally large area: the whole world. Since this study requires data for South Africa, the image was clipped so that it only covered the country of interest using a GADM (2020) geopackage of the country's administrative boundary. The command "*Generate points (pixel centroids) inside polygons*" was used to obtain the pixel centroids for an image (raster) inside a polygon (vector) to produce a vector whose longitude and latitude points could be calculated using the "*Field*" calculator. The image was used as the raster layer, and the administrative boundaries of the South African provinces were used as the vector layer. The Point Sampling Tool and vectors were used with a set of points as a "*Layer containing sampling points*" in order to select the latitude and longitude of the vector and Band 1 of the raster as "*Layers with fields/bands to get values from*". This was repeated by going to the "Fields" tab and renaming Band 1 as a night-time light then going back to the "General" tab to save the "*Output point vector layer*" as a .csv file. The final step was to use a buffer algorithm to merge data with DHS clusters to produce a .csv file with latitude and longitude coordinates along with a night-time light intensity indicator for each DHS cluster and values for luminosity. A total of 6,513,751 images were downloaded. Figure 4.3 shows the night-time lights data that were extracted for South Africa and Gauteng night-time lights merged with the DHS clusters. High luminosity can be identified in major cities in various provinces such as Bloemfontein, Cape Town, Durban, Johannesburg, Nelspruit, Polokwane, Port Elizabeth, Pretoria, and Rustenburg.



Figure 4.3: South African night-time luminosity data (left) and night-time data merged with DHS coordinates for Gauteng province (right)

4.3.3 Daytime satellite imagery

With the Google Cloud API and secret key, a total of 113,215 images were downloaded with a scale of 17 and pixel resolution of 1.25m². Figure 4.4 shows a sample image from the downloaded dataset. The daytime satellite imagery data was downloaded for locations in the 2016 DHS coordinates. The images are 400 × 400 pixels.



Figure 4.4: Sample Google satellite image for South Africa downloaded from (Google 2020)

4.3.4 OpenStreetMap

After the initial OSM data processing, the relevant data points are exported and converted to csv files. The OSM database consists of a collection of vector data objects that include point features (nodes), lines, and polygon features. QGIS is used to match each database table and convert the attributes of interest into PostgreSQL files in order to allow greater and faster manipulation of spatial data. The noise introduced while jittering DHS location data to protect respondent confidentiality will likely introduce bias in the spatial analysis, and therefore the extraction of the OSM features was done after setting a radius buffer zone of up to 2 km and 5 km for urban and rural areas, respectively, with each area centred on the cluster locations. The OSM features that were extracted included roads, buildings, POIs, railways,

waterways, transport, and land use. As a method to convert the OSM features into the required columns in the required files, an algorithm named “Distance to the nearest (hub) points” in QGIS was used to find the distance of each cluster to buildings, POIs, railways, and roads. The remaining columns were obtained using the “Buffer” algorithm in QGIS with the clusters as input and then using the “Join attributes by location (summary)” algorithm with the buffered clusters as input. Pre-processing for road feature extraction was performed using Zhao and Kusumaputri's (2016) technique where they got the type of road, the total number of roads, or lengths of roads in the cluster. For buildings, the total count, the average area, and the proportion of the area covered in a cluster were obtained. Hundreds of areas with POIs were identified, including the total within each cluster and the proximity to the POI within the cluster. Figure 4.5 shows the road network, POIs and buildings extracted from OSM and visualised through QGIS.

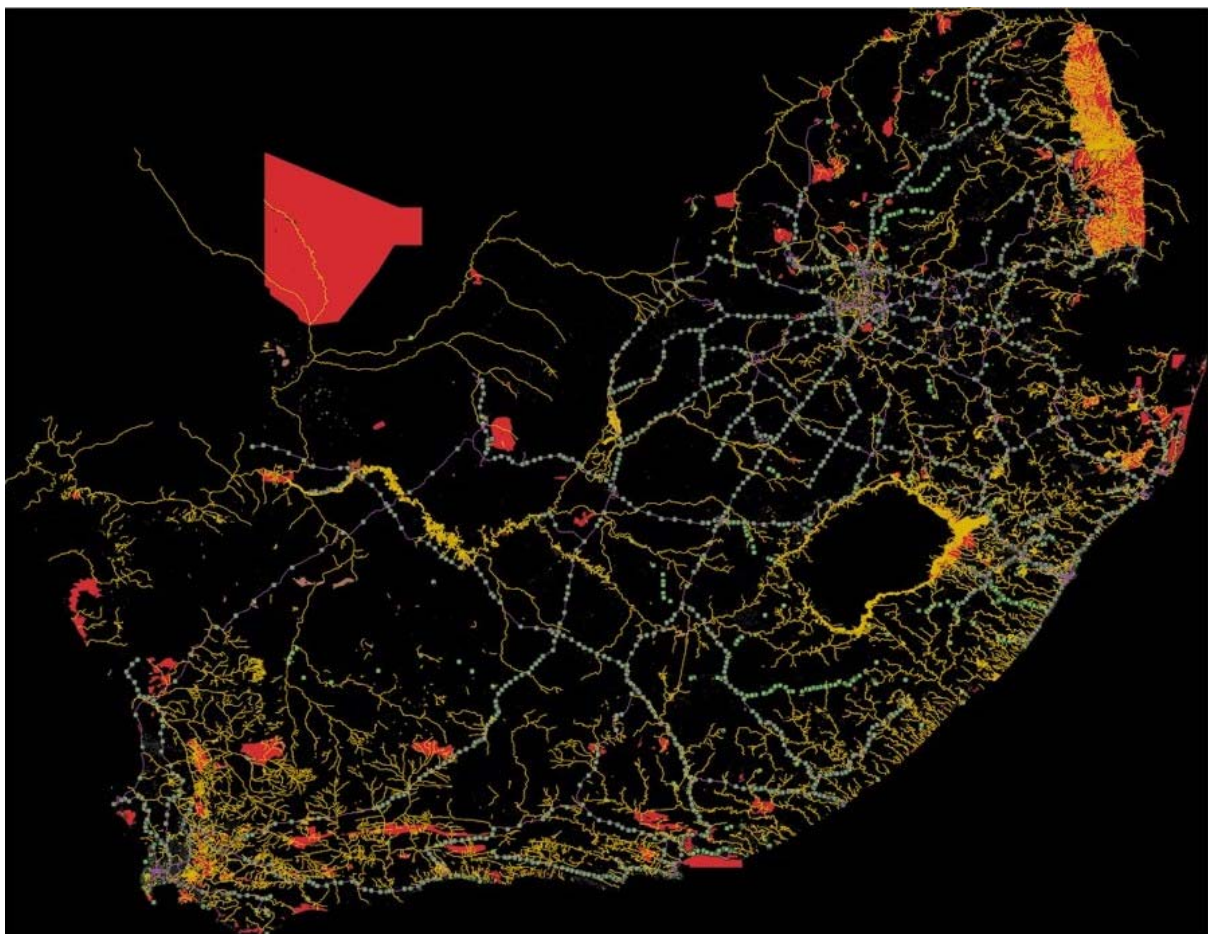


Figure 4.5: Visualised OSM data

After obtaining attributes of the layers produced by algorithms with the help of pandas (a Python-based software library for data manipulation and analysis), this research pivoted the data, conducting the necessary renaming to produce wider categorical features. After exporting roads, buildings, and POI data, the researcher created categories based on OSM guidelines (OpenStreetMap 2020). Data were analysed according to the categories presented in Table 4.3.

Table 4.3: OSM feature categories

Buildings	Residential, Industrial, Commercial, Education, Health, Administration, Damaged, Undefined
Roads	Primary, Paved, Residential, Unpaved, Trunk, Unknown
Points of Interest	Food, Education, Emergency, Finance, Healthcare, Entertainment, Tourism, Historic, Natural, Shop, Leisure, Others

The basis of the analysis for this model is data containing the road network, buildings, POIs and night-time light data for South Africa extracted from OSM and VIIR. A general overview of South African data merged with DHS clusters generated using QGIS is shown in Figure 4.6.

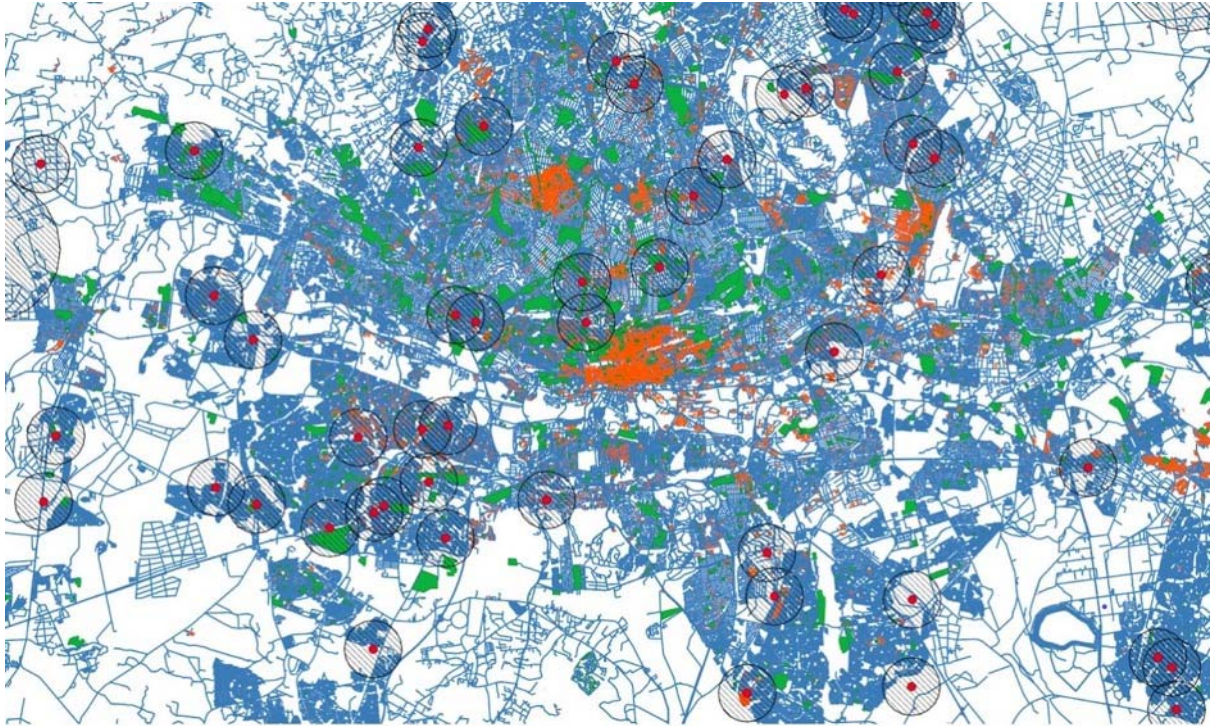


Figure 4.6: QGIS OSM datapoints showing roads (blue), POI (green), buildings (orange), and DHS clusters (red dots)

OSM data were pre-processed using QGIS, which is based on a graphical user interface (GUI) and involves multiple repetitive tasks which are tedious. The task was therefore time-consuming and took three weeks to complete. Furthermore, QGIS has limited support for machine learning models. A more efficient method of processing OSM data would have been to use Geopandas. However, this option was discovered after most of the data had already been processed.

4.4 Experimentation

A literature review on the machine learning techniques used to predict socio-economic well-being was conducted. The review included the use of various cost-effective geospatial datasets to estimate regional socio-economic well-being in DHS household survey clusters. The study used a positivist research paradigm and design science approach to deduce how well the models perform compared to each other based on single or merged data sets. The following section discusses the development steps for the models and the experiments conducted to evaluate them. The aim of the

experiment is to show that cost-efficient and readily accessible data can be used to provide more granular data for poverty estimation than other approaches.

4.4.1 Experiment setup

Machine learning models were implemented to simulate and run experiments. Python 3.8 programming language was used to implement the deep learning network for transfer learning and classification tasks. The transfer learning and satellite imagery pipeline models were implemented using a virtual machine instance, as computer vision and neural network architecture require a lot of resources due to fast computation capacity, efficient matrix multiplication, convolution, and parallelism requirements where normal CPUs would require days to complete. An AWS instance was used because it was available freely unlike GCP which was not freely available and would require a subscription to use GPU. The AWS instance had the specifications outlined in Table 4.4.

Table 4.4: AWS instance specifications

Instance	g3.4xlarge
GPUs	1
vCPU	16
Memory (GiB)	122
GPU Memory (GiB)	8
GPU	NVIDIA Tesla M60 GPU

The scripts were developed and deployed in GCP using Google Colab. Python is an open-source high-level programming language. Among the advantages of Python is that it comes with an extensive list of standard libraries that are suitable for many fields, including all the tasks performed in this study. During implementation, the libraries were downloaded and used. In addition to being free and open source, Python was selected because it is portable (i.e., it can run on either Windows or Macintosh safely) and provides rich modules and functions. The software, tools, and libraries that were used for this study are listed in Table 4.5, including the versions of the software and how they were used.

Table 4.5: Software and library versions

Software / Library	Version	Use / Description
QGIS	3.10	Open-source application that supports the viewing, editing, and analysis of geospatial data.
Python	3.8	An interpreted, high-level, and general-purpose programming language.
matplotlib	3.0.2	Embed plots into applications as an extension of numpy.
pandas	1.0.3	Data manipulation and analysis.
seaborn	0.9.0	Works with matplotlib to provide a high-level interface for drawing attractive and informative statistical graphics.
numpy	1.16.0	Support for large, multi-dimensional arrays and matrices from feature extraction, along with a large collection of high-level mathematical functions to operate on these arrays.
pandas	0.24.0	Data manipulation and analysis.
torchsummary	1.5.1	View visualisation of the model and useful for debugging the network.
torchvision	0.2.1	Image loading, transformation and common architectures for computer vision, pre-trained weights.
tqdm	4.30.0	Progress bar library.

When conducting predictive and exploratory studies that analyse geospatial and satellite imagery data, free cloud services that also offer free Graphical Processing Units (GPUs) can be leveraged. This study used Google Drive, Google Earth Engine, Google Colab, and Amazon Web Services (AWS). Google Colaboratory (Colab), a Google-Cloud-hosted Jupyter notebook service, was used to develop the models, as it offers a free Python development environment that can be used to write and execute code through a web browser, and is also highly integrated with Google Drive (Google 2020). These tools allow researchers to combine executable Python scripts, images, and rich text in a single document, and, generally speaking, they are easy to set up, access, and share. With Colab, one can harness the full power of popular Python libraries. Due to the size of the images and the computing resources required to process satellite imagery, a virtual machine was set up to run on AWS Cloud Platform: this allowed for accelerated classification and testing of solutions. AWS is available from the researcher's office and, therefore, was a cheaper option compared to running Google's Cloud Platform (2020, p. 1), which requires a subscription to access GPU

tools. Once compiled and verified, the training sets were uploaded to Google Cloud to facilitate the experimental analysis of poverty indicators within multiple mixed-use (urban and suburban) landscapes in South Africa. The models for this study were implemented and deployed on Google Colab.

4.4.2 Satellite imagery approach

This approach will help identify poverty and locate regional-specific issues that will provide policy makers with accurate decision-making information for poverty alleviation. This will be done by creating a model that uses computer vision to extract features from daytime satellite images using night-time satellite images as a proxy task. In order to use daytime satellite images to predict cluster wealth, features must be extracted first. The images are a combination of red, green, and blue (RGB) colours with each pixel encoded as an integer between 0 (no light) and 255 (brightest). Figure 4.7 highlights the steps required to carry out the experiment and the simulations for this approach.

