Spatial Information Research
https://doi.org/10.1007/s41324-024-00584-y

Advancing COVID-19 poverty estimation with satellite imagery-based deep learning techniques: a systematic review

Shruti Mishra[1] · Sandeep Kumar Satapathy[2] · Sung-Bae Cho[2] · Sachi Nandan Mohanty[3] · Sweeti Sah[4] · Shweta Sharma[4]

Received: 9 September 2023 / Revised: 23 April 2024 / Accepted: 23 April 2024

Abstract

In today's world, where the global population is expanding at an unprecedented rate, addressing the challenge of poverty has become more critical than ever before. In the wake of the COVID-19 pandemic, the issue of poverty has taken on renewed urgency as communities worldwide grapple with the socioeconomic fallout. Amidst the difficulties posed by the COVID-19 crisis, innovative approaches are required to address the evolving nature of poverty in the context of a pandemic-stricken world. This study systematically reviews the usage of machine learning (ML) and deep learning (DL) methods for COVID-19 poverty estimation using satellite imagery, emphasizing the need for innovative approaches due to the growing global population and poverty levels. It assesses how ML and DL leverage diverse data sources, such as mobile phone records, satellite imagery, and household surveys, to identify poverty indicators and enhance analysis precision. This study identifies challenges, including data availability and model biases, and suggests future directions focusing on dynamic models and multidimensional COVID-19 poverty assessment. It highlights the implications for spatial information science, advocating for improved data integration and model transparency to support effective COVID-19 poverty alleviation policies.

Keywords Poverty estimation · Machine learning · Deep learning · Mobile phone data · Satellite imagery · COVID-19 impact

✉ Sachi Nandan Mohanty
sachinandan09@gmail.com

Shruti Mishra
shrutim2129@gmail.com

Sandeep Kumar Satapathy
sandeepkumar04@gmail.com

Sung-Bae Cho
sbechoo@yonsei.ac.kr

Sweeti Sah
sweetisah3@nitkkr.ac.in

Shweta Sharma
shweta.sharma@nitkkr.ac.in

1 Centre for Advanced Data Science, Vellore Institute of Technology, Chennai, Tamil Nadu, India

2 Department of Computer Science, Yonsei University, 50 Yonsei-ro, Sudaemoon-gu, Seoul 03722, South Korea

3 School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh, India

4 Department of Computer Engineering, NIT Kurukshetra, Kurukshetra, Haryana, India

1 Introduction

Elimination of poverty is a major objective of the United Nations Sustainable Development Goals. However, the evolution of COVID-19 has resulted in a major setback in the quest, with widespread poverty rising for the first time in decades [1]. A study reveals a substantial rise in poverty due to the COVID-19 pandemic with the count of people falling into poverty increasing from 71 million to 88 million [2].

Machine Learning (ML) and Deep Learning (DL) have been an intriguing and intelligent approach to creating diverse fields for developing various statistical algorithms to predict COVID-19 cases [3]. However, measuring poverty presents a unique set of challenges, ranging from the selection of the data to the methods employed for predicting and forecasting. Researchers have continued to predict poverty using economic modeling where accuracy strongly depends on the model's assumptions [4]. However, recent developments have shown that using ML and DL can be a promising alternative to traditional approaches for COVID-19 poverty

Published online: 07 May 2024

estimation [5]. In recent times, researchers have been gathering image data from satellites and utilizing DL techniques to gauge COVID-19-related poverty [6, 7].

Thus, this survey presents a crucial examination of the intersection between the United Nations Sustainable Development objective and the worsening of poverty levels due to the COVID-19 pandemic. This study underscores the complexity of defining and measuring poverty, which varies significantly across regions due to differing thresholds and indicators. This study emphasizes the importance of continuous and updated data collection for effective COVID-19 poverty analysis, spotlighting the evolution from traditional census and household surveys to innovative deep learning techniques. These methodologies promise enhanced predictive power and analytical depth by utilizing diverse data sources, including surveys, satellite images, and mobile network data. Such advancements aim to overcome the limitations of conventional poverty measurement tools by providing a more dynamic and comprehensive understanding of COVID-19 poverty's scope and scale.

This survey aims to systematically review the application of ML and DL techniques in estimating and analyzing global poverty, with a focus on advancements introduced by satellite imagery in the context of the COVID-19 pandemic. By evaluating the potency of these methodologies, this survey seeks to illuminate the potential of ML and DL techniques in refining poverty estimation processes, thereby contributing to more informed and effective poverty alleviation policies.

The key aspect of this survey involves understanding and examining various ML and DL techniques along with various classification and regression models for COVID-19 poverty estimation. This study also presents satellite imagery data for temporal analysis of poverty dynamics and enhances spatial analysis capabilities to map socio-economic conditions accurately. This survey provides wide geographical coverage by including remote and inaccessible regions to ensure that data collection is not limited by physical accessibility. This study enables researchers to track changes and trends in COVID-19 poverty over time, assess the impact of interventions, and understand seasonal variations.

2 Overview

2.1 Approaches to poverty

A complex phenomenon, poverty is influenced by several factors that can be studied from different perspectives. Measuring poverty is not a simple task, as it requires an in-depth analysis of each factor that influences it and varies from country to country. There are several ways to measure poverty depending on the type of base information used. Poverty can be classified into two types, the first is objective and the other is subjective [8]. Objective poverty research makes use of information gathered through direct observation, like household income and expenditure. Subjective poverty research is based on an individual's perception of their situation.

