International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Poverty Detection Using Satellite Imagery

Govinde Gowda Y , Greeshma TC , Siddeshwar

Department of Computer Science and Engineering

Rajiv Gandhi Institute of Technology

Banglore-560032 , Karnataka

---***---

Abstract - *:* As the universe finds it challenging to define poverty, the world bank views poverty as anyone living below $2 per day. Government and international organizations are working to eradicate this poverty. The study reviews some research on satellite images on the prediction of poverty through the concept of CNN. The study split the datasets satellite images of four (4) countries: Nigeria, Mali, Malawi, and Ethiopia obtained from Kaggle in 90% for training with 15% of it for validation and 10% for testing. The datasets were analyzed using CNN, VGG16, ResNet50, which shows that the VGG16 model performs better than the other two models with the validation accuracy of 94%, while CNN has 91%. ResNet has the lowest validation accuracy of 62%. 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 the prediction of poverty is much easier, faster, and less expensive compared to more the prediction becomes more accessible. 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 of satellite images in detecting crime, road traffic, agricultural soil, and the like.

Key Words: Convolutional Neural Network, Residual Network, Satellite Images, Architectural Neural Network, Support Vector Machine, Visual Geometry Group.

1.INTRODUCTION

The detection of poverty is a critical task in both developed and developing regions, as it informs policy-making, resource allocation, and humanitarian efforts. Traditional methods of poverty detection have relied heavily on census data, surveys, and economic indicators, which can be time-consuming, costly, and prone to inaccuracies. With the advent of technology and the widespread availability of satellite imagery, machine learning algorithms, and data analytics tools, there has been a shift towards more innovative approaches for poverty detection. This abstract explores the current state-of-the-art techniques in poverty detection, including satellite imagery analysis, social media mining, mobile phone data analysis, and machine learning models. It discusses the advantages and limitations of these methods, as well as their potential for scalability and accuracy. Furthermore, it examines the ethical considerations and privacy concerns associated with the use of personal data for poverty detection. Finally, it suggests future directions for research and emphasizes the importance of interdisciplinary collaboration in addressing the complex issue of poverty detection.

There is a need to measure the welfare of people and nations. The world's population should be counted, measured, weighed and evaluated .Behind these rather raw statements hides a more humanistic perspective where measuring and counting people are in support of humanitarian and developmental efforts, targeting, mapping and monitoring people at risk of food insecurity, famine, poverty and disease .The effort is supported by the 17 Sustainable Development Goals along with their 169 targets that were adopted by member states of the United Nations as part of the 2030 Agenda (United Nation). To balance the agenda's economic, social and environmental aspects, more timely, reliable and appropriate ways of collecting and interpreting information on a broad range of human development outcomes are needed . Traditional approaches such as household surveys, often rich in detail but infrequent in time and space, especially in the poor regions of the world, have served the development community with data for a long time are now augmented with new methods and new types of data. The new approaches are predominantly digital in nature, for example, using cell phone data, harvesting social media including Internet, crowd sourced data, imagery including google street view and other forms of remotely sensed imagery, for measuring welfare and poverty and monitoring the progress towards the attainment of these SDGs. Many are computationally intensive and advanced and many times qualify as 'big data'.

For close to a decade now, the most interesting approach in this area of research has been the combination of satellite imagery (SI) with different machine learning (ML) algorithms, including deep learning (DL) for the estimation of human outcomes. When put to work, wealth and poverty can be estimated from a single satellite image almost as good as achieved from surveys. A recent review of 12 studies evinces that the methodology can predict, for example, the Demographic and Health Survey (DHS) welfare asset index with R^2 between 0.45 and 0.80 .