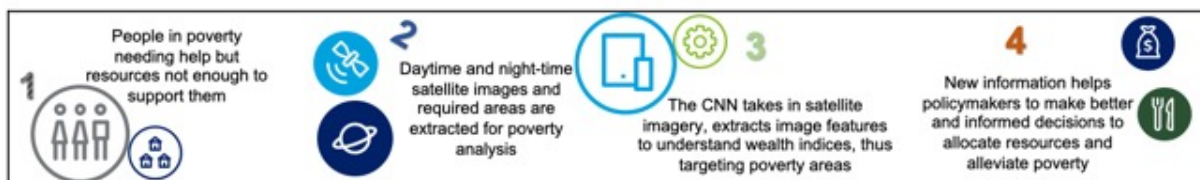


Figure 4.7: Satellite imagery approach steps

The second and third processes are related to this experiment stage and therefore discussed in detail in the following sections.

4.4.2.1 Transfer learning

In order to speed up the training and reduce the amount of data needed to train and extract features from satellite images, models that already exist and have learnt generalisable weights pre-trained from existing datasets can be used (Yamashita et al. 2018, p. 620), and then fine-tune the fully connected network layer to suit the case under study. For the research under study, a VGG16 CNN architecture for fixed

feature extraction was selected as the baseline and pre-trained on the ImageNet dataset to assign labels to a large set of images (Deng et al. 2009, p. 254). The advantages of the VGG over other networks are that it has a simple architecture. The CNN model was trained to recognize 1000 different class labels reduced to five night-time luminosity intensity classes using a well-known technique referred to earlier in the literature as transfer learning (Pan & Yang 2010, p. 1345).

Due to the structural variation between the VGG16 model (224 × 224 pixel images) and the input subset (400 × 400 pixel images), the convolutional architecture proposed by Xie et al. (2016, p. 3931) and improved upon by Jean et al. (2016) was adopted. It involves replacing some set of fully connected layers of the network with one that is randomly initialized. In this way, the model will not lose any information even if it is fed images of any size, compared to cropping or scaling. The input images were transformed by subtracting the mean across each individual pixel before being fed into the CNN. Network fine-tuning was introduced by applying an Adam adaptive learning rate optimiser. The model was trained to extract feature vectors as output. In the current study, transfer learning resolved the challenges associated with capturing, testing, and revising classification datasets relative to poverty assessment based on satellite images from South Africa.

As presented in Table 4.5, there were 113,215 images of size 400 × 400 with RGB spectral bands, and to prevent overfitting, the images were divided into 0.9 and 0.1 sets which were used for training and validation, respectively. Night-time class imbalances were handled by implementing up-sampling and down-sampling for the three classes—low, medium, and high—in order to balance the nightlights data set to avoid bias in the results. The night-time light characteristics of the distribution consist of summary descriptive statistics discussed in Section 3.9.5 within each DHS cluster. Summary numbers of VIIR images for the night-time model, including the distribution between training and validation sets, are presented in Table 4.6.

Table 4.6: Satellite imagery and distribution between training and validation sets

	Total	Training set	Validation set
Low	50,055	45,117	4,938
Low medium	30,257	27,148	3,109

Medium	21,598	19,444	2,154
High medium	6,844	6,122	722
High	4,461	4,062	399
Total	113,215	101893	11322

Visualising what features the CNN regards as the most important when encoding information into the feature vectors is important in order to confirm whether the training of the night-time intensity task used during transfer learning guides the network to learn the important features. Figure 4.8 shows the sample images that were saved during the training to validate the model classification of the images based on the five luminosity intensity classes.



Figure 4.8: Examples of daytime satellite images corresponding to light intensities

This is achieved by visualising the activation maps of the fifth convolutional layer of the black box, which result from the convolutions (Zeiler & Fergus 2014, p. 819).

4.4.2.2 Estimating night-time light intensities

This study aimed not to use daytime satellite imagery directly for feature extraction using CNN, but instead to use them in lieu to predict night-time lights and then to use the predicted values for wealth estimation. Night-time indicators were used as a proxy task for temporal socio-economic activities for which traditional data: would take a long time to compile, are of poor quality, or are not available for inaccessible areas (Henderson et al. 2012, p. 1025; Mellander, Lobo, Stolarick, & Matheson 2015, p. 15) were subsequently classified relative to intensity levels on a 5-point Likert scale: 1 = low, 2 = moderately low, 3 = medium, 4 = moderately high, and 5 = high brightness

bins—by fitting a one-dimensional Gaussian mixture model to the relative frequencies of the night-time light intensity values from the sample in order to simplify the tasks, which are the five luminosity intensity classes discussed in Section 3.6.2. Since we used only five classes of bins for the night-time luminosity, the number of outputs in the last layer of the pre-trained model was changed from 1,000 to 5. Figure 4.9 shows how the model splits the images between the training and validation set in addition to the allocation to specific night-time luminosity classes.

```
Size of training set: 101893
low                45117
low medium        27148
medium            19444
high medium        6122
high               4062
Name: label, dtype: int64

Size of validation set: 11322
low                4938
low medium         3109
medium            2154
high medium        722
high               399
Name: label, dtype: int64
```

Figure 4.9: Data split between training and validation

Feature embeddings that are useful for wealth prediction were used as input into ridge regression and random forest models. Similarly, the asset wealth index and GPS location labels from the DHS for South Africa for 2016 were used. Wealth index scores were combined with images centred on household locations. However, the coverage in the surveys is sparse compared to the large amount of satellite image data. From the learning curve in Figure 4.10, it is apparent that the loss in the training and validation sets starts decreasing in the first few iterations and becomes more stable after 200,000 iterations. It was decided to take a snapshot of the model at the 200,000th iteration, since there was no significant change in the loss after that iteration. This occurred after having performed the optimization until one million iterations.



Figure 4.10: Learning plot

Following the approaches of Simonyan, Vedaldi, and Zisserman (2014, p. 4) and Jean et al. (2016, p. 792), class saliency maps were generated based on an image class to visualise the night-time light classification model of the fifth convolutional layer. Each image has its own saliency. The saliency is calculated from the amount of the class score derivative, indicating the influence of each pixel on the class score. The saliency maps are generated using backpropagation passes through the trained CNN, and examples are shown in Figure 4.11.

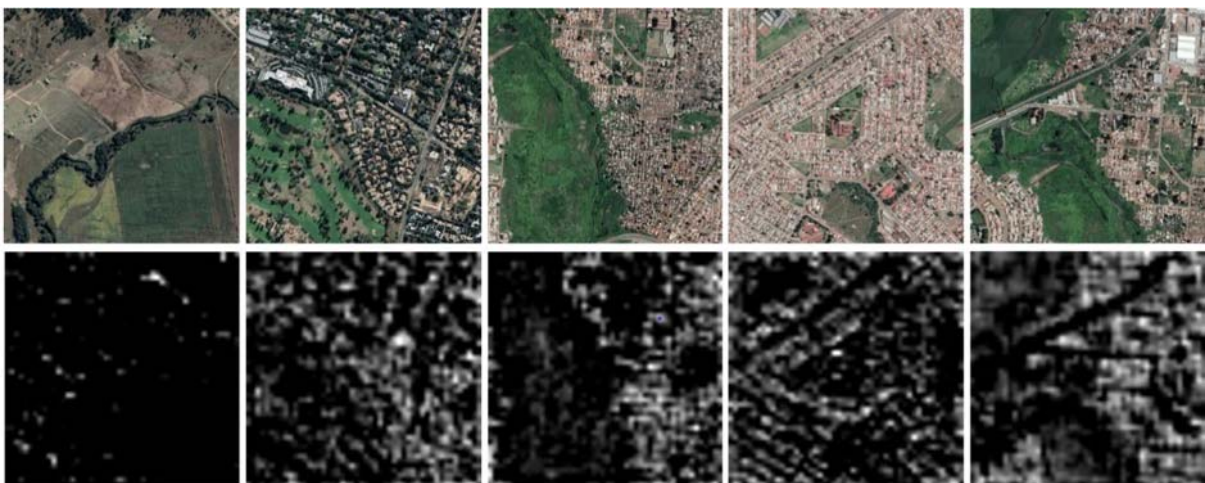


Figure 4.11: Night-time classification saliency maps generated from satellite imagery

From Figure 4.9, it can be observed that the model correctly identified pixels related to the presence of buildings and roads as important features. In contrast, pixels related to vegetation were linked to areas with lower luminosity classes. In the left column, an isolated house was identified regardless of its size. The identification and classification of buildings as an important feature made sense as they tend to produce more light, both from inside and outside, and hence they would be classified as a distinguishing feature when determining the night-time light intensity class.

4.4.2.3 Final feature vectors

The final output of the feature was the 4,096-character feature vector for each cluster which was used for further analysis. From the set of fully connected layers that have 4,096 neurons, a 4,096 one-dimensional vector of activations was extracted. These were optimised into a softmax to distinguish the night-time luminosity levels for each image. This was repeated for all images within a cluster. Finally, the mean for each cluster was calculated from the one-dimensional arrays of the total cluster. In order to predict the socio-economic well-being, the regression models were fed with the cluster-level feature vectors. As in previous studies by Head et al. (2017) and Jean et al. (2016), regression models were used to learn the wealth distribution rate of a cluster from the image feature extracted by the CNN model and stored in the last convolutional layer.

4.4.3 Crowd-sourced information approach

In this approach, an OSM model was implemented for wealth estimation using only OSM data. The model involved extracting quantifiable features from OSM geospatial data. The goal was to extract features that have the most predictive power regarding wealth estimation. OSM data contains metadata associated with roads, buildings, and POIs. QGIS routines were used to extract and calculate the required features that are included around each of the DHS surveyed clusters. Radius buffers were set to 2km in urban areas and 5km in rural areas to capture the surrounding elements. In addition, the buffers were also included to compensate for noise, as DHS

data are randomly displaced to protect the confidentiality of survey participants. Regional economic activities can be illuminated through OSM features and therefore can be used to train a model to discover paramount features to help with wealth level estimation. By conjecture, an affluent neighbourhood will have, for example, tarred roads, residential areas, and public infrastructure within proximity. On the contrary, a poorer neighbourhood will have low-cost housing structures, fewer public facilities, or narrow roads. Therefore, combining these data will present an opportunity to determine features that can help predict the poverty distribution of a survey cluster in South Africa. This approach does not capture qualitative differences in geospatial features, such as the nature of the building or the type of road of the same length. During pre-processing, buildings were classified, and the classification results were used to calculate the input parameters within each cluster. Road characteristics were processed using the same technique as described by Zhao and Kusumaputri (2016). Lastly, POIs were identified within each cluster and the researcher counted the total number within a certain proximity to the area.

A performance comparison was carried out for the regression models that were deployed on each type of OSM feature, both individually and combined with night-time lights statistical results to predict socio-economic well-being. Given the lack of relevant poverty indicators, it was determined that a specific and targeted object class would need to be developed that could be integrated into the training construct in order to focus in depth on both the OSM features and the luminosity-derived features to test predictive performance. Zhao et al. (2019) acknowledge that the overall performance of their experiment was deficient and that it failed to achieve the transparency needed to improve future predictability. Instead, features from mixed data sources could bolster the model performance. By combining data with quantifiable and qualitative features in 10km × 10km grid cells, Zhao et al. (2019) proposed that the analysis could be used to train models more effectively, complementing learnt object characteristics and extending the data set towards specific contexts, and applications to other countries (Zhao et al. 2019). For the current study, regression models were trained using data that had merged the features of night-time lights and OSM within each cluster as input.

4.5 Feature extraction

The vector of features for each cluster needed to be extracted before training the regression models. The features were extracted from satellite imagery downloaded for South Africa using the Google Maps API. Deep feature extraction from satellite imagery was achieved by implementing a CNN architecture that uses deep learning approaches to learn hierarchical representations of data through non-linear transformations. CNNs are designed for image input and encompass convolution operations to encode translational invariance of image features, which is the capability to ignore positional shifts of the target in the image: for example, the application should still recognise a feature across an image whether it slightly faces up or down (rotational) or is slightly moved in the image (locational) (Shin et al. 2016, p. 133). Naturally, a convolution operation involves sliding filters across the image where each filter is a template to match a certain image feature. Deep learning models are referred to as such because they comprise various layers of operations. CNN models in particular have a front-end component that executes low-level feature learning from image data and then other layers that learn higher-level features and carry out the classification task. They define a mapping from input tensors to feature vectors using the following formula:

$$f_{\theta} : \mathbb{R}^{w \times h \times d} \rightarrow \mathbb{R}^n \quad (4 - 1)$$

where θ represents the parameters of the CNN, $w \times h \times d$ represents the dimensions of the input, and n is the feature vector dimensionality. The feature vectors are used as input to a classifier for the well-being predictions. Therefore, CNN architectures summarise the complex image input into a feature vector for prediction. Vector features often represent complex compositions of the lower-level features extracted by the initial layers and can include textures and objects (Zeiler & Fergus 2014, p. 823).

The deep learning architecture model was first used in 2013 and is among the most prevalent for image recognition when using labelled daytime satellite imagery for training. Although using night-time lights to train a CNN model is a good idea, night-

time luminosity is a fundamentally flawed measure because you cannot easily differentiate areas that have night lights but have no proximity to essential services *and* have bad road networks, such as rural areas, for example. This is especially true for South Africa compared to other sub-Saharan African countries, as access to electricity is better. Therefore, a hypothesis to merge data from various sources to measure both proximity and luminosity would be able to mitigate the challenge. Using night-time lights for transfer learning using convolutional network parameters which were used for this study are shown in Figure 4.12.

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 64, 400, 400]	1,792
ReLU-2	[-1, 64, 400, 400]	0
Conv2d-3	[-1, 64, 400, 400]	36,928
ReLU-4	[-1, 64, 400, 400]	0
MaxPool2d-5	[-1, 64, 200, 200]	0
Conv2d-6	[-1, 128, 200, 200]	73,856
ReLU-7	[-1, 128, 200, 200]	0
Conv2d-8	[-1, 128, 200, 200]	147,584
ReLU-9	[-1, 128, 200, 200]	0
MaxPool2d-10	[-1, 128, 100, 100]	0
Conv2d-11	[-1, 256, 100, 100]	295,168
ReLU-12	[-1, 256, 100, 100]	0
Conv2d-13	[-1, 256, 100, 100]	590,080
ReLU-14	[-1, 256, 100, 100]	0
Conv2d-15	[-1, 256, 100, 100]	590,080
ReLU-16	[-1, 256, 100, 100]	0
MaxPool2d-17	[-1, 256, 50, 50]	0
Conv2d-18	[-1, 512, 50, 50]	1,180,160
ReLU-19	[-1, 512, 50, 50]	0
Conv2d-20	[-1, 512, 50, 50]	2,359,808
ReLU-21	[-1, 512, 50, 50]	0
Conv2d-22	[-1, 512, 50, 50]	2,359,808
ReLU-23	[-1, 512, 50, 50]	0
MaxPool2d-24	[-1, 512, 25, 25]	0
Conv2d-25	[-1, 512, 25, 25]	2,359,808
ReLU-26	[-1, 512, 25, 25]	0
Conv2d-27	[-1, 512, 25, 25]	2,359,808
ReLU-28	[-1, 512, 25, 25]	0
Conv2d-29	[-1, 512, 25, 25]	2,359,808
ReLU-30	[-1, 512, 25, 25]	0
MaxPool2d-31	[-1, 512, 12, 12]	0
Conv2d-32	[-1, 4096, 2, 2]	75,501,568
ReLU-33	[-1, 4096, 2, 2]	0
Dropout-34	[-1, 4096, 2, 2]	0
Conv2d-35	[-1, 4096, 2, 2]	16,781,312
ReLU-36	[-1, 4096, 2, 2]	0
Dropout-37	[-1, 4096, 2, 2]	0
AvgPool2d-38	[-1, 4096, 1, 1]	0
Conv2d-39	[-1, 5, 1, 1]	20,485
Softmax-40	[-1, 5, 1, 1]	0

Figure 4.12: Sixteen-layer VGG16 model with parameters for this study

Transforming satellite imagery into inputs that can be understood and processed by the computer model involves converting them into numerical format. Key to this effort is recognising that a normal Red, Green, and Blue (RGB) image comprises three independent channels that correspond to varying levels of RGB light intensity represented by pixel values ranging from 0 (which means that no colour appears) to 255 (which is the highest colour present in a pixel). Therefore, the pixel values are representations of satellite imagery that are translated into values that are fed into the model.

