FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

POVERTY PREDICTION BY USING DEEP LEARNING ON SATELLITE IMAGES

Sabeer SAEED

Master's Thesis

DEPARTMENT OF SOFTWARE ENGINEERING

JUNE 2022

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Software Engineering

Master's Thesis

POVERTY PREDICTION BY USING DEEP LEARNING ON SATELLITE IMAGES

Author
Sabeer SAEED

Supervisor
Prof. Dr. İbrahim TÜRKOĞLU

JUNE 2022
ELAZIG

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Software Engineering

Master's Thesis

Title:	Poverty Prediction By Using Deep Learning on Satellite Images
Author:	Sabeer SAEED
Submission Date:	20 May 2022
Defense Date:	27 June 2022

THESIS APPROVAL

This thesis, which was prepared according to the thesis writing rules of the Graduate School of Natural and Applied Sciences, Fırat University, was evaluated by the committee members who have signed the following signatures and was unanimously approved after the defense exam made open to the academic audience.

		Signature
Supervisor:	Prof. Dr. İbrahim TÜRKOĞLU Firat University, Faculty of Technology	Approved
Chair:	Prof. Dr. Muhammed Fatih TALU İnönü University, Faculty of Engineering	Approved
Member:	Assist. Prof. Dr. Ferhat UCAR Fırat University, Faculty of Technology	Approved

This thesis was approved by the Administrative Board of the Graduate School on

.... / .... / 20 ....

Signature

Prof. Dr. Kürşat Esat ALYAMAÇ
Director of the Graduate School

DECLARATION

I hereby declare that I wrote this Master's Thesis titled "Poverty Prediction By Using Deep Learning on Satellite Images" in consistent with the thesis writing guide of the Graduate School of Natural and Applied Sciences, Firat University. I also declare that all information in it is correct, that I acted according to scientific ethics in producing and presenting the findings, cited all the references I used, express all institutions or organizations or persons who supported the thesis financially. I have never used the data and information I provide here in order to get a degree in any way.

27 June 2022

Sabeer SAEED

PREFACE

Predicting poverty gives a pathway for government, non-governmental and other institutions to make economic policy decisions. This thesis explores the emergence of deep learning helps this poverty prediction easily using the satellite images, making the poverty estimation quickly, comfortably, and less expensive compared to the traditional house-to-house surveys. As remote sensing keeps on getting more attention in terms of investment & innovation and deep learning attracts more in research and development funding, we are looking forward to seeing many on poverty prediction using deep learning on satellite images.

If the work could only predict poverty using three deep learning models, such as CNN, VGG16, and ResNet50, covering only four African countries using satellite images from the Kaggle website. The challenge was running the algorithm as it could not afford high computing resources. It had to run many times as google collab could not allow the dataset's images to be executed. But my friend Khalid Sani helped me with computer resources.

What motivated the study was the fact that a study like this makes an impact on the development of a country in making decisions on economic policies. I got the topic idea from my Supervisor, Prof. Dr. Ibrahim Turkoglu, one of the most outstanding scholars I have ever met. I want to use this medium to thank my supervisor, friends, families, government (NITDA), and Firat University Elazig.

Sabeer SAEED
ELAZIG, 2022

TABLE OF CONTENTS

ABSTRACT

Poverty Prediction By Using Deep Learning on Satellite Images

Sabeer SAEED

Master's Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences

Department of Software Engineering

June 2022, Page: x + 35

Abstract- As the universe finds it challenging to define poverty, the World Bank consider poverty as anyone living below $2 per day. Governments and international organizations are working to eradicate this poverty. In this study, some research on satellite imagery on poverty prediction through the concept of CNN (Convolutional Neural Network) is reviewed. The study considered satellite images of Nigeria, Mali, Malawi and Ethiopia from Kaggle. While 90% of the dataset was used for training, the remaining 10% was used for testing. Datasets were analyzed using CNN, VGG16 and ResNet50 and it was observed that the VGG16 model outperformed the other two models with a validation accuracy of 94%. CNN had the second with 91% and ResNet had the lowest validation success with 62%. The rise of high-resolution satellite imagery containing comprehensive data of patterns, features and landscapes of regions or countries can be applied to determine the economic lifestyle of people or nations. Applying satellite imagery to poverty estimation would be easier, faster and cheaper. This study suggests that large satellite images should be available for each region or country. In future studies, researchers should focus on satellite imagery to apply it to the prediction and detection of poverty and crime, road traffic, agricultural land and similar practices.

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, and Convolutional Neural Network.

ÖZET

Uydu Görüntülerinde Derin Öğrenme Kullanarak Yoksulluk Tahmini

Sabeer SAEED

Yüksek Lisans Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Yazılım Mühendisliği Anabilim Dalı

Haziran 2022, Sayfa: x + 35

Dünya yoksulluğu tanımlamakta zorlanmaktadır. Dünya Bankası yoksulluğu günde 2 doların altında yaşayan herkes olarak görmektedir. Hükümetler ve uluslararası örgütler bu yoksulluğu ortadan kaldırmak için çalışmaktadırlar. Bu çalışmada, CNN kavramı aracılığıyla yoksulluğun tahmini üzerine uydu görüntüleri üzerine yapılan bazı araştırmalar gözden geçirilmiştir. Yapılan çalışma Kaggle'dan Nijerya, Mali, Malawi ve Etiyopya'ya ait uydu görüntülerini veri setleri olarak ele almıştır. Bu veri setinin %90 oranındaki kısmı eğitim, %10 oranındaki kısmı test verisi olarak kullanılmıştır. Veri setleri CNN, VGG16 ve ResNet50 kullanılarak analiz edilmiş ve VGG16 modelinin %94 doğrulama başarım değeriyle diğer iki modelden daha iyi performans gösterdiği gözlemlenmiştir. CNN %91 ile ikinci, ResNet'in ise %62 oranıyla en düşük doğrulama başarısına sahip olduğu görülmüştür. Bölgelerin veya ülkelerin desenleri, özellikleri ve manzaralarının kapsamlı verilerini içeren yüksek çözünürlüklü uydu görüntülerinin yükselişi, insanların veya ulusların ekonomik yaşam biçiminin belirlenmesi için uygulanabilmektedir. Uydu görüntülerinin yoksulluk tahminine uygulanması daha fazla tahmine göre daha kolay, daha hızlı ve daha ucuz olmaktadır. Bu çalışma, her bölge veya ülke için büyük uydu görüntülerinin bulunması gerektiğini önermektedir. Gelecek çalışmalarda araştırmacılar yoksulluğu ve suç, yol trafiği, tarım alanları ve benzeri uygulamaların tahmin ve tespit edilmesinde uygulamak için uydu görüntülerine odaklanmalıdırlar.

Anahtar Kelimeler: Yapay Zeka, Makine Öğrenimi, Derin Öğrenme ve Evrişimli Sinir Ağı.

LIST OF FIGURES

Page

LIST OF TABLES

Page

SYMBOLS AND ABBREVIATIONS

Abbreviations

AI	: Artificial Intelligence
ANN	: Artificial Neural Network
CNN	: Convolutional Neural Network
DL	: Deep Learning
ML	: Machine Learning
ResNet50	: Residual Network 50
VGG16	: Visual Graphic/Geometry Group 16

1. INTRODUCTION

The term Artificial intelligence (AI) was initiated around the 1950s when a 'fistful' of pioneers from the emerging sector of computer science began questioning whether computer machines could be developed to 'think,' an investigation field whose implication is nevertheless explored nowadays. A succinct description of the practice can be as: *an effort to automatically replace intellectual functions frequently performed by Mankind*. As such, AI is a global area consisting of machine learning and deep learning, but that also consists of different methods that do not envelop any kind of learning.

Early chess game applications, for example, solely surrounded hardcoded instructions commands designed by programmers and could not regard as machine learning. For an adequately lengthy time, many scholars and practitioners faithfully accepted that mankind-level artificial intelligence could be ascertained by possessing programmers coding an adequate set of huge explicit command instructions for influencing knowledge. This type of method is perceived as *symbolic AI*, and it was the popular pattern model in Artificial Intelligence from the 1950s to the year near-ending of 1980s. It attained its top or apex popularity during the emergence of the *expert systems,* which boomed around the year the 1980s [1].

Computer Scientists are attempting to adopt Artificial Intelligence (AI) to make machines acquire knowledge as Mankind's brain. In Computer Science, AI refers to the study of 'intelligent agents': any device that understands its ecosystem and takes decisions that improve its likeliness of profitably accomplishing its objectives. Unacademically, the subject "artificial intelligence" is used when a device can do tasks that are Mankind equivalent to other Mankind's minds, like 'learning' or 'understanding' and 'problem-solving.' Learning is a refreshing feature of machines. Therefore, machine learning is a subarea of AI. Scientists have been taking steps since the 1950s in machine learning.

The last few decades show terrific endeavors are made in machine learning enhancements that result in more anticipations from machines. Deep learning is an action toward this dimension, a subfield of machine learning [2]. In general, Artificial Intelligence, Machine Learning, and Deep Learning are related as follows: "Deep-Learning refers as a subtype of Machine-Learning, and Machine-Learning refers as a subtype of Artificial Intelligence." In what manner is that? Is it easy, isn't it? This segmentation can or not be as possible as the natural laws, but it is globally admitted [3]. Machine Learning was one of the advancements of AI, which increased AI with the ability to extract features from raw data and understand and learn from experience.

Satellite Images are pictures snapped using satellite technology, satellite technology is a concept designed to be mounted in orbit between celestial bodies, purposely to relay data such as satellite images to earth. They are used for disaster management, estimation, agricultural survey, etc. Through the innovation of deep learning, we have the means to predict poverty using satellite images. The traditional way of surveying to predict poverty involves going to house-to-house, household-to-household to get the necessary information to determine the poverty.

The latest trend of satellite images allows us to potentially extract socio-economic details in a fast and relevantly cheap way. Satellite pictures hold underbuilding and agricultural details that, in a prior investigation concentrated on countries like Afghanistan and Botswana, were displayed to be purposeful for determining geo-regional socio-economic data. Researchers [6] studied to provide a survey reaction concern on wealth and crime were determined with supervised machine learning structural designs via their image patterns features obtained from their space reference coinciding satellite pictures [6].