This study focuses on studies at the intersection between ML/DL/TL, SI and poverty analysis. The SIML (Satellite Image Machine Learning) methodology combines some of the recent achievements from computer science and object recognition research and applies them to the field of human development research. At a conceptual level, the SIML approach shares many similarities with well-known applications such as learning different object categories from imagery, for example, distinguishing dogs from cats in photographs. For several reasons, it is more complicated to train algorithms to estimate, for example, poverty from imagery. This may be partly because labelled training data are less abundant for human development targets than for everyday objects typically found in image databases such as ImageNetn or AlexNet. Another is that poverty and welfare and how to measure them are strongly contested issues a fact that is not yet adequately reflected in recent works on SIML. Still, this is a rapidly growing area of scholarship.

The conceptualization and measurement of poverty are a complex endeavour. The increasing success rate of the SIML approach in this domain is influenced by the kind of indicator or measure of poverty that is being targeted. In this regard, it is important to note the different forms and manifestations of poverty or welfare so that in measuring or predicting it, both the so-called 'soft' measures (income, nutrition, food consumption, literacy rates, etc.) and the 'hard' measures (mainly physical assets) are taken into consideration or at least acknowledged. The use of the latter in machine vision processes would be expected to yield higher accuracy than the former in poverty prediction. That is, physical indicators of welfare such as roofing quality and type, infrastructure and farm sizes can more effectively be detected and classified from high-resolution SI than the quantity of meat consumed by a household in the last week. This constrains its applicability to certain types of poverty, which are place based and have a physical manifestation, rather than those measures that are transient in character.

The present study introduces this paradigm of SI/ML/poverty analysis to a wider audience in the humanities, the social sciences and the development community. For anyone interested in this topic, the available literature is difficult to navigate and filled with discipline specific notations. The aim of this study is to quantitatively synthetize the findings and methods in the field, from the perspective of a potential user.

Body of Paper

Aim and Objectives

1. 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.
2. The technique can serve as a means of estimating poverty using satellite images from countries in Africa as data for development.
3. 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.
4. 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.

II. LITERATURE REVIEW

With the help 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 architectural 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. Piaggesi [2], in his work with his colleagues, investigated the economic livelihood prediction process in the urban ecosystem of two developed nations. They presented that the applied 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 [2,5,13].

Pandey, Agarwal, and Krishnan [3] 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 [3,9].

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 [16]. Their 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 [4, 11, 16].

(Yeh et al., 2020) deep learning method also perhaps 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. 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 [4, 11].

(Kondmann, Zhu., 2020) Results outline that pioneering approaches that map poverty from satellite images with deep learning may struggle to capture trends in economic development over time [5]. Thoroughly validating these results in other countries and with other imagery is necessary to communicate the robustness of this weakness [5, 12].

(Head et al., 2017) The research 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 [6]. 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) [6, 13].

Angelini and Colleagues' work advances their previous study into picture-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 [17].

The work of Engstrom and his research partners [18] shows results that spatial and spectral patterns did adequately well on their own at elaborating economic livelihood, with adjusted R2 number 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 onthe-ground objects [18, 6, 13].

Irvine, Wood, and McBee [19], 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 [19].

In the paper [20], 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. They 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 [20, 8].

From their World Bank publication, Engstrom, Hersh, and Newhouse [26] 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.

Okaidat and study peers [27] 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. The data that Okaidat and peers [27] 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.

They applied Convolutional Neural Networks (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 [27].

III. MATERIALS AND METHODOLOGY

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.

Our review method can be described as integrative rather than systematicThis body of knowledge, mixing preprints, working papers, technical reports, peer-reviewed papers and conference papers with contributions of various disciplines is notoriously difficult to capture with one single approach. We one of the first publications to examine the application of ML and SI for measuring poverty and economic well-being—as the benchmark and narrative for our analysis. On this basis, papers completed prior to 2014 and do not apply ML to study socioeconomic wellbeing from SI were excluded from the study. We include literature from published journal articles, grey literature such as working papers and validation studies that have clear empirical application; that is, we excluded reviews of literature. We however limited our inclusion criteria for the year of publication or completion of the drafts of the grey literature. For study design, we included any study that sufficiently describes the application of AI, ML and DL on SI. On selection criteria based on population and geographical location, we had no restrictions, meaning that studies from high-, middle- and low-income countries were eligible for inclusion. For thematic focus, we included studies that explicitly describe or propose either conventional or new ways of measuring the welfare or poverty levels of populations or proxies for doing so within social science disciplines.

