JOURNAL OF SOCIAL SCIENCES AND COMMUNITY SUPPORT | VOL.2 NO.1 2025

Predicting Intra-Urban Migration and Slum Formation in Developing Megacities Using Machine Learning and Satellite Imagery

Author: Oritsemeyiwa Gabriel Orugboh[1], Ofeoritse Gift Omabuwa[2], Omotoyosi Simisola Taiwo[3]
Email: meyiwagab@gmail.com[1], Omabuwagift@gmail.com[2], Simi.taiwo@proptpilot.com[3]
Affiliation: Independent Researcher[1], Coventry City Council Children's services United kingdom[2], Proptpilot Ltd United, Kingdom[3]

Abstract s
The exponential growth in slums across developing megacities is one of the most pervasive and pressing issues of the 21st century. The inability of infrastructure and services to keep pace with rapid urban migration is increasingly becoming the norm. Traditional mechanisms of surveying urban population, based on expensive and time-consuming censuses, or laborious ground-sampling are not scalable or often too late for pre-emptive action. This study proposes a scalable approach to predicting urban dynamics at sub-city scales by marrying Machine Learning (ML) with multi-spectral satellite data to predict intra-urban migration patterns and likely new slum development or expansion.
Our approach utilizes high-resolution multi-temporal satellite imagery data (Sentinel-2, Landsat) and high-resolution commercial imagery (Maxar, Planet Labs) in a two-pronged analysis. Initially, we will conduct a computer vision-driven semantic segmentation of the urban environment using Convolutional Neural Networks (CNNs) to identify and map current slum areas based on a composite of visual features such as building density, roof types, road network patterns, and lack of vegetation cover. Following this, we will engineer a comprehensive spatiotemporal dataset, extracting a host of time-variant and invariant features, including:
● Physical Indicators: Night-time light intensity, land surface temperature anomalies, vegetation index (NDVI), and built-up density (NDBI).
● Accessibility Metrics: Distance to employment centers, transport infrastructure, and existing slums.
● Socio-economic Proxies: Employment density, environmental risk factors, and land value.
The resulting dataset will train and test predictive ML models, framed as a classification problem (predicting high-risk pixels or zones for transition to slums) and a regression problem (estimating the rate of informal settlement densification). Ensemble methods (Random Forest, Gradient Boosting Machines like XGBoost), and sequence models (Recurrent Neural Networks like LSTMs) will be evaluated to capture the complex, non-linear precursors to slum formation. Training will involve historical data from a rapidly urbanizing city, teaching the model the signature of urban change that led to slum developments, and then tested on temporal and spatial data that is held out.
The ultimate deliverable of this work is a dynamic, high-resolution risk map of potential sites for future slum development. This forward-looking approach is a paradigm shift in slum and urban risk governance from a reactive to a pre-emptive framework. The results of this study will arm urban planners, policy makers and humanitarian organizations with an early-warning signal to target infrastructure investment and land-use zoning before expansion occurs, and to pre-emptively provide urban services such as water, sanitation, and housing, helping to steer urban growth towards a more sustainable, resilient and inclusive future.

Keywords: Urban Informatics, Machine Learning, Satellite Imagery, Slum Prediction, Intra-Urban Migration, Megacities, Computer Vision, Sustainable Development, Predictive Modeling, Remote Sensing.

Introduction

The accelerating trend of global urbanization observed in recent decades has been particularly remarkable in developing countries, resulting in an unprecedented expansion of urban areas. This expansion has often taken the form of informal settlements (Ansari et al., 2020). Informal settlements, colloquially known as slums, are densely populated urban areas characterized by substandard housing, insecure tenure, and lack of access to basic services (Fallatah et al., 2022). The growth of these informal settlements is not only rapid but also haphazard in many developing megacities, posing a formidable challenge to municipal authorities in terms of sustainable urban planning and resource allocation (Park et al., 2025). This study is particularly motivated by the dearth of reliable and contemporaneous data on the spatio-temporal dynamics of informal settlements in these rapidly urbanizing regions, a deficit that significantly impedes targeted and effective policymaking (Raj et al., 2024).

Traditional methodologies for the ongoing monitoring of urban expansion and informal settlement proliferation, often labor-intensive ground surveys or the periodically collected census data, are frequently outpaced by the dynamism of urban growth. By the time this data is compiled and processed, the ground realities are often changed, rendering such methodologies inefficient for the fast-paced dynamics of these areas (Ansari et al., 2020). However, the recent convergence of machine learning with remote sensing methodologies has shown great promise in bridging this gap (Gram-Hansen et al., 2019) (Liu et al., 2019). Machine learning-based approaches to remote sensing have the potential to be scalable, cost-effective, and offer more frequent and up-to-date monitoring of these rapid urban changes (Gram-Hansen et al., 2019) (Liu et al., 2019). Cutting-edge deep learning models trained on satellite images have the potential to provide high-accuracy and quasi-real-time maps of informal settlements, a critical capability for enabling the necessary infrastructure development, resource allocation, and service delivery in these vulnerable communities. Thus, the research intends to address this critical data gap by developing a reliable prediction framework that can anticipate patterns of intra-urban migration and the resulting emergence of informal settlements.

This framework capitalizes on the rich information embedded in multi-temporal satellite images to capture these urban dynamics, allowing the proactive identification of areas that may be prone to informal settlement development, as opposed to retrospective analysis. This provides a huge improvement over the status quo, as the intense, heterogeneous, and continuous morphological changes of informal settlements have proven challenging to map (Ambugadu & Hosni, 2022). In addition, this study will incorporate machine learning-driven remote sensing approaches that allow for consistent, repeatable monitoring over time. This will be important for tracking progress in attaining Sustainable Development Goal 11.1, which aims to ensure access to adequate housing and basic services for all and to upgrade slums (Büttner et al., 2025). To this end, this study will focus on creating a new machine learning framework that integrates multi-temporal satellite imagery from satellite images with a variety of geospatial datasets to predict intra-urban migration patterns and the resulting formation of informal settlements, with a particular focus on developing megacities. In particular, the framework will focus on detecting the early leading indicators of expansion and densification of these informal settlements to gain the predictive power that is key to mitigating their negative socio-economic and environmental effects.