As the technique of spatial resolution of satellite images keeps on trending, the abstraction of poverty between urban regions has constantly been incrementing in the previous decade with a significant concentrate on the abstracting of the existence/nonexistence of slums as cited in [7]. As a module of their research framework [7], a remote sensing structural design has been initiated that has enabled to recognize what a slum views like in images.

While this concept enables us to reason what a slum might view like within satellite images, basic visually translation slum against non-slum regions in qualitative spatial resolution satellite images can be challenging because of varying local explanations of slums as well as varying morphologies of slums in several settlings [7].

The idea of digital picture processing also enables satellite image information to be significantly applied for several applications. Supervision of natural catastrophic events by locating floods, wildfires, earthquakes, and fault lines using the satellite images and remote sensing data approach have been applied as essential tools to evaluate the destruction related to them. Satellite images initiate a main technique in designing and strategizing for real estate. Satellite images can also provide explanative data on the ground and analysis situs [8].

Satellite image data categorically night captured image data can be applied for the purpose to develop a fully spatial regional people activities information like size of road and distribution of building data as they have a great interrelation with satellite night light data. With the help of night light satellite image, prediction measures of mankind socio-economic affairs which is an important aspect is achievable. Adjusting the prediction of people's income of regions with low status and unreasonable survey information can be achieved through this technique. Adopting the conventional statistical models with satellite images data has led to overshadowing this issue [8].

1.1. Purpose of Thesis

The universe finds it difficult actually to define what poverty means. Some view it as a lack of essential things to make life comfortable, some percept it as hunger and lack of livelihood. Others see it as any group of people who lacks the basic amenities of life to live their lives, such as no access to school, shelter, food, healthcare services, and water resources, etc. The World Bank refers to poverty as any person living below $2 per day is regarded as a person living below poverty.

The African continent is categorized as the continent with the highest rate of people living below poverty. One of the features that cause poverty is the deficiency in good education. It elevates people to earn income, thus overflowing poverty. These are the reasons many countries budget large shares to go to education. Deficiency of jobs security or job growth is another prime cause of poverty because jobs security elevates people to access good incomes [9].

Conflicts and warfare are other big reason that causes poverty. When a country is at war, the war interrupts productivity and affects the nation's Gross Domestic Product, which demolishes that country's economy. Families' livelihood will be extremely tough and leading to displacement from one place to another in search of survival.

Also, whether or climate change can affect the livelihood of people; changes in weather cause floods (and erosion), drought (below-average rainfall), and severe storms. These natural disasters or emergency phenomena source the fall of prosperous countries down and emerging countries down further, the regaining over becoming extremely hard. Drought leads to inadequate or lack of farming to produce crops that serve as food to both people and animals, which results in extreme hunger. Too much rain leads to erosion which washes away farm crop products and floods that destroy lives, houses, and properties.

Government, non-governmental organization (NGO), an international organization needs to know the status and condition of people, especially people living in poverty, to make policies and predictions. Accumulating the exact information, which involves physically going and conducting an extensive survey and analysis, is quite challenging and expensive. Having the technique of spatial distribution of development in developing countries is significant but expensive [9] due to the nature of the countries but not costly compared to the survey.

The challenges of traditional surveys sometimes aftermath to the unavailability of rate and certify data on poverty, an innovative technique has been suggested that uses the high and medium resolution satellite imagery data of different geo-regions to provide us estimate figures of poverty in different geo-regions of the country. Satellite imagery data has been applied in diverse instances to solve the problems of development and survey data in nations with less-developed and developing economies to compute an estimate of poverty at a huge ratio or scale [9].

Aim and Objectives

- ✓ The prediction of poverty in Africa using satellite images is very significant when dealing with socio-economics to understand the level of people as evidence to apply for government social programs.

- ✓ The technique can serve as a means of estimating poverty using satellite images from countries in Africa as data for development.
- ✓ Hence, the result to be ascertained in this research study will provide increased algorithm results that can be credible and reliable for predicting the region's poverty and livelihood using satellite images.
- ✓ The study will contribute to the algorithm's performance in the segmentation and prediction of poverty and livelihood in Africa will be found to be an increased algorithm compared to the existing tradition of the art of survey used for the same issues.

The availability of and accessibility to satellite imagery data proposes an exemplary quicker, less expensive, and riskless approach for collecting socio-economic information compared to the typical survey collection approach [10]. But currently, the technique artificial neural network (ANN) of machine learning combined with satellite imagery data have been the latest trends in quantifying economic livelihood status. This technology was due to the emergence of different algorithms models used to recreate from the machine learning communities.

The machine learning researchers use a Convolutional Neural Network (CNN) to extract and detect feature patterns from day-to-day satellite images and apply these patterns to compute or quantify economic livelihood status. In 2015, Kaiming He et al. suggested a new network architecture known as Residual Neural Network (ResNet). It is adopted due to its residual cognition to train deeper networks. And overshadow the challenges of decreased learning efficiency and incapability to increase the accuracy rate because of the deeper network depth [11].

Many neural network algorithms emerged to increase the speed and accurate rate and make the estimation of economic livelihood safer and more accessible. This research explores the capability of quantifying economic livelihood status using elements of a geographical region such as roads, electricity during the night, water, agricultural cover, and manufactured building structures only via satellite imagery data and showing its an advantage over physical survey.

The result of the economic status of a geographical area is beneficial to non-governmental, governmental agencies, and international organizations to make policies and decisions. Also, the method can be additionally applied to predict the possibilities of drought, flood, agricultural productivity, road/dam developmental status, and structural-developmental status of a geographical area.

The satellite images can enable housing and population census, especially in a situation where the region or area is difficult to go or reach. Any part or region recognized as chaos or war means that it is not a safe area to go to or a ‘no going area.’ To know the housing and population status of that region or location, the application of satellite imagery can be beneficial/helpful in these circumstances to predict or estimate the number of buildings, farms, and population of the given region.

A times there are some sets of people or regions with a culture that does not allow for foreigners or new faces to be allowed entry into their territory, community, or houses to survey the economic livelihood status of the region. So, with the help of CNN and satellite images, such economic livelihood status survey can be conducted without stepping into such territory all in the computer systems.

1.2. Literature Review

The study also reviews on satellite image data which are mostly night captured image data that can be applied to develop a fully spatial regional people activities information like length of the road, number of buildings, and distribution data. They have a significant interrelation with satellite night light data. With the help of nightlight satellite images, prediction measures of humankind's socio-economic affairs, which is a critical aspect, are achievable—adjusting the prediction of people's income in regions with low status and unreasonable survey information can be achieved through this technique [9, 17]. This literature review chapter represents the review of other authors with similar research topic papers.

Piaggesi [9], in his work with his colleagues, they applied satellite images to investigate the economic livelihood prediction process in the urban ecosystem of two developed nations. They presented that the laid procedures for economic livelihood mapping in materials can also use low settings in this boundary. Categorically, they have precept that an algorithm pre-trained on the ImageNet dataset can elaborate on the target, an essential fraction of the variance with no adjust-tuning routine or proxies [12].

The concept of satellite imagery data and datasets, the determination of population economic well-being status will be effective and efficient, especially when such information is needed urgently. As everyday research on artificial neural networks keeps attracting more research grants, we will be seeing the emergence of new models or algorithms to support the prediction of poverty using satellite imagery data [10].

Wu and Tan's article used Chongqing, China, to exemplify the application of the ResNet50 neural network model by analyzing it in geo-regional economic research. Numerous experiments reveal the impact of satellite imagery and machine learning. They can apply classification or segmentation learning of sunset or nighttime light data using the ResNet50 model, which can be applied for geo-regional financial-economic status calculation. Wu and Tan's approach outperforms the direct use of sunset/nightlight imagery data to predict economic status. Moreover, the 'Squeeze-and-Excitation' roof/blocks are added into the Resnet50 model. The outcomes are also improved, which displays that the module can better increase the execution of the model and extract pattern features that better represent the economic level of the geo-region [14].

The work of Engstrom and his research partners [7] shows results that spatial and spectral patterns did adequately well on their own at elaborating economic livelihood, with adjusted R^2 metric starting from decimal 0.46 to 0.54. After all, they also discover the spatial autocorrelation in the framed procedural residuals, which shows that significant explanatory variables are reducing or absent from the models. But, it is not surprising, especially when considering the complicated nature of urban economic livelihood. It is more significant than an essential precursor and rise of the spatial set-up of on-the-ground objects [13].

Irvine, Wood, and et al. [15], in their analysis of some selected sub-Saharan African geo-regions, image-derived patterns provide essential data for estimating survey answers across a range of questions. The achievement is commonly most substantial with questions on infrastructure, like accessibility to electricity, clean water, shelter, healthcare service and, sewage disposal. Social behaviors can also involve questions, but they act only slightly greater than chance. Compared to their results from the earlier study conducted in Afghanistan, the achievement in this work is less compelling [15].

In the paper [8], Das, Chhabra, and Dubey viewed that 'from applying the first standard techniques of gathering data on paper, to using technology that was not yet really explored in this specific domain, the goal was usually alike: Reduction of Global Poverty.

Das, Chhabra, and Dubey outline that this can be performed only with a correct poverty map of the earth. Figuring observations from numerous researches, it is clear that satellite imagery information mixed with different approaches studied for this paper or otherwise looks like the best means by which the universe needs to move forward and solve this significant global problem. The accumulated solutions can be helpful in policy-making by policymakers to develop frameworks that can work actively at all classes or levels [16].

Okaidat and study peers [17] regard that 'number one objective of sustainable development goal (SDG) is to overshadow poverty.' The scholars primarily outline procedures to recognize the spatial distribution of economic livelihood. After all, the typical process of tracing census data by working around rural geographic areas consumes too much time. On the other side, the accessibility of high-resolution satellite images that incorporate a large amount of data about pictures and landscape patterns can permit or correlate with economic livelihood status activity [17, 18].

