Mapping Socioeconomic Conditions Using Satellite Imagery: A Computer Vision Approach for Developing Countries

Arslan Arshad  
National University of Science and Technology, Pakistan

Junaid Zulfikar  
Shanghai Jiao Tong University, P.R China

Muhammad Hassan Zaib  
Bahria University, Pakistan

Ahsan Khan  
Federation University, Australia

Muhammad Jahanzeb Khan (✉️ jahanzeb@nevada.unr.edu)  
University of Nevada - Reno

Research Article

Keywords: Socioeconomic mapping, Satellite imagery, Computer vision, Developing countries

Posted Date: September 6th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3318082/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: Competing interests: The authors declare no competing interests.
Mapping Socioeconomic Conditions Using Satellite Imagery: A Computer Vision Approach for Developing Countries

Arslan Arshad¹, Junaid Zulfikar², Muhammad Hassan Zaib³, Ahsan Khan⁴, Muhammad Jahanzeb Khan⁵*

¹Careem, Uber Technologies Inc, Dubai, United Arab Emirates.
²Department of Computer Science, Shanghai Jiao Tong University, Shanghai, P.R China.
³Department of Computer Science, Bahria University, Islamabad, Pakistan.
⁴Department of Computer Science, Federation University, Sydney, Australia.
⁵*Department of Computer Science, University of Nevada - Reno, Reno, Nevada, United States.

*Corresponding author(s). E-mail(s): jahanzeb@nevada.unr.edu;
Contributing authors: arslan.arshad@careem.com;
junaid_zulfikar1@sjtu.edu.cn; mhasanzaib@gmail.com;
ahsan15500@gmail.com;

Abstract

Pakistan is a developing country with more than a quarter of its population living under the poverty line. Data relating to the socioeconomic conditions of the population is scarce and sporadic. Government, NGOs, and other international organizations perform door-to-door surveys to collect data, but these can be expensive and time-consuming to conduct. Currently, statistics about poverty in Pakistan exist only as estimates from field surveys, which are often unofficial and limited. This lack of reliable information leads to ineffective policy decisions. This is one of the biggest challenges in fighting issues such as poverty. Reliable information is thus needed for the unified and concerted direction of development activities. This thesis aims to build a computer vision system that can automatically extract information about poverty from publicly available high-resolution satellite imagery. The proposed system output is heat maps indicating poverty
and development levels. These heat maps can be overlaid on digital maps for easy visualization. We propose to use transfer learning techniques to extract indicators such as geographical and man-made structural features from high-resolution satellite imagery. The trained model learns to filter and distinguish between various terrains and man-made features, such as highways, buildings, and farmlands. We show that these learned traits are quite useful for mapping socioeconomic variables and even come close to matching the prediction data with field survey data. For a developing country like Pakistan, this thesis is useful because the approach uses publicly available data and is a scalable and inexpensive alternative to traditional surveys. The results from our thesis can aid policymakers and NGOs in distributing their funds to areas that are most deserving and enacting and assessing policies more effectively.

Keywords: Socioeconomic mapping, Satellite imagery, Computer vision, Developing countries

1 Introduction

For governments, one of the most important functions is to create policies that solve their country’s problems. These policies require precise data and information as the policy can affect millions of people positively or negatively, for which the government is responsible. Therefore accurate data is of very high importance when formulating such policies. Unfortunately gathering data at a large scale for millions of people is not a small task and requires substantial resources of time and finances. However, recent advancements and innovations in the field of Big Data are opening up new ways of measuring important factors that can help in policymaking.

The eradication of poverty is the first of the United Nations Sustainable Development Goals [2]. An accurate assessment of poverty and related human living standards outcomes are important to making the correct decision for sustainable development. However, in many developing countries, it is difficult to find accurate indicators of economic well-being. Typically, such measurements are made through household surveys, which are costly and time-consuming to carry out across wide geographies. Also, these surveys do not contain feedback from every household, scaling up these manual surveys to get information about the socioeconomics variable of more regions would likely be very expensive and not possible for most of the developing countries [13].

One of the most important problems that governments face is predicting the rate of poverty which is essential to creating the right policies. According to the World Bank measuring poverty is extremely difficult as the collection of detailed data on households is time-consuming and expensive. However, it is necessary to measure poverty as the poverty line is used to determine the eligibility for federal, provincial, or local aid. It also helps in understanding where the governments should focus on investments in infrastructure as well as creating more opportunities.

Let us take an example of Thailand where overall GDP per capita has risen by 3.35 percent every year during the previous few decades. Before the COVID-19 pandemic, the tremendous economic growth of the country coincided with a decline in
poverty rates of households, which went from 61.41 percent in 1988 to 5.04 percent in 2019. However, In rural regions, where 6.76 percent of households are regarded to be poor, there are still significant pockets of poverty [26]. Furthermore, the COVID-19 infection may hinder long-term attempts to eradicate poverty. As a result, monitoring poverty is a crucial task for the nation’s development professionals. The Household Socioeconomic Survey (HSES), which gathers information on the income of households every two years, provides the foundation for the study of poverty statistics. The survey’s sample size produces estimates that are only somewhat reliable when offered at the provincial and national levels, but they are often insufficiently comprehensive to produce reliable estimates at a more granular level. In contrast, because of the high expense, increasing the survey representative and frequency sizes is frequently unsuitable [34]. Researchers and development professionals have examined different methodological approaches due to the need for more timely and detailed poverty data that can be used to target resident characteristics that have the greatest need for intervention. Small area estimate (SAE) approaches, for instance, have been frequently used to enable analysis at levels more granular than are feasible when working with surveys alone. SAE methods integrate census, survey, and other types of organizational data. The acknowledged poverty maps are not always practical, though, as SAE needs census data that is not always available. Small location estimate (SAE) techniques have been widely used to enable analysis at finer scales than are feasible when using surveys alone. Using SAE methods, census and other administrative data are combined with survey data. However, these techniques are very expensive and don’t always yield the right information. Therefore, for countries like Thailand and those similar to Thailand, a prediction technique using satellite images is a much-needed approach.

The typical way of measuring poverty is difficult as it involves measuring the income, education status, health standards, living environment and many others, for millions of households. Even conducting a census has not been easy for many countries and it takes years before data becomes available. It is very hard to measure the different socio-economic variables, so countries usually perform small-scale surveys to record poverty which is a very weak approach yielding inaccurate data, resulting in bad policy making. A cost-effective quick and repeatable way of measuring poverty is required to solve these challenges.

For the majority of countries, there is a dearth of reliable information on the socioeconomic needs of individuals, including fitness indices, consumption expenses, and wealth holdings. On-site surveys are a common component of traditional methods for managing or gathering this data, but they may be expensive and time-consuming. On the other hand, remote sensing data, such as high-resolution satellite images, are becoming increasingly available. Technology-based solutions have previously been employed effectively for basic satellite image testing from resource-poor nations to overcome the scarcity of socio-economic data at high granularity.