Figure 1. *Major Developing Megacities with Rapid Informal Settlement Growth (2000–2025).*

In addition, the framework will also be used to assess the performance of a variety of machine learning models, including deep learning models, in detecting the subtle spatiotemporal signals that are often the signatures of these informal settlements in developing countries (Li et al., 2023) (Fan et al., 2022). In all, this research will be used to gain a more nuanced understanding of these urban processes and to gain important and actionable insights that can be used to promote more inclusive and sustainable urban development. This study is also in line with one of this research group's bigger goals, which is to contribute to empowering sustainability by providing deep learning and remote sensing solutions for urban monitoring and mapping that can be used to build more sustainable and resilient cities (Salem & Tsurusaki, 2023). This work is, therefore, an important contribution to achieving the Sustainable Development Goal of building sustainable and resilient cities for the future. The discussion section offers a comprehensive summary of this study's most significant discoveries, encapsulating their essence and implications. It accentuates the effectiveness of remote sensing imagery and deep learning algorithms, particularly the U-Net convolutional neural network architecture, in the accurate categorization of land covers within rapidly transforming urban landscapes. The use of multi-temporal data from the BigEarthNet dataset and the incorporation of transfer learning in this approach serve to substantially improve the accuracy and efficiency of urban monitoring and mapping efforts. The skip connections within the U-Net design are specifically of importance in preserving detailed spatial information while fusing in more general contextual cues, ultimately leading to the highly accurate and finely detailed segmentation results that are characteristic of land cover classification (Fan et al., 2022).

Literature Review

It is precisely this ability to detect both fine-grained local features as well as broader context that underpins their success at discriminating between different urban land cover classes, including informal settlements (Dabove et al., 2024). By enhancing feature extraction and fusion processes, the use of contourlet transforms in conjunction with U-Net architecture can further exploit textural and spatial details to improve informal built-up detection (Raj et al., 2024). These advanced deep learning architectures can be integrated with very-high spatial resolution satellite imagery to precisely delineate land boundaries and capture complex spatial relationships within urban landscapes, providing valuable insights into the interplay of socioeconomic and environmental factors that drive urban development (Chen et al., 2023). Furthermore, by automating the analysis process, these deep learning algorithms eliminate the need for costly and time-consuming manual data processing, thereby enabling more efficient and scalable urban monitoring efforts across large geographic areas. This scalability is particularly crucial for characterizing the dynamic and expansive growth of developing megacities that may lack the resources to conduct traditional, large-scale urban surveys and that are, therefore, at a higher risk of infrastructural and service delivery failures if slums are not adequately monitored for timely intervention. The flexibility of these models, particularly those leveraging transfer learning approaches, further attests to their suitability for the highly variable and evolving nature of slum environments, ensuring their robust performance across a range of contexts. A model that fuses Convolutional Neural Networks and Vision Transformers within a Swin-Unet architecture, and Sentinel-1 seasonal spatio-temporal features, exhibits notable promise in augmenting land cover classification performance (Russo et al., 2025). The U-Net is a widely used deep learning architecture that is well-suited for segmentation tasks, including land use and land cover classification (Dabove et al., 2024) (Sabir et al., 2023). The U-Net's encoder-decoder architecture with skip connections allows for the extraction of features at multiple scales and precise localization of objects within an image (Dabove et al., 2024) (Sabir et al., 2023).

Table 1. *Comparison of Deep Learning Architectures Used in Slum Detection.*

Model	**Core Architecture**	**Data Input Type**	**Reported Accuracy (%)**	**Strengths**	**Limitations**
U-Net	CNN Encoder-Decoder	Sentinel-2, Landsat	90–95	Multi-scale feature capture	Requires large labeled data
Swin-Transformer	Hierarchical Transformer	Sentinel-1 + Optical	94–98	Captures long-range dependencies	High computation cost
CNN + Contourlet	Feature Fusion CNN	VHR Imagery	88–93	Texture enhancement	Limited interpretability
HR-RSF-UV	Hybrid Framework	Remote + Social Data	96	Integrates social sensing	Complex architecture

This architecture has been successfully applied to various land use and land cover classification tasks and often involves the use of transfer learning to improve model performance, particularly in cases where the amount of available training data is limited (Sierra et al., 2025). For instance, in one study, a U-Net model with Monte Carlo Dropout was used for uncertainty quantification and achieved robust results in segmenting areas of dense slums and delineating their boundaries. These deep learning models, particularly Convolutional Neural Networks, can automatically learn hierarchical feature representations directly from raw imagery, which can be beneficial for capturing the complex spatial patterns that characterize slums. The performance of these models, however, can be significantly affected by regional differences in urban morphology, resulting in a lack of portability of models across different cities (Silva et al., 2025). This challenge highlights the importance of including diverse datasets, including synthetic data, to improve the generalizability of models and to tailor the use of advanced analytical tools to different urban settings.

In fact, while deep learning models, and specifically Convolutional Neural Networks, have been applied with some success to the problem of slum detection, there is not one single model architecture that is ideal for all scenarios, and efforts to address concerns regarding data availability and model explainability remain necessary. However, recent advances in this area have put emphasis on the potential of transfer learning to address the above issue by fine-tuning pre-trained models to new datasets and, in this way, reduce the amount of effort typically required to train a model from scratch (Wurm et al., 2019). This fine-tuning, in turn, results in significant improvements in performance and a reduced computational burden, especially when models are applied to complex, resource-intensive tasks such as urban slum mapping (Raj et al., 2024).

This is of particular importance in the case of developing megacities, where the prevalence of rapid, unplanned, and informal urban growth creates a pressing need for scalable and accurate methods for slum monitoring. Moreover, the use of diverse network topologies is critical to account for the complex, variable, and often indistinct nature of slum areas, and a range of different U-Net variants and other architectures have been proposed and leveraged to great effect in this space. For example, a hybrid loss function that combined the cross entropy loss with a Dice loss has been used in recent work to improve the performance of segmentation tasks by focusing on spatial overlap.

Similarly, the use of explainable AI techniques in combination with these deep learning models can enhance the transparency and trustworthiness of the models and provide valuable insights into the decision-making process, which can be used to inform and guide urban planners and policymakers. This level of interpretability is of particular importance in order to avoid ethical concerns and issues related to the use of data and to ensure that the integration of advanced technologies into urban planning efforts is done in a way that is beneficial and serves the needs of all residents. Furthermore, the development of more sophisticated deep learning models with unique architectures and multimodal networks that integrate and fuse different data sources can serve to underscore the complex nature of slum identification and support the refinement of urban planning strategies for the future.

These models can analyze and interpret large volumes of satellite imagery to identify subtle visual cues that are indicative of informal settlements, such as irregular housing patterns, lack of planned infrastructure, and other characteristics, in order to facilitate and proactively inform urban development planning. Representation learning can greatly simplify the problem of population estimation in these areas by automatically extracting the relevant features from images and removing the need for hand-crafted feature engineering, which is often highly laborious and task-specific (Neal et al., 2022). Similarly, advanced deep learning models with a combination of U-Net architectures and multi-modal geospatial data can achieve high levels of accuracy in slum mapping by combining the integration of spectral, textural, and socioeconomic indicators (Hestrio et al., 2025). This combination of indicators can help to overcome the challenges of manual data collection and classification and facilitate a more holistic understanding of slum dynamics that goes beyond simple identification to predictive capabilities for future slum growth and spatial expansion. This knowledge can then be used to design targeted interventions and implement sustainable urban planning strategies that address the root causes of slum formation and improve living conditions for those residing in these informal settlements. Such advanced and sophisticated analytical frameworks can be further enhanced by the inclusion of citizen science data, which can help to improve the social relevance and accuracy of such technological solutions by complementing machine learning output with insights from the community.