These categories can be further broken down based on the scale of reference used to measure poverty. Absolute poverty is a scenario where a person's basic needs like food and clothing are not sufficient. Absolute poverty is omnipresent; the concept can be seen in all countries and communities [9]. Measuring absolute poverty is a herculean task and is extremely difficult owing to the inconsistent and dynamic nature of data. Relative poverty studies a particular society. To put this concept into perspective, a person is considered to be affected by poverty if they are disadvantaged in comparison to the rest of their society. This could be financially or through other related factors.

For instance, assume the case of two countries A and B. Country A considers people to be in poverty if they have an annual income of less than \$1000, whereas Country B has a bar of \$3000 annually. A poor person in country B may not be categorized as poor in country A. Rather than being considered a static concept, poverty is something that has a different definition from country to country [10].

2.2 Measurements of poverty used for machine learning tasks

2.2.1 Poverty line

Poverty lines classify people into two categories- "poor" and "not poor" depending on which side of the bar they find themselves on. These lines are usually measured or determined by monetary factors like annual household income [11]. The conventional approach to setting a poverty line is to determine it as the minimum basic income and expenditure required to meet primary human needs. While it is a relatively simple approach that makes it suitable for quick measurement of poverty and comparison across different countries, there are a significant number of problems with the same [12]. A rigid income level does not always accurately represent different factors that may account for poverty apart from income. It does not possess the ability to represent health, education, housing, and other social factors. These measurements do not account for the number of years that people live under the poverty line, and can thus be misleading. Furthermore, poverty lines cannot distinguish or adjust differences in the price of living between rural and urban areas.

Researchers have studied the poverty line disparity in various countries by applying Machine Learning models to country-specific data. A study in Jordan [13] used household expenditure and income survey data to determine and supervise the poverty statuses of Jordan households. The work described results in an assessment of the chance of a household being poor, the poverty rate at a point in time, and changes in poverty based on the poverty line. A challenge to be noted here was the presence of unbalanced data, which could affect the training-testing ratio during model evaluation. The authors handled this by using random oversampling [14], which involved duplicating the data in the minority category to increase the volume.

Static machine learning models rely on pre-existing datasets for prediction and therefore face challenges due to evolving statistics, necessitating frequent updates and retraining. In contrast, dynamic machine learning models offer a solution by adapting to new data in real time through continuous learning and updates. These models are particularly suited for poverty prediction, given the annual updates in poverty data by organizations such as the World Bank. However, the high costs associated with gathering training data, exemplified by the $934 million spent on the 2021 USA national census, pose significant challenges for the development and performance of these models in poverty prediction efforts [15–17].

2.2.2 Household surveys

Since the establishment of the Living Standards Measurement Survey (LSMS) by the World Bank in 1979, household surveys have played a crucial role in measuring poverty around the globe [18]. They are currently the most reliable source of socio-economic data. The LSMS program focuses on involving data users in the design of surveys that can help assess the status of poverty in a country. In this section, the evolution of these surveys and the impact they have in a data-driven world are analyzed in detail.

Over the last century, developing and developed countries have been conducting national surveys dedicated to assessing the socio-economic status of their population regularly. One notable survey is the National Sample Survey [19] which has been in existence since 1950 and is conducted by the Government of India every year. Well-designed and properly executed surveys may be representative of a population's demographic characteristics like age, sex, and income, as well as ethnic and cultural composition [20]. However, generalization to the source population is dependent on the sampling frame and method. For instance, household surveys are usually restricted to the civilian population.

In 1996, Lanjouw and Lanjouw [21] proposed a sturdy technique to monitor poverty using household survey data from various sources or countries. They demonstrated the idea of a "food poverty line" and a "final poverty line" which allows for comparing headcount rate calculations from different surveys [22]. Measuring poverty through household surveys has its own set of challenges and limitations [23]. These surveys require a high expenditure and are difficult to update and maintain. It is not feasible to conduct them every year in most developing countries, especially without substantial extrinsic support, and the continuity of monitoring is lost. The number of households considered to be poor may only make up a fraction of the population which could hinder the generation of meaningful estimates of poverty. Data needs to be representative and cover different areas. Surveys tend to provide a narrow outlook on poverty since it is relevant to a particular period. Measuring poverty and tracking it requires analyzing changes over time and calculating increases or decreases in household income. Household surveys and questionnaires used in literature and the topics covered across different countries like Ethiopia, Nigeria, Africa and other global south countries [24–32] have been taken into consideration for examining how they have evolved changed over a period from 2007 to 2008 in terms of food, fuel and financial crisis based on the profile of undernutrition among children.

2.2.3 Multidimensional poverty index (MPI)

To overcome some of the limitations of household surveys [33], the Oxford Poverty and Human Development Initiative (OPHI) [34] developed a modern international measure of poverty called the Multidimensional Poverty Index (MPI) [35]. The MPI is measured using various socioeconomic indicators differing from health and education to income. MPI can be calculated as follows [36]. If H is the proportion of persons in multidimensional poverty and A is the average proportion of indicators in which households are deprived across the sample, then the measure of multidimensional poverty is shown in Eq. (1):

$$P = H * A \tag{1}$$

This index goes beyond a traditional focus to assess the nature and poverty intensity at an individual level. Deprivations are measured directly and are used extensively as an analytical tool. In the subsequent sections, literature using the MPI data to predict poverty using machine learning will be discussed.