Transfer learning has become the most popular approach to speed up the training and deployment of models where pre-trained CNN architecture weights and biases are repurposed and leveraged for a different task as a baseline. This aids the learning process for tasks that are similar to the original use case, thereby reducing the computing time, data, and resources needed. This research therefore employed a VGG16 model pre-trained on ImageNet as a generic image feature representation. Proposed by Simonyan and Zisserman (2015), VGG16 has stacks of convolutional and fully connected layers and accepts a fixed input image size of 224×224 pixels, which is smaller compared to the 400×400 pixel images from Google. It consists of a max pooling layer to reduce the size of the feature maps. The 400×400 input produces an output of size $2 \times 2 \times 4096$ by setting a pooling layer, which is similar to down-sampling, that strides across individual feature maps. There are two hidden layers, each with 4,096 nodes which represent the scores of four overlapping quadrants of the image for 4,096 features. The cluster scores are then averaged to obtain a 4,096-dimensional feature vector optimised to distinguish between the five levels of night-time light luminosity. The learned representations of the feature vector are averaged into a single vector, which in turn is used as input for the classification regression model to predict socio-economic indicators. The output layer has five nodes representing the five classes of night-time luminosity with a SoftMax activation.

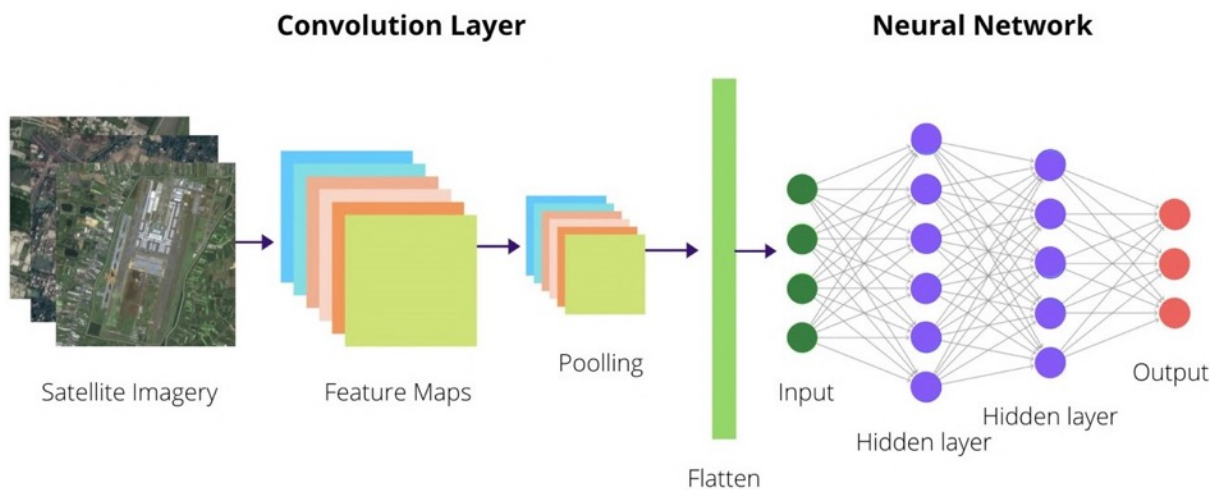


Figure 4.13: Model of the CNN architecture

Satellite images were split into two sets so that 90% of the images were used for model training and the remaining 10% for validation. Training data were augmented using random horizontal mirroring, and CNN was initialised using an Adam optimizer.

4.6 Chapter summary

Comparing the performance of the regression models trained on various features (both separately and combined) enabled the selection of the best performing regression method. The experiments helped to determine whether the predictive performance of the models would bolster performance when used with multiple data sources. Furthermore, regression models were built, trained, and tested while carrying out the implementation experiments. In the next chapter, the results of the trained models are analysed and discussed by assessing the performance of the models using their respective evaluation metrics, as discussed earlier in Section 3.9.

Chapter 5. Discussion and analysis

5.1 Introduction

The evidence presented in the preceding chapter has confirmed several important findings. This chapter summarises the findings regarding the performance of the models that were implemented in an attempt to answer the research questions presented in the first chapter. First, the poverty levels in South Africa are sufficiently high to warrant persistent and targeted analysis of time-series evidence, allowing comparisons to be drawn between multiple models and analyses. Second, the application of a night-time luminosity model for poverty analysis has partial validity, with particular advantages in rural or suburban landscapes in South Africa. However, this study has revealed several important insights that need to be further discussed in relation to the prior academic investigations in this field and the theoretical propositions afforded by multi-dimensional poverty proxies and remote sensing capabilities.

An attempt is made to link the performance analysis of the models with the research questions and to identify the models that perform the best in the research. This diversity is important in investigating the validity of the research results. Elements that may impact the validity of research experiments, in addition to discussions about how to limit their impact, are addressed in this chapter. The evaluation of the performance of an algorithm can sometimes be ambiguous due to inherent variance as a result of ambiguities in lexical, syntactic, semantic, and metonymy, for example. The following sections will outline and discuss these findings, drawing on previous research in this field to triangulate the evidence and illuminate practical opportunities for improving future analyses.

5.2 Satellite imagery approach

The model for this approach was developed in order to test the performance of satellite-based deep learning models in South Africa. The main task was to predict South African socio-economic indicators in a cluster of households from the ground

truth data set. The ground truth dataset had 11,083 households grouped into 750 unique clusters. The average latitude and longitude locations of the households were maintained within each cluster with noise of up to 5km in each direction in order to preserve the anonymity of individual households. As a result of the significant noise in the available input data, the tiles were obtained within a radius surrounding each cluster centroid, which corresponds to a pixel in the night-time dataset. Regression models were trained and evaluated using a five-fold nested cross-validation scheme. The results of deep learning methods presented by previous investigations in this field, including those of Jean et al. (2016, p. 792) and Pfeifer et al. (2018, p. 20) provided the methodological basis for the exploratory analysis of poverty in South Africa through GIS-based remote sensing and machine learning techniques. Predictably, the range of results encountered by these previous studies varied, offering R^2 indications that were regionally distinct and varied quantitatively depending upon the machine learning method and image quality (e.g., Uganda: 0.69, Nigeria: 0.68, Malawi: 0.55, Tanzania: 0.57 Rwanda, 0.75) (Jean et al. 2016, p. 22). The current study achieved a comparative R^2 outcome of 0.55, thus confirming that the techniques employed are not only comparable to other approaches, but they extend the prior research and classification techniques of researchers like Jean et al. (2016, p. 792) and Head et al. (2017, p. 4) towards an improved classification framework.

Specifically, the current study developed novel pre-processing techniques that considered the uniqueness of the South African urban landscape, an environment that is characterised by high urban density and high urban sprawl amongst low-income citizens and low urban density, enclavisation amongst higher-income citizens. Drawing upon the socio-cultural roots of apartheid, it was possible to predict variations in the visual representation of poverty in South Africa, considering the monumentality of specific landmarks such as highways, factories, and waterways in relation to the localisation of high- and low-income enclaves (Statistics South Africa 2019b, p. 3). Yet, Pfeifer et al. (2018, p. 32) had previously confirmed that regional development in South Africa had impacted upon the localisation of income, wealth, and development to key areas affected by the World Cup. Therefore, by selectively targeting urban areas where improved structures, new buildings, and social service facilities had been constructed during this large-scale refurbishment initiative, it was possible to extend

the work of Pfeifer et al. (2018, p. 32) and chart post-developmental effects on poverty over time.

The study was able to achieve a wealth index R^2 score of 0.54. It can be seen from the findings in Table 5.1 that the regression methods do not generalise to other socio-economic indicators with the same accuracy as wealth: for example, the random forest result for education score is $R^2 = 0.16$, access to electricity has $R^2 = 0.07$, and access to water has $R^2 = 0.27$. Similar to the findings of Head et al. (2017) and Tingzon et al. (2019, p. 425), the scores for years of education completed, access to electricity, and access to water are low as they do not have a clearly defined relationship with night-time lights.

Table 5.1: Satellite imagery regression results comparison table

Model	Ridge Regression		Random Forest	
	R^2	RMSE	R^2	RMSE
Wealth index	0.532	68207	0.541	68207
Education completed	0.156	3.586	0.153	3.584
Access to Electricity	0.060	0.205	0.072	0.233
Access to Water	0.256	9.202	0.274	9.024

The performance of the regression models was analysed using R^2 and RMSE performance metrics. After comparing the metrics, the regression model which performed the best was selected (this is discussed in detail in Section 5.4). The metrics are good performance indicators because they consider how well the different models fit the data based on the relationship between the response and independent variables as well as the overall prediction ability. The variables and metrics were detailed in Section 3.9.

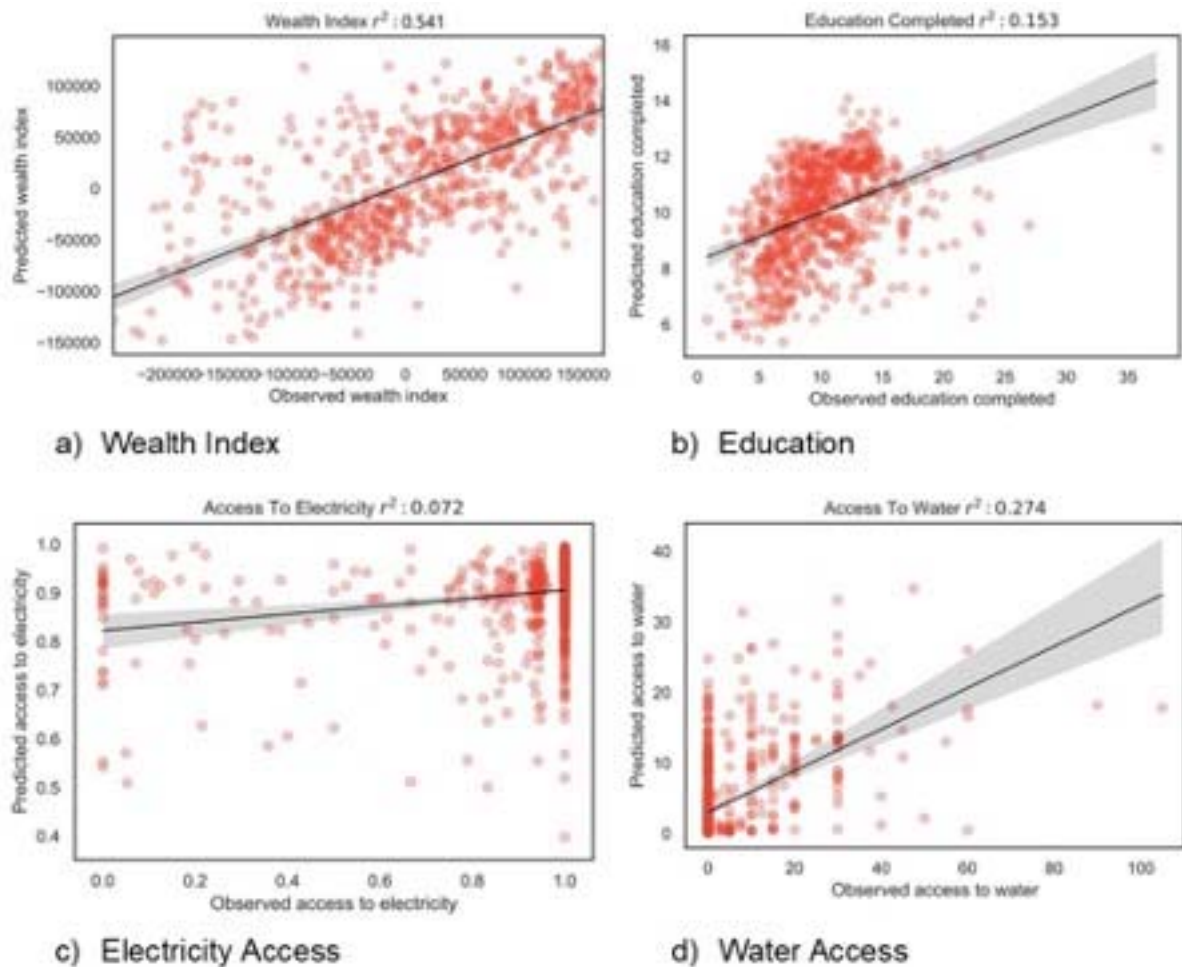


Figure 5.1: South Africa ground-truth wealth index and the model poverty prediction results for the study

A central advantage of the night-time lights model adopted for this study was its ability to extrapolate a gradient interpretation of regional luminosity, mirroring the evidence presented by Jean et al. (2016, p. 792) in relation to urban and rural spaces and the multi-class model developed by Tingzon et al. (2019, p. 428) to differentiate night-time light intensity across varied urban densities. As a model for charting population density or highlighting centres of commerce, night-time light density analysis offers a valuable tool for evaluating variations throughout extended periods of time or following significant developmental improvements or changes. Engstrom et al. (2017, p. 19) have confirmed, however, that night-time luminosity is not a viable indication of poverty in most urban settings because of the heterogeneity and interconnection of various socio-economic classes. In such urban environments, the poor and wealthy are likely to live in close proximity as the urban space itself evolves and adapts to changing population flows and residential patterns.

5.3 Crowd-sourced information approach

To assess and predict regional poverty, the current study adopted a set of random forest and ridge regression object-oriented feature detection training sets. For visual sanity checks, Spearman and Pearson rank correlation coefficients were implemented to evaluate the potential relationship between the features and the target variables. By focusing on specific features (e.g., road, building, point of interest), variance could be calculated in relation to poverty indicators at both low-level inclusion and higher granularity inclusion levels. By integrating all three characteristics of the OSM features, including night-time lights, OSM features, and binary regional indicators, the accuracy of the wealth prediction model increased significantly to 0.55. A summary of the results for the two regression models is shown in Table 5.2. Further subdivisions of key socio-economic indicators related to education, electricity access and water access were found to predict poverty across R^2 of 0.17, 0.08, and 0.27, respectively with water access being the best predictor. The model trained on OSM features alone achieved an R^2 of 0.44. Although data on household durable goods were extrapolated from Statistics South Africa (2019b, p. 1) dataset, the ability to correlate household consumption with GIS-based poverty indicators was restricted by a lack of intra-household congruity and insights. The poverty prediction approach for this research was optimised to predict asset-based wealth as measured by DHS. Other indicators that included access to hospitals and access to education that incorporated proximity dimensions and were found to have a similar positive 0.48 for access to electricity or water, since the features of the night-time lights could be correlated within the statistical regression model. The correlation could be due to the nature of the task, as assets can be visible from satellite images. The overall results achieved also fall in line with MPI measurement, where standards of living indicators are related to asset ownership and accessibility to essential services.

Table 5.2: Comparison table of OSM and VIIR data regression results

Model	Ridge Regression		Random Forest Regression	
	R^2	RMSE	R^2	RMSE
Wealth index	0.518	76852	0.546	68207
Education completed	0.169	3.548	0.155	3.580

Access to electricity	0.065	0.231	0.075	0.232
Access to water	0.242	9.280	0.268	9.085

This study confirms that the accuracy and overall performance of the OSM-night-lights-hybrid model are statistically comparable to those of the satellite-based transfer learning model. For policymakers and researchers in development planning, this evidence offers a significant advantage due to the publicly accessible OSM databases and the emerging transparency of government-based Statistics South Africa (2019b, p. 1) evidence. With the accuracy of the predictive models almost producing similar results, the applicability of these findings to future research in this field yields tangible cost savings over field-based research or those satellite-based datasets requiring pay-for-access in order to download daytime satellite images (Fatehkia et al. 2020, p. 13). Additional improvements to poverty or wealth indicators through model training and AI interpretation of key structural and spatial features will predictably improve the accuracy and structural consistency of such analyses, as existing models can be critically compared with future datasets and changing regional geographies.