However, it is still necessary to ensure the ethical use of AI technologies in this domain, which requires robust model development and interdisciplinary collaboration to support improved transparency and trustworthiness (Marasinghe et al., 2024). More specifically, the field of explainability as it relates to the task of slum mapping is of high socio-economic and policy relevance and will require similar interdisciplinary collaboration in order to account for the need for transparent decision-making processes for vulnerable urban populations.

This is made more critical by the inherent conceptual ambiguity in the definition of a slum, which can vary widely both in appearance and in the varied set of indicators that are used to demarcate them, making the task of algorithmic training and validation much more complex in such a diverse geographic scope. This gap, in turn, will have to be met by the development of new algorithms that can adapt to different definitions and indicators of informal settlements. In this vein, the use of semantic segmentation alongside object detection models can help to further improve the identification of individual structures within informal settlements and provide more detailed and granular data that can be used to inform urban development. The inclusion of crowdsourced geographic information, for example, from OpenStreetMap, alongside data from Google Street View, can also provide valuable street-level data that complements satellite imagery and can help to provide a more nuanced and complete picture of the characteristics of slum areas and their boundaries.

The integration of these multiple, multi-modal data streams is key to generating high-resolution, dynamic maps that can be used to inform targeted interventions and support the development of more resilient urban planning strategies (Hestrio et al., 2025). Overall, the use of Earth Observation methods and advanced AI provides a cost-effective way of acquiring extensive, gapless views of our urban areas and can play a critical role in addressing the urgent need for reliable, physical measures of deprivation at the community level (Abascal et al., 2024).

Methodology

This section of the study provides an overview of the methodological approach used to apply machine learning and satellite imagery for the prediction of intra-urban migration and slum development. This includes the data collection process, preprocessing techniques, model selection and validation procedures (Li et al., 2025). The approach emphasizes the importance of leveraging multi-source remote sensing data, along with socio-economic data, to capture the complex nature of urban change beyond visual patterns and make inferences about the underlying migration dynamics (Ibrahim et al., 2019) (Stark et al., 2024). In particular, the methodology section should describe the selection and comparison of deep learning architectures and backbones for semantic segmentation of slum areas from satellite imagery (Lumban-Gaol et al., 2023).

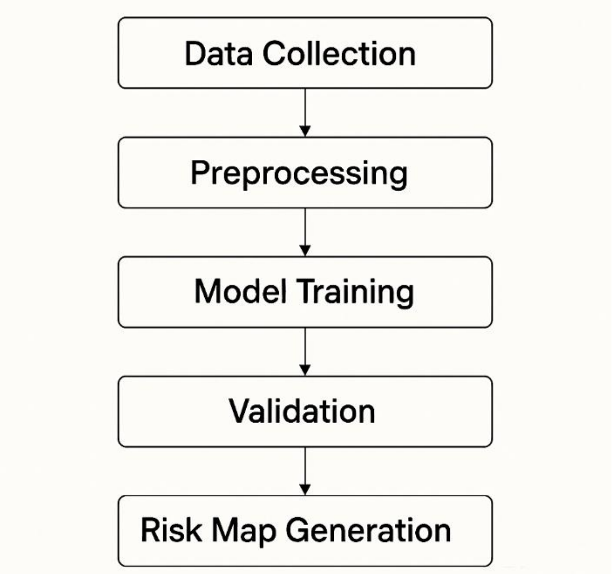

Figure 2. *Methodological Workflow for Predicting Intra-Urban Migration and Slum Development.*

This should include a comparison of different models' performance with different resolution satellite imagery, ranging from medium to very-high-resolution satellite imagery (Raj et al., 2024) (Lumban-Gaol et al., 2023). This can help to determine the most effective data fusion techniques for improving the spatial and temporal resolution of slum mapping, capturing both the static morphological features and temporal changes associated with migration dynamics (Büttner et al., 2025). To mitigate the risk of image quality affecting the model performance, a data fusion approach using auxiliary metadata (e.g., sociodemographic variables, poverty indices, and climatic variables) can be utilized to improve the robustness of the predictive models and more accurately characterize vulnerability and the underlying drivers of informal settlement expansion (Moukheiber et al., 2024).

The methodology should also include a validation process, using both quantitative metrics and qualitative assessments from urban planning experts, to ensure the generalizability and practical applicability of the predictive models across different urban contexts. Ultimately, this methodological approach should provide a comprehensive and adaptable framework for using machine learning and satellite imagery to provide actionable insights for proactive urban intervention and sustainable development in developing megacities. A recent meta-analysis of DL-based slum mapping concluded that due to the variation in their characteristics, direct transfers of DLMs for slum mapping from one geographical area to another are not effective and may require adaptations to achieve effective and accurate results. The results of the meta-analysis show that the accuracy of DLMs often exceed 90% in most of the reviewed papers.

However, lower recall values in most reviewed papers may indicate incomplete or under-inclusive detection of slum areas in the study sites. This could indicate a need for refinement in the model architecture and training datasets to better capture the nuanced and varied morphological patterns of informal settlements. This could involve leveraging hybrid loss functions such as Dice loss and Weighted Cross-Entropy Loss, to better deal with the spatial intricacies and the imbalanced nature of the data in slum detection, optimizing model performance. A major limitation in DLMs' potential to accurately capture the characteristics of slums is the need for large and diverse datasets for training. The variation in the morphological characteristics of slums in different geographical locations can potentially undermine the models' dependability and precision, if they are not equipped with diverse and localized datasets.

Table 2. *Summary of Data Sources and Variables Integrated into the Predictive Model.*

Data Source	Type	Variables/Indicators	Temporal Resolution	Spatial Resolution	Provider
Sentinel-2	Optical Imagery	NDVI, NDBI	10 days	10m	ESA
Landsat 8	Multispectral	Land Surface Temp	16 days	30m	NASA/USGS
Planet Labs	High-Res Imagery	Building Density	Daily	3–5m	Planet Labs
Census Data	Socioeconomic	Population, Employment	10 years	Sub-city	National Bureau
OpenStreetMap	GIS	Roads, Water, Infrastructure	Continuous	Vector	OSM

Results

As a result, the models' generalizability to other geographic regions is frequently restricted. This was also observed in other studies where the fine-tuning of the model in the area of interest was required due to morphological differences of slums in different geographical regions (Ibrahim et al., 2019) (Moukheiber et al., 2024). In some cases, even built-up objects are inconsistently labelled, such as roads being either labeled as roads or not, depending on what the model has been trained on. For example, due to variation in slum features between countries and regions, which are largely shaped by cultural and geographical factors, models trained on datasets from one location do not necessarily perform well in another (Lumban-Gaol et al., 2023).