We gathered papers from multiple sources using different search words, phrases and topics related to the subject of the study. We focused on the use of SI or data, prediction of socioeconomic welfare indicators within the timeframe specified earlier. Since our interest was on both peer-reviewed papers and grey literature, we did not restrict our search to any specific search engines. However, we accessed papers on the former from Google Scholar and ScienceDirect. Our final database from this search comprised 60 papers from peer-reviewed journal articles, preprints, conference presentations and working papers and other grey literature.

A. Architecture Workflow

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.

Country	Number of Images			
	Training	Validation	Testing	Total
Ethiopia	90%	15% of train	10%	8,590
Malawi	90%	15% of train	10%	12,700
Mali	90%	15% of train	10%	14,759
Nigeria	90%	15% of train	10%	11,551
Total	Train % total	15%	10% total	47,600

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.

B. 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.

Dataset: The datasets were satellite images from [1] Kaggle website captured using satellite technology, as the details of the satellite pictures are shown in the table below. The 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 I. SHOWING THE TOTAL NUMBER OF IMAGES AND THEIR TRAINING AND TESTING SETS

Table I. The 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 traintest split is 90:10%

C. 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

Fig. 1 Showing CNN Classifier [8]

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.

By feeding each model with the augmented images, 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., convolutional neural network (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.

Data preparation

Some studies report multiple results for the same metric used in analysing the target outcome, making it difficult to record results for such studies as a single variable. The multiple results may be estimated and reported for different datasets, models, study locations, years or a combination of these. To identify the actual performance of models, where separate results are presented for training and validation datasetswe recorded the results for the latter. For results of different models, we recorded the results of the model with the highest precision (e.g.). Where separate results are reported for satellite and ground truth data, the former is used Where several models are run to observe control effects, the results of the full model are used .

A number of studies looked at several target outcome variables. Where studies report multiple indicators, we captured the main target welfare outcome of interest specified or inferred by the authors. Unless otherwise specified, we report the composite target where there is one among the targets of a study. These are usually related to indices of poverty, inequality and related measures or proxies of welfare. And in the absence of such explicit targets as described, we capture data on the target closest or related to the indices of economic wellbeing. The primary targets include asset wealth index, poverty rates, socioeconomic status and slum mapping. Others target socioeconomic indicators (which were rarely captured in our study because of our focus on the main indicators), including access to electricity, NTL, access to water, access to a toilet, educational attainment, monetary income, body mass index and .These are used as proxies for measuring poverty and inequality and yet are not readily observable from daytime SI.

We define target as the outcome variable or what each paper tries to estimate or predict. The target indicators are measured at different levels: individual, household, neighbourhood, village, enumeration area, and so forth. Some authors used multiple indicators for measuring their poverty/welfare outcome variable. In such cases, we extracted the main target outcome reported in such papers for the purpose of our regression model. These were referenced in the title of the study, reported in the abstract or in the main text of the papers.

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.

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 dropouts. 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. 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 crossentropy 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.

Our review method can be described as integrative rather than systematic This body of knowledge, mixing preprints, working papers, technical reports, peer-reviewed papers and conference papers with contributions of various disciplines is notoriously difficult to capture with one single approach. We used of the first publications to examine the application of ML and SI for measuring poverty and economic well-being—as the benchmark and narrative for our analysis. On this basis, papers completed prior to 2014 and do not apply ML to study socioeconomic wellbeing from SI were excluded from the study. We include literature from published journal articles, grey literature such as working papers and validation studies that have clear empirical application; that is, we excluded reviews of literature. We however limited our inclusion criteria for the year of publication or completion of the drafts of the grey literature. For study design, we included any study that sufficiently describes the application of AI, ML and DL on SI. On selection criteria based on population and geographical location, we had no restrictions, meaning that studies from high-, middle- and low-income countries were eligible for inclusion. For thematic focus, we included studies that explicitly describe or propose either conventional or new ways of measuring the welfare or poverty levels of populations or proxies for doing so within social science disciplines.