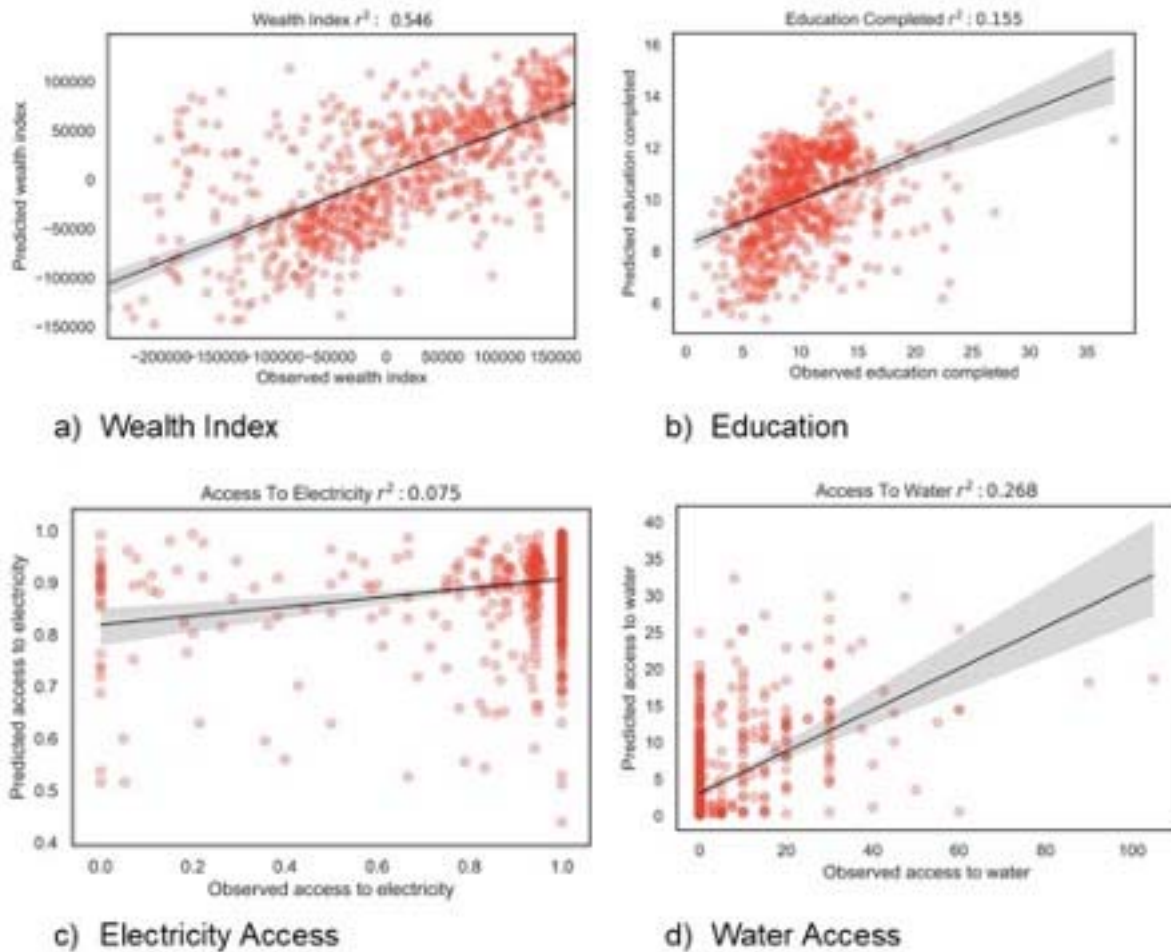


Figure 5.2: Results of the South African ground truth wealth index and poverty prediction model

However, the OSM features have some drawbacks when used to predict poverty. Amongst the disadvantages is the fact that OSM can be regarded as semi-automated, which means it requires assessment to measure the accuracy and completeness of the data before use: it is volunteer-curated. Therefore, it is highly dependent on the input of domain-experienced surveyors to map up-to-date geospatial specific features. For countries like South Africa, this can be easily achieved. In some settings, certain information could be missing, making the dependability of this data source a challenge. Finally, it was noted during pre-processing that the OSM dataset contains a large number of different features, with multiple attributes pertaining to the same building type or POI, for example, thus requiring extensive data cleaning and feature engineering.

5.4 Features and further evaluation

Several evaluation techniques were used to measure the performance of the models (see Section 3.9). These are discussed in the following sections. In addition, the usual train/test splitting process used cross-validation, which could potentially lead to different training and testing data distributions. As discussed in Section 3.9.3, 5-fold cross-validation was used to mitigate the problem.

5.4.1 R-squared

The machine learning models adopted for estimation were ridge regression and random forest, which were checked to see how the results would change if one of them was used. R^2 , a statistical measure representing the proportion of variance for the dependent, as explained by an independent variable in a regression model was used to evaluate the performance. Comparing the R^2 results of the satellite imagery and crowd-sourced information approaches prediction accuracy in Tables 5.1 and 5.2 respectively from the ridge and random forest regression, both approaches did well, although random forest was generally more accurate in all estimations except for Education Completed indicators. Random forest estimation is a prominent machine learning algorithm that is commonly adopted because of its strengths. Among the strengths of random forest are that it has precise learning algorithms, operates well on large datasets, manages thousands of input variables, provides estimates on variables that are critical in classification, and is capable of estimating missing data (Breiman & Cutler 2002). Therefore, it means that Random Forest is a good predictor of poverty using data of the type discussed in this study, though not in all cases. Table 5.3 shows the results for the models that were tested for this study.

Table 5.3: R^2 output comparison

Model / Indicators	OSM Features Only	VIIRs Data	OSM+VIIRs	Satellite Imagery
Wealth Index	0.322	0.375	0.546	0.541
Education Completed (Years)	0.141	0.162	0.174	0.153
Access to Electricity	0.077	0.115	0.103	0.072
Access to Water (Minutes)	0.177	0.218	0.264	0.247

5.4.2 Root Mean Square Error

Comparison of the RMSE, which uses sample and population values to measure the differences between the predicted values and the values observed from the computational methods (Tables 5.1 and 5.2) shows that the random forest method produced the lowest RMSE value. In particular, the satellite imagery model generated the highest RMSE.

5.4.3 Descriptive statistics

To enable the summation of large amounts of data in limited numerical values that highlight only important data features, descriptive statistics were used as part of the model development and evaluation. The models that were developed involved several datasets, some of which were images or numbers that were counted by the system and needed to be displayed as a means of sanity checks and evaluation to ensure the progress of the model and to display some of the results of the deep learning 'black-box.' The success in extracting the features with minimum effort and without relying on expert-labelled data suggests that the transfer learning method is able to generalise to the socio-economic measurement and prediction problem. From the results of night-time features, summary statistics and histograms were produced for the luminosity pixels within each cluster. In order to train the satellite imagery model on night-time luminosity data, it was required to obtain the characteristics of the distribution of the different values of night-time luminosity per DHS cluster. The list of luminosity values was processed and the bins were grouped according to the mean, maximum, minimum, median, coefficient of variation, standard deviation, skewness, and kurtosis. Data was saved as a .csv file, and Figure 5.3 shows the sample results.

```

nightlights_summary = nightlights_summary_df['stat'].apply(pd.Series)
nightlights_summary['DHSCLUST'] = nightlights_summary_df['DHSCLUST']
nightlights_summary.rename({0:'mean', 1:'max', 2:'min', 3:'median', 4:'cov', 5:'std'

```

```
nightlights_summary.head()
```

	mean	max	min	median	cov	std	skewness	kurtosis	DHSCLUST
0	3.071989	11.44415	0.0	2.392455	6.956913	2.637596	1.507983	1.947498	1.0
1	0.467845	3.07438	0.0	0.470080	0.208338	0.456440	1.815148	6.905500	2.0
2	0.867598	3.38449	0.0	0.820385	0.451345	0.671822	1.085238	3.004943	3.0
3	0.012949	1.92717	0.0	0.000000	0.015556	0.124722	13.209361	187.920936	4.0
4	0.123967	0.98906	0.0	0.000000	0.035462	0.188314	1.866225	3.510677	5.0

Figure 5.3: Night-time lights summary data

The Gaussian mixture model was used to map the luminosity values to the night-time class bins of low, low, medium, medium, high medium, and high. Table 5.4 shows how the data was allocated to the different bins.

Table 5.4: Gaussian mix model image bin assignment

Bin	Number of images	Percentage
Low	50055	0.442123
Low medium	50055	0.267253
Medium	21598	0.190770
High medium	6844	0.060451
High	4461	0.039403
Total	113215	100

Figure 5.4 shows the average distribution of the night-time light intensities within each cluster.

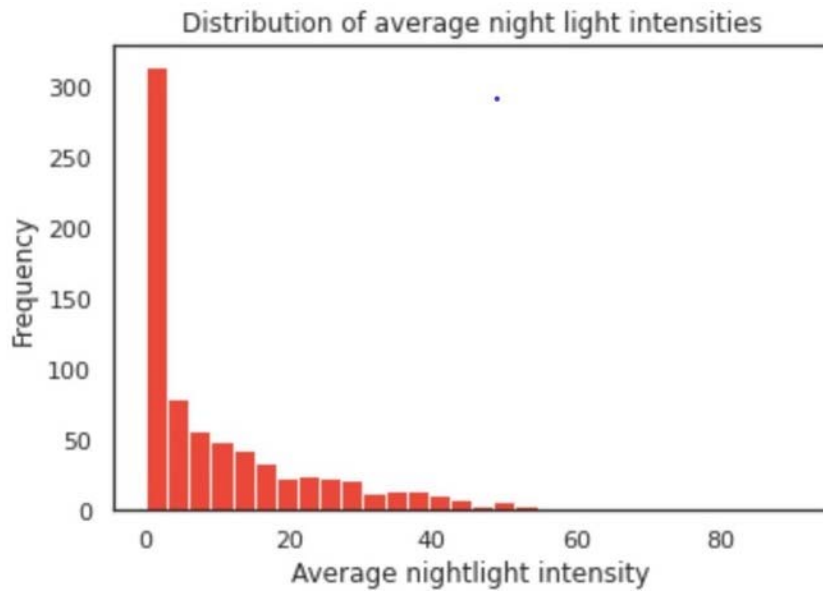


Figure 5.4: Average night-time light distribution per cluster

The transfer learning approach used five night-time light intensity classes which were obtained by fitting a mixture of five Gaussian distributions to the relative frequencies of the night-time intensity values instead of the three used by Jean et al. (2016, p. 790) to improve the model performance. The five classes were adopted by monitoring the night-time light intensity histograms, as in Figure 5.5, for the South African DHS survey locations that insinuated five dominant night-time light modes. The Gaussian mixture model provided a principled approach to bin data.

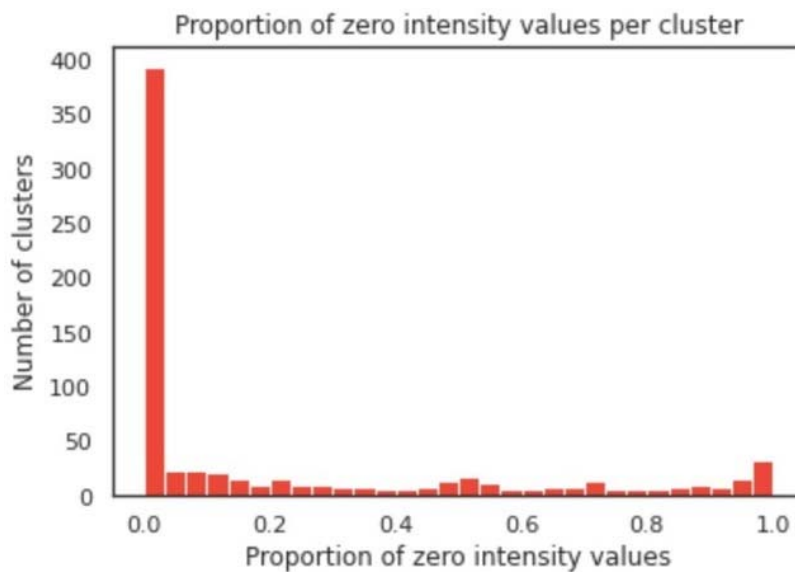


Figure 5.5: Distribution of the light intensity at night per cluster

The relationship between wealth index, completed education, access to electricity, and access to water socio-economic indicators at the cluster level and the mean night-time light intensities per cluster was computed and is summarised in Figure 5.6.

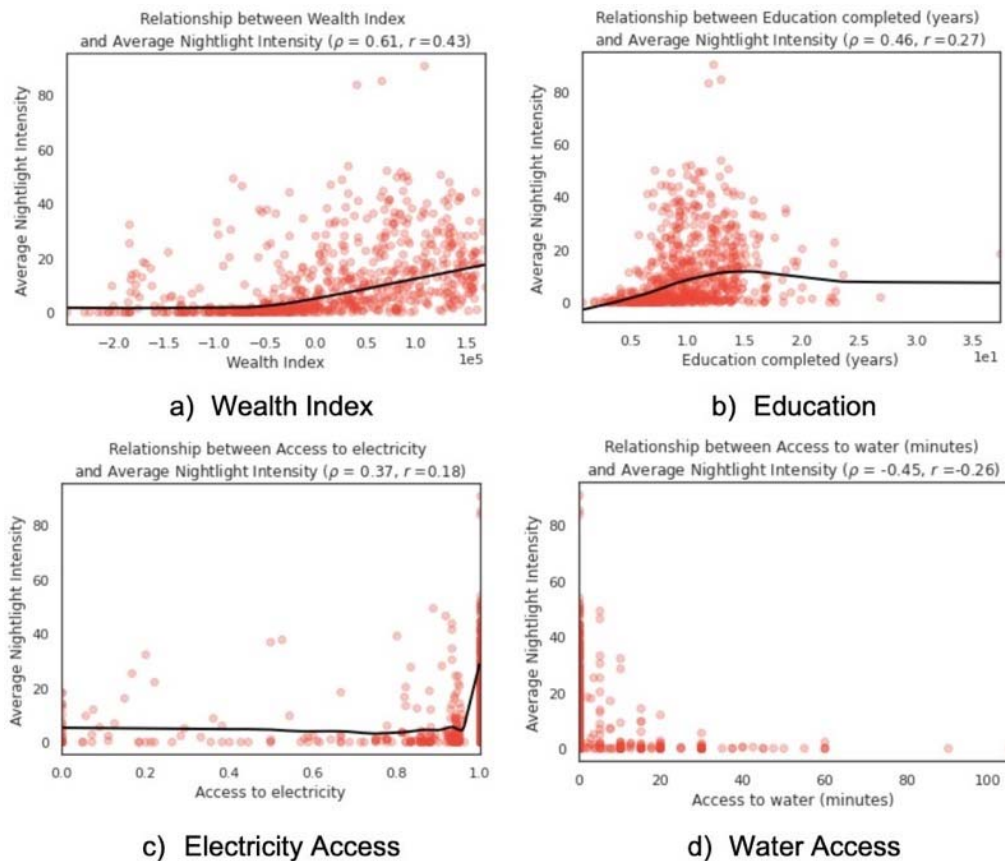


Figure 5.6: Relationship between Night-Time Light Luminosity and DHS Indicators

The dark lines represent a LOWESS (Locally Weighted Scatterplot Smoothing) curve, which is a regression fitted to the data.

A t-SNE plot is a visualisation of the clusters in the data, namely, clusters for embeddings. Embeddings are high-dimensional vectors that encode information regarding distinctive attributes in a picture. For this research, attributes were pulled from satellite images that were responsible for night-time lighting (e.g., roads, buildings, etc.). Not only was the number pulled, but we also extracted additional information that the deep learning model could parse. All the information was then encoded in the vectors as shown in Figure 5.7.