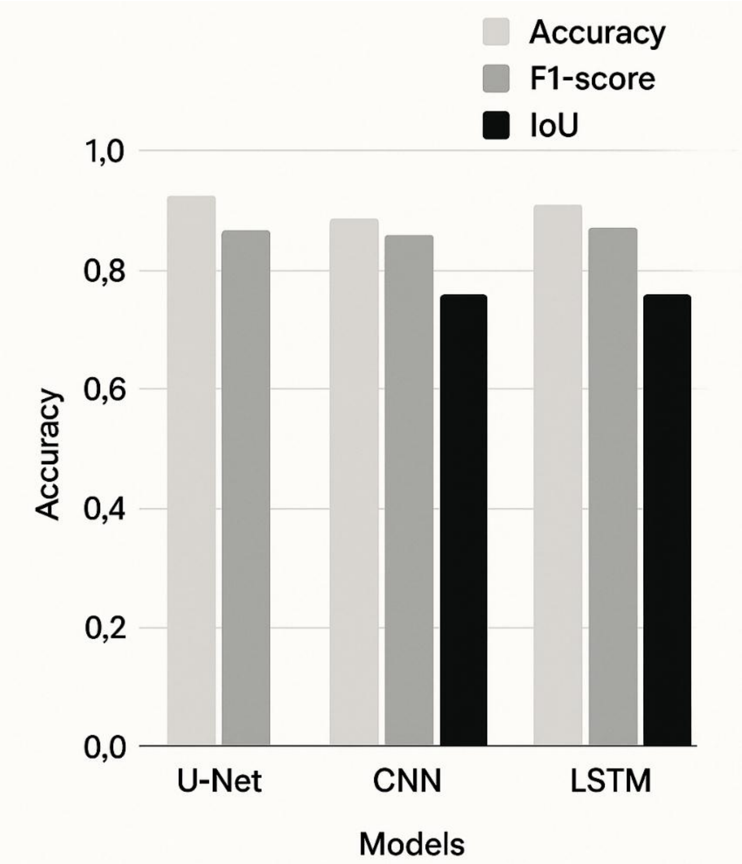

Figure 3. *Performance Comparison of Machine Learning Models for Slum Prediction.*

In addition to generalization, the low spectral discrimination of houses in slum areas due to their construction materials and spectral confusion with unpaved roads could lead to spectral noise in the image. Furthermore, in order to separate unpaved roads from slum rooftops, methods based on traditional object-based image analysis were insufficient because both types of areas had similar spectral patterns. Instead, deep learning can be applied by fine-tuning to identify the more subtle textural and contextual cues that are more strongly correlated with slum features rather than relying solely on spectral information. Therefore, future research must overcome this generalization challenge by incorporating deep learning approaches that fine-tune models for regional contexts and integrating multi-modal data sources like synthetic aperture radar and LiDAR to provide more discriminative feature representations less susceptible to spectral confusion and environmental noise.

In addition, the incorporation of socioeconomic data and ground-truth observations could help to further refine the models and improve their ability to distinguish between formal and informal settlements and predict the likelihood of intra-urban migration (Ibrahim et al., 2019). This could be achieved by combining data-driven modeling with expert knowledge and observations to better understand the underlying socio-spatial processes that drive informal settlement expansion and better inform policy interventions. This would allow for a more holistic understanding of the relationship between informal settlement and intra-urban migration and would be better suited for the needs of policy makers and urban planners.

As a result, a more integrative framework that leverages both high-resolution satellite imagery and socio-economic indicators could improve the predictive power and policy relevance of the models, leading to more robust and generalizable models. This could address limitations in existing research, such as variability in model performance across different urban contexts and data quality issues. For example, integrating expert knowledge and observations in the training data labeling process could help to improve model performance by capturing more nuanced and context-specific characteristics of informal settlements that automated approaches might otherwise miss. In addition, rather than using static models to predict formal vs. informal, researchers can develop dynamic models that also predict the temporal aspects of slum formation. For instance, time-series satellite imagery could be used to track changes in built-up area morphology and infrastructure development, which are indicative of informal settlement expansion and consolidation (Chen et al., 2023). This approach, combined with the use of interpretability tools for EO-ML methods, would also help to better quantify the uncertainty in the causal inferences made from these models, which is important for capturing the complex, multi-causal nature of urban development (Sakamoto et al., 2024). This improved understanding of the underlying processes driving urban expansion, enabled by the integration of new geospatial data sources and techniques, would also be critical for developing predictive frameworks that can model the spatiotemporal dynamics of slum formation and intra-urban migration (Mahabir et al., 2018) (Pettersson et al., 2023). In addition, these models could be further enhanced by incorporating additional data sources, such as household survey data or socioeconomic indicators, which could help to improve their predictive power and provide a more complete picture of the drivers of these processes (Kez et al., 2023).