Plans for evaluating socio-economic indicators using geospatial data, such as nightlights and satellite photography, have been developed as a result of recent breakthroughs in machine learning methodologies and techniques and the enormous quantity
of remote sensing data that has been collected. Such methods have been used to evaluate or get around poverty at a high ranking in underdeveloped countries, as well as to get around laws or a lack of result data.

Perhaps the only technique that is both efficient and affordable that can provide information on a worldwide ranking is remote sensing, particularly satellite images. Commercial advantages are predicted to give sub-meter resolution images worldwide within 10 years at a fraction of current costs. A plethora of information for bearable development might be provided by this degree of material and spatial solution. Unfortunately, the positive lack of development in this raw data makes it difficult to scale the extraction of useful insights.

In this research, we overcome this challenge of the dataset by using a sequence of transfer learning with a fully connected convolutional neural network model (CNN). Using a CNN model helps us identify the high-level features from satellite imagery. Some related work in this domain is already done to extract features from the Overfeat network [28] using the ImageNet model, however, ImageNet is useful for learning low-level features such as edges. So we used the VGGF_BN to overcome this challenge and on the second transfer learning step, where nighttime light intensities are used as a proxy for predicting the socioeconomic variables similar to Addison et al. [1, 23]. Using this approach, our model was able to detect man-made infrastructure such as railway tracks, roads, buildings, vehicles, and some others as we explain in the following sections.

1.1 Problem Statement

Poverty prediction is still a domain challenge as it requires several variables such as wealth index, food consumption, household income, and many others. The wealth index, a metric composite assessment of a household’s cumulative living standard, is this study’s key finding. Economic surveys are held across the globe by various governments at times to collect data on multiple socioeconomic variables at an exceptional cost. Furthermore, this labor-intensive collection of data is not definitive. Whereas more precise, real-time, and economically viable data acquisition is possible by automating the wealth index prediction method using satellite imagery. Automation could prove to be cost-effective and more reliable for policymakers in achieving sustainable development. Unfortunately in most developing countries, this type of research is rarely seen, likewise in Pakistan policymakers are more dependent on the Pakistan Bureau of Statistics rather than automated data collection methods. The data of the Pakistan Bureau of Statistics is primarily based upon outdated data acquisition methods which mainly use smaller and ambiguous footprints for data collection hence resulting in insufficient and unrealistic information. Eventually, automated prediction of socioeconomic variables is seemingly more viable in solving poverty-related problems.

1.2 Prior State of the Art

Governments depend on manual surveys to be conducted every alternate year to get insightful data related to socioeconomic variables, Jean et al. [12] in 2016 proposed a
methodology to automate this process and developed a model to predict poverty using satellite imagery. Their study has shown the potential to predict poverty using satellite imagery. Different studies have been done after that on the topic and improved the methodology significantly over time.

Jean et al. [12] proposed a transfer learning model to train deep learning for predicting nightlight bin classes and extract features from the trained deep learning model. In the next step they trained a ridge regression model to predict the consumption per food item, using the data from the World Bank about the living standard measurement. This research was able to achieve 73% accuracy [12]. Xie et al. [31] extended Jean’s work for five African countries.

In 2020 Kumar et al. [3] proposed a method to detect objects in satellite imagery using the YOLO V3 dataset and trained their model. Using this training data, the author fine-tuned the model and tried to predict the number of different objects detected in the satellite imagery. A dataset was created based on this information and a machine learning model was trained to predict the consumption of food items per household. Using this approach the author was able to improve the work done by Jean et al. [12] by 35%.

This shows the significant progress in this field but they have their limitations, which include the availability of high-resolution satellite imagery and well-conducted surveys to train the model. In our method, we are trying to improve their proposed methodology for Pakistan with the dataset currently available.

1.3 Proposed Method

In my study, I will be focusing on making this happen for Pakistan. For Pakistan a good reliable dataset was the challenge as the publicly available dataset on the World Bank website was 23 years old. I have researched a number of datasets and chose the Demographic and Health Surveys (DHS) dataset for Pakistan. They have the latest dataset available for the year 2017 and also have the geo-variables with the surveys as well.

The dataset provided by DHS is quite comprehensive, I have used the household recode surveys and geo variable data from their dataset. DHS do random survey in different regions in Pakistan. They have made 560 clusters to do these surveys where each cluster represents an administrative unit, and in each cluster, they have done 20-30 random household surveys. They have captured more than 4000 socioeconomic variables for each household which gives an understanding of wealth, health, education, and other important insights about the households.

In our proposed methodology, we have processed this dataset to calculate the aggregate information on the cluster level so we can process it, to get the aggregate values on the cluster level we took the mean of variables that are important for our work, such as Wealth Index (HV271) and Wealth Index Factor (HV271). This gives us a high-level picture of each cluster’s wealth assets. We downloaded the 30 images per cluster which got us more than 18000 images.

All of these images were downloaded from the Google Maps API in high resolution with a zoom level of 16. Images with low nightlight intensity were discarded as we will be training the model to predict the nightlight bin and then extract the features, for the
deep learning we had tried different CNN models and finalized to use the VGGF_BN as it gives the maximum accuracy while predicting the nightlight bin classes and extract the high-level features from the image. From the trained model we have extracted the features to train the regression model, for regression, we have tried different regressors. Gradient Boosting regressor performs well for the Pakistani dataset. In the end, we were able to predict the Wealth Index at the cluster level.

1.4 Contributions

The main contribution of this thesis includes providing real-time updates about the poverty index based on daylight satellite images. Predicting the poverty index can help policymakers to make decisions for sustainable development. Although these updates depend on the frequency of satellite image updates they still provide data automation and can give up-to-date information.

The key impacts of this research study are as follows:

1. **Reduce the cost of manual survey**: Currently to acquire information related to poverty and other socioeconomic variables, manual surveys take place which are very labor intensive and cost a huge amount to government national budgets.

2. **Real-time information about different economic parameters** Our model can provide up-to-date information from satellite imagery. Although it depends on the frequency of satellite images getting updated from the satellite to the publicly available API it still provides up-to-date information every week, as most images get refreshed weekly.

3. **Data automation** Our model will provide data automation and predict up-to-date information based on satellite imagery. This will significantly reduce the cost of manual surveys.

   Our model can be applicable to these entities.

   1. Pakistan Bureau of Statistics (PBS)
   2. Ministry of Planning and Development
   3. World Bank and other monetary firms.
   4. Other policymakers

   Right now most departments depend upon manual surveys to be conducted each alternate year, These manual surveys do not have information for all regions and there is the possibility of human error as well.