We gathered papers from multiple sources using different search words, phrases and topics related to the subject of the study. We focused on the use of SI or data, prediction of socioeconomic welfare indicators within the timeframe specified earlier. Since our interest was on both peer-reviewed papers and grey literature, we did not restrict our search to any specific search engines. However, we accessed papers on the former from Google Scholar and ScienceDirect. Our final database from this search comprised 60 papers from peer-reviewed journal articles, preprints, conference presentations and working papers and other grey literature.

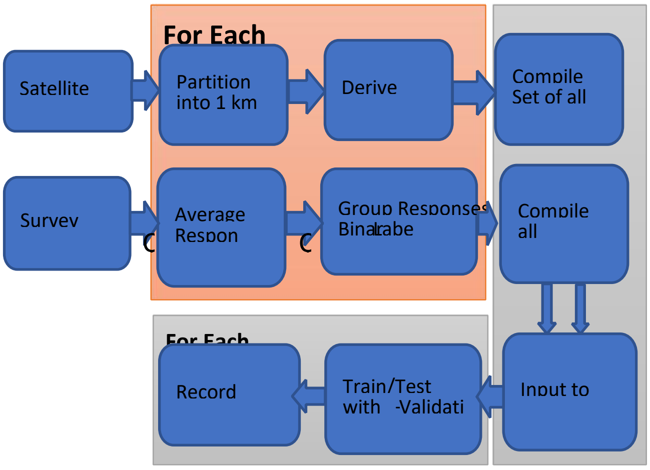

Fig. 2 Showing CNN satellite image and survey Procedure

TABLE II. SHOWING THE MODELS PARAMETERS

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

IV. FINDINGS

Finding is the chapter that displays the result of the study by showing any findings or result from the work and development activity that can be initiated or improved product, design, process, invention, innovation, and advancement in any procedure, approach, methodology, technique, apparatus or the machine.

The study examines the dynamic nature of movements into and out of poverty over a period when poverty has fallen substantially in India. The analysis identifies people who escaped poverty and those who fell into it over the period 2005 to 2012. The analysis identifies people who escaped poverty and those who fell into it over the period 2005 to 2012. Using panel data from the India Human Development Survey for 2005 and 2012, we find that the risks of marginalized communities such as Dalits and Adivasis of falling into or remaining in poverty were higher than those for more privileged groups. Some, but not all of these higher risks are explained by educational, financial, and social disadvantages of these groups in 2005.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

People escape poverty differ from those that push people into it and that the strength of their effects varies.

The multi-class classification of the study was done using the three models, namely CNN, ResNet, and VGG16.

Figure3 shows the results that CNN linear model does okay 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.

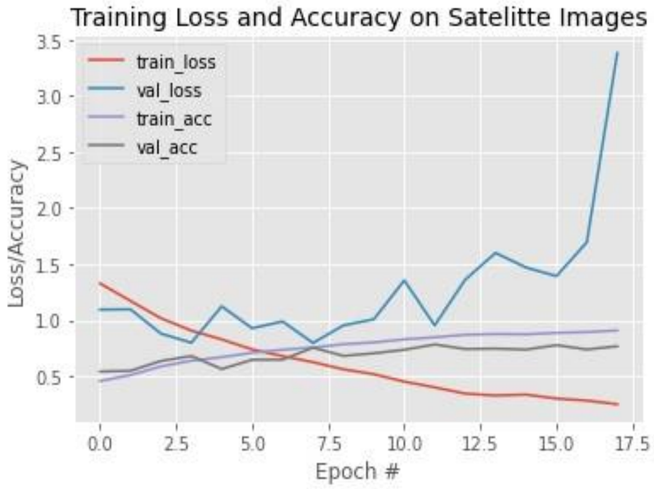

Fig. 3 Showing the performance of CNN model loss and accuracy.