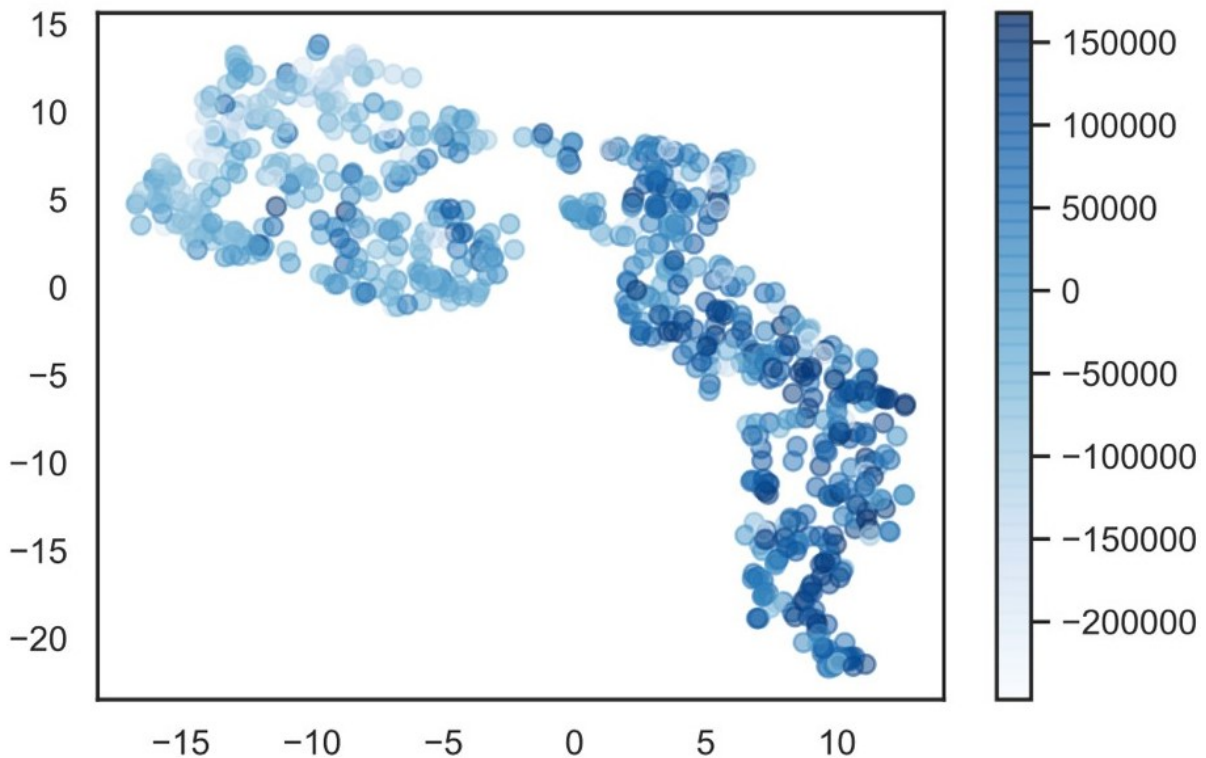


Figure 5.7: t-SNE plot for night-time lights

The t-SNE plot demonstrates the internal structure of all the embeddings as multidimensional data. The main purpose is to check if the internal structure has some data clustering and if the result is a very tight ball, which means that the embeddings could not pick up any distinctive features from the data, which is purely a sanity check. The results of this study show that more light points are closer and more blue ones are separated, so the vectors responsible for low intensity are close to each other, for example.

Regression graphs were used to visualise the relationship between the independent variables and the indicators used plots prediction in this study. Figures 5.1, 5.2 and 5.5 are some of the plots that were generated to show the relationship between the socio-economic indicators from the DHS ground truth data and the predicted values. After merging the data for buildings, roads, POIs, and night-time lights, it was important to investigate their impact on the predictive power of the model using the Spearman-Pearson correlation. Pearson benchmarks the linear relationship, whilst Spearman benchmarks the monotonic relationship. The most important features are shown in Figure 5.8. Of the important features listed, most consist of POIs such as schools, healthcare facilities, or leisure. Paved roads were of great importance. The

presence of the feature in a cluster was considered for this study, and it would be ideal in the future to consider whether the distance to the POI influences the level of poverty.

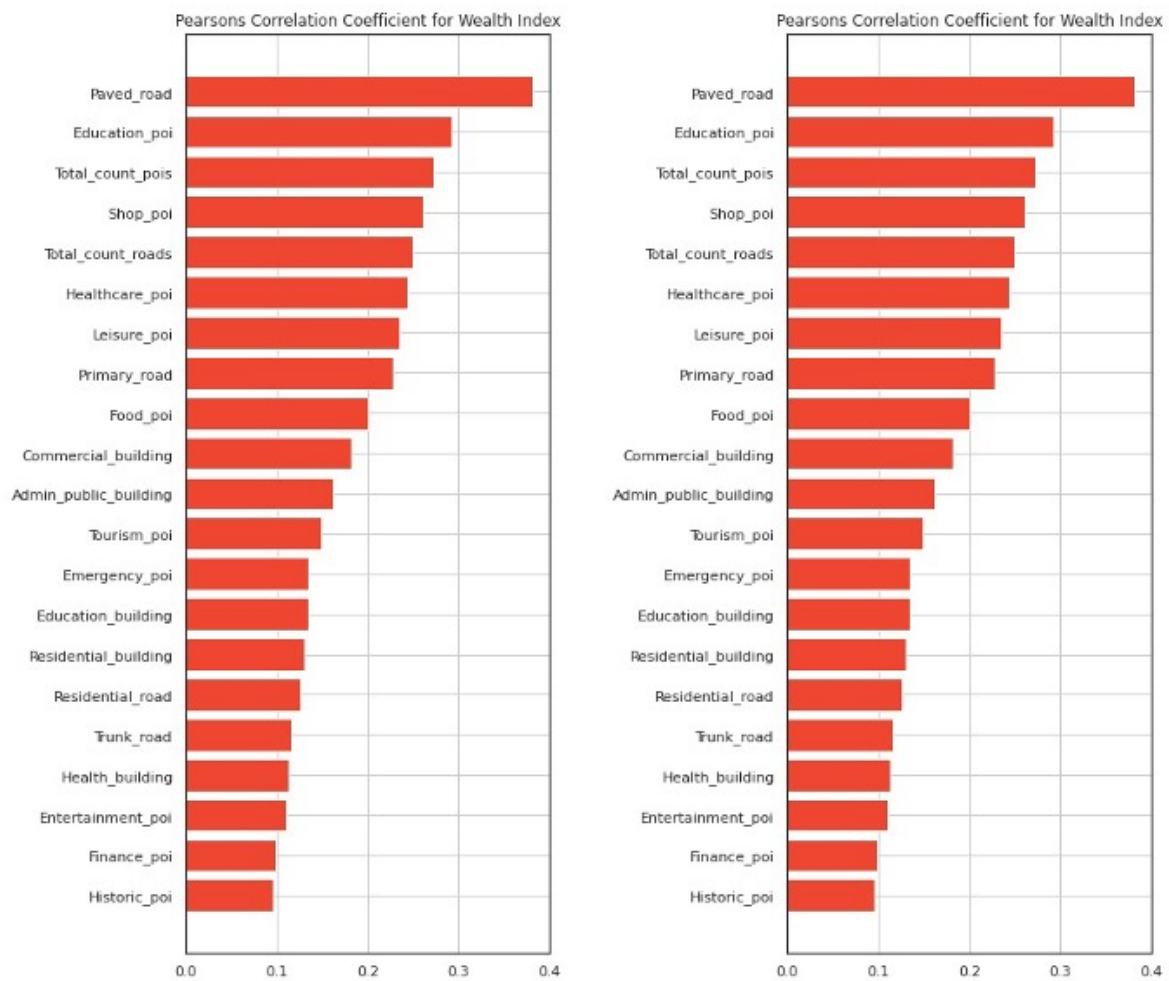


Figure 5.8: Spearman and Pearson's correlation for important features

5.5 Poverty mapping

Binary regional indicators were used as predictors of a wealth index score. Using a fine-grained poverty map that uses the 2016 DHS survey wealth index data for comparison, a qualitative poverty map was generated at the district level using the predicted wealth index. The wealth index score was compared with data from the DHS survey and then used as a proof of concept to reconstruct poverty maps at the district level, as shown in Figure 5.8. These are computed by grouping the estimated wealth values at the cluster level. From this figure, it can be seen that the final estimated wealth model was able to produce a distribution of poverty in South Africa comparable to the map obtained using the data from the DHS survey. In Figure 5.9, the dark green to dark red color gradient is used to show the intensity of poverty in a district. Dark green represents that the location has a higher wealth index, whilst dark red indicates an exceedingly low wealth index.

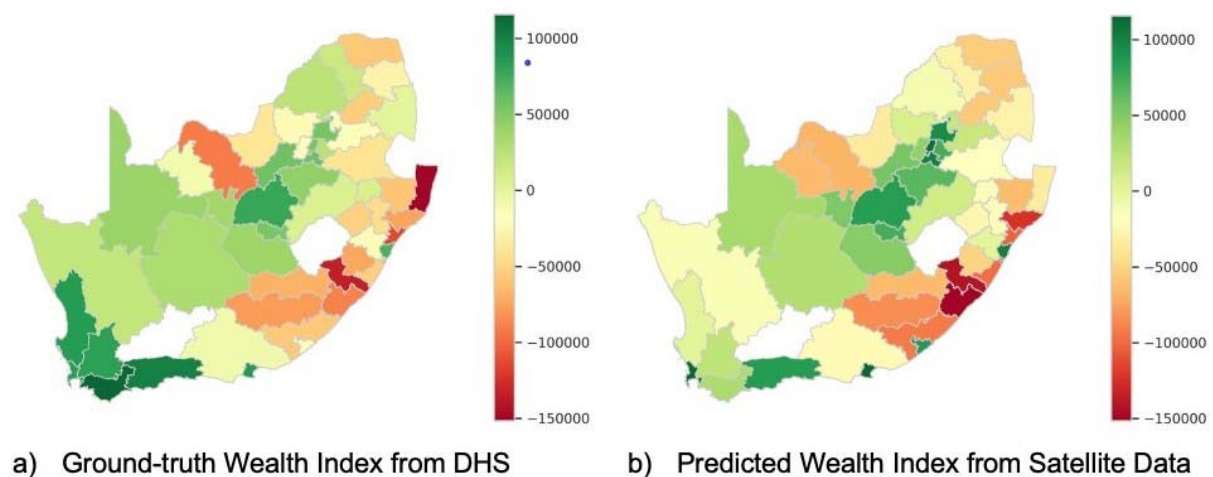


Figure 5.9: District-level ground truth and predicted wealth indices

It appears that the highest poverty is in the districts of the Eastern Cape, KwaZulu-Natal and Limpopo. The maps have some differences that may be due to the temporal disparity between the DHS and satellite imagery data years or inaccuracies in either map. Another comparison as visualised in Figure 5.10 shows the most impoverished regions of South Africa, again confirming that they are concentrated around high-density urban centres, as citizens continue to live in segregated enclaves close to their familial homes under apartheid (South Africa Gateway 2019, p. 1).

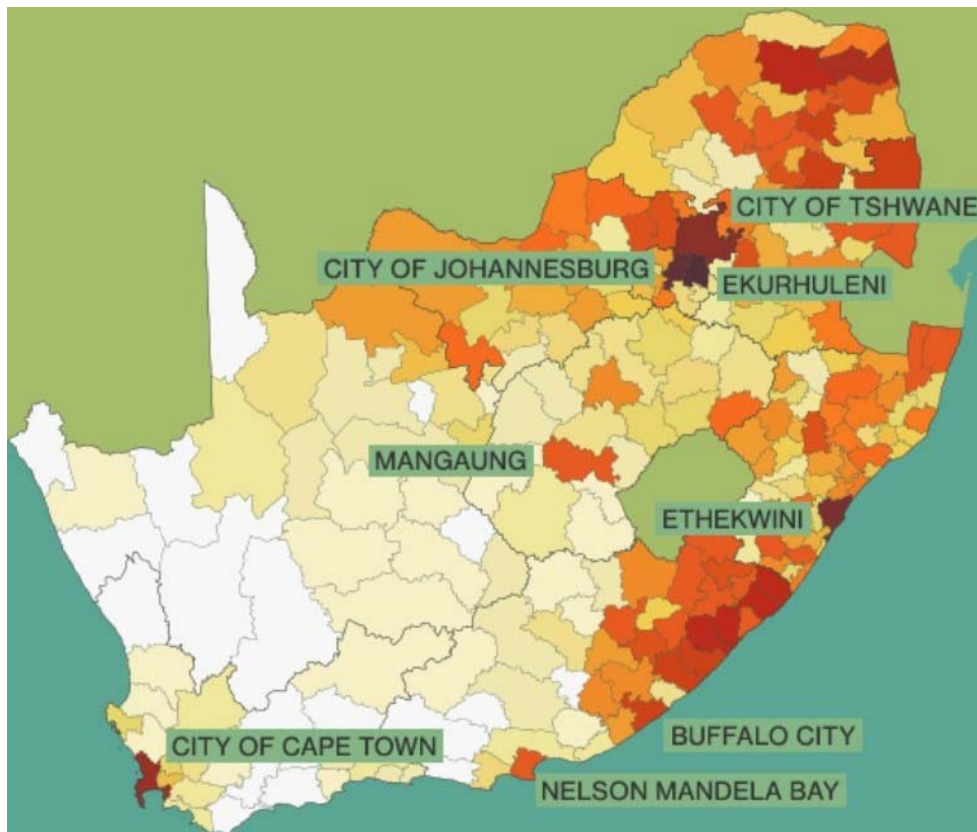


Figure 5.10: Regional localisation of domestic poverty (South Africa Gateway 2019, p. 1)

Prior research in this field conducted by Pfeifer et al. (2018, p. 4) shows that positive infrastructural gains following the 2010 World Cup in South Africa have had measurable effects on regional poverty. With a decline in unemployment of more than 1.3% in areas central to the event infrastructure and further improvements across city spaces where transportation infrastructure was upgraded or installed to meet visitor needs during the 2010 FIFA World Cup, the night-light evidence modelled by Pfeifer et al. (2018, pp. 27–8) suggests similar effects across urban spaces. Drawing on such insights and the practical advantages of remote sensing and GIS mapping, the current investigation was carried out to critically evaluate transformative patterns of poverty at various urban levels. By applying multi-layered data collection and analysis to the diverse regions of South Africa’s urban centres, this study has narrowed the insights of this model to specific developmental relationships, providing meaningful outputs that can be applied to poverty analysis in the future.

5.6 Reliability, validity, and generalisability

The central comparability indicator for this study was based upon the incidences of wealth in the ground truth for South African clusters. By comparing income-class specific differentials across urban spaces, large-scale geographic luminosity comparisons were drawn through satellite-based night-time spatial GIS modelling. The degree of reliability of the model was calculated and indicated a lower degree of methodological validity. Specifically, the OSM and night-time lights analysis was best suited to comparative regional assessments, allowing a pixelated distillation of daytime poverty data and night-time comparison to demonstrate a high degree of correlation and poverty-oriented indications. While the training process allowed the machine learning execution to detect gradient pixilation with relative accuracy, the translation of these scalar dimensions (e.g., low luminosity, moderate luminosity) into poverty indicators was contingent upon other regional socio-economic conditions (e.g., housing density, regional income levels, access to services).

One of the problems with reliability and validity in a nation like South Africa where peri urban sprawl has led to informal development and unplanned housing construction is that traditional, physical landmarks and indicators of regional stratification are missing. For example, in Cape Town, where peri-urban construction is directly separated from high income residential enclaves via highways, the arc of neighbourhood construction follows the roadway, not a planned residential map or a grid model that might otherwise inject specific facilities (e.g., medical, education) into the landscape itself. In the research conducted by Fatehkia et al. (2020), a similar incongruity was observed in relation to the varied distribution of wealth across multiple island chains and archipelagos that resulted in the direct impacts of the physical landscape of the associated built environment.