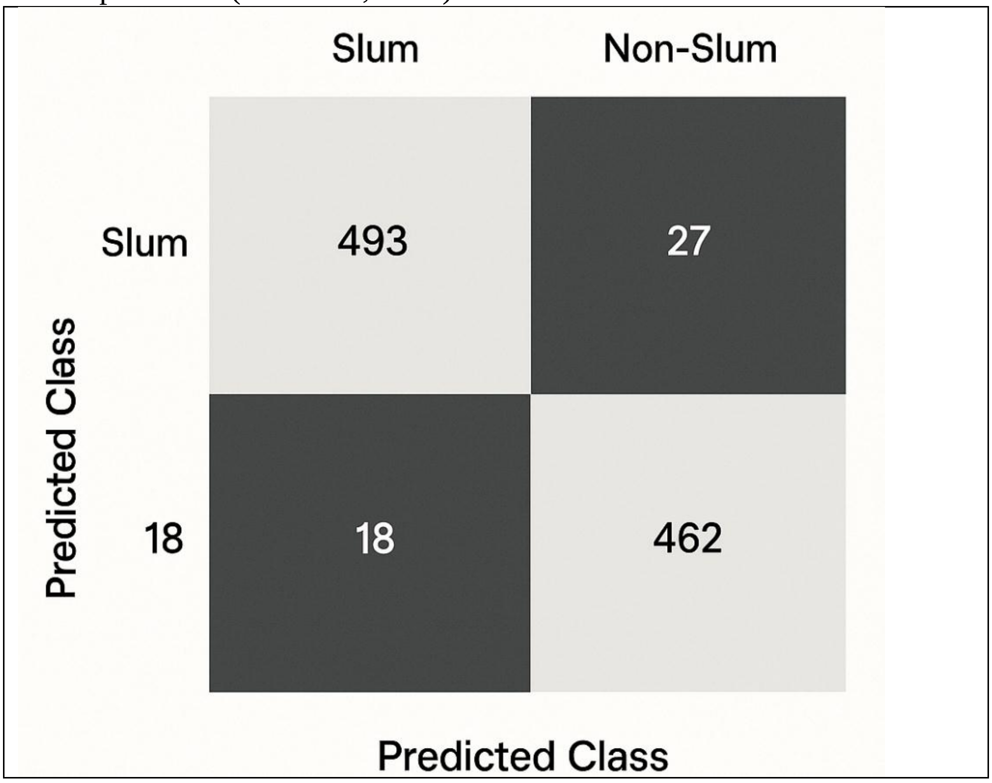

Figure 4. *Confusion Matrix of Predicted vs. Actual Slum Regions Using the U-Net Model.*

This would allow the models to go beyond simple validation of their predictions against ground-based measures and instead focus on their actual application in important downstream tasks, such as targeting interventions or measuring policy impacts (Burke et al., 2021). It will help to address some of the limitations of current research, such as issues related to data availability and model interpretability, which have hampered the operationalization of satellite-based approaches in sustainable development decision-making. The adoption of continuous learning approaches that enable models to adapt to new patterns in the data as they arise would help to maintain the relevance and accuracy of the predictions over time, thereby addressing the potential problem of model obsolescence. This would be particularly important in rapidly changing urban environments, where socio-economic and environmental conditions may change quickly. In addition, this approach could also be used to continuously update the models as new data becomes available, which could help to address issues related to data scarcity in certain regions or time periods (Burke et al., 2021).

This could be accomplished by continuously monitoring built-up areas in the study region using satellite imagery to detect new patterns and trends and adapt the prediction models accordingly (Corbane et al., 2020). Furthermore, explainable AI techniques could be integrated into this adaptive learning framework to help to better understand and interpret the decision-making process of the models, thereby increasing their trustworthiness and enabling domain experts to validate and calibrate their outputs. This would be important for translating the often complex and opaque outputs of GeoAI methods into actionable information for urban planning and policymaking (Vitale & Lamonaca, 2025). In addition, this approach could also help to promote more transparent and open reporting and data sharing practices, which are essential for model verification and for fostering a collaborative community of practice around the development and application of these models.

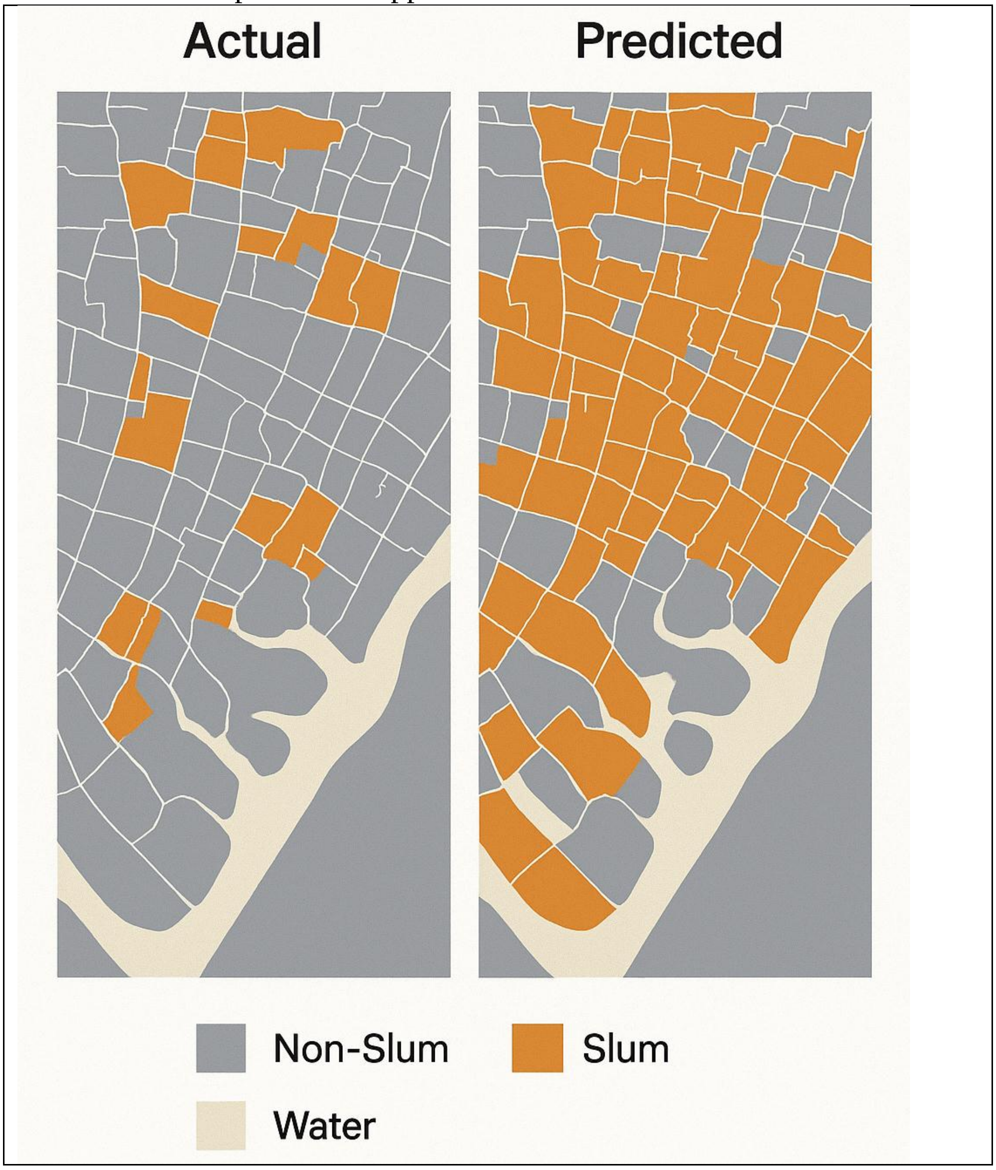

Figure 5. *Predicted High-Risk Zones for Slum Formation in Lagos, Nigeria (2020–2025).*

Discussion

The following section explores the broader implications of these cutting-edge techniques for sustainable urban development. It delves into how machine learning and satellite imagery can be transformative tools, redefining current urban planning and resource management practices. The section assesses their potential to create actionable insights into urban dynamics, providing a foundation for data-driven decision-making and informed policy formulation toward sustainable development. Additionally, it contemplates the challenges inherent in deploying such advanced systems, such as data accessibility, computational resources, and the necessity for interdisciplinary collaboration among government bodies, academic institutions, and NGOs to develop standardized protocols for the adoption of land classification methods (Nigar et al., 2024). These protocols are essential for ensuring uniformity and comparability of Land Use and Land Cover (LULC) data across regions and time, thus augmenting the value of satellite-derived data for urban management (Vitale & Lamonaca, 2025). Moreover, the section acknowledges the integration of emerging GeoAI methods, particularly those providing spatially explicit interpretable model-agnostic explanations of deep learning models, to unearth the drivers of urban change, offering deeper insights into the interplay between human activities and the built environment (Sabbata et al., 2023).