Those images and landscapes of Okaidat and their workgroup work can be applied with deep learning in an operative way to estimate the distribution of poverty quicker, simpler, and cheaply. The data that Okaidat and Peers used in their project contains three datasets that involve satellite images for three nations: Malawi, Ethiopia, and Nigeria. They used 30% of every dataset for testing and 70% for training. They also use 20% of the training set for validation.

Okaidat and Peers work applied CNN classifier with particular architecture to classify or segment the satellite images of Malawi, Ethiopia, and Nigeria into three groups related to three countries with various poverty levels. Class 0 is correlatedly assigned to Ethiopia, with the lowest economic livelihood status; class 1 was connectedly set to Malawi with the intermediate financial position. And class 2 was correlatedly assigned to Nigeria with the highest economic livelihood status.

Pandey, Agarwal, and Krishnan [10] proposed a two-step method for estimating poverty in rural geo-regions of India from satellite imagery data. Firstly, they train a multi-task amply convolutional model to determine three developmental signs. The leading resource is the roof, source of nighttime lighting, and access to drinking water from satellite imagery. They observed that the multi-task amply convolutional model automatically understood symbolic features, like roads, settlements, agricultural lands, and water bodies. Secondly, they train a model to estimate the economic livelihood status (straightway indicator of poverty) using the first model's computed developmental parameter/value results.

From their World Bank publication, Engstrom, Hersh, and Newhouse [18] question how well economic livelihood status derived from satellite imagery predicts poverty and which position is most significant? They examine these questions using a research segment of 1,291 villages in Sri Lanka, connecting parameters of economic well-being level with features or patterns obtained from high-resolution satellite imagery.

The result of their work shows a commendably strong correlation between satellite imagery obtained indicators and economic livelihood status. Basic linear models elaborate from 60 to 61 percent of the parameters in poverty and average log per capita consumption [18]. The authors also added that parameters measuring building ratio or density, built-up geographic area, and shadows are the robust estimators of variation in poverty or economic well-being in rural and urban geographic regions.

As predicted, the range and lushness of agriculture or vegetation are inversely correlated with economic well-being in rural geographic areas and slightly positively correlated with financial income in urban geographical regions. They recommend that plants, trees, and other vegetation be a luxury in urban geographic regions [18].

In 2020 Yeh viewed deep learning method also best regarded as a way to enlarge rather than replace traditional survey efforts, as local training data can frequently further increase model performance, and because other key living well-being outcomes often measured in surveys such as how wealth is shared between households, or among families within rural areas are harder to obtain in imagery [11]. Likewise, they could also use their method to measure other key outcomes, including consumption-based poverty metrics or other essential livelihoods directional such as health results [20, 21].

In 2020 Kondmann and Zhu’s work results outline that pioneering approaches that map poverty from satellite images with deep learning may struggle to capture trends in economic development over time [12]. Thoroughly validating these results in other countries and with other imagery is necessary to communicate the robustness of this weakness [12, 22].

In 2017 Head presents a preliminary evaluation of the globalization ability of satellite-based approaches for predicting human development after replicating past studies that reestablished the potential for such techniques to estimate asset worth based in Rwanda [13]. They explained that the same method could not be trivially interpreted into evaluating other “softer” development outcomes (like health outcomes and source to clean drinking water) with the same correctness in other countries (precisely like Haiti and Nepal) [13].

The research uses Chongqing, China, to determine the application of the ResNet50 deep learning model in geographic, economic livelihood study. A sequence of studies outlines can apply the segmentation understanding of night-time light data using ResNet50 architecture for geographical economic livelihood indicator prediction [14].

Angelini’s work advances their previous study into image-based models of economic livelihood situations using satellite imagery. Their study outcomes for three African countries are similar to their earlier studies of three different African countries using high-resolution public imagery [6].

To use example of Okaidat and Colleague work results show that 'spatial and spectral' patterns do essentially suitable on their own at expatiating poverty, with settings of R^2 values [17]. Irvine and his associate recognized various fields for future research that could increase the model performance. They asserted an excellent sense of possible patterns recognizable from satellite imagery data [15].

Piaggesi and Colleagues analyze from the use of the traditional approach of gathering data on books to applying the CNN model with the same aim, which is to eradicate poverty [9]. They remarked that this could ascertain with the correct economic livelihood map of the earth [8]. With the help of deep learning using satellite images, we can estimate development indication conditions. These possibilities occur through the gathering of satellite images as datasets with a computer application that allows the user to assess or evaluate the development state of affairs of a geographical area or region [16].

Applying models that understand geographic diversity, like ensemble approaches or deep learning approaches, will probably increase accomplishment for this action. Another element is the time variation at that the capturing satellite images were snapped, which can cause noise in the self-standing variables towards geographical areas or regions. This issue may impact specific indexes like counting vehicles, which can change primarily based on the day the satellite imagery was snapped [18, 27].

Engstrom and Peers, while showing that satellite imagery and CNN results are very motivating, more analysis recommends warning when extending estimations into geographically adjoining regions or areas. According to the occurrence of poverty, the standardized error values rank from a fourth fraction to one-half of the economic livelihood (poverty) rates. The potential obstacle to extension is geographic heterogeneity in the correlation between indexes and well-being or welfare [18].

Author Agarwal and Colleague [27] work Metric system of agriculture also executes regarded ‘yeartime’ variation, which can confuse extension estimates to adjoined areas or regions. This system recommends that some indexes specifically put up to bias when predicting space and that the time and date of the image is a significant deliberation when considering high spatial resolution estimation using satellite-based indexes. The researchers hope that this kind of thing will improve as the reusing ratio of satellites increases.

The work of Irvine and his partners overview the parameters to determine economic livelihood indicators through physical structures, wellbeing, differentiation in amount of income toward the population. Those economic livelihood indexes correlate to points in public and socio-economic capital. Irvine and colleagues were able to outline some of the critical economic livelihood indexes that evolved from empiric studies in the social sciences recommend correlation between the economic livelihood index and political/social factors [15, 27]:

- ✓ Economic that is not equal adversely affects the foundation and maintenance of democratic workout.
- ✓ Tribes' differences decrease public goods victuals.
- ✓ Lower Gross Domestic Product (GDP) adversely affects democracy and general goods victuals.
- ✓ Practicing democracy improves entire wealth.
- ✓ Lower attention of households leads to lower manufacturing of public goods.

Given these features within economic livelihood and socio-political progression, the authors [27] hope to explore suitable substitutes for socio-economic livelihood circumstances through these image patterns. Sometimes to determine the economic livelihood of a region or area, we need to look at their democratic system, governmental affairs system, their customs & values, their GDP, and rate of access to qualitative education healthcare services.

To ascertain that is to find answers for questions like what type of leadership affairs do the practice? Is it a monarchy? Is it democracy? If it is a monarchy, do they have a parliament? Does their economic policy reach the masses from top to bottom, whereby everyone has access to qualitative education, healthcare services, source of income? If it is a democracy, what type of economic system do they run? Is socialism or capitalism, or is it both? Does every eligible person have access to vote? Does their monetary policy reach the masses from top to bottom, whereby everyone has access to qualitative education, healthcare services, source of income?

Engstrom and colleagues while trying to conclude their report titled 'Poverty from space: using high-resolution satellite imagery for estimating economic well-being,' they asserted that conventionally, given the expensive cost of carrying out surveys adequately huge to produce correct statistics figures for small geographic areas, providing not-large region economic livelihood (poverty) predicts needs combining a well-being survey with a collection of census or inter-census conventional survey [18].

The arc of collecting conventional census and inter-census survey data is expensively and infrequent. The survey information is also mostly scattered with an essential latency, making it hard to examine local living status changes quickly. The results of Engstrom and colleague work demonstrate that indexes derived from satellite imagery produce correct estimations of economic local-level livelihood (poverty) and well-being when combined with survey data. By and huge, the requirement relations are of referent signs and magnitudes.

2. MATERIALS AND METHODOLOGY

This chapter represents how and with what the study was conducted, which will enable the work to evaluate the appropriateness of the study methods & materials, reliability, and validity. The research methodology format structured the data collection through which the identified variables and their correlation to the study are explained in this chapter. While the research materials format the substance or sample that were applied for the study.

In terms of materials, satellite images have been generally used in various sectors, like classification and detection. But satellite image processing needs a more preprocessing approach than typical classification or detection, as satellite images mainly depend on the procedure used. Therefore, researchers, both academia and industries, are committed to developing satellite image processing to maximize performance features like classification, preprocessing, detection and segmentation, etc.

Artificial Neural Network (ANN), the foundation of the Deep Learning (DL) algorithm, has been used in satellite images for almost/over a decade. Before DL emerged, researchers focused on using Support Vector Machines (SVM) and ensemble classifiers such as Random Forest (RF) for image classification and detection. The materials used for the study are all discussed in this chapter. But to highlight them for the study in terms of performing the poverty prediction using satellite images are ANN such as (CNN, ResNet50, and VGG16), satellite images, image processing techniques, and a vast number of published academic kinds of literature from recognized journals.

The machine learning approach is the effective methodology needed for transforming these enormous amounts of unstructured satellite images into organized predicted ground status. The study adopted satellite images and a machine learning approach tailored to predict poverty using satellite images. This approach can improve the excellent segmentation of land use, poverty status, forest cover, and population unit, thereby enabling decision-making and research studies. The study obtained huge algorithmic possessions in model training & testing, connecting to leading artificial neural networks, via algorithmic validations that get the opportunity that satellite pictures are snapped from a fixed length and seeing angle from and capture repeating features and objects.

The thesis method (s) or approaches and materials are to ascertain the research's goal, such as approaches/theoretical/simulation, which aim to outline data/results that can verify the hypotheses. A short explanation of the parameters or things to be measured, estimated, predicted, or something to do with it is described. The satellite images were run through machine learning algorithms using a python programming language. It took a lot of hours in running the algorithm. The satellite images consist of the four African countries Nigeria, Mali, Malawi, and Ethiopia.