3 Data description

This section delves into a thorough examination of the fundamental data collection methods pivotal for machine learning endeavors, alongside an analysis of published datasets.

3.1 Data collection for machine learning tasks

Data collection for machine learning tasks encompasses three primary methodologies for classification or prediction [37]. The first is Data Acquisition, wherein new data is obtained through various means such as scraping, collection, augmentation, or generation tailored to a certain use case. The next step includes labeling the acquired raw data, which is essential for training machine learning models effectively. Lastly, the process also involves improving existing data to enhance its quality and relevance.

In the pursuit of understanding socio-economic development, the availability of robust data indicators plays a pivotal role. Several researchers including Joliffe [38] introduced an innovative approach to track poverty using supplementary data sources beyond traditional household surveys. However, comparing data across different countries presents unique challenges, necessitating consistent data sources [39]. Given the diverse economies and operational frameworks worldwide, classifying countries under a unified poverty line is intricate. To address this, the World Bank's International Comparison Program (ICP) employs Purchasing Power Parities (PPP) as a standardized measure to gauge economic price level disparities globally [40]. Notably, the ICP computed annual PPPs for 176 countries in 2017, facilitating cross-country economic comparisons.

3.2 Benchmark and published datasets

Published datasets [41] provide a standard of measurement and comparison of poverty. These datasets have been collected through studies by organizations such as the World Bank, the Oxford Poverty and Human Development Initiative, and various national governments. These datasets vary in terms of features and indicators but have a common goal to make poverty easier to analyze. However, gathering data to measure poverty on a global scale presents significant challenges, particularly due to the scarcity of data in developing countries [42]. Thus, the World Bank has developed several datasets that are regularly updated and made available through its Open Data and MicroData platforms [43]. Another major source of poverty-related data is PovCalNet [44], which presents a country-wise view on features related to poverty, such as poverty gap, poverty line, population, and other relevant indicators. Estimating and measuring poverty at a country level is often carried out through the country's government through national surveys or household questionnaires [45]. For example, Table 1 showcases a sample survey designed with questions specific to Afghanistan.

4 Methods

This section shows the methods employed by researchers for COVID-19 poverty prediction.

4.1 COVID-19 poverty prediction and classification using ML

Several ML techniques such as Naive Bayes, K-Nearest Neighbors, Decision Trees, Random Forest, and gradient boosting are commonly applied by researchers for COVID-19 poverty classification, which are discussed as follows:

4.1.1 Naive bayes

The Naive Bayes classifier, illustrated in Fig. 1., stands out as one of the remarkably utilized supervised machine learning algorithms for classification tasks due to its ability to simplify complex problems. Its applications span a broad range of areas, including disease prediction, educational technology, malware detection, and fake news identification [46]. Particularly in the context of COVID-19 poverty classification using survey data, Naive Bayes is known for its high accuracy levels. The classifier's effectiveness in scenarios with limited training data, such as those often found in developing countries, is a key factor in its success. For instance, in a study conducted in Bantul, the Naive Bayes algorithm was employed to categorize families as poor or not based on indicators such as food availability, clothing, and housing conditions.

Table 1 Sample of Afghanistan NRVA survey

Question	Type of attribute
In the last summer cultivation season, how many jeribs of irrigated land did you or your household rent out?	Numerical
How much has the household spent on Tahwiz/Shoyest (health talismans) over the past 12 months?	Numerical
According to the recent prices, how much do you think you could get if you sold all of *kilim, satrangi, namad, fash* (other carpet production)?	Numerical

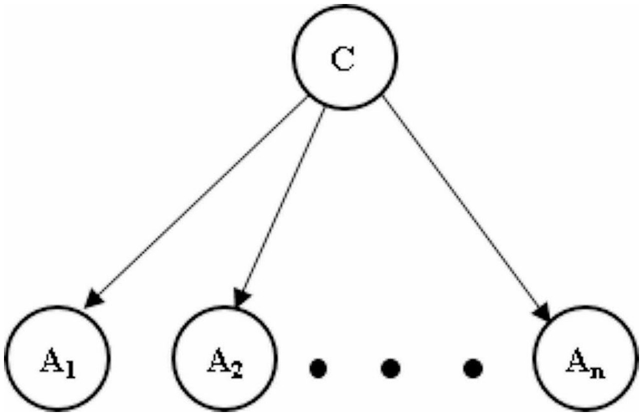


Fig. 1 Naive Bayes classifier

4.1.2 K-Nearest neighbor (KNN) and decision tree

The K-Nearest Neighbors (KNN) algorithm is ranked among the top algorithms for both classification and regression tasks [47]. A unique feature of KNN is that it retains all training data and only formulates a classification decision when required to classify a new sample, basing its predictions on the labels of the nearest neighbors to the test sample. KNN has been applied in studies for identifying COVID-19 poverty levels, providing valuable insights for governments in pinpointing groups in need of aid. However, the performance of KNN can be adversely affected by noisy data and missing values, which poses a challenge when dealing with inconsistent poverty data in developing countries.

On the other hand, the Decision Tree algorithm is notable for its capability to manage data without the need for normalization, and its robustness against missing values, making it particularly well-suited for COVID-19 poverty prediction tasks [48]. Techniques based on the ID-3 algorithm for poverty mapping offer considerable benefits, such as providing nearly unbiased estimates with reduced standard errors. Despite these advantages, the decision tree's limitation lies in its inability to predict continuous outcomes, which restricts its applicability in modeling the distribution of poverty across different regions.