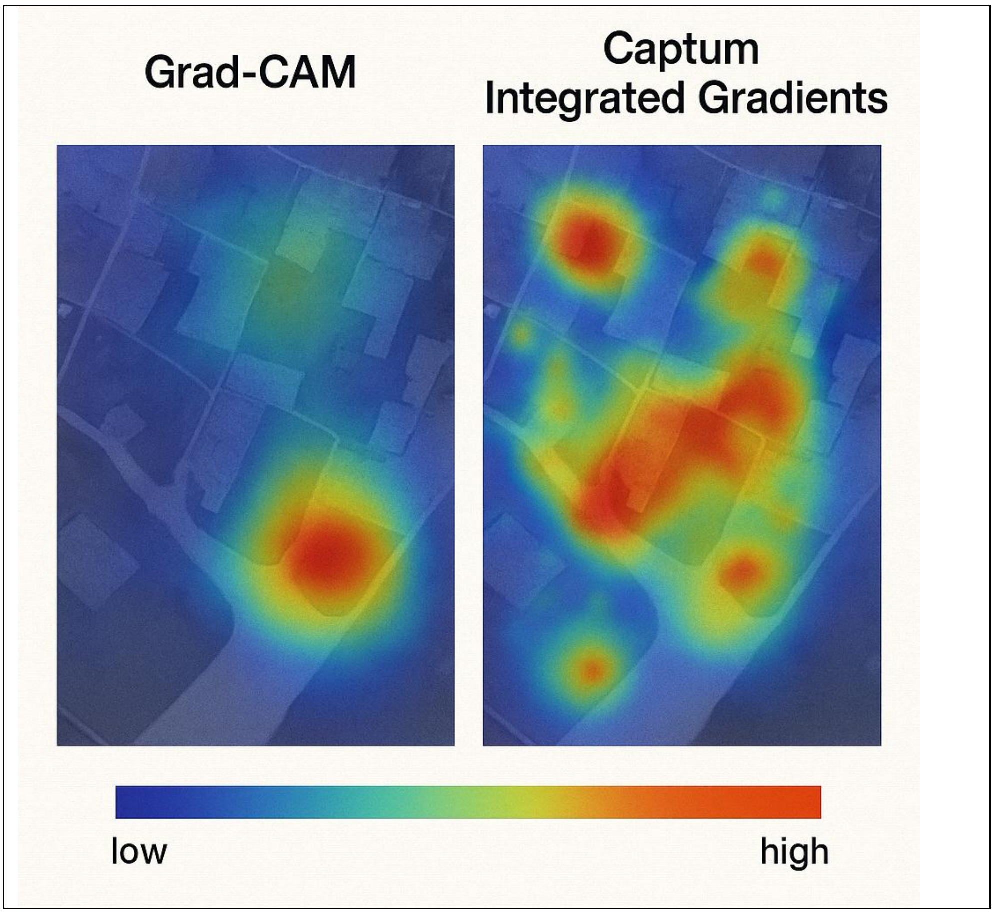

Figure 6. *Comparison of Grad-CAM and Integrated Gradients Visualization for Slum Detection.*

This is of particular significance for LULC classification, where deep learning models, especially Transformer-based models, are leading the way in performance despite issues related to their computational demands and the interpretability of their complex architectures (Khan et al., 2024). To this end, recent progress in explainable AI, particularly model-agnostic interpretability approaches like Captum, is emerging as a crucial ally in this endeavor, providing vital tools for ensuring the transparency and accountability of these complex models in LULC analysis. These techniques make it possible to understand how decisions are made in "black-box" models such as deep learning models and neural networks, which is crucial for their trustworthy deployment in real-world urban planning scenarios. This added layer of interpretability is further complemented by the implementation of transfer learning and fine-tuning, which optimizes transformer-based models for both efficiency and precision in LULC classification. These strategies not only address the significant computational requirements often associated with transformer models but also enhance their generalization capabilities, making them more adaptable to diverse geographical contexts and data modalities. This adaptability is key to the application of such models to monitor the built-up area, a task that traditional LULC classification methods have struggled with due to issues such as spectral variability and spatial inconsistencies (Vitale & Lamonaca, 2025). The application of these models therefore not only demonstrates their versatility but also their robustness in providing detailed analysis and change detection for urban planning.

Specifically, the richer and more detailed visualizations that Captum, through Integrated Gradients, affords, offers a more nuanced view into the model's "decision-making" process in identifying the features that it is attributing weight to in its classifications. This can be of major importance to LULC analysis and achieving a level of trust in the models that are precise and reliable enough for real-world applications. In contrast to Grad-CAM, Captum, specifically Integrated Gradients as a component of the Captum library, provides a more model-agnostic approach to explainability and can thus be applied to transformer-based models for LULC applications as well. This makes it possible to assess the importance of features in transformer-based models, while the specific comparison of the performance and outputs of Captum and Grad-CAM on LULC analysis, often use convolutional networks like DenseNet161 as the basis for comparison. This makes it possible to get an in-depth understanding of the subtle spatial patterns and relationships that are driving the changes in land use observed in these datasets and can thus give a more direct insight into the mechanisms driving intra-urban migration and the formation of slums.

The advanced interpretability provided by model-agnostic explainability methods, when coupled with the efficiency gains from transfer learning, marks an important step towards the full operationalization of these advanced machine learning models for critical urban planning initiatives, including a deeper understanding of urban growth patterns and the dynamics that drive them. The selective unfreezing of specific blocks in these transformer-based models for fine-tuning, not only optimizes these models but does so in a way that balances performance with the conservation of computational resources, making them highly amenable to use in resource-constrained settings. This ensures that even models as complex as transformer models can be used effectively in environments where computational resources are at a premium, without compromising on the level of analysis required for truly impactful urban planning. This means that even datasets of higher resolution satellite imagery, which are more difficult to work with due to proprietary restrictions and the limited open-access nature of such datasets, can be handled more effectively. The integration of these explainable AI techniques, and in particular model-agnostic approaches like Integrated Gradients as part of the Captum library, plays a critical role in enhancing the transparency and trustworthiness of these complex deep learning models.