2.1. Satellite Imagery

Sometimes refers as remote sensing, is the viewing of the earth by satellite technology or high-flying airplane to capture data like images of the earth. Satellite Images are picture snapped using satellite technology, satellite technology is a concept designed to be mounted in orbit between celestial bodies, purposely to relay data such as satellite images to earth. An example of the where satellite images can be explored is google earth but the Satellite Image of this study was obtained from Kaggle [28].

Dataset: The dataset images involve a large scale of information concerned with landscape patterns related to livelihood operation and recognize some main factors like the source of water, road, roof building, building, and farmlands. The image is in original pixels size of 256x256 Red Green Blue (RGB), i.e., color images. The datasets images are in four African countries: Malawi, Ethiopia, Mali, and Nigeria. Mali has 14,759 images, Ethiopia 8,590 images, Malawi 12,700 images, and Nigeria 11,551 images. Therefore, the total images of 47,600.

Table 2.1 The total number of images used

Countries	Number of Images
Ethiopia	8,590
Malawi	12,700
Mali	14,759
Nigeria	11,551
Total	47,600

Table 2.1 shows the total number of images that were used for the study. The directory will comprise two folders - training set and test set, each comprising four folders having images containing agriculture, buildings, roads, and water. Each image covers an area of approximately a 6km radius. The train-test split is 90:10%, example of the satellite images are displayed in the Figure 2.1, Figure 2.2, Figure 2.3, and Figure 2.4. below.

Figure 2.1. An example satellite image of Mali

Figure 2.2. An example satellite image of Ethiopia

Figure 2.3. An example satellite image of Malawi

Figure 2.4. An example satellite image of Nigeria

2.2. Satellite Image Processing Techniques

Image processing is the kind of signal processing technique that accepts images as input and transforms it into output pattern feature relevant to that input image. ANN is a technique used in executing operations on images with the aim of extracting useful features such as information or image enhancement. We have segmentation, preprocessing, transforming, and correction types of processing techniques.

Segmentation creates a pixel-wise mask for every object in the image and hence is not only needed for the detection of multiple objects but also to understand the shape of the object. Region-based segmentation for detecting buildings and edge-based segmentation for seeing the other factors are used.

The study model is implemented in two pre-trained models (ResNet50, VGG16) and a developed CNN model from the beginning to classify/segment the high-resolution satellite pictures into four groups of classes as four ranks of poverty: very high, high, medium, and low poverty. Nigeria represents very high poverty status, Malawi represents low poverty status, Mali represents medium poverty, and Ethiopia represents intense poverty.

The satellite images processing can be performed using a ridge regression model by applying image vector exhibition as estimator variables and socio-economic income levels as the target. Hyperparameters (parameter of a prior) are predicted with 10-fold cross-validation by applying as distance measurement (metric) calculation coefficient R2 and measuring it on separate validation sets (as specified in all the experiments). The two processing techniques can be:

Image level prediction was made using patterns of images-values embedded as the sequence of estimators and aligned to the aim target level high-sampled according to self-standing resolution position or level (region, roads, census record, and the likes). The image is not placed on top of another (superimposition) between images, and it helps us avoid the data escaping among the training and test sets in the learning procedure.

Due to the fine-bushed attribute of the collection of images, in this scenario, it is plausible to maintain the model on levels that prosper or not prosper to rural areas. And then, the study overviews the level of occupation of region areas that may affect the final execution of the model and also view whether if a general training increases the prediction only in rural spaces.

Cluster-level Prediction was made using estimator variables for each cluster to determine the average of image-level embeddings calculated on the images that are the portion of the cluster itself. It is a similar approach applied by previous scholars. As adopted from [9] who reject spatial clusters or groups that are not above ten images for regression work.

2.3. Deep Neural Networks

The neural network (or artificial neural network or ANN) is derived from the human neurological idea of neurons. The neuron is the connection structure of cells in the brain. So neural network works like how neuron works [29, 30, 31]. But for the benefit of this study, let just not make it complex by simply see Neural Network in a simple way that applies a waterfall of many layers of non-sequential processing units for feature uprooting or extraction and transformation [2, 30, 31].

Therefore, CNN, is also kind of neural network known as LeNet architecture. Before the emergence of CNN, practitioners applied a standard densely layer for computer vision, and which is challenging and less effective than CNN. The fundamental difference between a densely joined neural network layer and a CNN layer architecture is that: Dense layers 'perceive' global patterns in their input feature extraction space (using datasets like MNIST digit, prints feature consisting of all pixels). Whereas CNN architectural layers' perceive' local patterns extraction in applying images [24, 25].

So, deep learning works on triple layers. The first layer serves as input (which acts similarly to dendrites of the brain). The second which is the hidden layer prepares the input (similar to the soma & axon of the brain). Thirdly, the output layer transforms the computed task (similar to dendrite terminals of the brain) [2, 29 30]. The architectural structure figured a hierarchical, topological, and formidable feature representation as shown in figure 2.5.

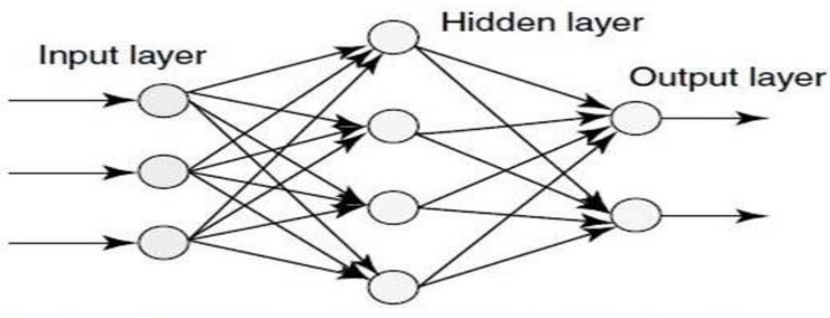

Figure 2.5. Architecture View of Deep Learning.

2.3.1. Convolutional Neural Network

Sometimes called LeNet architecture begun by LeCunet al. in the year 1998 in their paper, 'Gradient-Based Learning' applied to 'paper written work recognition.' Cutting edge of ML. [2, 29]. CNN is one of the Deep Learning architectures that are largely recognized for making computer vision effectively. Before the emerging of CNN, practitioners applied normally densely layer for computer vision which is difficult and less effective compare to CNN [2, 30].

The basic dissimilarities between a densely joined neural network layer and a CNN layer architecture is that: Dense layers 'perceive' global patterns in their input feature extraction space (using datasets like MNIST digit, patterns feature consisting all pixels), whereas CNN architectural layers 'perceive' local patterns extraction in terms of applying of images. So, in this kind of issue, the patterns feature found in small 2D widows of the inputs [1, 31].

The attractive fields that CNN dominate consist mostly image-operation and character (especially handwritten) recognition for the use of postal code clarifications. Looking at the architectural layer in the figure below, prior layers are applied for detecting the pattern features such as edges and the past layers are applied for the re-synergizing of features to generate the maximum level characteristics of the input followed by the segmentation. Pooling is done after the segmentation, which moderates or mitigates the quantifiability of the extracted features [29, 30].

CNN has great accuracy and increases the works of the system because of its exclusive features nature such as the likes of its local connectivity and apportioned weights. It performs much greater than any other deep learning technique. It is one of the most popularly used deep learning architectures as contrasted to others. Figure 2.6. represents the working structure of CNN with the flow of data from the inputs layer, convolutional layer, pooling layers, hidden layers, and the outputs layer [30, 31].

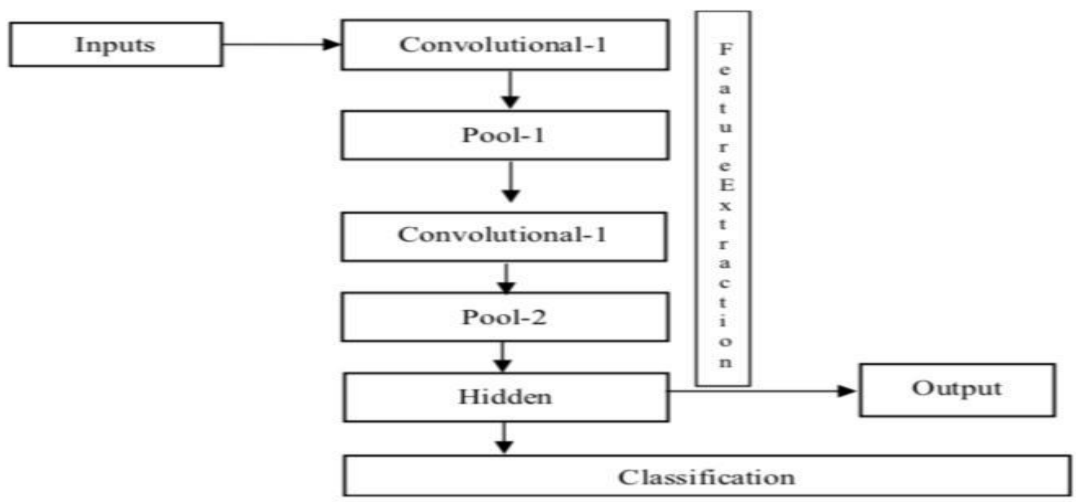

Figure 2.6. The architectural view of CNN.

The attractive fields that CNN dominate consist mostly image-operation and character (especially handwritten) recognition for the use of postal code clarifications. And now it is used for classification of satellite images to predict poverty. Looking at the architectural layer in the figure below, prior layers are applied for detecting the pattern features such as edges and the past layers are applied for the re-synergizing of features to generate the maximum level characteristics of the input followed by the segmentation. Pooling is done after the segmentation, which moderates or mitigates the quantifiability of the extracted features [29, 30, 31].

2.3.2. Visual Graphic Group 16 Network:

VGG-16 was initiated by the technologist at the Visual Graphics Group from Oxford which is shape like a pyramid. Its concept has the lower layers that are wide and the upper layers that are deep. VGG successfully consists of 19 layers. Because of the 16 numbers of layers that the VGG contains as layers it is known as VGG16. The VGG-16 is a type of deep CNN model network introduced through the publication by Simonyan and Zisserman in 2015. In the work of [32], the concept of the Visual Geometry Group (VGG) examines the influence of improving the depth of the CNN on its correctness or accuracy. The experts use architecture with a low number of convolution filters of length 3 x 3 that reveals a remarkable achievement in terms of comparison with the state-of-the-art parameters. Their work outcome was among the submitted ones during the 2014 ImageNet challenge [29, 32].