4.1.3 Logistic regression and random forest

Logistic regression [49], a machine learning algorithm suitable for binary target variables as depicted in Fig. 2, has been widely used for micro-level COVID-19 poverty assessment. This approach allows for the analysis of both internal and external factors that may affect a household's likelihood of being categorized as poor. In a notable study conducted in Kenya, researchers utilized the Demographic and Health Surveys (DHS) data and applied Principal Component Analysis to develop an asset index. This logistic regression model successfully demonstrated a correlation between the DHS survey results and its findings, highlighting an inverse relationship between the level of education and poverty rates.

The Random Forest classifier, as shown in Fig. 3, operates through a collective approach involving numerous decision trees [50]. Each tree within the ensemble contributes a vote towards classifying a given sample, and the class acquiring the majority of votes is selected as the ultimate prediction by the model, akin to a democratic voting process. This algorithm excels in identifying the significance of variables, making it particularly adept at examining how various factors influence COVID-19 poverty levels. When applied to census data, Random Forests not only enhance the accuracy of COVID-19 poverty status classification but also underscore the importance of variables such as employment hours, educational attainment, and gender in determining poverty status. Utilizing Random Forests to extract relevant features for classification in developing countries offers insights into the shifting dynamics of economic activities over time.

Fig. 2 Logistic regression

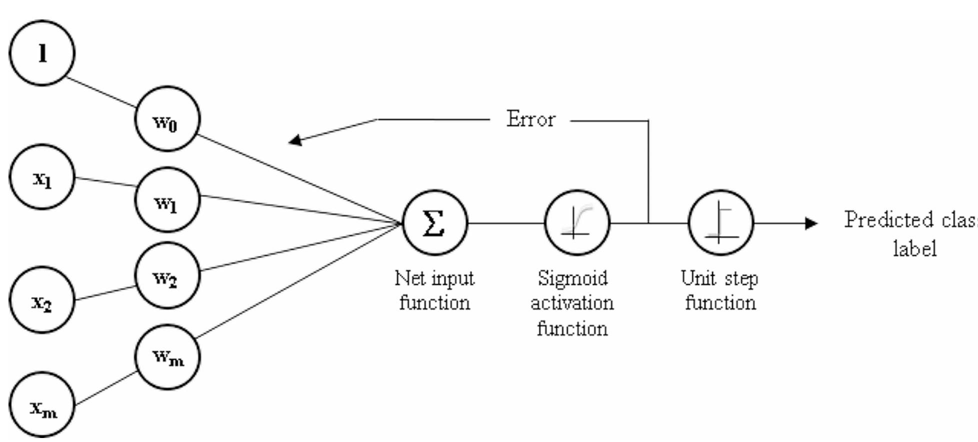
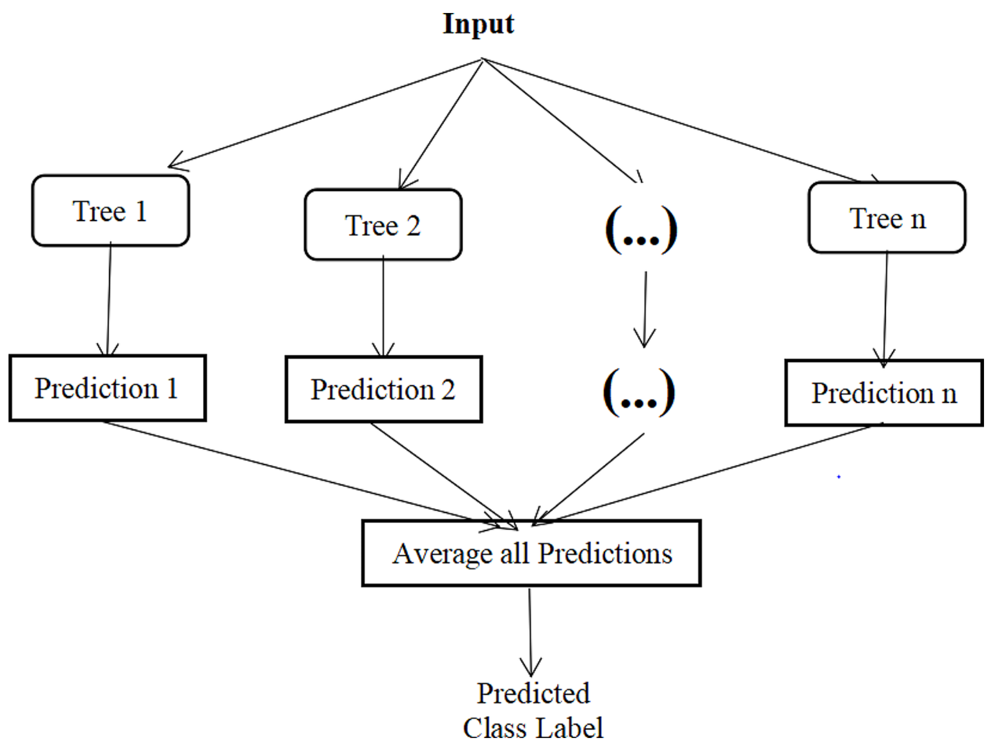


Fig. 3 Random forest



4.2 Poverty mapping using mobile data

Over the past two decades, mobile phones have become widespread due to their convenience and efficiency. This rise in mobile phone usage has led to an important increase in the production of metadata [51]. Utilizing this metadata to predict the well-being of mobile phone users offers a novel approach to monitoring COVID-19 poverty and implementing targeted interventions. This method facilitates the global and regional mapping of COVID-19 poverty distributions by leveraging data on communication patterns, social networks, travel behaviors, and consumption and expenditure histories captured through mobile phone usage.