2 Related Work

In this section, we are going to discuss the recent work done on this topic, Predicting poverty using satellite imagery. Daytime imagery is becoming a large practical source of information on welfare. Advanced techniques of Deep Learning such as Convolutional Neural Networks (CNN) have the ability to classify/detect the objects such as buildings, roads, cars, crops, and roof type [15]. These objects can be strongly related to wealth and poverty.
Different studies have been done to predict poverty using satellite imagery, [12] proposed transfer learning method to predict the consumption per household using high-resolution satellite imagery. They have done work for Uganda and then extended their work for 5 African countries, [24] proposed the method where the author uses lower-resolution (30m) multi-spectral satellite data to train a CNN to predict the African wealth asset. Performance of their model was comparable with [12]. These two papers show the potential in this field and their results were accurate and comparable. [9] proposed the method to use the Google Street View image data to train the deep learning model and predict the race, income, education, and voting pattern.

In [8] author did the study to identify items and textural cues from the high-resolution satellite imagery and predicted the wealth index. In [3] author proposed a method to detect the objects from satellite imagery and then train the regression model against it to predict the consumption per household. This study focused on one African country "Uganda" and they have improved the work done by Jean et al [12] by 35%.

2.1 Transfer learning using high-resolution satellite imagery and nightlight data

This approach was presented by [12] in 2016. The motivation behind this study was to reduce the expenditure and utilize remote sensing data which is becoming one of the large sources of information, and satellite imagery is available at a much cheaper cost.[27] shows the promising result of ImageNet to detect the object and find the pattern. The author purposed to use the transfer learning approach to predict poverty. In this study, they have used the fully connected convolutional network for spatial examination of arbitrary size inputs [21, 30].

Fig. 1 Left: Predicted poverty probabilities at a fine-grained 10km 10km block level. Middle: Predicted poverty probabilities aggregated at the district level. Right: 2005 survey results for comparison (World Resources Institute 2009).[31]
Jean et al used the Living Standards Measurement Study (LSMS) survey conducted in Uganda by the Uganda Bureau of Statistics between 2011 and 2012 (Uganda Bureau of Statistics 2012). This data contains a survey of 2716 households which are grouped into 643 clusters. Each cluster represents a geographical area of 10X10 KM centering the main town and they download 100 images for each cluster. Also, each household is assigned a binary class whether it is under poverty or not based on the consumption per household. This information is used to find the status at the cluster level.

The author has used the ImageNet and VGG for the deep learning model to train to predict the nightlight bin classes, the nightlight is very important data that gives insight about the availability of lights in an area at night time. In this study, they have tried to predict the nightlight bin class using the deep learning and after training the deep learning model they extracted the features from 2nd last layer which give an integer array of the feature vector for each image. They have used that data to train the regression model to predict poverty against extracted features. Poverty is defined as households that earn less than 1.98 per day.

The ridge regression model was used to predict the poverty, 4 shows the results where most left image is the predicted poverty on block level of 10X10 KM area.

2.2 Predicting poverty using low-resolution satellite imagery

Perez et al [24] proposed the method to use low-resolution satellite imagery to reduce the cost further which is incurred by acquiring high-resolution imagery from different providers, the cost of acquiring the high-resolution imagery is a fraction of the cost of a manual survey but using the low-resolution imagery can further reduce the costs. The author purposed to use low-resolution images that are freely available on the internet, for this study author used the data from the Landsat 7 satellite. Images from Landsat 7 satellite have 9 spectral bands, the author used both low-gain and high-gain thermal bands ranging from 60 meters per pixel to 15 meters per pixel.

As proposed by [12] they have used transfer learning, they bin the nightlight intensities in 3 bins (0, 1, 2) where 0 is low intensity and 2 is high intensity. They trained the CNN model to predict the nightlight bins from the Landsat 7 satellite imagery. They did the training with 60 epochs and optimized the weights from the epochs which helped them to achieve maximum accuracy during the validation. Subsequently, they used the trained deep learning model to extract the features and train the machine learning model to predict the socioeconomic variables of the DHS survey.

As in [12] they were able to extract the high-level features and predict the wealth index against it, but as these images were very low level their study was able to predict the socioeconomic variable more accurately at higher level images as compared to images with greater zoom.

3 Predicting poverty using Google Street View images

In [9], the Author proposed the approach to use the publically available dataset of images collected by Google Street View cars instead of satellite imagery. They have gathered 50 million images, in these images only the exterior of houses, vehicles, and
As per [22] 90% of Americans have a car in their house and the make and model of the car can give information about the household wealth status. In their proposed methodology, they have trained the model to detect the vehicles and link them with the socio-economic variables. CNN model can not only detect the make and model of cars from the image but other significant information as well. In the first step, they collected 50 million images from Google Street view car cameras of 3,068 zip codes and 39,286 voting precincts spanning 200 US cities.

22 million unique vehicles were detected from the dataset using CNN [15, 17], and with this model, they were able to categorize the vehicles into 2,657 different fine-grained categories as shown in 2, after that they have partitioned the dataset into subset: A training set and Test set. They have used the election voting data to train the model and evaluate the model on test data. After this, they extracted the features from deep learning and trained the logistic regression machine learning model to estimate race and education level. And ridge regression to predict the average household income and voter preference in a cluster based on the collection of vehicles.
For the median household income, their $R^2$ was 0.82, and the accuracy to predict the voter preferences was 85%. Their result was significantly good but the availability of Google Street View is not available for all countries. It is available for some developed countries and can be used to link with socioeconomic variables and can generate promising results.

4 Predicting poverty using object detection in satellite images

[3] presents an interpretable computational framework to predict poverty at the local spatial level by implementing high-resolution satellite images. Images are collected from areas that have favorable ground data on consumption for benchmark-related purposes. The imagery object detector is trained on a public dataset called xView[16] which focuses on location-specific training and provides an elaborate general object detection model.

The retrieved elements may then be used to predict support wealth and consumption expenses across five African nations [24] train a CNN to forecast African asset worth using 30m multispectral satellite images, yielding comparable results [12]. Upon analysis, the Uganda dataset seems to be the most desirable and has favorable features for new model training.

The next stage is to apply this detector to high-resolution images that are taken over rural areas (mostly measured in an existing georeferenced household survey). [3] claims to have outperformed on performance benchmark based on the General classifier for object detection and their technique is intuitively interpretable.

The model performance shows significant improvements however contrary to [3] claims there are certain drawbacks of the technique that is adopted. The model has been specifically trained in Uganda shown in Figure 3, which has publicly available poverty statistics and geolocation-based accurate results for measuring consumption benchmark through object detection [14] but this is not replicable to all poverty-ridden locations as poverty is substantially a household or individual based, and geolocation images are not considered to be optimal features for model training.