2.3.3. ResNet50:

ResNet50 means Residual Network 50 that applies residual component unit to spotted on one through the other and generate an effective and accomplish node-to-node network, so multiple residual layers are able of gathering a trained network. ResNeXt It an advancement of ResNet with novel and improved architecture and with increased performance. The word ResNet50 stands for the residual network model that has 50 numbers of layers. When ResNet50 is compared to VGG-16, it shows that ResNet50 has an extension unique nature mapping ability. ResNet estimates the delta needed to complete the last estimation from one layer to another. ResNet adjusts the disappearing gradient issues by giving a possible shortcut line for the gradient to pass through. The unique nature mapping applied provides the model with a way to bypass a convolutional network (CNN) mass or weight layer if the current layer is unnecessary. The ResNet50 approach helps us prevent over-fitting issues to the training set as it exists of 50 layers [32].

2.4. Poverty Analysis

The World Bank defines poverty as any Humankind living below $2 per day is regarded as a person living below poverty. Some scholars view poverty as a lack of essential things to make life comfortable, some percept it as hunger and lack of livelihood. Poverty is measured in any group of people based on their accessibility to basic amenities of life to live their lives, such as no access to school, shelter, food, healthcare services, water resources, etc. With satellite images we can apply deep neural network approach to detect those life basic amenities features measure poverty status of a region or country.

Sometimes called social impact analysis, it can be view a procedure to evaluate the extent of existence and social significant policy reformation on the economic status of various classes of population or regions, specifically on the most vulnerable people or poor people. Governmental or non-governmental organization are persistently trying to get the correct answer to development issues, and this consist of analyzing and making several policy options. The survey is a conventional instrument for obtaining data for poverty analysis. Still, nowadays, technology allows us to conduct the data gathering with satellite images and ML to predict poverty like the work of authors [15].

The authors [15], while trying to determine economic livelihood, socio-political, socio-cultural of their research target, applied to qualify survey data and satellite imageries. They tried to determine the correlation between parameters acquired from the traditional survey and patterns inferred from remote-sensed or satellite imageries data. Their goal is to develop models that estimate the parameters of specific indexes using the patterns features originated from the satellite imagery.

The techniques are described into five steps. The technique steps are:

1. **Evaluate the survey data**: This step affirms the “gold standard” for what we want to estimate from the satellite imagery-derived parameters. Through the evaluated data we can use it to make predictions of poverty
2. **Satellite Imagery Evaluation**: Using image processing techniques and computer vision by setting the conditional parameters
3. **Defining the Model**: Mathematical analysis of the correlation between the traditional survey-based indexes and satellite imagery-derived pattern features succumb participant models.
4. **Model Validation and Examination**: Analyzing the observed traditional survey-based indexes to model estimations compute the model’s execution.
5. **Application of the model**: The outcomes of this procedure is a model for estimating or computing the conventional survey-based indexes based on satellite imagery-derived patterns features. The model can easily be used on an epic scale to estimate related social indexes.

3. POVERTY PREDICTION ON SATELLITE IMAGES

This is the application of satellite images and machine learning to remotely determine the poverty of a region or country. Satellite imagery gives new information on the socioeconomic status of wealth and poverty. Some study predicted poverty status using the satellite images were taken during the night used houses with electric lights as a household living above poverty, while places that do not have the electric light is categorized as houses living in the poverty.

The concept machine learning technique is the significant approach required for predicting these vast amounts of unorganized satellite images into organized predicted information indicating poverty status data. The study applied satellite images and a deep learning approach purposely to predict or estimate poverty status using satellite images. This approach can also improve the better segmentation of land use, poverty status, forest cover, and population unit, thereby enabling decision-making and research studies.

This study obtained huge algorithmic possessions in model training & testing, connecting to leading deep neural networks, through algorithmic validations that get the benefits that satellite images are snapped from a fixed length and seeing angle from and capture repeating features and objects. Through the innovation of deep learning, we have the means to predict poverty using satellite images. The traditional way of surveying to predict poverty involves going to house-to-house, household-to-household to get the necessary information to determine the poverty.

The latest trend of satellite images allows us to potentially extract socio-economic details in a fast and relevantly cheap way. Satellite pictures hold underbuilding and agricultural details that, in a prior investigation concentrated on countries like Afghanistan and Botswana, were displayed to be purposeful for determining geo-regional socio-economic data. Researchers [6] studied to provide a survey reaction concern on wealth and crime were determined with supervised machine learning structural designs via their image patterns features obtained from their space reference coinciding satellite pictures [6].

3.1. The Study Area and Context

The African continent is categorized as the continent with the highest rate of people living below poverty. The study participant countries are Nigeria, Malawi, Mali, and Malawi. One of the features that cause poverty is the deficiency in good education. It elevates people to earn income, thus overflowing poverty. The datasets images are in four African countries: Mali has 14,759 images, Ethiopia 8,590 images, Malawi 12,700 images, and Nigeria 11,551 images. Therefore, the total images of 47,600 as it is shown below.

Table 3.1. The total number of images and their training and testing sets

Country	Number of Images		
	Training	Testing	Total
Ethiopia	90%	10%	8,590
Malawi	90%	10%	12,700
Mali	90%	10%	14,759
Nigeria	90%	10%	11,551
Total	Train % total	10% total	47,600

Table 3.1 displayed the images directory comprises two folders - training set and test set, each comprising four folders having images containing agriculture, buildings, roads, and water. Each image covers an area of approximately a 6km radius. The train-test split is 90:10%.

3.2. Proposed Approach of the Study

Several approaches were available for the satellite image, but the classification approach was best suited for predicting poverty using satellite imagery. We have classification, transforming, and correction techniques. In proposed approach can provide qualitative and quantitative information cheaper than a typical survey. The proposed classification using the deep learning (CNN) approach plays a vital duty in exploiting, extracting, and transforming worthy information from vast satellite images.

Classifier: Convolutional Neural Network is applied to extract features such as edges, shapes, corners, and pixel intensities. The output of the trained CNN model will be one of the four classes – agriculture substances, building roof, road, water. Figure 3.1 shows the architectural view of how the CNN classification operates.

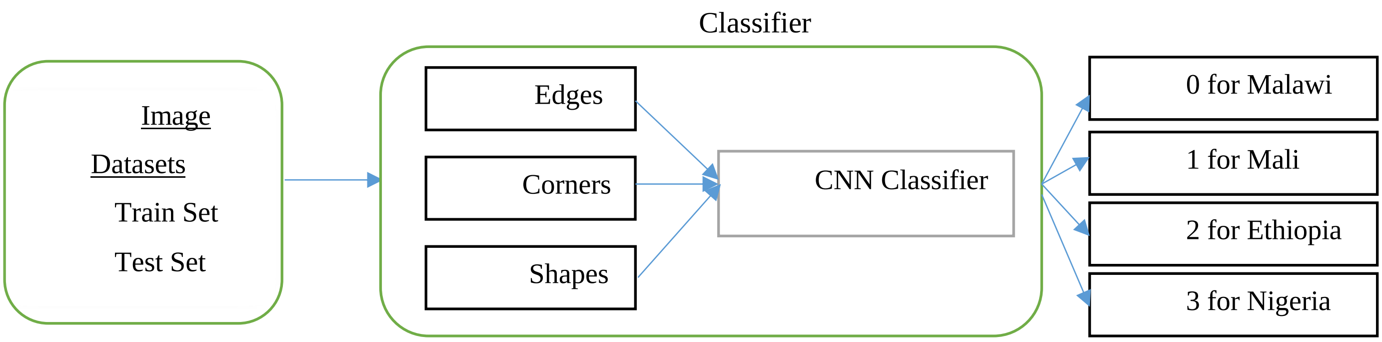

Figure 3.1. CNN Classifier

From the above Figure 3.1 shows the architectural view and operation through the following steps.

Step 1: Convolution Operation a filter passes over the image after scanning, and then ReLU Layer will pick up

Step 2: Pooling "spatial variance."

Step 3: Flattening Takes flattens and transforms the output

Step 4: Pooling happens after the flattening phase

Step 5: Full Connection Takes the inputs from the feature analysis and applies weights.

Step 6: SoftMax and; Cross-Entropy Optimization Classification and Bounding.

3.3. Data Acquisition and Preprocessing

The techniques used in data acquisition and preprocessing are divided into five steps which were explained in the previous chapter. The technique steps are:

Evaluate the survey data: Evaluation of the survey participants generates indexes of local attitudes about economic livelihood status and governance. This step affirms the "gold standard" for what we want to estimate from the satellite imagery-derived parameters.

Satellite Imagery Evaluation: Applying image processing approaches and computer vision supervision to extract related pattern features from the satellite imagery and then build local conditions parameters.

Defining the Model: Mathematical analysis of the correlation between the traditional survey-based indexes and satellite imagery-derived pattern features succumb participant models. The over roll satellite imagery covers precisely the geographic area or region as the traditional survey data and the geo-location of survey answers is the connection between the two data sources.

Model Validation and Examination: The model validation needs a different set of traditional survey and satellite imagery data that is grasped back (i.e., separated) during the process of model development. Applying this separated set of data can determine conventional survey-based indexes and satellite imagery-derived pattern features. Analyzing the observed traditional survey-based indexes to model estimations compute the model's execution.

Application of the model: The outcomes of this procedure is a model for estimating or computing the conventional survey-based indexes based on satellite imagery-derived patterns features. The model can easily be used on an epic scale to estimate related social indexes

3.3.1. Satellite Data

Satellite data, also known as satellite imagery, is learned as data describing the space and the planets especially the earth generated by human-made satellite technology mounted in their orbits. One of the most valuable aspects of satellite data is monitoring the earth through Earth Observation (EO) to provide information about the planet Earth's surface and weather status. Through satellite image classification, we can apply this information provided as satellite imagery for analysis and prediction of poverty, disaster management, economic survey, decision making, information extraction, and agricultural and mining research.