Recent advances in this field have explored the integration of machine learning techniques with extensive call detail records (CDRs), offering a cost-effective alternative to traditional census and survey methods. Applying multivariate time-series models to CDR data reveals trends that support effective COVID-19 poverty monitoring and forecasting, which assists in identifying and aiding vulnerable groups.

Furthermore, selecting key features from CDRs is a viable strategy for COVID-19 poverty prediction, treating it as a classification issue. Techniques such as Support Vector Machines (SVMs) and Random Forests have been successful in providing precise COVID-19 poverty assessments based on mobile records.

4.3 Poverty mapping using satellite imagery

Satellite imagery capturing the Earth's surface at night reveals the glow of cities and towns, offering a unique perspective on human activity through the observation of nightlights. This method, known as remote sensing using nocturnal luminosity, enables the monitoring of human endeavors both at specific moments and over time. Consequently, images of night-time lights have become a valuable tool for delineating global socioeconomic trends, including shifts in economic activity and emissions of greenhouse gases. The correlation between the intensity of night-time illumination and economic dynamics has been well-documented, suggesting that areas with brighter night-time lights are often more economically active.

Elvidge et al. [52] have demonstrated through their research that night-time lights offer a robust and precise indicator for gauging development, proving to be an effective measure for various development-related variables. In the ambiance of developing countries, where the collection of ground data through surveys and similar methods faces challenges of consistency and accessibility, the luminosity captured from space provides a dependable alternative.

Moreover, researchers explored nightlight images from the VIIRS on the Suomi NPP satellite to determine the economic consequences of COVID-19 in India [6]. They developed a framework for analyzing the VNP46A1 radiance dataset and examined the correlation between night light intensity, electricity consumption, and GDP. The analysis revealed a significant link between these factors and accurately predicted a 24% GDP contraction in FY2020Q1, closely matching the official figures. This suggests that nightlight and electricity consumption are effective indicators for measuring the economic effects of sudden disruptions like the COVID-19 pandemic.

4.4 Poverty mapping using deep learning

Pioneering studies [53–59] employed DL techniques, especially Convolutional Neural Networks (CNNs), for poverty mapping and prediction using satellite imagery. These studies showcase the application of transfer learning, ResNet-18, VGG, GoogleNet architectures, and more, on datasets ranging from Google Static Maps API and DMSP-OLS Night-time Lights to high-resolution Landsat and Planet imagery. The research demonstrated the potential of satellite images in understanding socioeconomic conditions in regions like Africa, Mexico, and Thailand, highlighting the effectiveness of CNNs in capturing and analyzing data for poverty assessment. Each study contributed to advancing the methodology for poverty prediction, offering insights into economic well-being and poverty distribution through innovative ML approaches.

Additionally, Table 2 summarizes research that uses satellite imagery and employs deep learning techniques to estimate and map poverty levels, enhancing the clarity and accessibility of the reviewed studies in this domain.

5 Discussion

The review of ML and DL techniques for COVID-19 poverty estimation in this study highlights the revolutionary potential of sophisticated analytics in comprehending the dynamics of poverty. Various researchers [53–59] applied ML and DL techniques for COVID-19 poverty estimation. Remarkably, the use of satellite images [54, 58] for temporal and spatial analysis represents a breakthrough in precisely mapping and tracking socio-economic circumstances, particularly in isolated and unreachable regions. Researchers use satellite imagery to estimate poverty levels to achieve accuracy close to that of the costly and labor-intensive data collected manually on the ground [57].

Considering separate poverty rates with satellite imagery for urban and rural areas is beneficial in calculating the fraction of households living in poverty [55]. Using mobile phone datasets from developing countries and applying semi-supervised learning by incorporating multi-view graph convolutional networks shows remarkable performance in poverty research [56]. Night-time lights facilitate its use as an objective measure for tracking and assessing COVID-19 poverty and help in the decision-making procedure and distribution of resources to areas most in need [52]. With the availability of high resolution time-series of day-time imagery, the researchers evaluate the ability of the transfer learning to forecast changes in economic well-being overtime at the specific locality [53].

Our study helps to understand how poverty changed during the pandemic. Researchers can benefit greatly from the ability to examine poverty trends across time using such methods since it makes it possible to establish timely, efficient, and targeted initiatives for reducing poverty. This survey pushes the limits of poverty estimating techniques by emphasizing the crucial role that combines technical innovations with conventional socio-economic research in addressing the complex difficulties provided by COVID-19.