Measuring poverty statistics over the imagery set has shown significant results and the model retained a high-value Pearson score but that does not qualify as a uniform matrix. The model predicts consumption expenditure from high-resolution satellite images and proposes an efficient, solvable, and transferable approach that connects object detection and regression. This model reaches village-level consumption cost in...
Uganda, even when the provided areas are influenced by noise and the general number of labels is small, also shows that our components achieve favorable results even with simple linear regression model strategies.

Their results offer a good method for developing interpretable poverty predictions for significant livelihood results, even in environments with limited training data strategies.

5 Predicting poverty for city area using transfer learning

Piaggesi et al [25] proposed the method with a pre-trained Convolutional Neural Networks (CNN) model and fine-tuned the model with the five cities in North and South America’s satellite imagery. The author has used the ResNet [10] and VGGF for the deep learning. These methods are used to tackle the problem of lack of development data and large-scale poverty estimation in developing countries [7]. Subsequently, the author applies a machine learning approach similar to the one developed by Jean et al. [12] for predicting economic outcomes in 5 African countries from satellite imagery. In particular, they explored the chances of predicting household income at various levels in a city. Keeping this in view, they examined the Metropolitan Area of Santiago in Chile and the other five big cities in the USA: Los Angeles, Philadelphia, Boston, Chicago, and Houston. Their work tackles three main research questions:

1. Is it possible to extend machine learning methods, previously applied to resource-poor settings, to estimate poverty levels in a city of a developed country?
2. Given different aggregation levels in a city – usually corresponding to different administrative subdivisions – can a model train on a lower spatial resolution yield information about a more granular aggregation level?
3. What is the out-of-sample predictive power of such a model, when tested on a new city? Their main findings related to the above questions are:
   (a) Starting from pre-trained deep computer vision models they used regression to show the feasibility of predicting household income in a city, urban areas showed the best results. They show that to achieve good performances, there is no need to fine-tune existing models or to leverage proxy variables (such as night-time light data).
   (b) Just considering municipalities in which the regressor is better predictive about the target, they can also improve the estimation in more fine-grained levels of aggregation.
   (c) Testing their regression model on a new city which has never been seen before by the algorithm, showed better results if compared to a null model.
6 Predicting housing inequality and urban poverty using high-resolution satellite imagery

Li et al. [19] did the work to predict the urban poverty and housing inequality for the Jiangxia and Huangpi suburbs of Wuhan. Authors used a different machine learning model with the original approach presented by [12] and tried to improve the model.

China has made great strides in eradicating poverty in recent years, with the World Bank reporting a 646 million decrease in the number of severely poor people between 1993 and 2013 [5]. On the basis of new poverty features, China is still grappling with a major poverty issue [18]. Poverty’s growing urbanization is one of its most notable characteristics. China’s economy has steadily shifted from central planning to a market-oriented economy since its reform and opening up in the 1980s. With a growth of nearly 39.45 percent in the urban population between 1978 and 2016, China saw tremendous urbanization expansion during this time, leading to the largest population resettling in history [20, 33].

The Jiangxia and Huangpi suburbs of Wuhan, China, were the prime focus of the author’s research. Wuhan, the provincial capital of Hubei and a strategically important metropolis in China is located at the confluence of the Yangzi and Han Rivers in the eastern part of the Jianghan Plain. With a total population of 1.13 million in 2017 and a rate of urbanization of 45.52 percent in 2016, its regional gross domestic product (GDP) is around 70.25 billion Chinese yuan (CNY). One of Wuhan’s six suburbs, Jiangxia is the southern entrance to the city and has a total land area of 2018.3 km² with a resident population of 0.91 million as of 2017. Jiangxia is seeing a rise in population and urbanization because of its geographic importance and proximity to...
the Wuhan metropolitan region. Wuhan’s urban periphery Huangpi has a total size of 2256.7 km².

In order to ensure the relative homogeneity of regional built-up regions and effective resource allocation by policymakers, this study evaluates the link between urban poverty and the image characteristics of built-up areas determined from remote sensing images at the lowest administrative level. Urban communities are served by the neighborhood committee or village committee, which is the lowest administrative level in China. Some neighborhood/village committee units in Jiangxia have been divided as a result of the statistical unit of population data; as a result, there are 408 spatial units in Jiangxia and 653 spatial units in Huangpi, which make up the analytical region of this study.
The following data are among the research findings used in this study:

1. The study used a set of Google Earth (GE) data collected in 2016 that consists of 3-band multispectral stacks (R: red wavelength, G: green wavelength, and B: blue wavelength) with a spatial resolution of 4.09 m as shown in fig 5. After data pre-processing, GE imagery combines satellite, aerial, and Street View photos; among them, the satellite imagery includes QuickBird, Landsat, and WorldView, while aerial data is mostly acquired from for-profit organizations.

2. The local department submitted a land cover dataset and the boundaries of the administrative division from the Geographical Information Monitoring in 2016.

3. The local department is the source for the 2016 census of population and the sparse population data of neighborhood and village committees. Since each dataset contains information on its geographic coordinates, we may spatially merge them. In this study, urban poverty was defined as the proportion of the poor to the entire population of the region, or poverty incidence (PI), which is frequently used to define regional poverty.

For each machine learning model, the best model performances are chosen by varying the parameters and quantity of sampling functions, which is specified as "set.seed" in the R program. According to the validation indication for the regression findings, the $R^2$ for Jiangxia’s model performance varies from 0.3492 to 0.5341, and the $R^2$ for Huangpi’s model performance ranges from 0.4231 to 0.5324. With $R^2$ values of 0.5341 and 0.5324, respectively, the results demonstrate that among the examined regressions, the SVR method best captures Jiangxia and Huangpi’s performance.

In the previous section, we elaborated on the different methods and approaches to predicting poverty using satellite imagery. Almost all methods are based on transfer learning due to their high performance and ability to optimize the solution. Different authors proposed different methodologies in transfer learning techniques, And in different studies, different parameters of socioeconomics have been used to relate it with poverty.

Different work was done for Urban areas and Rural areas and predicting poverty from high-resolution satellite images. They propose different solutions. Some of the work uses consumption expenditure [3, 12] and some is used to predict the Wealth Index (WI) [25]. Recent transfer learning techniques have been developed for the analysis of satellite imagery that makes use of various data-rich proxy studies as well as recent serious learning advancements.

Jean et al [12] model reach village-level consumption costs on food in Uganda, even when the provided areas are influenced by noise and the general number of labels is small, their work achieves favorable results even with simple linear regression model strategies. The author offers a good method for developing interpretable poverty predictions for significant livelihood results, even in environments with limited training data strategies.