The help of satellite image classification allows us to extract information on remote sensing images, analyze vegetation varieties like foresters & agriculture, and analyze rural & urban areas to examine several land uses in regions. The diagram shown in figure 3.2. below explain how the satellite images were used to predict poverty in those selected African countries.

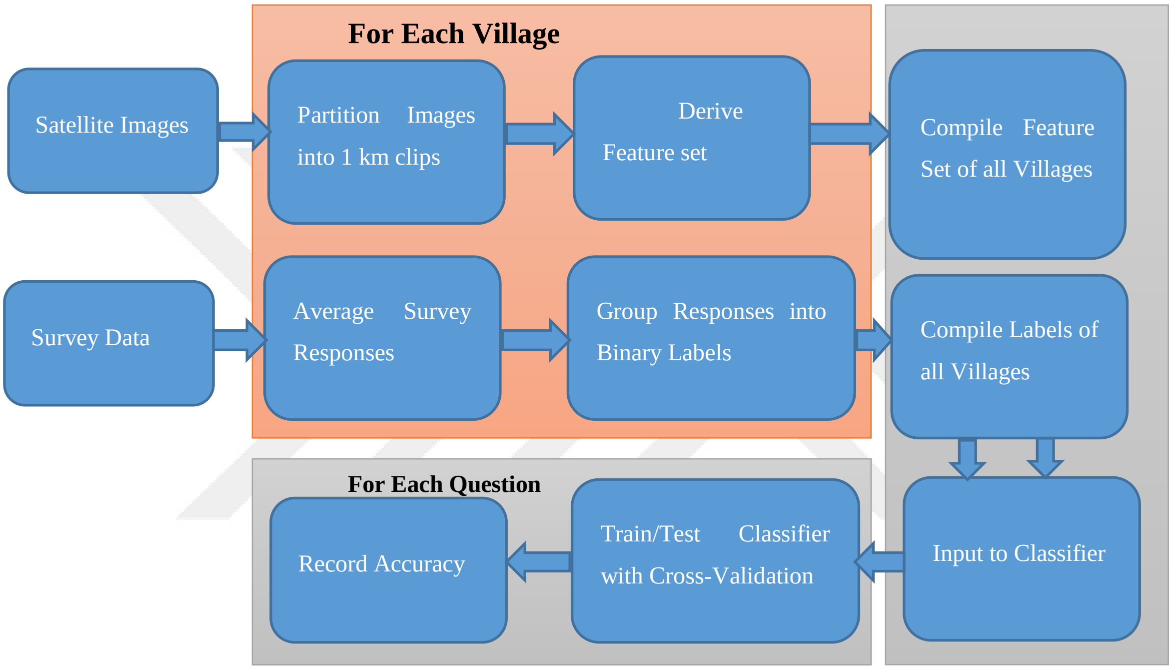

Figure 3.2. CNN satellite image and survey Procedure

Figure 3.2. Displaying the process of estimation or prediction of poverty using satellite images with the help of deep learning classification technique.

3.3.2. Image Preprocessing

By feeding each model with the augmented imageries, then the CNNs algorithm will automatically understand to extract many features that recognize some primary factors. Such as roads, buildings, vegetation and & farmlands, water resources, and building roof, discover relatively image patterns and make predictions of the livelihood well-being or poverty level.

The data preprocessing techniques was with the help of the Keras library frame on top of TensorFlow using a python programming language. This technique allows for the easy execution of the architectural neural network, i.e., CNN for image segmentation. The input layers accept in standard image size as already stated there rescaling regular of occurrence are divided by 255 across the dataset to make easy accessibility of the neural network. It has an import layer acceptable for RGB picture height, width, and depth. Eight (8) batch sizes were initialized as datasets generators' sizes. There was an adjustment of the size of the training set in adjusting the model overfitting and enhancing the generalization.

Table 3.2. The models parameters

Hyper-Parameters	**Models**		
	CNN	VGG16	ResNet50
Learning-Rate	1e-4	1e-4	1e-5
Drop-out	0.5	0.5	0.5
Hidden-layer activation function	ReLU	ReLU	ReLU
Output-layer activation layer	Softmax	Softmax	Softmax
Epochs	200	200	200
Size of batch in training	8	8	8
Size of batch for validation	8	8	8

Table 3.2. above shows the applied parameters while running the algorithm or model from the learning rate, drop-out, hidden layer, output layer, activation function, etc.

3.4. Classifications

Also known as segmentation, means the abilities to determine a problem in an appropriate way, to charter for its solution and to picture when new data is required and procedure to get it. Categorically, classification recognizes items by separating them into different set of classes which consist of examining the scaled features. We have different types of classification techniques, some are ANN, RF, binary classification, multi-label classification and the likes. In the study, CNN, VGG16 and ResNet50 from ANN models were used for classification. These models are described below:

3.4.1. CNN

The study has eight (8) convolutional layers, two (2) connected layers, and an output layer with padding and the activation function. The first layer has 32 filters and a 3x3 canal size with padding. Then the layers were followed by the activation function used in all the convolution layers, ReLU. ReLU allows averting exponential growth in the prediction needed to execute the neural network. The second also has 32 filters with a 3x3 canal and batch normalization.

After the second convolutional layer, the study applied max pooling to reduce the specialized dimension of the previous layer's output, then a drop-out of 0.25 was applied as regularization. The third and fourth convolutional layer has 64 filters of 3x3 canal size with padding followed with batch normalization and max-pooling layer and a drop-out (the drop-out is just like the previous).

The fifth and sixth convolutional layer has 128 filters with a 3x3 canal size followed by batch normalization, max pooling, and drop out. The seventh and eighth convolutional layer has 256 filters with a 3x3 canal size followed by batch normalization, max-pooling layer, and drop out. After the final 8th convolutional layer, the dense or fully-connected layer was passed to the fully connected layer.

The first and second dense layers have 1024 nodes, followed by activation function 'ReLU' and batch normalization, followed by a drop-out layer in the second dense or fully connected layer. The output or Softmax layer has four nodes that indicate 'very high,' 'high,' 'medium,' and 'low.' The Softmax function is applied for multiple class segmentation problems where group membership is needed more than one-two groups/class labels.

Experimentation: Adam optimization ie-4 (0.0001) was also applied as the learning rate with the decay of ie-4. They are all optimized by the learning rate by the number of epochs defined. As the class was multiple groups, a categorical cross-entropy was applied since it is not in binary mode. Early stopping was initialized into the model to monitor overfitting, and image batch sizes of eight (8) were set as model generators. Both the train and validation images were passed to the network as generators.

3.4.2. VGG16

VGG16 is a pre-train CNN architecture that consists of 16 layers, and the input layer accepts an image's shape of 256 × 256 x 3 and the weight of Imagenet. Then followed base pair of convolution layers with its average pooling size of 4 x 4. Then flatten function, which serves for converting data into a one-dimensional array for accepting into the near layer—followed by 4096 dense or fully-connected layers with ReLU activation function with the drop-out of 0.5. Then the Softmax activation function or/with output layer.

The concept of the VGG model is according to ANN with the can filter size 3 x 3 as the datasets are not that compatible with the VGG16 algorithm set. A fine-tuned concept was applied to the model by adjusting some layers as frozen and applying global average pooling layers to the architecture that assisted in revising the parameters. In trying to codify the VGG16 network to reduce overfitting, a dense of 4096 neurons and ReLU activation function was applied and 0.5 drop-outs. The multiclass function is four (4) from the four (4) neurons. That is the reason why the softmax was involved as an activation function.

Experimentation: VGG16 is a pre-train CNN architecture that consists of 16 layers, and the total Epochs of 200 were applied with a learning rate set as 1e-4(0.001). Adam optimization 1e-4 was used as the learning rate with the decay of 1e-4. All the optimization by the learning rate is divided by the total number of epochs defined. The epoch was reduced to 100 during the experiment but was stopped at 16 epochs. The early stopping technique was initialized into the model to monitor overfitting, and image batch sizes of eight (8) were set as model generators. The VGG network training continues from 96 epochs to 100. As the class was multiple groups, a categorical cross-entropy was applied since it is not in binary mode. Hyperparameters rates for every network model were regarded as the same technique. Both the train and validation images were passed to the network as generators.

3.4.3. RESNET50

ResNet50 architecture is on the style on large arranged in stacked residual quantity and applies a jump connection approach to resolve the disappearing/destroying gradient challenge. Fully named residual network 50 was initialized using the shape 256 x 256 using 3 x 3 canal filters layer as size. The convolutional average pooling size of 7 x 7 was initiated, followed by flattening, which is used to transform the data into a one-dimensional array for inserting it into the following layer. Then fully connected or dense layer of 4096 with ReLU activation function. They were followed by the drop-out of 0.5, Then a dense or fully-connected layer of 2048 with ReLU activation function with the drop-out of 0.5, followed by 1024 dense layer with ReLU activation function with the drop-out of 0.5, then the fully connected or dense softmax activation function.

Experimentation: Epochs of 200 were initialized with the learning rate of 1e-4 (0.0001), Adam optimization with learning and decay of 1e-4 are divided by the epochs defined. The callback monitors were used then adjusted the learning rate by remaking it 1e-5 (0.00001). Also, the same with the learning rate and the decay rate is divided by the epochs defined. Categorical cross-entropy was applied to maintain the multiple class. The model was trained using 200 epochs, and the early stopping method was used; as a result, the Resnet50 model stopped after executing 59 epochs.

3.5. Classifications Results

The multi-class classification of the study was done using the three models, namely CNN, ResNet, and VGG16. The idea of deep learning is not just coming to stay to be solving our modern-day problems only, but it keeps on increasing in terms of its capacity as a lot of funds have been invested in its research and development every day. This study combined the latest conventional survey of the four African countries, Nigeria, Mali, Malawi, Ethiopia, and the high-resolution satellite images of the same countries.

Figure 3.3. Results that CNN linear model performed well on the training data with 91% accuracy as the accuracy from validation is 71%. The training and validation accuracies are shown in the Figures below.