6 Conclusion

This research reviews the application of ML and DL techniques to COVID-19 poverty estimation, highlighting the evolution of data sources from household surveys to mobile and satellite data. It discusses the adaptability of models to dynamic socio-economic factors influenced by the COVID-19 pandemic and underscores the significance of incorporating behavioral and psychological aspects into poverty estimation. One of the key contributions of this research is the demonstration of the efficacy of ML and DL in overcoming the drawbacks of traditional poverty estimation methods, which often rely on outdated or sparse data. The integration of diverse data sources such as satellite imagery has enhanced the accuracy of poverty maps and facilitated a more nuanced understanding of poverty dynamics. This approach allows for real-time analysis and monitoring, which is crucial for the timely implementation of poverty alleviation strategies. However, the study acknowledges the inherent challenges in applying ML and DL to COVID-19 poverty estimation. Data availability, quality, and representativeness pose significant hurdles, alongside ethical considerations regarding privacy and data security. Moreover, the potential biases embedded within ML and DL algorithms necessitate rigorous validation and calibration to ensure the reliability of the outcomes.

Future research focuses on developing models that are sensitive to these varied dimensions, aiming to offer actionable insights for effective poverty alleviation policies, thus contributing to the broader goal of sustainable development.

Table 2 Summary of key works in poverty mapping using deep learning techniques

Reference	Author(s)	Title	Application	Model(s) Used	Data	Description of the method
[53]	Neal Jean, Marshall Burke, Michael Xie et al.	"Combining satellite imagery and machine learning to predict poverty"	Poverty Prediction	CNN on ImageNet	Google Static Maps API, DMSP-OLS Night-time Lights Time Series	Proposes a transfer learning approach to determine poverty in African countries, using a CNN to forecast nighttime light intensities corresponding to input daytime satellite imagery.
[54]	Christopher Yeh, Anthony Perez, Anne Driscoll, et al.	"Using publicly available satellite imagery and deep learning to understand economic well-being in Africa"	Estimating economic well-being	CNN with ResNet-18 architecture	Landsat 5,7,8 satellite images	To cover areas where survey data may not be available, this work trains a CNN on high-resolution satellite imagery after constructing an asset wealth index for African countries
[55]	Boris Babenko, Jonathon Hersh, David Newhouse, et al.	"Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, With an Application in Mexico"	Poverty Mapping	CNN with VGG and GoogleNet architectures	Planet and Digital Globe imagery	Introduces the concept of poverty mapping for municipalities in Mexico and can account for 61% of poverty variation
[56]	Muhammad Raza Khan, Joshua E. Blumenstock	"Multi-gcn: Graph convolutional networks for multi-view networks, with applications to global poverty"	Poverty Prediction	Graph CNN	Phone Logs (East Africa)	Explores and develops a novel multi-view graph convolutional network with a significant improvement from existing literature in the poverty prediction task
[57]	Shailesh M. Pandey, Tushar Agarwal, Narayanan C Krishnan	"Multi-Task Deep Learning for Predicting Poverty from Satellite Images"	Poverty Prediction	Fully convolutional CNN	Google Static Maps API	This work employed a fully convolutional CNN for the prediction of poverty by detecting
[58]	Asian Development Bank- Arturo Martinez Jr, Marymell Martillan, et al.	"Mapping the Spatial Distribution of Poverty using Satellite Imagery in Thailand"	Poverty Mapping	CNN with Ridge Regression	Landsat 8 images, Sentinel data	Thoroughly explores and analyzes poverty indicators through the use of a CNN on data for the region of Thailand. Uses a "cyclical learning rates" approach to avoid trial and error when finding the optimal values and schedule of global learning rates
[59]	Zhaozhuo Xu, Zhihan Jiang and Yicheng Li	"Poverty Prediction by Selected Remote Sensing CNN Features"	Poverty Prediction	CNN converted from VGG-F, pre-trained on ImageNet	DHS Survey data (5 African countries) and corresponding night light intensity data	The work uses regression models like linear, ridge regression, and XGBoost to test selected economic features. A forward search was employed effectively as a feature selection method.

Declarations

Conflict of interest The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

References

1. United Nations (The Sustainable Development Goals Report 2021). 2021, Accessed: 2021-01-20, https://unstats.un.org/sdgs/report/2022/.
2. Sumner, A., Ortiz-Juarez, E., & Hoy, C. (2022). Measuring global poverty before and during the pandemic: a political economy of overoptimism, Third World Quarterly, vol. 43, pp. 1–17, https://doi.org/10.1080/01436597.2021.1995712.
3. Sah, S., Surendiran, B., Dhanalakshmi, R., & Yamin, M. (2023). Covid-19 cases prediction using SARIMAX Model by tuning hyperparameter through grid search cross-validation approach, Expert Systems, 40, pp. 1–21, https://doi.org/10.1111/exsy.13086.
4. Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata, Science, vol. 350, pp. 1073–1076, https://doi.org/10.1126/science.aac4420.
5. Dash, S., Chakravati, S., Mohanty, S. N., Patnaik, C. R., & Jain, S. (2021). A deep learning method to forecast Covid-19 outbreak. *New Generation Computing*, *39*(2), 437–461. https://doi.org/10.1007/s00354-021-00129-z. ISSN: 02883635