In spite of the lack of reliability and validity in relation to night-time lights and urban poverty, these findings offer generalisable machine learning techniques that can be adopted in future investigations regarding urban density, poverty patterns, and government development strategies. For example, Mveyange (2016, p. 15) and Babenko et al. (2017, p. 3) applied the CNN-based machine learning solution to the evaluation of tangible relationships between access to key government services and regional poverty. Similarly, Engstrom et al. (2017, p. 7) revealed the classification advantages of object-based training, suggesting that varying patterns of usage,

varying property characteristics, and vehicle density can all be used to generalise socio-economic wealth (or poverty). The current study has confirmed the viability of such object-oriented AI training, demonstrating both the advantages and limitations of feature granularity as a range of urban characteristics have been assessed and modelled in relation to South Africa.

Similarly, Engstrom et al. (2017, p. 7) revealed the classification advantages of object-based training, suggesting that varying patterns of usage, varying property characteristics, and vehicle density can all be used to generalise socio-economic wealth (or poverty). This study has confirmed the viability of such object-oriented AI training, with roads and buildings being the best indicators, as they showed a variance of 45 - 50% demonstrating both the advantages and limitations of feature granularity as a range of urban characteristics have been assessed and modelled with relation to South Africa.

5.7 Chapter summary

This chapter has discussed the research and experimentally evaluated the performance of transfer learning and OSM geospatial methods to predict socio-economic well-being in South Africa in relation to the research objectives and questions set out in the first chapter. Central to the evidence captured during the course of this study is a positive, affirming relationship between earlier research conducted by Jean et al. (2016, p. 792) and Head et al. (2017, p. 1). By confirming the transferability of these combinative, daytime-night-time-OSM comparisons to the South African landscape, this investigation has demonstrated the broader potential for remote sensing, crowd-sourced data, and machine learning across global poverty studies and socio-economic assessments. For South Africa, these findings have direct implications, supporting future research beyond the costly and time-consuming limitations of *in situ* fieldwork. The experimental results revealed that the proposed approach that merges night-time lights data and OSM features produced the most optimal results. By continuing to capture visual relationships via satellite and crowd-sourced images, it is possible to track the socio-economic impacts of government investment, development, and policymaking over time. Regression models were

deployed, and the statistical significance of the results was assessed using R^2 and RMSE. The following chapter will conclude these findings and outline future paths for further research in this field and alternative applications of these techniques.

Chapter 6. Conclusions and recommendations

6.1 Introduction

In the previous chapters, this research compared approaches based on transfer learning and OSM geospatial features to predict socio-economic well-being of clusters within the DHS-surveyed areas in South Africa. The research involved investigating the effectiveness of combining computer vision, satellite imagery, night-time luminosity data, and crowd-sourced geospatial information to provide granular and reliable poverty estimates while overcoming the challenges of laborious, expensive, and time-intensive methods required to conduct household surveys on the ground. This final chapter concludes the research implemented by presenting an overview of the findings, highlighting the achievements and contributions of this multidimensional study, and the potential applicability of the proposed experimental model. The research had one main research question further broken down into three sub-questions (listed in Section 1.5) to realise the research objectives (summarised in Section 6.2). This chapter concludes the research with recommendations for future research to address the limitations of the study in an effort to alleviate poverty.

6.2 Overview of the study

The study drew findings by establishing a reasonably strong socio-economic well-being estimation, predictive accuracy was yielded by using readily accessible night-time lights and volunteered geographic information datasets. Approaches to poverty estimation using machine learning and leveraging these datasets yield a wealth estimation R^2 of 0.56 from the random forest regression model compared to an R^2 value of 0.53 for the satellite imagery approach. This study included a literature review of the problems with machine learning and computer vision models that have been adopted to date in order to determine the limitations of reliable poverty data from developing countries. This information can affect decision makers and help them make informed policy decisions and effectively allocate appropriate resources to the neediest areas. The information garnered helped to re-establish the research objective

and show relevance of the study. An artefact was designed, developed, and evaluated using design science research methodology. It was demonstrated in the form of experiments using South African data. The artefact design process was evaluated using documented procedures to ensure research rigour. Evaluation was performed in iterations by running the models until similar results were reported to ensure maximum performance and utility for the model.

6.3 Revisiting the problem statement and research questions

The main objective of this study was to develop and test an integrative approach to combine publicly available and freely accessible geospatial data, as well as to use machine learning capabilities for the estimation of socio-economic well-being (see Section 1.6). From the extant design science literature, the main end result of a research is the development of an artefact such as methods, models, or instantiations (see Section 3.7). Thus, the major contribution of this study is the artefact that was developed (see Chapter 4). As stated earlier, Google restricts imagery download API requests to 25,000 per day and will start charging for additional downloads while high-resolution from DigitalGlobe, for example, are proprietary, expensive, and non-redistributable. Alternatively, the OSM and night-time lights datasets are publicly available, cost-effective, and update frequently, making them reliable to complement the time-lagging and expensive *in situ* surveys as solutions for diagnosing, modelling, and interpreting the conditions of systemic exclusion in order to improve policymaking and target future interventions. In the process of developing the artefact, the dissertation answered the research question and three sub research questions identified at the beginning of this study (see Section 1.5). The findings and conclusions pertaining to each specific research question, answer, and findings are summarised in the remainder of this section.

Which technique is suitable for conducting remote sensing of satellite imagery, and how can this approach be adapted to analyse poverty in South Africa?

To test whether the model would improve upon the use of either satellite imagery or merged night-time lights and geospatial data in estimating the wealth index, regression models were deployed, and the outcomes of the models were evaluated.

The first approach was a model that used daytime satellite imagery with night-time lights as a proxy for transfer learning. The results of this approach were 4,096-dimensional feature vectors that were fed into a regression model along with DHS socio-economics data to estimate wealth indices. The second approach was also based on regression models that used DHS socio-economic indicators along with night-time lights data and crowd sourced OSM information, first separately and then combined. The results of the approaches were compared to determine the performance of the models. Compared to the results of the models that used night-time lights and daytime satellite imagery for wealth estimation, the night-time lights and OSM data model accuracy results were slightly higher. This suggested that a model can be improved in accuracy and performance by using data of different types, where an R^2 of 0.55 was achieved from cost-effective datasets. Taking into account the amount of time required to process daytime satellite imagery data in addition to computing resource requirements, the OSM crowd-sourced information and night-time lights data approach was more efficient and accurate than the daytime satellite imagery approach with a computationally cheap and easy-to-implement approach.

What are the challenges that affect machine learning capabilities and their application to satellite-based imagery of varied regional cityscapes?

Satellite-based model performance, particularly in settings beyond where they were trained, is perhaps the most common and important concern for researchers and policy makers interested in potential applications in sustainable development. For privacy reasons, the GPS coordinates of the DHS survey were displaced by up to 5 km in all directions, thereby introducing noise into the training data. The performance of models can be degraded due to noisy training data in two possible ways. To begin with, the ability of the model to learn features in imagery that are predictive of the outcome of interest can be diminished. Second, and more subtly, the model may learn the relevant features but perform poorly in predicting test data precisely because the test data have noise. The latter concern would lead studies to understate the true performance of the model. Noisy datasets are increasingly being employed for model development, which requires researchers to contend with the dual challenges of not overfitting to noise and not understating the performance of a model with respect to reality. Some of the challenges noted from the literature review include blooming of night-time lights, reliability and consistency associated with field-level investigations,

data not readily accessible (expensive), defining economic labels, lack of available ground truth data, overfitting, generalisability problems, the black box problem, possibility to granularly assess within the lowest levels of clusters, heterogeneous influences, and transferability to other regions.

What best practices and techniques can be applied to architect predictable remote sensing capabilities in future poverty assessment studies?

Integrating multiple sources of data is an effective way to derive maximum benefit from conventional and innovative machine learning algorithms. Studies on the feasibility of incorporating satellite imagery and geospatial data do not aim to replace conventional methods of measuring poverty and wellness. Rather, satellite imagery analysis, crowd-sourced geospatial data, and conventional sources of poverty data can be blended to complement official statistical data to provide a level of granularity that sheds light on a country's true spatial distribution of socio-economic disadvantage. These studies can help address some of the limitations associated with traditional poverty estimation techniques. In addition, they can help validate the findings produced using traditional methodologies, which can serve as a means to build trust in the poverty statistics compiled by the statistical offices.

6.4 Summary of contributions

Research contributes to the existing literature on poverty estimation by developing a prediction model that uses publicly available datasets and computer vision techniques. By doing so, the objective is to analyze the results and assess the feasibility of the solution to serve as a means of reinforcing confidence in poverty maps compiled using this innovative methodology in South Africa, a feat that is currently being thoroughly explored globally. It will serve as an excellent starting point to adopt machine learning techniques without incurring significant financial costs upfront. This model is a new contribution to the South African context as no other similar studies have been conducted to the best of my knowledge. This has the potential to interest policymakers and development agencies to understand the spatial distribution of poverty at zero data cost and aid the use of national statistics data for poverty intervention policies and programs.

6.5 Concluding remarks

Although South Africa is colloquially viewed as a developmental success on the African continent, the longstanding legacy of apartheid and the persistent effects of robust socio-economic inequality indicate that developmental incongruities perpetuate poverty and restrict widespread development. Although government officials and domestic welfare agencies have undertaken to calculate the extent and impact of poverty in the South African landscape, the high cost, timeliness, and unreliability of such field-based research techniques have created significant gaps and inaccuracies in reporting output. There is an unrelenting necessity to find reliable data for socio-economic well-being estimation in lieu of statistical survey data. Consequently, an alternative solution was needed in order to diagnose, model, and interpret systemic exclusion conditions in order to improve and inform policymaking, allowing government programmes to meet a wider range of needs and social solutions based upon specific regional gaps. This study recognised the emerging advantages of remote sensing technologies, explored a variety of previous techniques and approaches, and developed a viable technical solution that could be applied to South Africa. Remote sensing data, geographic information, and night-time luminosity data can reflect some of the characteristics associated with poverty and have been proposed by several researchers.

The primary research question answered over the course of this investigation focused on how night-time analysis of luminosity and crowd-sourced geographic information could be used to determine the relationship and predict socio-economic variations. At the beginning of this study, it was predicted that there was a negative correlation between luminosity and wealth estimation that could be measured by applying regional density assessments and welfare statistics to the outputs of a night-time luminescence assessment. In order to capture this relationship, however, it was essential for this methodology to consider poverty in terms of key socio-economic indicators including density, access (e.g., education, healthcare, transportation, water), and household consumption. A night-time model would have been inadequate on its own to identify the specific features (e.g., hospitals, schools) needed to assess

the relationship between luminosity and socio-economic deprivation. As a result, multiple databases, including OSM geographical information, were used to assess and identify a training set of structural characteristics that could then be applied pixelated to the night-time assessment.

By implementing the deep learning approach and accurately training a CNN AI-based application to recognise asset-based wealth and poverty via remote sensing, this study has successfully and reliably adapted the techniques outlined in previous poverty studies, including Jean et al. (2016, p. 792) and Xie et al. (2016, p. 3935). The ubiquity of publicly accessible GIS-based data resources via online channels from both global agencies such as NOAA and crowd-sourced data from OSM ensured that this study compared geographical domains across similar scalar and resolution-controlled models. However, the evidence captured during this research process suggests that remote sensing is neither a trivial solution to government poverty surveys, nor a simplistic appropriation of technological advances in the field of machine learning and geospatial analysis. Instead, a systematic, structured, and selective procedure has emerged from the confluence of this and other studies in this field that reveals significant opportunities for poverty modelling and analysis in the future.

The study found that, in general, the merging of OSM features such as roads, buildings, and POI with night-time lights gives slightly better results than the satellite imagery approach. The features collectively explained an R^2 of 55% for wealth estimation, surpassing the results of daytime satellite imagery alone (54%) and OSM alone (44%), which indicates that the models have better generalisability. Furthermore, the approach showed great potential as an alternative cost-effective approach in supplementing existing tools of governments and policy makers using free and publicly available datasets (in addition to its simplicity). Compared to the approach of using daytime satellite imagery to estimate socio-economic well-being, the merged night-time lights and OSM crowd-sourced information model produced results whose accuracy was higher. It can be used to efficiently measure, monitor, and analyse poverty rates against policies implemented to mitigate poverty effects.

As an extension of other research conducted in relation to South Africa (Pfeifer et al. 2018, p. 20), this study demonstrates a strong correlation between urban

improvements and night-time luminosity. However, the overall results in South Africa were surprisingly lower than the values of other studies in the sub-Saharan region. The metrics revealed that luminosity is relevant, but South Africa is the most developed economy in sub-Saharan Africa with the highest access to electricity (Sarkodie & Adams 2020, p. 1). Despite having very high income inequality in the subregion and the world (Todes & Turok 2018, p. 1), South Africa had 86%, with 85% for rural areas and 87% for urban areas as of 2014 (Sarkodie & Adams 2020, p. 1) which limits the effectiveness of electricity or night-time lights as a poverty proxy. Owing to this, some clusters that could show poverty from daytime satellite imagery would probably be assigned a high luminosity class. At the same time, due to the structurally variegated nature of the South African urban landscape, it is unlikely that luminosity itself can adequately serve as a proxy for poverty. Specifically, this research has determined that the persistent consequences of apartheid in the form of socio-economic exclusion and inequality have created a sort of aggressive divide that positions wealthy, high-income, high-net-worth residents within one or two streets of the high-density slums and impoverished areas where opportunity-seeking urbanites continue to reside. As a result, it was concluded over the course of this study that night-time luminosity would only offer a valuable comparative mechanism in order to assess the affective influence of regional development and improvements on resource availability (e.g., electricity) and usage.

Alternatively, this research confirms previous findings presented by Fatehkia et al. (2020) that when combined with other daytime indicators and texture-based or structural assessments, it becomes possible to efficiently classify regional populations according to service access, resource propagation, and comparative poverty indicators (e.g., socio-economic, demographics). This combinative approach draws upon the full, comparative potential of the machine learning capabilities, the accuracy of the incumbent training set, and the multi-model overlay of daytime and night-time images. Regionalisation of resource classifications, for example, can be used to distinguish between health services and educational resources in targeted urban areas, highlighting the relationship between access and impoverishment over time. While much of the prior research in this field has been presented as a form of proof-of-concept approach to remote sensing, the current study validates prior hypotheses

regarding the capacity to assess regional poverty accurately and consistently through remote sensing capabilities and techniques.

Geospatial variances and regional socio-economic inequalities are widespread in South Africa and present a valuable, comparative opportunity for future research in this field and for additional comparative insights as they relate to the visual-economic relationships manifesting in urban areas. Whilst government efforts to extend the reach of infrastructure and the inclusion of all citizens in healthcare and education have fulfilled a range of developmental objectives in the past two decades, the juxtaposition of wealthy and poor and the extent of this extreme inequality suggest that there is a lack of congruity in the support being provided by these policies and the effects of such support on the regional populations. This study concludes that remote sensing can be used effectively in the future to model and analyse changing patterns of urban development, whilst additional modifications to the training set and various socio-economic indicators will continue to improve the overall accuracy of the CNN and machine learning solutions derived from both daytime and night-time images.

6.6 Recommendations and future research

A central objective of this study was to recommend an alternative to the *in-situ* survey-based poverty assessment in developing nations. Evidence collected during the course of this study has revealed that the proposed approach is a viable high-fidelity solution to several challenges facing domestic policymakers as they seek to identify regional inequalities and systematic dysfunctionality in urban landscapes using machine learning models combined with crowd-sourced and remote sensing data. Based upon these techniques, the following future research is recommended, the extension of which will involve model testing and validation on an international scale.

Recommendation 1) Improved privacy guidelines: Future studies can be done to improve privacy guidelines. Whilst prediction has become increasingly granular and accurate, there is a need to use precisely georeferenced ground truth data to train or validate models without undermining the privacy of household survey participants.