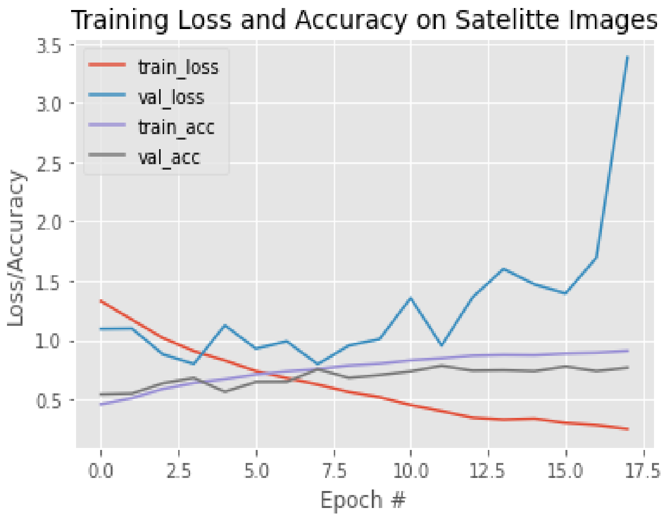

Figure 3.3. The performance of CNN model loss and accuracy.

The training loss curves are down while the validation loss is inverse to the training loss in the CNN model, as shown in Figure 3.3.

In terms of VGG16 shown in Figure 3.4., the training accuracy of 94%, while its validation accuracy is 86%. The training and validation accuracies are shown in the Figures below.

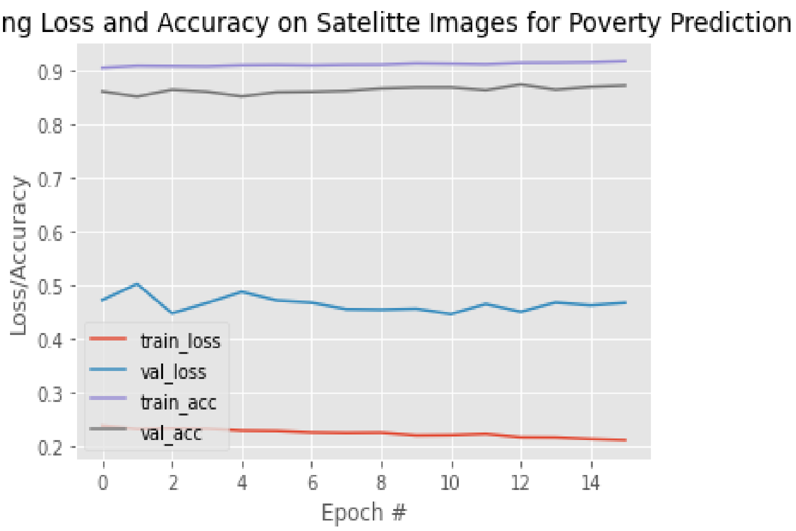

Figure 3.4. The performance of VGG16 model loss and accuracy.

The VGG16 also has variations between the training and validation losses, as shown in Figure 3.4.

The ResNet50 model in Figure 3.5 shows that the training accuracy is 62% while its validation accuracy of 58%. The training and validation accuracies are shown in the Figures below.

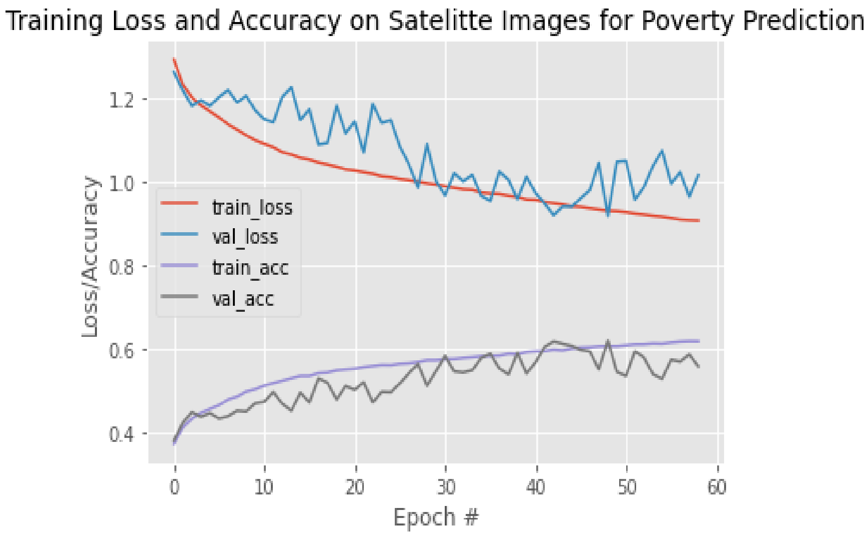

Figure 3.5. The performance of ResNet model loss and accuracy.

The ResNet50 model shows that the training and validation losses are on a closed curve, as shown in Figure 3.5.

Now let look at the full results as follows,

Table 3.3. The CNN model results

Countries	Precision	Recall	F1-Score	Support
Nigeria	0.70	0.72	0.71	1154
Mali	0.83	0.64	0.72	1475
Malawi	0.74	0.82	0.78	1270
Ethiopia	0.75	0.83	0.79	1717
	Accuracy		0.75	5616
Macro Average	0.76	0.75	0.75	5616
Weighted Average	0.76	0.75	0.75	5616

Table 3.3. represent that the CNN model performs averagely well with an accuracy of 0.75. It also shows that Ethiopia has the highest F1 score while Nigeria has the lowest, with 0.71. CNN performance is better than the ResNet model.

Table 3.4. The VGG16 model results

Countries	Precision	Recall	F1-Score	Support
Nigeria	0.87	0.77	0.82	1154
Mali	0.81	0.91	0.86	1475
Malawi	0.88	0.88	0.88	1270
Ethiopia	0.93	0.92	0.92	1717
	Accuracy		0.87	5616
Macro Average	0.87	0.87	0.87	5616
Weighted Average	0.88	0.87	0.87	5616

The Table 3.4. shows that the VGG16 has a performance accuracy of 0.87 which means it performs better than the two models. The country with the lowest F1-score of 0.82 while the country with the highest F1-score of 0.92. The table shows that the VGG16 model performs better than all other models.

Table 3.5. The ResNet model results

Countries	Precision	Recall	F1-Score	Support
Nigeria	0.53	0.59	0.56	1154
Mali	0.64	0.65	0.64	1475
Malawi	0.56	0.59	0.57	1270
Ethiopia	0.74	0.66	0.70	1717
	Accuracy		0.62	5616
Macro Average	0.62	0.62	0.62	5616
Weighted Average	0.63	0.62	0.63	5616

The Table 3.5. shows that the ResNet performed weak with an accuracy of 0.62. ResNet from the previous work of other authors shows that the model is not suitable for such kind of analysis.

The countries were classified where zero (0) was assigned to Malawi as a country with the lowest economic status, one (1) was assigned to Mali as a country with a medium economic status, two (2) was assigned to Ethiopia as the country with the high economic status, while three (3) was assigned to Nigeria as the country with the very high economic status. The classification was applied using the convolutional neural network, one of the vital models of deep learning techniques.

4. CONCLUSION AND DISCUSSION

The idea of knowing the economic livelihood of people or region by government, international organization, or non-governmental organization to make policy or decision that uplift the living standard status of the people. The manual survey of going house to house, household to household in either the rural or urban areas, is time-consuming and expensive.

The rise of high-resolution satellite images that contain extensive data of regions or countries' patterns, features, and landscapes can be applied to determine the economic livelihood of people or nations.

The application of satellite images to predict poverty is much easier, faster, and less expensive than the conventional survey. The study deploys datasets from the website Kaggle with a 5.2 GB size that contain images of 4 countries: Nigeria, Mali, Malawi, and Ethiopia. The datasets were split into 90% for training sets and 10% for testing sets were applied.

The countries were classified where zero (0) was assigned to Malawi as a country with the lowest economic status, one (1) was assigned to Mali as a country with a medium economic status, two (2) was assigned to Ethiopia as the country with the high economic status, while three (3) was assigned to Nigeria as the country with the very high economic status. The classification was applied using the convolutional neural network, one of the vital models of deep learning techniques.

The CNN model performs averagely with a percentage of 75%, which shows that it is better than ResNet. The VGG16 model was the model that performs best compared to the CNN and ResNet, with an accurate performance of 87%. While the ResNet was the model with the lowest version of 62% compared to the other two models. The VGG16 was the best model with the highest performance, CNN with medium performance, and ResNet was the lowest performance.

The study came up with the average result of the countries' economic livelihood status (poverty). Which result can be helpful for financial budget allocation, household statistics, per capita income, policy, and decision making by either governmental or non-governmental organizations.

Recommendation: The idea by international organizations, government, and non-governmental organizations are making efforts to eradicate poverty. There is always a need for conducting analysis or surveys to predict economic livelihood for those organizations or governments to create a financial, monetary policy decision. This poverty survey can be a traditional survey or the trends of satellite images with CNN. The conventional survey is recognizable as expensive and time-consuming.

As poverty differs from one region to another, each country or territory requires its satellite imageries to make predictions of economic livelihood for those regions. This study recommends the need for the availability of large satellites images for every region or country. Future researchers should focus on satellite images to predict poverty and the application in detecting crime, road traffic, agricultural soil, and the like.

High-resolution satellite images can be used not just to predict poverty but also for agricultural crop prediction or estimation, crime detection, mineral resources detection, flood, and erosion preview to make it few for this study. Therefore, this study recommends capturing the image according to the specific goal season. The reason is that if the image of a specific geographic area that their people used to migrate seasonally is snapped, then the probability of having bias is high because the satellite image can be photographed at the season that they migrate.

So, scientists or engineers should capture the satellite images at the appropriate season that the people are in their geographic area where they are available. Also, in terms of using agriculture to analyze crop production, the engineers should snap the satellite image during the rainy season if the people do not do irrigation farming. But if the people practice both irrigation and non-irrigation farming, then the satellite imageries can be snapped anytime.

Enough of these high-resolution satellite images should be snapped at the accurate season for the satellite images to perform accurately by avoiding bias. The more the images are available for any country, the more the country performs more analysis for their decision-making and economic livelihood policies. The satellite images can help governments avoid or handle natural disasters such as water floods or erosion.