6. Dasgupta, N. (2022). Using satellite images of nighttime lights to predict the economic impact of COVID-19 in India. *Advances in Space Research*, *70*, 863–879. https://doi.org/10.1016/j.asr.2022.05.039.
7. Minetto, R., Segundo, M. P., Rotich, G., & Sarkar, S. (2021). Measuring Human and Economic Activity From Satellite Imagery to Support City-Scale Decision-Making During COVID-19 Pandemic, *IEEE Transactions on Big Data*, 7, 1, 56–68, doi: https://doi.org/10.1109/TBDATA.2020.3032839.
8. Instituto Nacional De Estadistica Poverty and its Measurement, Accessed: 2021-01-17.
9. Eskelinen, T., Poverty, A., & Chatterjee, D. K. (Eds.). (2011). *Encyclopedia of Global Justice*. Springer. https://doi.org/10.1007/978-1-4020-9160-5_178.
10. Deonandan, R. (2019). Defining poverty: A Summary of competing models. *Journal of Social and Political Sciences*, *2*, 17–21. https://doi.org/10.31014/aior.1991.02.01.44.
11. Greeley, M. (1994). Measurement of poverty and poverty of measurement. *IDS Bulletin* 25.
12. Kakwani, N. (2003). Issues in Setting Absolute Poverty Lines, Asian Development Bank.
13. Alsharkawi, A., Al-Fetyani, M., Dawas, M., Saadeh, H., & Alyaman, M. (2021). Poverty classification using machine learning: The case of Jordan. *Sustainability*, *13*(2021), 1412. https://doi.org/10.3390/su13031412.
14. Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020). Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results, pp. 243–248, https://doi.org/10.1109/ICICS49469.2020.239556.
15. Kara, A., Nikolic, M., Olteanu, D., & Zhang, H. (2021). Machine learning over static and dynamic Relational Data. *International Conference on Distributed and Event-based Systems*, *ACM*, 160–163. https://doi.org/10.1145/3465480.3467843.
16. Griffin, P., & The World Bank., *New Political Economy* 11 (2006), https://doi.org/10.1080/13563460600991028.
17. United States Census, U.S Census Bureau's Budget Fiscal Year 2021, Accessed: 2021-01-20.
18. Alkire, S., Roche, J., & Sumner, A. (2013). Where Do the World's Multidimensionally Poor People Live? Working Papers 61, Oxford Poverty & Human Development Initiative (OPHI).
19. Indian Statistical Institute, The National Sample Survey. (1953). General Report 1. First Round: October 1950 - March 1951. *Sankhyā: The Indian Journal of Statistics (1933–1960)*, *13*(1/2), 47–214. http://www.jstor.org/stable/25048165.
20. Sadana, R., Mathers, C. D., Lopez, A. D., Murray, C. J. L., & Iburg, K. (2001). Comparative Analyses of More than 50 Household Surveys on Health Status, World Health Organization, GPE Discussion Paper Series 15.
21. Lanjouw, J., & Lanjouw, P. (1997). Poverty Comparisons with Noncompatible Data: Theory and Illustrations, Policy Research Working Paper, The World Bank.
22. Lanjouw, P., & Ravallion, M. (1996). How Should We Assess Poverty Using Data from Different Surveys? World Bank Joint Publication 3.
23. Lao Statistics Bureaua, & Bank, W. (2020). Poverty Profile in Lao PDR: Poverty Report for the Lao Expenditure and Consumption Survey 2018–2019.
24. Nandy, S., Daoud, A., & Gordon, D. (2016). *Examining the changing profile of undernutrition in the context of food price rises and greater inequality* (Vol. 149, pp. 153–156). Social Science & Medicine.
25. Daoud, A., Jordán, F., Sharma, M., Johansson, F., Dubhashi, D., Paul, S., & Banerjee, S. (2023). *Using Satellite images and deep learning to measure Health and Living standards in India, Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-life measurement* (Vol. 167, pp. 475–505). Springer. 1.
26. Shankar, K., Mohanty, S. N., Yadav, K., & Gopalakrishnan, T. (2021). Automated COVID-19 diagnosis and classification using convolutional neural network with fusion-based feature extraction model. Cognitive Neurodynamics, 16, Issue 1, doi.org/10.1007/s11571-021-09712-y. ISSN: 1871-4099.
27. Jerzak, C. T., & Johansson, F. (2023). A. *Daoud. Integrating Earth Observation Data into Causal Inference: Challenges and opportunities*. Cornell University.
28. Jerzak, C. T., Johansson, F., & Daoud, A. (2023). Image-based Treatment Effect Heterogeneity, Cornell University.
29. Jerzak, C. T., Johansson, F., & Daoud, A. (2023). *Estimating Causal effects under Image Confounding Bias with an application to poverty in Africa*. Cornell University.
30. Balgi, S., Pena, J. M., & Daoud, A. (2022). Personalized Public Policy Analysis in Social Sciences using causal-graphical normalizing flows. *Proceedings of the AAAI Conference on Artificial Intelligence*, *36*(11), 11810–11818.
31. Shome, D., Kar, T., Mohanty, S. N., Tiwari, P., Muhammad, K., AlTameem, A., Zhang, Y., Saudagar, & A. K. J. (2021). COVID-transformer: Interpretable COVID-19 detection using vision transformer for healthcare. *International Journal of Environmental Research and Public Health*, *18*(1), 1–14. https://doi.org/10.3390/ijerph182111086. ISSN: 1660-4601
32. Daoud, A., & Johansson, F. (2019). Estimating Treatment Heterogeneity of International Monetary Fund Programs on Child Poverty with Generalized Random Forest, SocArXiv Papers.
33. Wolff, E. N., & Distribution, W. (2015). International Encyclopedia of the Social & Behavioral Sciences (Second Edition), Elsevier, pp. 450–455, https://doi.org/10.1016/B978-0-08-097086-8.71017-8.
34. Alkire, S., Roche, J., & Sumner, A. (2013). Where Do the World's Multidimensionally Poor People Live? Working Papers 61, Oxford Poverty & Human Development Initiative.
35. Alkire, S., Santos, M. E., & Index, M. P. (2010). Oxford Poverty & Human Development Initiative (OPHI).
36. Seth, S., & Alkire, S. (2013). Measuring and Decomposing Inequality among the Multidimensionally Poor using Ordinal Variables: A Counting Approach, OPHI Working paper 68.
37. Roh, Y., Heo, G., Whang, S. E., & Perspective (2021). *IEEE Transactions on Knowledge and Data Engineering* 33(4), 1328–1347, doi: https://doi.org/10.1109/TKDE.2019.2946162.
38. Jolliffe, D., & Prydz, E. B. (2016). Estimating international poverty lines from comparable national thresholds. *J Econ Inequal*, *14*, 185–198. https://doi.org/10.1007/s10888-016-9327-5.
39. Kaminska, O., & Lynn, P. (2017). Cross-country comparisons where Countries Vary in Sample Design: Issues and solutions. *Journal of Official Statistics*, *33*, 123–136. https://doi.org/10.1515/jos-2017-0007.
40. World Bank, International Comparison Program (2021). Accessed: 2021-01-16, https://www.worldbank.org/en/programs/icp#1.
41. Dekker, R. (2006). The importance of having data-sets, 27th IATUL Conference.
42. Elahi, A. (2008). Challenges of data collection in developing countries - the Pakistani experience as a way forward. *Statistical Journal of the IAOS*, *25*(1,2), 11–17.
43. World Bank, World Bank Open Data (2021). Accessed: 2021-01-16, https://data.worldbank.org/.
44. Sah, S., Surendiran, B., Dhanalakshmi, R., Mohanty, S. N., Alenezi, F., & Polat, K. (2022). Forecasting COVID-19 pandemic using prophet, ARIMA, and hybrid stacked LSTM-GRU models in India. *Computational and Mathematical Methods in Medicine*, *2022*, Article ID 1556025. https://doi.org/10.1155/2022/1556025. ISSN: 17486718, 1748670X