Piaggesi et al [25] work examined the predicting poverty in two developed countries’ urban environments, the author proposed methods to be used for poverty mapping in areas with limited resources. Ayush et al. [3] proposed a method where instead of predicting the nightlight bins classes from a deep learning model, the Author purpose
to detect the object from satellite imagery and extract the information related to the number of concrete-made buildings in the image, number of vehicles, road infrastructure. The author used the YOLO V3 dataset [16] to train the model to detect the objects, Their method shows significant improvements from the original paper [12].

We see the potential for improvement in it, as due to the poor quality of the survey conducted latest cutting edge technology can further overcome the problem which is caused by the noise, And we can improve the overall performance of the methods. Also, we will mainly focus on making this work for Pakistan and improving the methodology.

This section includes the process of the methodology used to predict poverty against the satellite imagery of Pakistan. It is divided into three parts: Data Acquisition, Feature Extraction using Deep Learning, Train, and Test sets. Details are given below.

7 Data Acquisition

There are different types of data available for Socioeconomic variables and this data is provided by different bodies. So relevant dataset was available as discussed below.

1. World Bank data for Pakistan World Bank provides a micro dataset for the household surveys which include different parameters including consumption per household etc. Pakistan has a dataset available for 1997 which was 2 decades old dataset. So we will not be using this dataset in this study.

2. HIES dataset by Pakistan Bureau of Statistics They provide a household dataset by randomly interviewing the households in an administrative unit. Their questionnaire includes different types of socio-economic questions they gather. This dataset is available for an alternative year for Pakistan. This data was not linked with geolocation so it will require adding the latitude and longitude for each cluster.

3. Pakistan Standard of Living Measures dataset by Pakistan Bureau of Statistics This dataset was relevant to the income of households and the distribution of expenses.

4. DHS Living Standard Measures survey of Pakistan. DHS is a chapter of the United Nations and it compiles data provided by the Pakistan Bureau of Statistics, They add a different calculated variable based on other variables. Also, they have the Geolocation dataset with each cluster. In this study, this dataset is used as it includes the Wealth Index which will be explained in the next points.

For our problem we need the updated data that can be trained with the latest available satellite imagery, So as the other sources provide incomplete and old data except for DHS we choose the DHS-provided dataset. DHS LSMS dataset of the year [2017-2018] was chosen with geolocation information for each cluster.

7.1 Dataset Explained

In this research, the DHS LSMS dataset was chosen with the year [2017-2018] with geolocation information of each cluster. DHS Dataset includes different types of data from different types of surveys, our focus will be on household recode which gave us information about the Wealth Index (WI) a metric composite assessment of a
household’s cumulative living standard [29], and other variables to understand the wealth distribution of that cluster. Table 1 shows the different variables available in the Demographic and Health Surveys (DHS) household-recode dataset. Not all the parameters are required for our purpose but those variables can be used to link with the poverty of households, We will be using the WI.

Table 1 OVERVIEW OF DHS Household

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV000</td>
<td>Country Code and Phase</td>
</tr>
<tr>
<td>HV001</td>
<td>DHS Cluster Number which represent a geo area.</td>
</tr>
<tr>
<td>HV002</td>
<td>Household number.</td>
</tr>
<tr>
<td>HV105</td>
<td>Members of the household’s ages discrete numbers.</td>
</tr>
<tr>
<td>HV106</td>
<td>highest level of schooling obtained discrete numeric.</td>
</tr>
<tr>
<td>HV107</td>
<td>highest academic year finished discrete numeric.</td>
</tr>
<tr>
<td>HV108</td>
<td>Education Completed in one academic year discrete numeric.</td>
</tr>
<tr>
<td>HV109</td>
<td>Educational success discrete numeric.</td>
</tr>
<tr>
<td>HV110</td>
<td>NA - Member is still Student</td>
</tr>
<tr>
<td>HV206</td>
<td>Has electricity discrete numeric.</td>
</tr>
<tr>
<td>HV207</td>
<td>Has radio discrete numeric.</td>
</tr>
<tr>
<td>HV208</td>
<td>Has television discrete numeric.</td>
</tr>
<tr>
<td>HV209</td>
<td>Has refrigerator discrete numeric.</td>
</tr>
<tr>
<td>HV210</td>
<td>Has bicycle discrete numeric.</td>
</tr>
<tr>
<td>HV211</td>
<td>Has motorcycle/scooter discrete numeric.</td>
</tr>
<tr>
<td>HV212</td>
<td>Has car/truck discrete numeric.</td>
</tr>
<tr>
<td>HV213</td>
<td>Main floor material discrete numeric.</td>
</tr>
<tr>
<td>HV214</td>
<td>Main wall material discrete numeric.</td>
</tr>
<tr>
<td>HV215</td>
<td>Main roof material discrete numeric.</td>
</tr>
<tr>
<td>HV216</td>
<td>Number of rooms used for sleeping.</td>
</tr>
<tr>
<td>HV270</td>
<td>Wealth index combined discrete numeric.</td>
</tr>
<tr>
<td>HV270A</td>
<td>Wealth index factor score combined (5 decimals) contain numeric.</td>
</tr>
<tr>
<td>HV271</td>
<td>Wealth index factor score for urban/rural (5 decimals) contain numeric.</td>
</tr>
<tr>
<td>HMl1</td>
<td>Number of mosquito bed nets discrete numeric.</td>
</tr>
<tr>
<td>HMl1a</td>
<td>Number of mosquito bed nets with specific information discrete numeric.</td>
</tr>
<tr>
<td>HMl2</td>
<td>Number of children under a mosquito bed net the previous night.</td>
</tr>
</tbody>
</table>

Table 2 OVERVIEW OF DHS Household

<table>
<thead>
<tr>
<th>HHID</th>
<th>Country</th>
<th>DHSCLUST</th>
<th>....</th>
<th>HV270A</th>
<th>HV271</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PK7</td>
<td>1</td>
<td>...</td>
<td>...</td>
<td>-139014</td>
</tr>
<tr>
<td>1 2</td>
<td>PK7</td>
<td>1</td>
<td>...</td>
<td>1</td>
<td>-123202</td>
</tr>
<tr>
<td>1 3</td>
<td>PK7</td>
<td>1</td>
<td>...</td>
<td>1</td>
<td>-121264</td>
</tr>
</tbody>
</table>
| ...   | ....    | ....     | .... | ..     | ...
| 580   | PK7     | 580      | .....| 3      | -11448 |

The Wealth Index is based on the main components analysis of characteristics that are simple to see from the surveyor’s point of view, such as bicycles, phones, and the availability of water. However, Head et al. [11] have demonstrated that this approach does not generalize in the same way that other development metrics predict having access to clean water and a number of other health indicators.
The consumption index from the Living Standards Measurement Survey (LSMS) is one example of a different metric that is not as predictive as asset-based indices. Overall, forecasts of poverty based on satellite imagery may currently account for more than half of the fluctuation and occasionally as much as 85% of the poverty as assessed by surveys.