Producing guidelines to navigate these issues will help with the accuracy of model estimation, which is a critical issue.

Recommendation 2) Factor-based degradation analysis applying varimax rotation and statistical assessment of critical poverty indicators to GIS-based visual mapping: Central to these empirical findings was the revelation that some dimensions have greater accuracy than others when predicting poverty. For example, night-time luminosity is likely to have high value effects on the prediction of socio-economic improvement across rural locations when applied to longitudinal analyses over time. Similarly, housing density in urban spaces can be used as a proxy for welfare needs, regional socio-economic equity, or government welfare programme efficiency. Based upon these findings, it is predicted that by conducting a factor-based discrimination analysis at multiple points of influence, a more accurate and rigorous regional model can be developed for the prediction of poverty.

Recommendation 3) Textural assessment of rooftop construction materials versus night-time luminosity effects: Central to the effectiveness of night-time luminosity analysis is the ability to reconcile the reflective characteristics of residential rooftops. This means a critical comparison of daytime housing textural features with night-time luminosity scores to establish test zones.

Recommendation 4) Differentiation between high-density urban poverty accuracy and lower-density, suburban luminosity models: A comparative assessment of the daytime and night-time density models in the current study revealed variations in poverty models that are indicative of distinctions between geospatial characteristics in urban landscapes. Therefore, additional analysis is needed to highlight whether it is possible to apply GIS-based machine learning and factor modelling to the assessment of poverty in urban / suburban high- and low-density spaces.

Recommendation 5) Model interpretability: Lastly, one barrier to entry and adoption of machine learning modelling is a lack of interpretability due to the black-box nature of deep learning models. Therefore, it is difficult for policymakers or officials to understand the behaviour of the models or interpret their results. Model interpretability

is an area of research that is greatly impacting the adoption of AI-assisted decision-making processes to come up with policies and intervention strategies.

6.7 Limitations

This research had some limitations. It was noted that dependence on incomplete and potentially unrepresentative data from OSM could have interpretation limitations. Whilst the derived characteristics can have wealth correlations, it does not follow that they have any causation effects. Therefore, caution should be exercised with interpretation claims unless the coverage and completeness of the datasets are investigated. Second, the data used for the study were acquired from slightly different time periods. The DHS survey dataset was conducted in 2016, while the night-time lights and satellite imagery data were collected in 2016. Unfortunately, it was not possible to download OSM data from 2016, and therefore the latest dataset was downloaded from that year. The differences from acquiring data in different timespans could result in reduced estimation accuracy and derail government initiatives that eliminate poverty as specified in the National Development Plan 2030. The ongoing initiatives could have improved the welfare of households since the last DHS survey or imagery was captured but not captured yet. Lastly, the DHS survey locations were not accurate due to the noise introduced, which could also affect the accuracy of the estimation of the models that were developed.

Despite these limitations, the research findings are valid and essential to advance the accuracy of poverty mapping, which presents great potential for policymakers, governments, and humanitarian organisations to better understand the spatial distribution of poverty and develop evidence-based targeted intervention strategies.

6.8 Chapter summary

This chapter has concluded these findings, which demonstrate the practical efficacy of machine learning and remote sensing in the diagnosis, assessment, and

interpretation of regional poverty indicators. For nations like South Africa, where socio-economic disadvantages are difficult to predict and likely to result in high-cost field research and surveys, the model proposed in this study offers substantial cost and time savings over other more intensive techniques. Based upon these findings, nighttime luminosity and crowd-sourced data were confirmed as a viable research technique for some urban analysis scenarios, particularly when there is a comparable model that can be tracked over time to determine the rate of change or developmental effects. Ultimately, these findings conclude that this technique is best suited to comparative, multi-factor models in which poverty can be tracked and analysed by applying multiple proxy dimensions to a given problem during the assessment. The recommendations for future research will seek to address the possible limitations of this model and will overcome the gaps associated with spatial heterogeneity in high-density urban landscapes.

References

- Alkire, S & Santos, ME 2013, *Measuring Acute Poverty in the Developing World: Robustness and Scope of the Multidimensional Poverty Index*, *SSRN Electronic Journal*, doi: 10.2139/ssrn.2296819
- Anderson, J, Sarkar, D, & Palen, L 2019, 'Corporate Editors in the Evolving Landscape of OpenStreetMap', *ISPRS International Journal of Geo-Information*, vol. 8, no.5, pp. 1–18, doi: 10.3390/ijgi8050232
- Arribas-Bel, D, Patino, JE, & Duque, JC 2017, 'Remote sensing-based measurement of Living Environment Deprivation: Improving classical approaches with machine learning', *PLoS ONE*, vol. 12, no.5, pp. 1–25, doi: 10.1371/journal.pone.0176684
- Ayush, K, UzKent, B, Burke, M, Lobell, D, & Ermon, S 2020, 'Generating Interpretable Poverty Maps using Object Detection in Satellite Images',
- Babenko, B, Hersh, J, Newhouse, D, Ramakrishnan, A, & Swartz, T 2017, 'Poverty mapping using convolutional neural networks trained on high and medium resolution satellite images, with an application in Mexico', *ArXiv*, no.Nips, pp. 1–4
- Barrington-Leigh, C & Millard-Ball, A 2017, 'The world's user-generated road map is more than 80% complete', *PLoS ONE*, vol. 12, no.8, pp. 1–20, doi: 10.1371/journal.pone.0180698
- Bency, AJ, Rallapalli, S, Ganti, RK, Srivatsa, M, & Manjunath, BS 2017, 'Beyond spatial auto-regressive models: Predicting housing prices with satellite imagery', *Proceedings - 2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017*, pp. 320–329, doi: 10.1109/WACV.2017.42
- Bertram-Hümmer, V & Baliki, G 2015, 'The Role of Visible Wealth for Deprivation', *Social Indicators Research*, vol. 124, no.3, pp. 765–783, doi: 10.1007/s11205-014-0824-2
- Bhattacharjee, A 2012, *Social Science Research: Principles, Methods, and Practices*,
- Blumenstock, J, Cadamuro, G, & On, R 2015, 'Predicting poverty and wealth from mobile phone metadata', *Science*, vol. 350, no.6264, pp. 1073–1076, doi: 10.1126/science.aac4420
- Borgen Project 2020, '5 Facts About Poverty in South Africa', , viewed 26 December

- 2020, <https://borgenproject.org/poverty-in-south-africa/>
- Breiman, L & Cutler, A 2002, 'Random forests - classification description', , viewed 17 January 2021, https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm
- Bryman, A 2012, *Social Research Methods*, , 4th edn, Oxford University Press, Oxford, UK
- Burgert, CR, Colston, J, Roy, T, & Zachary, B 2013, *Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys*,
- Carrig, MM & Hoyle, RH 2011, 'Measurement Choices: Reliability, Validity, and Generalizability', in A. T. Panter & S. K. Sterba (eds.), *Handbook of Ethics in Quantitative Methodology*, , pp. 127–157, Routledge, New York, doi: <https://doi-org.oasis.unisa.ac.za/10.4324/9780203840023>
- Chen, X & Nordhaus, WD 2011, 'Using luminosity data as a proxy for economic statistics', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no.21, pp. 8589–8594, doi: [10.1073/pnas.1017031108](https://doi.org/10.1073/pnas.1017031108)
- Deng, J, Dong, W, Socher, R, Li, L-J, Li, K, & Fei-Fei, L 2009, 'ImageNet: A large-scale hierarchical image database', in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, , pp. 248–255, IEEE, doi: [10.1109/CVPRW.2009.5206848](https://doi.org/10.1109/CVPRW.2009.5206848)
- Dresch, A, Lacerda, DP, & Antunes, JAV 2015, 'Design Science Research', in *Design Science Research*, , pp. 67–102, Springer International Publishing, Cham, doi: [10.1007/978-3-319-07374-3_4](https://doi.org/10.1007/978-3-319-07374-3_4)
- Engstrom, R, Hersh, J, & Newhouse, D 2017, *Poverty from Space: Using High-Resolution Satellite Imagery for Estimating Economic Well-Being*No. 8284,
- Fatehkia, M, Tingzon, I, Orden, A, Sy, S, Sekara, V, Garcia-Herranz, M, & Weber, I 2020, 'Mapping socioeconomic indicators using social media advertising data', *EPJ Data Science*, vol. 9, no.1, pp. 1–15, doi: [10.1140/epjds/s13688-020-00235-w](https://doi.org/10.1140/epjds/s13688-020-00235-w)
- Fayyad, U, Piatetsky-Shapiro, G, & Smyth, P 1996, 'The KDD process for extracting useful knowledge from volumes of data', *Communications of the ACM*, vol. 39, no.11, pp. 27–34, doi: [10.1145/240455.240464](https://doi.org/10.1145/240455.240464)
- GADM 2020, 'GADM', , viewed 30 November 2020,

- https://gadm.org/download_country_v3.html
- Geburu, T, Krause, J, Wang, Y, Chen, D, Deng, J, Aiden, EL, & Fei-Fei, L 2017, 'Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the United States', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no.50, pp. 13108–13113, doi: 10.1073/pnas.1700035114
- Geofabrik GmbH 2020, 'Geofabrik', , viewed 30 November 2020, <https://download.geofabrik.de/africa/south-africa.html>
- Gervasoni, L, Fenet, S, Perrier, R, & Sturm, P 2019, 'Convolutional neural networks for disaggregated population mapping using open data', *Proceedings - 2018 IEEE 5th International Conference on Data Science and Advanced Analytics, DSAA 2018*, pp. 594–603, doi: 10.1109/DSAA.2018.00076
- Ghosh, T, Anderson, SJ, Elvidge, CD, & Sutton, PC 2013, 'Using nighttime satellite imagery as a proxy measure of human well-being', *Sustainability*, vol. 5, no.12, pp. 4988–5019, doi: 10.3390/su5124988
- Goodfellow, I, Bengio, Y, & Courville, A 2016, *Deep Learning*, , MIT Press
- Google 2020, 'Cloud AI', , AI and Machine Learning Products, viewed 12 April 2020, <https://cloud.google.com/products/ai>
- Gregor, S 2002, 'Design Theory in Information Systems', *Australasian Journal of Information Systems*, vol. 10, no.1, pp. 14–22, doi: 10.3127/ajis.v10i1.439
- Head, A, Manguin, M, Tran, N, & Blumenstock, JE 2017, 'Can Human Development be Measured with Satellite Imagery?', in *Proceedings of the Ninth International Conference on Information and Communication Technologies and Development*, , pp. 1–11, ACM, New York, NY, USA, doi: 10.1145/3136560.3136576
- Henderson, JV, Storeygard, A, & Weil, DN 2012, 'Measuring economic growth from outer space', *American Economic Review*, vol. 102, no.2, pp. 994–1028, doi: 10.1257/aer.102.2.994
- Hevner, AR 2007, 'A Three Cycle View of Design Science Research', *Scandinavian Journal of Information Systems*, vol. 19, no.2, pp. 87–92
- Hevner, AR, March, ST, Park, J, & Ram, S 2004, 'Design Science in Information Systems Research', *MIS Quarterly*, vol. 28, no.1, pp. 75–105, doi: 10.2307/25148625
- Hillson, R, Alexandre, JD, Jacobsen, KH, Ansumana, R, Bockarie, AS, Bangura, U,

- Lamin, JM, Malanoski, AP, & Stenger, DA 2014, 'Methods for determining the uncertainty of population estimates derived from satellite imagery and limited survey data: A case study of Bo City, Sierra Leone', *PLoS ONE*, vol. 9, no.11, pp. 1–14, doi: 10.1371/journal.pone.0112241
- Holloway, J & Mengersen, K 2018, 'Statistical Machine Learning Methods and Remote Sensing for Sustainable Development Goals: A Review', *Remote Sensing*, vol. 10, no.9, pp. 1–21, doi: 10.3390/rs10091365
- Hu, T, Yang, J, Li, X, & Gong, P 2016, 'Mapping urban land use by using landsat images and open social data', *Remote Sensing*, vol. 8, no.2, pp. 1–18, doi: 10.3390/rs8020151
- Humanitarian OpenStreetMap Team 2020, 'Humanitarian OpenStreetMap Team', , viewed 30 November 2020, <https://www.hotosm.org/>
- Ilic, L, Sawada, M, & Zarzelli, A 2019, 'Deep mapping gentrification in a large Canadian city using deep learning and Google Street View', *PLoS ONE*, vol. 14, no.3, pp. 1–21, doi: 10.1371/journal.pone.0212814
- James, G, Witten, D, Hastie, T, & Tibshirani, R 2013, *An introduction to statistical learning*, , Springer, New York
- Jean, N, Burke, M, Xie, M, Davis, WM, Lobell, DB, & Ermon, S 2016, 'Combining satellite imagery and machine learning to predict poverty', *Science*, vol. 353, no.6301, pp. 790–794, doi: 10.1126/science.aaf7894
- Jerven, M 2014, *Benefits and costs of the data for development targets for the post-2015 development agenda*Data for Development Assessment Paper No. 16,
- Jonker, J & Pennink, B 2009, *The Essence of Research Methodology*, , Springer Berlin Heidelberg, Berlin, Heidelberg, doi: 10.1007/978-3-540-71659-4
- Karpathy, A 2020, 'Convolutional Neural Networks for Visual Recognition (Course Notes)', , viewed 1 February 2021, <https://cs231n.github.io>
- Kodratoff, Y 1999, 'Comparing Machine Learning and Knowledge Discovery in DataBases: An Application to Knowledge Discovery in Texts', in G. Paliouras, V. Karkaletsis, & C. D. Spyropoulos (eds.), *Machine Learning and Its Applications*, , pp. 1–21, Springer, Berlin, Heidelberg, doi: 10.1007/3-540-44673-7_1
- Kotsiantis, SB 2007, 'Supervised Machine Learning: A Review of Classification Techniques', in P. Kotsiantis, S.B., Zaharakis, I. and Pintelas (ed.), *Emerging artificial intelligence applications in computer engineering : real word AI systems with applications in eHealth, HCI, information retrieval and pervasive*

- technologies*, , pp. 3–24
- Kovalerchuk, B 2018, *Visual Knowledge Discovery and Machine Learning*, , vol. 144, Springer International Publishing, Cham, doi: 10.1007/978-3-319-73040-0
- Krizhevsky, A, Sutskever, I, & Hinton, GE 2017, 'ImageNet classification with deep convolutional neural networks', *Communications of the ACM*, vol. 60, no.6, pp. 84–90, doi: 10.1145/3065386
- Kuhn, TS 1970, *The structure of scientific revolutions*, , 2nd edn, University of Chicago Press
- Lam, D, Kuzma, R, McGee, K, Dooley, S, Laielli, M, Klaric, M, Bulatov, Y, & McCord, B 2018, 'xView: Objects in Context in Overhead Imagery', , pp. 1–16
- Law, S, Paige, B, & Russell, C 2019, 'Take a Look Around: Using Street View and Satellite Images to Estimate House Prices', *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no.5, pp. 1–19, doi: 10.1145/3342240
- Lillesand, T, Kiefer, RW, & Chipman, J 2015, *Remote sensing and image interpretation*, , John Wiley & Sons
- Long, Y, Gong, Y, Xiao, Z, & Liu, Q 2017, 'Accurate Object Localization in Remote Sensing Images Based on Convolutional Neural Networks', *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no.5, pp. 2486–2498, doi: 10.1109/TGRS.2016.2645610
- Ma, R, Wang, W, Zhang, F, Shim, K, & Ratti, C 2019, 'Typeface Reveals Spatial Economical Patterns', *Scientific Reports*, vol. 9, no.1, pp. 1–9, doi: 10.1038/s41598-019-52423-y
- Mahabir, R, Croitoru, A, Crooks, A, Agouris, P, & Stefanidis, A 2018, 'News coverage, digital activism, and geographical saliency: A case study of refugee camps and volunteered geographical information', *PLoS ONE*, vol. 13, no.11, pp. 1–28, doi: 10.1371/journal.pone.0206825
- Mellander, C, Lobo, J, Stolarick, K, & Matheson, Z 2015, 'Night-time light data: A good proxy measure for economic activity?', *PLoS ONE*, vol. 10, no.10, pp. 1–18, doi: 10.1371/journal.pone.0139779
- Michalopoulos, S & Papaioannou, E 2014, 'National Institutions and Subnational Development in Africa *', *The Quarterly Journal of Economics*, vol. 129, no.1, pp. 151–213, doi: 10.1093/qje/qjt029
- Mveyange, A 2016, 'Night Lights and Regional Income Inequality in Africa', , viewed 26 December 2020, <https://www.worldbank.org/content/dam/Worldbank/Feature>

Story/Africa/afr-anthony-mveyange.pdf

Nguyen, QC, Khanna, S, Dwivedi, P, Huang, D, Huang, Y, Tasdizen, T, Brunisholz, KD, Li, F, Gorman, W, Nguyen, TT, & Jiang, C 2019, 'Using Google Street View to examine associations between built environment characteristics and U.S. health outcomes', *Preventive Medicine Reports*, vol. 14, no.100859, pp. 1–11, doi: 10.1016/j.pmedr.2019.100859

NOAA/NCEI 2020, 'Earth Observation Group - Defense Meteorological Satellite Program, Boulder', , viewed 30 November 2020, https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html

NOAA 2020, 'VIIRS Day/Night Band (DNB) Nighttime Imagery', , viewed 26 December 2020, https://maps.ngdc.noaa.gov/viewers/VIIRS_DNB_nighttime_imagery/index.html

Oates, BJ 2006, *Researching Information Systems and Computing*, , Sage Publications Ltd.