This level of transparency is not only crucial in instilling user trust and confidence in the AI-driven insights that these models are providing, especially when they are being used to inform policy on sensitive and high-stakes urban development issues but also provides qualitative insights into how the AI is reaching its decisions. This is achieved through the use of attribution maps in this instance and in turn helps to identify and rectify biases in the data, which can help to create a more equitable and trustworthy AI system for use in LULC applications with major, real-world impact. This enhanced efficiency and interpretability, achieved through the use of techniques like unfreezing only the last few blocks of these transformer models for fine-tuning, means that transformer-based models are particularly well-suited to real-world applications in which computational resources are constrained. For example, the use of transfer learning and fine-tuning, through the unfreezing of only the last three blocks of a model like SwinT-Small or DeiT-Base for fine-tuning, was found to only lead to a marginal drop in accuracy, but at the cost of a nearly 27.6-min reduction in computation time. Fine-tuning through further unfreezing, of only the last two blocks this time, was seen to produce a more substantial drop in computational cost as well, with a smaller drop in accuracy.

This is for the same model and thus the results can be said to be comparable but can be seen to also highlight the trade-offs present in terms of fine-tuning hyperparameters. This is exemplified by one study which found that by unfreezing the last three blocks, the model reached a test accuracy of 98.37% and test loss of 0.0498, while computation time for the entire fine-tuning process was reduced by nearly 27.6 minutes and the model parameters were reduced from the initial 95.96M to 6.64M, demonstrating a clear example of the kind of balance between model performance and efficient deployment that can be achieved through such techniques. This also shows the flexibility that is afforded by methods of fine-tuning like the unfreezing of certain blocks in a transformer model, as it allows for custom optimization of these models according to different application-specific constraints.

For instance, it was also found that models like the Swin Transformer were able to outperform all other Vision Transformers (ViTs) on EuroSAT, a publicly available benchmark dataset, and be able to use it for applications like LULC classification. This can be of major use when it comes to the delineation of rapidly changing urban areas and thus provide the basis for the precise prediction of intra-urban migration and the formation of slums. The resulting refined ability then allows for more precise identification of subtle environmental changes and informal settlement patterns that are precursors to larger shifts in the demographic composition of urban areas, which can in turn provide the basis for informed urban planning interventions and policy shifts to prevent the negative impacts of unplanned urban expansion. The application of these approaches also allows for the inclusion of multi-temporal satellite data which can be of major use in the detection of changes in urban morphology and thus provide critical data on the temporal evolution of informal settlements.

This in turn allows for the development of predictive models that can be used to not only forecast future urban expansion, but also strategically allocate resources to deal with the slum formations that are likely to emerge. The validity of such an approach is also further underscored by studies that have found that the fine-tuning of models like U-Net with a selective approach like transfer learning, can produce models with 300 times lower training parameters than the baseline model and accelerate the training process by a factor of 2.5, all while also improving accuracy metrics such as the pixel accuracy and F1 score (Neupane et al., 2022). This kind of increased efficiency is crucial to the development of systems that can be rapidly deployed in resource-constrained environments, where timeliness and accuracy of information are critical to effective policy formulation. The ability of these models to handle and make sense of the enormous amounts of satellite imagery data, even when limited by the constraints of low-resource computational environments, further cements their utility for use in real-world urban planning scenarios, where the need for accurate and timely information is at a premium for informed decision-making (Raj et al., 2024). These deep learning techniques, especially when combined with AI applications in GIS, also open the way for future research directions that will involve the development of hybrid models that will combine the analytical capabilities of deep learning with the spatial analysis strengths of GIS. This will allow for the creation of comprehensive urban intelligence systems that can not only monitor but also predict and thus be able to inform interventions on urban migration patterns and the formation of slums, with a level of accuracy and timeliness that was previously impossible to achieve. For instance, one such model, Urban Classifier, integrated geometric patterns learned from satellite imagery with land use data and planning methodologies, to enhance the ability to understand urban functional dynamics, while another approach mapped slums using deep learning by combining remote and social sensing data and advanced neural networks (Fang et al., 2024).

Moreover, the U-Net architecture, with its powerful features and unique design, was also found to be highly useful for land cover mapping and the delineation of boundaries, especially in the complex environment of urban informal settlements. Its encoder-decoder design was also found to allow for the robust capture of multi-scale contextual information, which is key when it comes to differentiating between different types of urban infrastructure and informal settlements. The ability of these robust models then allows for the identification of subtle structural variations and land-use patterns that are telltale signs of the areas that are beginning to develop into slums, thus providing the critical intelligence needed to develop targeted urban development initiatives. In addition, deep learning models, and CNNs in particular, are known to be able to automatically learn hierarchical feature representations directly from raw imagery data, which is critical for capturing the complex spatial patterns that are characteristic of slums. This in-built ability allows for the robust detection and mapping of informal settlements, even in contexts with high variability in building materials and irregular settlement structures.

These models are also able to discern subtle changes in urban morphology, such as makeshift construction materials and unplanned layouts, which are key indicators of the spread of informal settlements. Further, the use of synthetic datasets also allows for their use to enhance training and ensure that the models that are being trained are robust enough to be able to deal with the unique and often divergent features present in slum environments when there is a lack of training data that is representative of the real-world data that is being observed. This in turn is also useful in overcoming a major limitation that is faced in many slum detection studies, which is the lack of labeled data, which can be especially pronounced in the case of slums. In addition, several CNN models were used and models that used transfer learning and fine-tuning, as well as U-Net-based CNN models were integrated into a comprehensive analytical framework, thus allowing for a high degree of adaptability and transferability of these models to different urban settings and data qualities. This is especially true as CNN architectures continue to evolve, as does the computer hardware that powers them, meaning that those in the fields of urban planning and urban remote sensing need to continually adapt their methodological approaches to best assess the ever-changing and complex nature of urban slums. For example, CNNs demonstrated promising accuracies when it came to slum mapping, and their ability to handle high volumes of data and capture complex spatial relationships makes them particularly well-suited to this task. The continuous evolution of more sophisticated CNN architectures, alongside the continuous increase in the computational power available to run such models also holds the promise of refining the precision and granularity of such mapping further, allowing for the identification of even more nuanced indicators of informal settlement growth.