In Table 2 Column HV270A contains numbers ranging from 1.5 where 1 is the poorest and 5 is the richest. In the same fashion, HV271 contains numbers ranging from -1,000,000 to 1,000,000 where the negative number represents the poor and the positive number represents the wealthiest.

DHS household-recode dataset contains 14,540 rows for a total available of 580 clusters. Where each cluster represents a geo area with some population in it. This makes 25 surveys per cluster. In Table 1 27 columns descriptions are shown, In total there are 4,552 columns for household-recode data only.

Every column contains different parameters related to the socioeconomics survey. There are different types of data provided by DHS which include household-recode, individual-recode, births-recode, couples-recode, geographics-data, geospatial covariates etc.

We use household records in the combination of geographics-data and geospatial-covariates. The geographics-data sample is shown in 3 which shows the latitude and longitude for every cluster number. HV001 exists in all of the surveys conducted by DHS, So this column can be used to link the data from different sources.

Other important data provided by DHS is geospatial-covariates which provide different information about each cluster e.g. minimum temperature, maximum temperature, rainy days, nighttime composite, etc. We have used the nighttime composite information from this dataset, the data sample is shown in Table 4

<table>
<thead>
<tr>
<th>Table 3 OVERVIEW OF DHS Geo Location Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHSCLUST</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>....</td>
</tr>
<tr>
<td>580</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4 OVERVIEW OF DHS Nightlight Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHSCLUST</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>....</td>
</tr>
<tr>
<td>580</td>
</tr>
</tbody>
</table>
8 Transfer Learning Model

The foundation of deep learning techniques is the autonomous learning of stacked, hierarchical data representations. Deep learning models are typically shown by deep feedforward neural networks. Convolutional Neural Networks (CNN) are created primarily for vision problems and contain convolutional operations over the input. Translation invariance is a crucial idea for finding meaningful characteristics in pictures, and convolutional filters are helpful for encoding it [4].

A series of convolutional and fully connected layers organized such that the output of one layer is the input of the following defines a CNN, which is a generic function approximator. According to Zeiler and Fergus [32], the initial layers of a neural network for image data often learn low-level characteristics like edges and corners whereas later levels learn high-level properties like textures and objects. A CNN functions as a mapping between tensors and feature vectors, which are then used as input by a final classifier.

Fig. 6 Transfer Learning Process

It is frequently possible to transfer the low-level and high-level properties that a CNN has learned from a source domain to enhance learning in a separate but similar target domain. We can transfer low-level characteristics, such as edges, and corners, from target issues with lots of data and learn new high-level features that are unique to the target problem. Learning new high-level features for target issues with little data is challenging. However, the feature representation that the CNN learned on the source job may also be applied to the target issue if the source and target domains are sufficiently comparable. For a variety of vision applications, deep features generated from CNNs trained on big datasets of annotated pictures have been employed quite well as generic features [6].

In this study, VGGF_BN 8 layers are used to extract the feature while training the model to predict the nightlight bin. VGGF_BN performs well to extract the high-level features e.g. Buildings and other information from the satellite images. We used this model to train while trying to predict the nightlight bin classes. Before the last
layer of the softmax classifier, we extract the high-level features from the model’s top layer. Our photos are the same size as the input of the VGGF fine-tuned model, but the pre-trained neural networks use images with a resolution of 224x224 pixels. In this instance, we obtained vector representations by averaging the feature vectors that were retrieved from each image’s four overlapping 224x224 quadrants. Before mapping, each picture is also rotated by multiples of 90 degrees and flipped horizontally and vertically to provide an enhanced dataset. Extracting these features before softmax will give 4096 feature vectors for each image.

These extracted features are then processed to calculate the feature on a cluster level. We take the mean of these features which gave us the features at the cluster level. As in fig 6 we tried to predict the Wealth Index HV271 using the regression models, we used Gradient Boosting regressor and Ridge regression. GBDT performs well for our dataset. VGGF_BN performs very well in extracting the features from the image. This approach lets us extract the hidden information from the satellite imagery and use this information to predict the Wealth Index which can be related to poverty.

![Fig. 7 Transfer Learning for Classification](image)

We are using the dataset from DHS which provides different socioeconomics variable data against the households but their data lacks the details about poverty. So we have analyzed the data and found the pattern in one of the columns Wealth Index Factor HV270 which shows the wealth status of a household. Possible values can be seen in table 9. So we have used this variable to classify the below-poverty households or above-poverty by defining a threshold. Figure 7 displays the flow diagram for the classification of poverty where we are assuming a threshold on HV270 $\geq 2.5$ as below the poverty line.

9 Pre Process Data

As explained in the above section we will be using a dataset by Demographic and Health Surveys (DHS) and their data contains a lot of socioeconomics information we will be focusing only on the variables containing information about the Wealth of...
households. In this dataset, we are using the Household Recode dataset which contains 20 to 30 random surveys for each cluster and we have geo-variable information on the cluster level. So we have to aggregate the information on the cluster level so we can process the geo-variable information along with it.

DHS household recode data was aggregated on the cluster level by taking the MEAN. We have tried different techniques to find the aggregate information but MEAN works well for us because values are negative as well and MEAN considers positive and negative values correctly.

Next, we merged the aggregated cluster level dataset with the geo-variable dataset (See table 3 which contains the information about latitude and longitude for each cluster, The Final dataset looks like as shown in table 5.

Table 5 OVERVIEW OF Processed Cluster Level Dataset

<table>
<thead>
<tr>
<th>HV001</th>
<th>HV270</th>
<th>HV271</th>
<th>nightlights</th>
<th>....</th>
<th>lat</th>
<th>lon</th>
</tr>
</thead>
<tbody>
<tr>
<td>258</td>
<td>1.928571</td>
<td>-159650.6071</td>
<td>0.003602475</td>
<td>...</td>
<td>24.2918</td>
<td>67.6844</td>
</tr>
<tr>
<td>256</td>
<td>1.142857</td>
<td>-104482.4643</td>
<td>0.037709849</td>
<td>...</td>
<td>24.3050</td>
<td>67.6244</td>
</tr>
<tr>
<td>259</td>
<td>1.333333</td>
<td>-1-74577.81481</td>
<td>0.023986174</td>
<td>...</td>
<td>24.3335</td>
<td>68.2587</td>
</tr>
<tr>
<td>....</td>
<td>....</td>
<td></td>
<td></td>
<td>....</td>
<td>....</td>
<td>....</td>
</tr>
<tr>
<td>488</td>
<td>2.347826</td>
<td>9811.608696</td>
<td>0.003534652</td>
<td>...</td>
<td>36.3597</td>
<td>74.8277</td>
</tr>
</tbody>
</table>

Next, we start to download the images for each cluster, we target to download 20 30 images per cluster in such a way each image contains the information with a zoom level of 16 and do not overlap with other images, we delete those image that has less information of nightlight intensity zero or very low because those images will not have any man-made infrastructure and we are not considering those images, all the images were downloaded using Google Maps Static API, it results in downloading total 18,541 images.