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.

In terms of VGG16 shown in figure 4, the training accuracy of 94%, while its validation accuracy is 86%. The training and validation accuracies are shown in the Figures below.

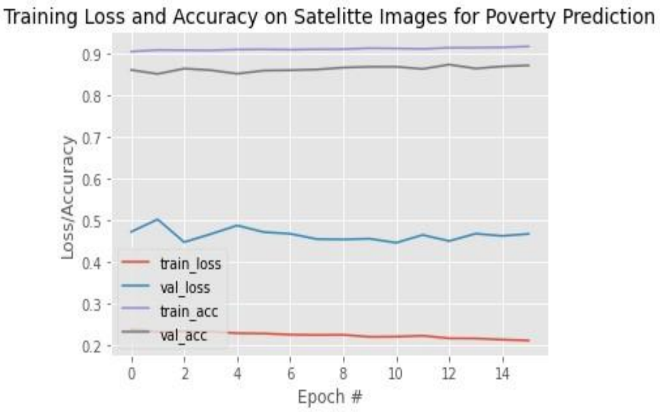

Fig. 4 Showing the performance of VGG16 model loss and accuracy.

The VGG16 also has variations between the training and validation losses, as shown in Figure 4 above.

The ResNet50 model in figure 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.

Hyper-Parameters	Models		
	CNN	VGG16	ResNet50
Learning-Rate	1e-4	1e-4	1e-4
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

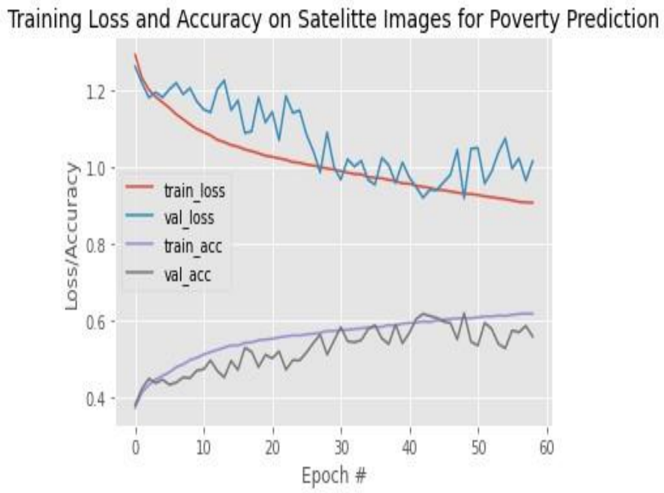

Fig. 5 Showing the performance of ResNet model loss and accuracy. The ResNet50 model shows that the training and validation losses are on a closed curve, as shown in Figure

TABLE III. SHOWING THE CNN MODEL RESULTS

Table III. 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.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

TABLE IV. SHOWING 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 above Table IV. 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 V. SHOWING 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 above Table V 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.

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

These **weighted average** were determined by multiplying every patterns prediction by the algorithm's model weight to produce the weighted total, then by dividing the value by the sum of the weights. The **macro average** was calculated straightforward using the normal averaging methods. The macro average F1 score is determined using the arithmetic mean (sometimes known as unweighted mean) of all the per class F1 scores. The countries were classified where zero was assigned to Malawi as a country with the lowest economic status, one was assigned to Mali as a country with a medium economic status, two was assigned to Ethiopia as the country with the high economic status, while three 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.