45. Gao, C., Fei, C. J., McCarl, B. A., & Leatham, D. J. (2020). *Identifying Vulnerable Households Using Machine Learning Sustainability* 12(15). https://doi.org/10.3390/su12156002.
46. Talingdan, J. A. (2019). Performance Comparison of Different Classification Algorithms for Household Poverty Classification, 2019 4th International Conference on Information Systems Engineering (ICISE), pp. 11–15, https://doi.org/10.1109/ICISE.2019.00010.
47. Wu, X., Kumar, V., Quinlan, R., Ghosh, J., Yang, Q., Motoda, H., Mclachlan, G., Ng, S. K. A., Liu, B., Yu, P., Zhou, Z., Steinbach, M., Hand, D., & Steinberg, D. (2007). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*. https://doi.org/10.1007/s10115-007-0114-2.
48. Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*, 81–106.
49. Wright, R. E., Logistic regression, L. G., Grimm, & Yarnold, P. R. (Eds.). (1995). Reading and understanding multivariate statistics, pp. 217–244.
50. Breiman, L., & Forests, R. (2001). Machine Learning 45, pp. 5–32, https://doi.org/10.1023/A:1010933404324.
51. Satapathy, S. K., Saravanan, S., Mishra, S., & Mohanty, S. N. (2023). A comparative analysis of multidimensional COVID-19 poverty determinants: An observational machine learning approach. *New Generation Computing*, *41*(1). https://doi.org/10.1007/s00354-023-00203-8. ISSN: 02883635
52. Elvidge, C., Safran, J., Tuttle, B., Sutton, P., Cinzano, P., Pettit, D., Arvesen, J., & Small, C. (2007). Potential for global mapping of development via a nightsat mission. *Geojournal*, *69*, 45–53. https://doi.org/10.1007/s10708-007-9104-x.
53. Jean, N., Burke, M., Xie, M., Davis, W. M., & Lobell, D. B. S. Ermon, combining satellite imagery and machine learning to predict poverty. *Science* 353(6301), pp. 790–794, https://doi.org/10.1126/science.aaf7894.
54. Yeh, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D., Ermon, S., & Burke, M. (2020). Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nature Communications*, *11*(2583). https://doi.org/10.1038/s41467-020-16185-w.
55. Babenko, B., Hersh, J., Newhouse, D., Ramakrishnan, A., & Swartz, T. (2017). Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, With an Application in Mexico, in: 31st Conference on Neural Information Processing Systems (NIPS), arXiv:1711:06323v1.
56. Raza Khan, M., & Blumensrandon mtock, J. E. (2019). Multi-GCN: Graph Convolutional Networks for Multi-View Networks, with Applications to Global Poverty, The Thirty-Third AAAI Conference on Artificial Intelligence.
57. Pandey, S. M., Agarwal, T., & Krishnan, N. C. (2018). Multi-Task Deep Learning for Predicting Poverty from Satellite Images, The Thirtieth AAAI Conference on Innovative Applications of Artificial Intelligence.
58. Asian Development Bank Mapping the spatial distribution of poverty using Satellite Imagery in Thailand, 2021, https://doi.org/10.22617/TCS210112-2.
59. Agarwal, N., Mohanty, S.N., Sankhwar, S., et al. (2023). A novel model to predict the effects of enhanced students' computer interaction on their health in COVID-19 pandemics. *New Generation Computing*, *41*, 635–668. https://doi.org/10.1007/s00354-023-00224-3