Next, downloaded images were passed to the trained convolutional neural networks (CNN) model to further fine-tune the model to predict the nightlight bins class. VGGF_BN [4] model used for this purpose, the model gets fine-tuned and extracts the high-level features from the satellite imagery, these features represent the man-made infrastructure as well as information about the agriculture in the satellite image. We extract the features from this model which gives us a total of 18,541 rows of extracted features of length 4096. To gain the features on the cluster level we take the mean of the features and in this way, we get the feature set for each cluster. This data will be used in the next step to train the machine learning model to predict the wealth index.

Table 5 shows the processed cluster-level dataset which was used to download the images and train the deep learning model. In this processed dataset important column is HV271 which has the number range from -1,000,000 to 1,000,000 where negative number shows the poorest and positive number shows the wealthiest household. We will be using the extracted feature against this column to train the regression models and will analyze the results in the next section.
10 Train and Predict the Wealth Index

As described in section 3.3 we have the extracted features from the satellite imagery using fully connected convolutional neural networks and then take the MEAN of these clusters to get the features on the cluster level. These features contain high-level features from satellite imagery. This process produces the processed data frame with a size of 580 X 4096.

We have used the processed data to train the machine learning model, Ridge regression and Gradient-boosted decision tree regression are used to predict the values. Gradient-boosted decision trees (GBDT) perform well with this data, due to limited data to work with, we have used cross-validation to resample the data to evaluate the machine learning model. We have done the 5-fold cross-validation for evaluating the machine learning model. $R^2$ 0.46 is achieved with this processed dataset while predicting the Wealth Index HV271. The $R^2$ is comparable with the original work done by Jean et al. [12] and Ayush et al. [3]. Due to limited data to work with, we have used cross-validation to resample the data to evaluate the machine learning model.

I have tried to predict other socioeconomic variables using the machine learning model as well. Our model was able to achieve the $R^2$ 0.44 for the Wealth Index Factor HV270. As both of the variables, HV270 and HV271 represent Wealth we can use the HV270 for the purpose as it shows better results with the model.

Figure 8 shows the variance between the predicted and actual data points with the $R^2$ 0.46. Although still there is a gap between the predicted and actual values which is due to the HV271 range is from negative one million to positive one million. And also Demographic and Health Surveys (DHS) add the calculated noise in the latitude and longitude to avoid privacy issues, Due to this added noise the results are not that much greater but still these results show the improvements.
11 Experiments and Prediction of Wealth Index

We have loaded the data and tried to predict the different socioeconomic parameters using our trained model. And displayed the result on Google Maps for comparison. These experiments display Wealth Distribution across the different regions of Pakistan.

11.1 Experiments

11.1.1 Regression Model

The proportion of variance between the predicted and actual values shows the $R^2$ score to be 0.46. Although it is not a significant improvement, it is still a comparable score and shows promising results on a higher level. The low score is due to limited data and DHS adds the calculated noise to the dataset to avoid the privacy issue. But still, this is comparable with the original work done for African countries, We are able to produce better results for Pakistan as well.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>$R^2$ Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV270</td>
<td>Wealth Index Factor</td>
<td>0.41</td>
</tr>
<tr>
<td>HV270A</td>
<td>Wealth Index Factor</td>
<td>0.30</td>
</tr>
<tr>
<td>HV271</td>
<td>Wealth Index</td>
<td>0.42</td>
</tr>
<tr>
<td>HV271A</td>
<td>Wealth Index</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 6 $R^2$ Score for different Wealth Index parameters with Ridge

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>$R^2$ Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV270</td>
<td>Wealth Index Factor</td>
<td>0.444</td>
</tr>
<tr>
<td>HV270A</td>
<td>Wealth Index Factor</td>
<td>0.346</td>
</tr>
<tr>
<td>HV271</td>
<td>Wealth Index</td>
<td>0.462</td>
</tr>
<tr>
<td>HV271A</td>
<td>Wealth Index</td>
<td>0.384</td>
</tr>
</tbody>
</table>

Table 7 $R^2$ Score for different Wealth Index parameters with GBDT

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>$R^2$ Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV270</td>
<td>Wealth Index Factor</td>
<td>0.09</td>
</tr>
<tr>
<td>HV270A</td>
<td>Wealth Index Factor</td>
<td>0.12</td>
</tr>
<tr>
<td>HV271</td>
<td>Wealth Index</td>
<td>0.10</td>
</tr>
<tr>
<td>HV271A</td>
<td>Wealth Index</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Table 8 $R^2$ Score for different Wealth Index parameters

In this study, we have done different experiments to predict different parameters using two regression models. Ridge regression results are shown in Table 6 shows...
the maximum $R^2$ achieved against the HV271, which is the Wealth Index. All of the experimented parameters belong to the Wealth Index.

Results shown in Table 7 are shown against the dataset which was processed by taking the average for each column to calculate the aggregate data at the cluster level. We have done the experiment with standard deviation and the results were not that good, Results are shown in Table 8

11.1.2 Classification Model

In this section, we have classified the clusters into the binary class either above poverty or below poverty. To do this we have set a threshold on $HV270$.

$HV270 \leq 2.5$

Where HV270 is a categorical variable and it contains the integer number as shown in table 9, we have processed the data to get the values on the cluster level by taking the mean, Which will make the HV270 values average number on cluster level. Our threshold set in the above equation will be used to assign the binary class to each cluster. In this way, we have assigned the binary class to each cluster which gave us 560 clusters with this information, the processed dataset is shown in table 5, this processed data will be used to train the ML model to predict the binary class.

Once the data is formed in as shown in 5 we trained the extracted feature to predict the binary classes using Logistics Regression. Our model gave us 73.8% accuracy. Figure 9 shows the confusion matrix and results of our model. We have used this trained model to predict poverty for different areas and compare it with the actual dataset. Section 4.3 shows the visualization for different areas of Pakistan.

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Poorest</td>
</tr>
<tr>
<td>2</td>
<td>Poorer</td>
</tr>
<tr>
<td>3</td>
<td>Middle</td>
</tr>
<tr>
<td>4</td>
<td>Richer</td>
</tr>
<tr>
<td>5</td>
<td>Richest</td>
</tr>
</tbody>
</table>

Table 9 Values of HV270 variable
11.2 Visualization of Regression Model

In this section, we will cover different visualizations of different regions and will discuss the results predicted vs. actual. Figure 10 demonstrates the actual Wealth Index and predicted Wealth Index on the map. We are showing the value in the range of -50,000 to 50,000 so we can have a clear picture of categorization in Wealth Index. On the high level, there are a few areas that are not predicted correctly but still, we can categorize the area based on these colors.