V. CONCLUSION

The present paper set out to analyse the state of the art at the intersection of the application of cluster of ML and DL tools on SI to predict poverty or welfare. Key findings from this nascent but rapidly growing field suggests the following. First, our finding that the relationship between the mean spatial resolution of the individual studies and their predictive power is not statistically significant was quite surprising. Conventional wisdom holds that higher resolution SI would contain more abundant information about the landscape and its features that could be correlated with economic activityIt then stands to reason that training datasets based on such higher resolution imagery would produce more accurate prediction and produce models that have higher predictive power.Our result suggesting a positive but statistically insignificant relationship between spatial resolution and accuracy has important implications. It suggests, for instance, that previous poor results achieved were down to other factors than the unavailability of higher spatial resolution satellite data per se. For researchers, this implies that going forward, additional resources would not need to be expended to acquire higher resolution imagery, which are often only commercially available at high cost and that publicly available SI would suffice in most cases.

While we do not find any evidence of a statistically significant effect that prediction performance using this approach increases over time, the contrary. It must be noted, however, that they assessed the performance of these approaches within the specific domains of smallholder agriculture, economic livelihoods, population and informal settlements. They also attribute the improving performance they measure to three main factors: more creative application of advances in computer visions, more abundant and higher quality SI, and more numerous and accurate training datasets. The latter jives with our findings, which suggest that the number of datasets used is positively and statistically significant for prediction performance. In this vein, the increasing proliferation of more accurate and higher quality training datasets portends well for this field of scholarship. Most studies in this area previously relied more heavily on NTLs datasets with coarse spatial resolutions (1 km/pixel) for estimating the level of welfare or development. The cluster of ML approaches recently applied in this intersection has proven to significantly improve predictions that could be achieved using NTL. For example, one of the most important findings from that NTLs tend to perform relatively poorly compared to daytime imagery in predicting asset wealth, largely because the former does not vary sufficiently in poor regions. The review also notes the limited downstream application, which it attributes, in part, to the novelty of the approaches and their lack of interpretability. With regards to the latter, explainable AI is the next rung in the ladder of applying ML to everyday social development issues such as poverty analysis. This requires transparency in model buildingWe argue that a necessary, even if insufficient, condition for the development of transparent, explainable and interpretable rather than black box ML models is adequate domain knowledge, which, in turn, requires co-option of development researchers and practitioners.

Other ML models might also contribute to better results in this field. Long Short-Term Memory (LSTM) networks, a class in the recurrent neural networks (RNNs) family, have been shown to be, capable of learning order dependence in sequence prediction problems. This quality makes them most useful in complex machine vision tasks such as predicting poverty from SI. However, thus far, LSTMs have seen limited application in this domain as none of our 60 reviewed papers employed this type of network; CNNs are the most common. Among our final 60 papers, 48% (29 of 60) of the studies employed some form of NN, with 21 of these studies using CNNs as the main model. The limited application of LSTM networks in studies at the intersection of poverty, SI and ML is not too surprising though, as CNNs are predominantly useful for spatial predictions while RNNs such as LSTM networks are more effective at capturing temporal predictions without suffering optimization hurdles, which tend to plague other RNNs, among other advantages, makes LSTM networks ideal candidate models for higher dimensional data analysis tasks such as handwriting recognition, and video and imagery analysis.

More recently, LSTM networks are being used to further enhance the already impressive prediction results achieved by traditional NNs. For example, LSTM models have been applied in the field of infectious diseases to predict the spread of the coronavirus in Bangladesh Similarly, in an innovative approach of combining CNNs and LSTMs, used a hybrid approach to achieve even better prediction results for surface erosion rates prediction at a significantly reduced time and computational costs. This holds great promise for the poverty-satellite data-ML domain of research, and we look forward to more studies adopting this hybrid ML approach in this niche area of research.