This study also recommends that some indexes can specifically put up to bias when predicting space and that the time and date of the image is a significant deliberation when considering high spatial resolution estimation using satellite-based indexes. The researchers hope that this kind of thing will improve as the reusing ratio of satellites imageries increases.

Future Research: The results of this study raise a host of issues for further research and put up an ongoing debate concerning the application of predictive or estimative approaches in public policy. One of the most nearly of these issues is the debate of whether high-resolution satellite indexes can replace with census data in a variation context and for several indexes.

Does the significant relationship between high-resolution satellite-based indexes and economic livelihood welfare extend to salary income quantified straight from a conventional disbursement survey? Second, it is substantial to learn better the range to which these outcomes derive to several social and ecological ecosystems, like the Africa region, the Middle-East region, and some portion of Asia.

There is no assurance that the estimative or predictive capacity of developing density, shadows, and other pattern features recorded will hold in all ecosystems. Another line of research may explain whether we can apply modifications in high-resolution satellite imagery to estimate or predict changes in economic livelihood status of well-being towards space and time. Economic

Livelihood (poverty) surveys are gathered using traditional means every three years, and the most current general predictions are generated with a three-year latency.

Therefore, the capability to “now-cast” computation of economic livelihood well-being status by often adding the latest modified high-resolution satellite imagery with the most current traditional survey-based estimation of economic livelihood (poverty) has strong potential. Another future research that can clear doubt on recognizing the excellent means to determine into adjacent geographical areas or regions that have not been dealt with by surveys.

Globally, the unavoidable improvements in the availability of high-resolution imagery and pattern feature recognition models, in connection with the motivating outcomes from this study, reflects that high-resolution satellite imagery will turn into a more worthy tool to assist governmental, non-governmental organizations, and other stakeholders better to know the spatial nature of economic livelihood poverty.

REFERENCES

[1] F. Chollet, "*Deep Learning with Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek*", MITP-Verlags GmbH & Co. KG, 2018.

[2] P. P. Shinde and S. Shah, "A review of Machine Learning Algorithm and Deep Learning Algorithm". I *International Conference on Computing Communication Control and Automation (ICCUBEA)*. IEEE Xplore, 2018.

[3] P. Kim, "MatLab Deep Learning with Machine Learning, Deep Learning, and Artificial Intelligence. USA: Springer Science", 2017.

[4] M. P. McBee, and others "Deep Learning in Radiology", *The Association of University Radiologist. Elsevir*, *25*(11), November, 2018.

[5] A. Luckow and Others "Artificial Intelligence and Deep Learning Applications for Automotive Manufacturing", *International Conference on Big Data (Big Data)*. *IEEE* 2018.

[6] Angelini, B., Crystal, M. R., & Irvine, J. M. (2019, October). Quantifying Socio-economic Context from Overhead Imagery. In *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)* (pp. 1-4). IEEE.

[7] Engstrom, R., Newhouse, D., Haldavanekar, V., Copenhaver, A., & Hersh, J. (2017, March). Evaluating the relationship between spatial and spectral features derived from high spatial resolution satellite data and urban poverty in Colombo, Sri Lanka. In *2017 Joint Urban Remote Sensing Event (JURSE)* (pp. 1-4). IEEE.

[8] Das, P. S., Chhabra, H., & Dubey, S. K. (2020, January). Socio Economic Analysis of India with High Resolution Satellite Imagery to Predict Poverty. In *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 310-314). IEEE.

[9] S. Piaggesi, L. Gauvin, M. Tizzoni, C. Cattuto, N. Adler, S. Verhulst, A. Young, R. Price, L. Ferres, and A. Panisson, (2019) "Predicting city poverty using satellite imagery," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019, pp. 90–96. Retrieved from https://openaccess.thecvf.com/content_CVPRW_2019/papers/cv4gc/Piaggesi_Predicting_City_Poverty_Using_Satellite_Imagery_CVPRW_2019_paper.pdf on December, 5th 2021

[10] Pandey, S., Agarwal, T., & C. Krishnan, N. (2018). Multi-Task Deep Learning for Predicting Poverty From Satellite Images. *Proceedings of the AAAI Conference on Artificial Intelligence*, *32*(1). Retrieved from https://ojs.aaai.org/index.php/AAAI/article/view/11416

[11] Yeh, C., Perez, A., Driscoll, A. *et al.* (2020) Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nat Commun* 11, 2583. https://doi.org/10.1038/s41467-020-16185-w Retrieved from https://www.nature.com/articles/s41467-020-16185-w.pdf on December, 5th 2021

[12] Kondmann, L., and, Zhu, X, X. (2020) Technical University of Munich (TUM), 80333 Munich, Germany German Aerospace Center (DLR), 82234 Wessling, Germany {lukas. kondmann, xiaoxiang.zhu}@dlr.de Retrieved from https://elib.dlr.de/ 137108/2/ camera_ ready.pdf on December 5th 2021

[13] Andrew Head, Mélanie Manguin, Nhat Tran, and Joshua E. Blumenstock. (2017). Can Human Development be Measured with Satellite Imagery?. In Proceedings of ICTD '17, Lahore, Pakistan, November 16–19, 2017, 11 pages. https://doi.org/10.1145/3136560.3136576 Retrieved from https://www.ictworks.org/wp-content/uploads/2018/12/neural-networks-satellite-imagery.pdf on December, 07th 2021

[14] Wu, P., & Tan, Y. (2019). Estimation of economic indicators using residual neural network resnet50. *IEEE International Conference on Data Mining Workshops, ICDMW, 2019-November*, 206–209. https://doi.org/10.1109/ICDMW.2019.00039

[15] Irvine, J. M., Wood, R. J., & McBee, P. (2017, October). Remote Sensing to Analyze Wealth, Poverty, and Crime. In *2017 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)* (pp. 1-9). IEEE.

[16] A. Shetty, A. Thorat, R. Singru, M. Shigawan and V. Gaikwad, "Predict Socio-Economic Status of an Area from Satellite Image Using Deep Learning," *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2020, pp. 177-182, doi: 10.1109/ICESC48915.2020.9155696.

[17] A. Okaidat, S. Melhem, H. Alenezi and R. Duwairi, "Using Convolutional Neural Networks on Satellite Images to Predict Poverty," 2021 12th International Conference on Information and Communication Systems (ICICS), 2021, pp. 164-170, doi: 10.1109/ICICS52457.2021.9464598.

[18] Engstrom, R., Hersh, J. S., & Newhouse, D. L. (2017). Poverty from space: using high-resolution satellite imagery for estimating economic well-being. *World Bank Policy Research Working Paper*, (8284).

[19] F. Chollet, "*Deep Learning with Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek*", MITP-Verlags GmbH & Co. KG, 2018.

[20] Engstrom, R., Hersh, J. S., & Newhouse, D. L. (2017). Poverty from space: using high-resolution satellite imagery for estimating economic well-being. *World Bank Policy Research Working Paper*, (8284).

[21] 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 preprint arXiv:1711.06323*.
Retrieved from https://arxiv.org/pdf/1711.06323 on December, 5th 2021

[22] Ledesma, C., Garonita, O. L., Flores, L. J., Tingzon, I., & Dalisay, D. (2020). Interpretable poverty mapping using social media data, satellite images, and geospatial information. *arXiv preprint arXiv:2011.13563*.
Retrieved from https://arxiv.org/pdf/2011.13563.pdf on December, 5th 2021

[23] Tingzon, I., Orden, A., Sy, S., Sekara, V., Weber, I., Fatehkia, M., ... & Kim, D. (2019, June). Mapping poverty in the Philippines using machine learning, satellite imagery, and crowd-sourced geospatial information. In *AI for Social Good ICML 2019 Workshop*. Retrieved from https://pdfs. semanticscholar.org/ 9d96/bbd1bab6f66015096336052bd86662e14c6d.pdf on December, 07th 2021

[24] P. P. Shinde and S. Shah, "A review of Machine Learning Algorithm and Deep Learning Algorithm". I *International Conference on Computing Communication Control and Automation (ICCUBEA)*. IEEE Xplore, 2018.

[25] P. Kim, "MatLab Deep Learning with Machine Learning, Deep Learning, and Artificial Intelligence. USA: Springer Science", 2017.

[26] Engstrom, R., Hersh, J. S., & Newhouse, D. L. (2017). Poverty from space: using high-resolution satellite imagery for estimating economic well-being. *World Bank Policy Research Working Paper*, (8284).

[27] Agarwal, p., Garg, N., & Verma, O. P. (2019). Predicting Poverty Index using Deep Learning on Remote Sensing and Household Data. *International Journal of Recent Technology and Engineering* (IJRTE) ISSN: 2277-3878, Volume-8 Issue-3, Retrieval Number: C3918098319/19©BEIESP DOI:10.35940/ijrte.C3918.098319 Retrieved from https://www.researchgate.net/profile/Pratibha-Singh_12/publication/346934816_Predicting_Poverty_Index_using_Deep_Learning_on_Remote_Sensing_and_Household_Data/links/5fd2fb3ca6fdcc697bf708a2/Predicting-Poverty-Index-using-Deep-Learning-on-Remote-Sensing-and-Household-Data.pdf on December, 07th 2021

[28] Datasets from https://www.kaggle.com/sandeshbhat/satellite-images-to-predict-povertyafrica

[29] A. Pannu, "Artificial Intelligence and Its Application in Different Areas", *IEEE International Journal of Engineering and Innovative Technology*, *Vol. 4*(10), April 2015.

[30] A. Dey "Machine Learning Algorithms: A Review", *International Journal of Computer Science and Information Technology*, *Vol. 7*(3), 2016.

[31] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar "A Survey of Deep Learning and its Applications: A new paradigm to Machine Learning", *Archives of Computational Methods in Engineering. June*, 2019. Available at https://doi.org/10.1007/s11831-019-09344-w

[32] Theckedath, D., & Sedamkar, R. R. (2020). Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50 Networks. *SN Computer Science*, *1*(2). https://doi.org/10.1007/s42979-020-0114-9