OpenStreetMap 2020, 'OpenStreetMap Wiki', , viewed 27 November 2020, https://wiki.openstreetmap.org/wiki/Main_Page

OSM 2020, 'OSM Completeness', , viewed 30 December 2020, <https://wiki.openstreetmap.org/wiki/Completeness>

Oxford Poverty and Human Development Initiative 2015, 'Alkire-Foster Method', , viewed 24 November 2020, <https://ophi.org.uk/research/multidimensional-poverty/alkire-foster-method/>

Pan, J & Hu, Y 2018, 'Spatial Identification of Multi-dimensional Poverty in Rural China: A Perspective of Nighttime-Light Remote Sensing Data', *Journal of the Indian Society of Remote Sensing*, vol. 46, no.7, pp. 1093–1111, doi: 10.1007/s12524-018-0772-4

Pan, SJ & Yang, Q 2010, 'A Survey on Transfer Learning', *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no.10, pp. 1345–1359, doi: 10.1109/TKDE.2009.191

Peppers, K, Tuunanen, T, Rothenberger, MA, & Chatterjee, S 2007, 'A design science research methodology for information systems research', *Journal of Management Information Systems*, vol. 24, no.3, pp. 45–77, doi: 10.2753/MIS0742-1222240302

Pfeifer, G, Wahl, F, & Marczak, M 2018, 'Illuminating the World Cup effect: Night lights evidence from South Africa', *Journal of Regional Science*, vol. 58, no.5,

- pp. 887–920, doi: 10.1111/jors.12410
- Pinkovskiy, M & Sala-i-Martin, X 2014, *Lights, Camera,... Income!: Estimating Poverty Using National Accounts, Survey Means, and Lights*NBER Working Paper Series No. 19831, , Cambridge, MA, doi:/10.3386/w19831
- Punch, KF 2013, *Introduction to social research: Quantitative and qualitative approaches*, , 3rd edn, Sage Publications, Los Angeles, CA
- Sahn, DE & Stifel, D 2003, 'Exploring Alternative Measures of Welfare in the Absence of Expenditure Data', *Review of Income and Wealth*, vol. 49, no.4, pp. 463–489, doi: 10.1111/j.0034-6586.2003.00100.x
- Sarkodie, SA & Adams, S 2020, 'Electricity access and income inequality in South Africa: Evidence from Bayesian and NARDL analyses', *Energy Strategy Reviews*, vol. 29, no.100480, pp. 1–11, doi: 10.1016/j.esr.2020.100480
- Saunders, MNK, Lewis, P, & Thornhill, A 2015, *Research Methods for Business Students*, , Pearson, Essex, UK
- Sen, A 1993, 'Capability and Well-being', in M. Nussbaum & A. Sen (eds.), *The quality of life*, , pp. 30–53, Oxford University Press, doi: 10.1093/0198287976.001.0001
- Shin, H-C, Orton, M, Collins, DJ, Doran, S, & Leach, MO 2016, *Organ Detection Using Deep Learning, Medical Image Recognition, Segmentation and Parsing*, Elsevier Inc., doi: 10.1016/b978-0-12-802581-9.00007-x
- Simon, HA 1996, *The sciences of the artificial*, , 3rd edn, MIT Press, Cambridge, MA
- Simonyan, K, Vedaldi, A, & Zisserman, A 2014, 'Deep inside convolutional networks: Visualising image classification models and saliency maps', in *2nd International Conference on Learning Representations, ICLR 2014 - Workshop Track Proceedings*, , pp. 1–8
- Simonyan, K & Zisserman, A 2015, 'Very deep convolutional networks for large-scale image recognition', *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–14
- Singh, K 2007, *Quantitative Social Research Methods*, , SAGE Publications, Los Angeles, CA
- South Africa Gateway 2019, 'Mapping poverty in South Africa', , viewed 18 December 2020, <https://southafrica-info.com/people/mapping-poverty-in-south-africa/>
- Statista 2018, 'Urbanization in South Africa', , Statista, viewed 18 April 2020,

<https://www.statista.com/statistics/455931/urbanization-in-south-africa/>
 Statistics South Africa 2019a, 'Five facts about poverty in South Africa', , Statistics South Africa, viewed 18 April 2020, <http://www.statssa.gov.za/?p=12075>
 Statistics South Africa 2019b, *Inequality Trends in South Africa*,
 Statistics South Africa 2020, 'Census', , Census South Africa, viewed 13 April 2020, http://www.statssa.gov.za/?page_id=3836
 Suel, E, Polak, JW, Bennett, JE, & Ezzati, M 2019, 'Measuring social, environmental and health inequalities using deep learning and street imagery', *Scientific Reports*, vol. 9, no.6229, pp. 1–10, doi: 10.1038/s41598-019-42036-w
 Tan, C, Sun, F, Kong, T, Zhang, W, Yang, C, & Liu, C 2018, 'A Survey on Deep Transfer Learning', in V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, & I. Maglogiannis (eds.), *Artificial Neural Networks and Machine Learning – ICANN 2018*, , pp. 270–279, Springer, Cham, doi: 10.1007/978-3-030-01424-7_27
 The Guardian 2016, 'Divided cities: South Africa's apartheid legacy photographed by drone', , The Guardian, viewed 18 April 2020, <https://www.theguardian.com/cities/gallery/2016/jun/23/south-africa-divided-cities-apartheid-photographed-drone>
 Tingzon, I, Orden, A, Go, KT, Sy, S, Sekara, V, Weber, I, Fatehkia, M, García-Herranz, M, & Kim, D 2019, 'Mapping Poverty in the Philippines Using Machine Learning, Satellite Imagery, and Crowd-sourced Geospatial Information', *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-4/W19, pp. 425–431, doi: 10.5194/isprs-archives-XLII-4-W19-425-2019
 Todes, A & Turok, I 2018, 'Spatial inequalities and policies in South Africa: Place-based or people-centred?', *Progress in Planning*, vol. 123, pp. 1–31, doi: 10.1016/j.progress.2017.03.001
 ul Hassan, M 2018, 'Convolutional Network for Classification and Detection', , viewed 13 December 2020, <https://neurohive.io/en/popular-networks/vgg16/>
 USAID 2020, 'The DHS Program', , viewed 30 November 2020, <https://dhsprogram.com/Who-We-Are/About-Us.cfm>
 Wagle, U 2008, *Multidimensional Poverty Measurement*, , vol. 4, Springer US, New York, NY, NY, doi: 10.1007/978-0-387-75875-6
 World Bank 2019, 'South Africa Overview', , viewed 26 December 2020, <https://www.worldbank.org/en/country/southafrica/overview>

- Worstell, T 2017, 'South Africa's Poverty Problem Won't Be Cured By Redistribution - Economic Growth Might Do It', , Forbes, viewed 18 April 2020, <https://www.forbes.com/sites/timworstell/2017/08/25/south-africas-poverty-problem-wont-be-cured-by-redistribution-economic-growth-might-do-it/#1d3771da2aa3>
- Xie, M, Jean, N, Burke, M, Lobell, D, & Ermon, S 2016, 'Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping', in *Thirtieth AAAI Conference on Artificial Intelligence*, , pp. 3929–3935, AAAI Publications
- Yamashita, R, Nishio, M, Do, RKG, & Togashi, K 2018, 'Convolutional neural networks: an overview and application in radiology', *Insights into Imaging*, vol. 9, no.4, pp. 611–629, doi: 10.1007/s13244-018-0639-9
- Yan, Y, Feng, CC, Huang, W, Fan, H, Wang, YC, & Zipf, A 2020, 'Volunteered geographic information research in the first decade: a narrative review of selected journal articles in GIScience', *International Journal of Geographical Information Science*, vol. 34, no.9, pp. 1765–1791, doi: 10.1080/13658816.2020.1730848
- Yang, Q, Zhang, Y, Dai, W, & Pan, SJ 2020, *Transfer Learning*, , Cambridge University Press, doi: 10.1017/9781139061773
- Ye, T, Zhao, N, Yang, X, Ouyang, Z, Liu, X, Chen, Q, Hu, K, Yue, W, Qi, J, Li, Z, & Jia, P 2019, 'Improved population mapping for China using remotely sensed and points-of-interest data within a random forests model', *Science of the Total Environment*, vol. 658, pp. 936–946, doi: 10.1016/j.scitotenv.2018.12.276
- Yu, X, Wu, X, Luo, C, & Ren, P 2017, 'Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework', *GIScience & Remote Sensing*, vol. 54, no.5, pp. 741–758, doi: 10.1080/15481603.2017.1323377
- Yue, P, Baumann, P, Bugbee, K, & Jiang, L 2015, 'Towards intelligent GIServices', *Earth Science Informatics*, vol. 8, no.3, pp. 463–481, doi: 10.1007/s12145-015-0229-z
- Zeiler, MD & Fergus, R 2014, 'Visualizing and Understanding Convolutional Networks', in *European conference on computer vision*, vol. 12, , pp. 818–833, doi: 10.1007/978-3-319-10590-1_53
- Zhang, X-D 2020, 'Machine Learning', in *A Matrix Algebra Approach to Artificial Intelligence*, , pp. 223–440, Springer Singapore, Singapore, doi: 10.1007/978-

981-15-2770-8_6

Zhao, L & Kusumaputri, P 2016, 'OpenStreetMap Road Network Analysis for Poverty Mapping',

Zhao, X, Yu, B, Liu, Y, Chen, Z, Li, Q, Wang, C, & Wu, J 2019, 'Estimation of poverty using random forest regression with multi-source data: A case study in Bangladesh', *Remote Sensing*, vol. 11, no.4, pp. 1–18, doi: 10.3390/rs11040375

Appendices

Appendix A: Ethical Clearance Certificate



UNISA COLLEGE OF SCIENCE, ENGINEERING AND TECHNOLOGY'S (CSET) ETHICS REVIEW COMMITTEE

4 September 2020

Dear Mr Musungwa

ERC Reference #: 2020/CSET/SOC/025

Name: Mr Tapera Musungwa

Student #: 44814526

**Decision: Ethics Approval from
4 September 2020 to
3 September 2023
(Humans involved)**

Researcher: Mr Tapera Musungwa
44814526@mylife.unisa.ac.za, +966 13 880 4662, +966 54 492 3963

Supervisor: Prof. Ernest Mnkandla
mnkane@unisa.a.c.za, 011 670 9059

Working title of research:

Exploring an ontology-guided semantic data integration framework to support the measuring of multidimensional urban poverty

Qualification: MTech

Thank you for the application for research ethics clearance by the Unisa College of Science, Engineering and Technology's (CSET) Ethics Review Committee for the above mentioned research. Ethics approval is granted for 3 years.

*The **low risk application** was expedited by the College of Science, Engineering and Technology's (CSET) Ethics Review Committee on 4 September 2020 in compliance with the Unisa Policy on Research Ethics and the Standard Operating Procedure on Research Ethics Risk Assessment. The decision will be tabled at the next Committee meeting for ratification.*

The proposed research may now commence with the provisions that:

1. The researcher will ensure that the research project adheres to the relevant guidelines set out in the Unisa COVID-19 position statement on research ethics

attached.

2. The researcher(s) will ensure that the research project adheres to the values and principles expressed in the UNISA Policy on Research Ethics.
3. Any adverse circumstance arising in the undertaking of the research project that is relevant to the ethicality of the study should be communicated in writing to the College of Science, Engineering and Technology's (CSET) Ethics Review Committee.
4. The researcher(s) will conduct the study according to the methods and procedures set out in the approved application.
5. Any changes that can affect the study-related risks for the research participants, particularly in terms of assurances made with regards to the protection of participants' privacy and the confidentiality of the data, should be reported to the Committee in writing, accompanied by a progress report.
6. The researcher will ensure that the research project adheres to any applicable national legislation, professional codes of conduct, institutional guidelines and scientific standards relevant to the specific field of study. Adherence to the following South African legislation is important, if applicable: Protection of Personal Information Act, no 4 of 2013; Children's act no 38 of 2005 and the National Health Act, no 61 of 2003.
7. Only de-identified research data may be used for secondary research purposes in future on condition that the research objectives are similar to those of the original research. Secondary use of identifiable human research data require additional ethics clearance.
8. No field work activities may continue after the expiry date 3 September 2023. Submission of a completed research ethics progress report will constitute an application for renewal of Ethics Research Committee approval.

Note

The reference number 2020/CSET/SOC/025 should be clearly indicated on all forms of communication with the intended research participants, as well as with the Committee.

Yours sincerely,



Mr C Pilkington
Chair of School of Computing Ethics Review Subcommittee
College of Science, Engineering and Technology (CSET)
E-mail: pilkiel@unisa.ac.za
Tel: (011) 471-2130



Prof. SJ Johnston
Head: Research and Graduate Studies
College of Science Engineering and
Technology (CSET)
E-mail: johnssj@unisa.ac.za
Tel: 011 670 9146



Prof. B Mamba
Executive Dean
College of Science Engineering and
Technology (CSET)
E-mail: mambabb@unisa.ac.za
Tel: 011 670 9230

Editorial Certificate

This certificate is to affirm that Editing Press Inc., comprising faculty and postgraduates from the Universities of Oxford and Cambridge, has edited, to the best of its members' abilities, the work entitled

A Machine Learning-guided Data Integration Framework to Measure Multidimensional Poverty

by

Tapera Musungwa

This certificate is issued without prejudice to the author on

23 February 2021.



Charles Anderson
Director of Academic Editing
Editing Press Inc.
www.editing.press

Appendix C: Turnitin Report

Revised Examiner Corrections

ORIGINALITY REPORT

16% SIMILARITY INDEX	14% INTERNET SOURCES	11% PUBLICATIONS	8% STUDENT PAPERS
--------------------------------	--------------------------------	----------------------------	-----------------------------

PRIMARY SOURCES

1	uir.unisa.ac.za Internet Source	1%
2	www.adb.org Internet Source	1%
3	www.int-arch-photogramm-remote-sens-spatial-inf-sci.net Internet Source	<1%
4	science.sciencemag.org Internet Source	<1%
5	doctorpenguin.com Internet Source	<1%
6	hdl.handle.net Internet Source	<1%
7	www.tandfonline.com Internet Source	<1%
8	link.springer.com Internet Source	<1%
9	arxiv.org Internet Source	<1%