Our finding of significantly higher predictive power of models that are based on visible features, the so-called ‘hard’ indicators, is instructive even if unsurprising given that ‘soft’ indicators such as income levels, expenditure or the quantity of meat consumed by a household, for example, are more difficult to estimate from an SI than the existence and size of buildings or quality of roofing in a scene. In this sense, these models may be grouped into feature-based algorithms—those that rely on quantifiable geospatial features such as the number

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

of building, length of road, number of junctions and image-based models—those that can recognize the qualitative characteristics of these features .The choice of either of these then comes down to the resolution of the satellite data available since lower resolution SI tends to provide more information about the spatial context such as whether the data are from a rural or urban landscape while higher resolution SI is more useful for extracting the qualitative characteristics of the features .This suggests that the increasingly more accurate results obtained by studies in this area of scholarship could be driven more by a combination of improving the spatial resolution of readily available SI data and, perhaps more importantly, the potency and effectiveness of new tools, and approaches as well as the computational power to implement these as contend. As an illustration, show the importance of 'hard' indicators such as infrastructure (rail tracks and bridges) as well as other physical features like vehicles, street lights and billboards as indicators of the existence of services and development, and by extension, welfare. It is little wonder that the recent trend is one of combining feature-based and image-based approaches in either sequential or complementary manner to predict welfare .

REFERENCES

[1] https://landsat.gsfc.nasa.gov/. Accessed: 2021-09-12.

[2]https://earthdata.nasa.gov/earth-observation-data/near-real-time/download-nrt-data/viirs-nrt. Accessed: 2021-09-12.

[3] https://data.worldbank.org/. Accessed: 2021-09-12.

[4https://developers.google.com/earth-engine. Accessed: 2021-09-12.

[5] A. Albert, J. Kaur, and M. C. Gonzalez. Using convolutional networks and satellite imagery to identify patterns in urban environments at a large scale. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 1357–1366. ACM, 2017.

[6] N. Audebert, B. Le Saux, and S. Lefèvre. Joint learning from earth observation and openstreetmap data to get faster better semantic maps. In EARTHVISION 2017 IEEE/ISPRS CVPR Workshop. Large Scale Computer Vision for Remote Sensing Imagery, 2017.

[7] T. Chai and R. R. Draxler. Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature. Geoscientific Model Development, 7(3):1247–1250, 2014.

[8] P. R. Dave and H. A. Pandya. Satellite image classification with data augmentation and convolutional neural network. In Lecture Notes in Electrical Engineering, Lecture notes in electrical engineering, pages 83–92. Springer Singapore, Singapore, 2020.

[9] dpicampaigns and fangweizhao. Goal 1: End poverty in all its forms everywhere. https://www.un.org/sustainabledevelopment/poverty/. Accessed: 2021-09-27.

[10] J. Han, D. Zhang, G. Cheng, N. Liu, and D. Xu. Advanced deep-learning techniques for salient and category-specific object detection: A survey. IEEE Signal Processing Magazine, 35(1):84–100, 2018.

[11] N. Jean, M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon. Combining satellite imagery and machine learning to predict poverty. Science, 353(6301):790–794, 2016.

[12] M. I. Jordan and T. M. Mitchell. Machine learning: Trends, perspectives, and prospects. Science (New York, 349(6245):255, 2015.

[13] S.-W. Kim, A. Heckel, S. McKeen, G. Frost, E.-Y. Hsie, M. Trainer, A. Richter, J. Burrows, S. Peckham, and G. Grell. Satellite-observed us power plant nox emission reductions and their impact on air quality. Geophysical Research Letters, 33(22), 2006.

[14] S. Li and W. Deng. Deep facial expression recognition: A survey. IEEE Transactions on Affective Computing, page 1–1, 2020.

[15] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár. Microsoft coco: Common objects in context, 2015.

[16] M. Mirza and S. Osindero. Conditional generative adversarial nets. Nov. 2014.