Fig. 10  Results for Whole Pakistan
Figure 10 shows the whole of Pakistan’s actual data points from the Demographic and Health Surveys (DHS) and predicted values from our transfer learning model. In this high-level figure, we can see the model was able to predict closer to the actual data and shows the potential. We will be analyzing each major region of Pakistan in the below sections.

In Figure 11, we can see the predicted and actual data comparison for South Punjab. As per the actual data, we can see the wealth distribution where points in yellow represent the wealthiest area and purple represents the areas that are not...
Fig. 13  Results for Lahore

wealthy. Actual data demonstrate that rural areas in south Punjab is not that wealthy but city area e.g. Multan, Faisalabad, and Quetta are wealthy as compared to their rural neighborhood. Predicted data also shows the same pattern with the $R^2$ 0.46. Although there are some errors that can be seen in this visualization still it shows the potential results for South Punjab and North Balouchistan. Similarly, in Figure 12 we can see the same results as for South Punjab where rural areas are not wealthy but the urban areas are shown as wealthy areas in actual data e.g. Karachi, Hyderabad. This data is limited so it does not contain the data of all of the areas but for the sake
of comparison, we have shown this data and predicted the same data points. We can see our model is able to predict comparable results with some exceptions.

Figure 13 shows the results for a big city of Pakistan Lahore. In this figure, we can see there is very little data from Demographic and Health Surveys (DHS) and it is not enough for a good analysis. In the surveys conducted for Lahore almost all clusters were shown as wealthiest and our model predicted the same for the given cluster points. We have run the model to predict the wealth index for the whole city in the shape of a 1X1 KM area. Figure 14 shows the predicted values of HV271 for the whole city. By seeing fig 14 we can see Johar Town is marked as wealthy and on the other hand areas around the walled city of Lahore are shown as not wealthy. Which looks to be accurate. Similarly, we can see the side areas of Lahore are marked as wealthy but as compared to other parts of Pakistan rural areas are not wealthy we can relate this to good resources available for south Punjab in rural areas compared to other parts of Pakistan.

Last but not least fig 15 shows the actual data points from DHS and predicted values, here we can see most of the part of Islamabad marked as wealthy as compared to the rest of Pakistan on the other hand the village in the mountains of Margalla is marked as poorer.

Figure 16 shows the results for Balouchistan, There were only 5 clusters for the south Balouchistan which our model predicted with accuracy.

11.3 Visualization of Classification Model

As explained in section 4.1.2, we have set the threshold on HV270 to classify the households in poverty or not. Sample values of HV270 are shown in table 9 so we classify the households as below the poverty line where HV270 ≤ 2.5, and above the poverty line where HV270 > 2.5. Using this assumption we have trained our machine learning model and tried to predict the binary class at the cluster level.
Fig. 16 Results for Whole Pakistan

Fig. 17 Results for Whole Pakistan

Fig 17 demonstrates the results for the whole of Pakistan. Where the right image shows the actual data points based on the assumption and the left image shows the predicted values from our classification model. In the figures, green points represent the clusters that are above the poverty line and red one represents the areas that are below the poverty line, in the figure we can see that North Punjab and South KPK is marked above poverty for most of the areas but for the South Punjab and Sindh actual data and predicted both show the areas living below poverty which demonstrate the accuracy of the model.
In fig 18 we can see the results for South Sindh where rural areas are mostly below poverty but on the other hand, some of the rural areas are living above poverty, and main cities like Hyderabad and Karachi also contain most of the areas living above poverty. Digging deeper to these main cities can reveal details about each individual area and their status about the wealth distribution, but as a whole, we can say that most of the area is above poverty in Karachi.

In fig 19 we can see the results for the whole city of Islamabad, Actual and predicted are almost the same, and also for Islamabad all areas in the city are showing to
live above poverty. Although there are some areas which live below poverty due to incomplete data of DHS, we only have the information about the areas which are above poverty. In fig 19 there is only one area marked living under poverty and that is a small village in Margalla hills named "Talhar".

12 Conclusion and Future Work

In this research, we have shown that satellite imagery can be used to automate the process of poverty prediction by calculating the wealth index. This can help governments plan and create policies for poverty reduction and alleviation strategies. As there is a scarcity of data and information available for Pakistan, our research gives an accurate picture of wealth distributions across Pakistan.

There is still a lot of gap for further improvements but our research has shown that we can automate the process of predicting poverty which can save costs while providing real-time information to administrative departments. This real-time information helps them in making policy-level decisions based on this dataset. This will save the cost of manual surveys as well as provide real-time information which is a much-needed improvement as these manual surveys were generally conducted in alternate years. Also, lack of the data available online was a challenge and it was difficult to find the relevant dataset.

This study used a novel approach to predict nightlight bins on the deep learning layer due to the availability of low-resolution images. Our model gave us 0.42 $R^2$ which is comparable with that of Jean et al [12] work. For Pakistan, we do not have the data for consumption per household but still, the wealth index gave us a picture of wealth distribution across the different regions of Pakistan.

The visualization of our model can be seen in Section 11.2 which shows us that our model works well in rural areas however in urban areas it has some imperfections but
still gives good results. By analyzing these visualizations we can see that our model has some limitations and it does not perform well in slum areas located in Big cities, as these are surrounded by large buildings. The noise in the DHS dataset is also one of the factors for this behavior.

This research can be further improved by using high-resolution images and instead of predicting nightlight bins, it should predict buildings, and road infrastructure and extract features during the prediction process. This will provide more details at the feature vector level and machine learning can show more accurate results. Also, this research can be linked with other parameters along with the feature vector which we extracted from the deep learning layer e.g. Urban or Rural, greenery ratio, and other factors that can be picked from the image.

We can also try to use other means of data such as IoT devices, CCTV cameras in public places, etc. We can find the trend between this data and socioeconomic variables, which can help us to fine-tune the existing transfer learning model and put this additional data as an auxiliary input to optimize the model.

We have predicted the socioeconomic variables related to wealth, but further improvements to predict the other socioeconomic variables using daylight satellite images can be done as well. Socioeconomic variables e.g. Literacy rate, health safety index, female education, clean water availability e.t.c show potential and they can be predicted using the transfer learning method as well. Also, future work can be done to enhance the dataset or combine the dataset of different providers to get maximum information that can be used to train the model for improved prediction results.

References


of the IEEE conference on computer vision and pattern recognition workshops, pages 806–813, 2014.