[17] Z. Nabulsi, A. Sellergren, S. Jamshy, C. Lau, E. Santos, A. P. Kiraly, W. Ye, J. Yang, R. Pilgrim, S. Kazemzadeh, J. Yu, S. R. Kalidindi, M. Etemadi, F. Garcia-Vicente, D. Melnick, G. S. Corrado, L. Peng, K. Eswaran, D. Tse, N. Beladia, Y. Liu, P.-H. C. Chen, and S. Shetty. Deep learning for distinguishing normal versus abnormal chest radiographs and generalization to two unseen diseases tuberculosis and COVID-19. Sci. Rep., 11(1):15523, Sept. 2021.

[18] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. PyTorch: An imperative style, high-performance deep learning library. Dec. 2019.

[19] A. Perez, C. Yeh, G. Azzari, M. Burke, D. Lobell, and S. Ermon. Poverty prediction with public landsat 7 satellite imagery and machine learning. 11 2017.

[20] A. Perez, C. Yeh, G. Azzari, M. Burke, D. Lobell, and S. Ermon. Poverty prediction with public landsat 7 satellite imagery and machine learning. Nov. 2017.

[21] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning, 2017.

[22] S. Ruder. An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747, 2016.

[23] E. Rusak, L. Schott, R. S. Zimmermann, J. Bitterwolf, O. Bringmann, M. Bethge, and W. Brendel. A simple way to make neural networks robust against diverse image corruptions, 2020. 13

[24] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. Imagenet large scale visual recognition challenge, 2015.

[25] M.-A. Schulz, B. T. T. Yeo, J. T. Vogelstein, J. Mourao-Miranada, J. N. Kather, K. Kording, B. Richards, and D. Bzdok. Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets. Nat. Commun., 11(1):4238, Aug. 2020.

[26] C. Shorten and T. M. Khoshgoftaar. A survey on image data augmentation for deep learning. J. Big Data, 6(1), Dec. 2019. [27] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15(56):1929–1958, 2014.

[28] V. Sze, Y.-H. Chen, T.-J. Yang, and J. Emer. Efficient processing of deep neural networks: A tutorial and survey, 2017.

[29] L. S. Tusting, D. Bisanzio, G. Alabaster, E. Cameron, R. Cibulskis, M. Davies, S. Flaxman, H. S. Gibson, J. Knudsen, C. Mbogo, F. O. Okumu, L. von Seidlein, D. J. Weiss, S. W. Lindsay, P. W. Gething, and S. Bhatt. Mapping changes in housing in sub-saharan africa from 2000 to 2015. Nature, 568(7752):391–394, Apr. 2019.

[30] B. Uzkent, E. Sheehan, C. Meng, Z. Tang, M. Burke, D. Lobell, and S. Ermon. Learning to interpret satellite images in global scale using wikipedia. May 2019.

[31] J. Wang, Q. Qin, Z. Li, X. Ye, J. Wang, X. Yang, and X. Qin. Deep hierarchical representation and segmentation of high resolution remote sensing images. In Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International, pages 4320–4323. IEEE, 2015.

[32] W. Wang, H. Cheng, and L. Zhang. Poverty assessment using dmsp/ols night-time light satellite imagery at a provincial scale in china. Advances in Space Research, 49(8):1253–1264, 2012.

[33] G.-S. Xia, X. Bai, J. Ding, Z. Zhu, S. Belongie, J. Luo, M. Datcu, M. Pelillo, and L. Zhang. Dota: A large-scale dataset for object detection in aerial images. In Proc. CVPR, 2018.

[34] M. Xie, N. Jean, M. Burke, D. Lobell, and S. Ermon. Transfer learning from deep features for remote sensing and poverty mapping. Sept. 2015.

[35] C. Yeh, A. Perez, A. Driscoll, G. Azzari, Z. Tang, D. Lobell, S. Ermon, and M. Burke. Using publicly available satellite imagery and deep learning to understand economic well-being in africa. Nat. Commun., 11(1):2583, May 2020. [36] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee. A survey of modern deep learning based object detection models, 2021.