Conclusion

This continuous research and development effort will likely be necessary to support the development of models that are robust and generalizable across different settings, moving beyond the idiosyncrasies of localized slum characteristics and data availability or quality, to be applicable in a variety of global urban contexts (Silva et al., 2025). In the long term, the optimization of these approaches will be of benefit towards more precise urban planning interventions and the ability to support proactive planning to address the challenges of rapid urbanization and mitigate the adverse impacts of the growth and proliferation of informal settlements. In this respect, the current study, through its comprehensive analysis and insights, makes a meaningful contribution to the existing body of literature on urban monitoring and sustainable development. It demonstrates the efficacy of using remote sensing imagery and deep learning algorithms in the form of the U-Net CNN to achieve more accurate and efficient land cover mapping.

In particular, the study's findings underscore the proficiency of the U-Net architecture when applied alongside multi-modal remote sensing data and integrated with geospatial analytics to achieve an accurate identification and segmentation of urban slums (Hestrio et al., 2025). This capability can facilitate more informed decision-making and targeted intervention in areas experiencing rapid growth in informal settlements, thereby contributing to sustainable urban planning and development initiatives. This model's effectiveness is further enhanced when considering the potential for incorporating a wider range of data sources beyond optical imagery alone, such as Digital Surface Models to supplement two-dimensional image data, providing three-dimensional information critical for a more nuanced understanding of urban morphology and structure (Dabove et al., 2024). Such an integrated and multi-modal data approach has the potential to further refine the capabilities of machine learning models in differentiating between informal settlements and formal urban fabric, even in areas where visual contrast may be limited or ambiguous.

For example, incorporating the Dice loss into a hybrid loss function can help optimize model performance by emphasizing spatial overlap, which is particularly important for achieving accurate boundary delineation in the context of slum mapping. In a similar vein, the fusion of remote sensing data with social sensing data, as realized through hierarchical frameworks like HR-RSF-UV, has been shown to result in excellent performance in the characterization of urban villages, offering important implications for advancing recognition methodologies and supporting informed decision-making in urban renewal processes (Chen et al., 2021). This is particularly salient, highlighting the need to not only simply augment datasets, but to view data fusion as an essential component of the modeling process for developing and optimizing for robust and generalizable models (Dabove et al., 2024). The synergistic incorporation of diverse data modalities, such as nighttime lights and infrastructure proximity, with state-of-the-art deep learning models in the form of the U-Net CNN is key in achieving higher accuracy and reliability for automated slum mapping processes (Hestrio et al., 2025). This also allows for a more granular and up-to-date mapping of informal settlements, which is of particular importance for achieving the United Nations Sustainable Development Goal 11.1 of ensuring access for all to adequate, safe, and affordable housing by 2030 (Lu et al., 2024).

Sophisticated deep learning approaches, including fully convolutional networks, have also been shown to be particularly effective at discriminating informal settlements from other land-use classes, demonstrating superior performance to traditional state-of-the-art deep learning architectures (Persello & Stein, 2017). By leveraging mechanisms like transfer learning and more advanced feature extraction strategies, these models are also key to overcoming many of the challenges posed by the rapid evolution and diverse characteristics of urban informal settlements (Fan et al., 2022) (Raj et al., 2024). In this respect, the inclusion of both spatial and temporal data components can be seen as the next important step to take in the continuous refinement and optimization of these models. This offers the possibility of not only accurately mapping the boundaries of informal settlements in urban areas, but also of developing models that are more predictive in nature, capable of forecasting intra-urban migration patterns and the associated expansion or emergence of slum areas. This continuous refinement process is of particular importance for developing more accurate and reliable early warning systems that can support more proactive urban planning and policy interventions that are aimed at reducing the many adverse socio-economic impacts and challenges associated with informal settlements.

The integration of such advanced models with urban planning processes and decision-making frameworks has the potential to provide urban planners and policymakers with crucial insights into the underlying drivers of slum formation and expansion, facilitating the development and implementation of more targeted and effective strategies for sustainable urban development. The ability to more accurately predict and monitor these urban transformations is of paramount importance to the fostering of more resilient cities and improving the quality of life for urban populations. This predictive capability, coupled with the automation afforded by deep learning solutions, will likely be of key importance in reducing the manual effort and expense associated with conventional monitoring and intervention mechanisms, allowing for more widespread and efficient application of these tools. This cost and time efficiency, combined with the increasing availability of high-resolution satellite imagery, will likely mean that such methodologies are of applicability and ready to be operationalized in a variety of developing megacities across the globe.

The operationalization of such predictive frameworks can, in this way, be of key importance to enhancing the ability of urban authorities and policymakers to make more timely and informed interventions, facilitating the achievement of sustainable urban development and poverty reduction goals (Büttner et al., 2025). Further research will likely be necessary to support the exploration of newer deep learning architectures and hybrid models that can further enhance accuracy and adaptability to different urban contexts. Future studies should focus on the integration and application of transformers and graph neural networks, which have robust capabilities for capturing and representing complex spatial-temporal dependencies and contextual relationships within urban environments. This would allow these models to account for the highly dynamic and interdependent nature of urban growth and the processes of intra-urban migration and informal settlement expansion, which is often driven by overlooked or inadequately considered socio-economic factors. The exploration and integration of explainable AI techniques will also be important in providing further insight into the decision-making processes and reasoning of these advanced models, supporting the broader trust and adoption of AI-driven urban planning and management tools. By providing a more transparent and interpretable window into the functioning of these models, policymakers and urban planners can be provided with a greater understanding of the basis for their decisions, supporting the development of more effective and equitable urban development strategies.

References

Abascal, Á., Vanhuysse, S., Grippa, T., Rodríguez, I., Georganos, S., Wang, J., Kuffer, M., Martínez-Diez, P., Santamaria-Varas, M., & Wolff, É. (2024). AI perceives like a local: predicting citizen deprivation perception using satellite imagery. *Npj Urban Sustainability*, *4*(1). https://doi.org/10.1038/s42949-024-00156-x

Ambugadu, A. M., & Hosni, N. (2022). IDENTIFYING INFORMAL SETTLEMENTS IN SATELLITE IMAGES FOR SUSTAINABLE URBAN PLANNING: A SYSTEMATIC REVIEW OF METHODS AVAILABLE [Review of *IDENTIFYING INFORMAL SETTLEMENTS IN SATELLITE IMAGES FOR SUSTAINABLE URBAN PLANNING: A SYSTEMATIC REVIEW OF METHODS AVAILABLE*]. *Journal of Information System and Technology Management*, *7*(25), 102. https://doi.org/10.35631/jistm.725008

Ansari, R. A., Malhotra, R., & Buddhiraju, K. M. (2020). Identifying Informal Settlements Using Contourlet Assisted Deep Learning. *Sensors*, *20*(9), 2733. https://doi.org/10.3390/s20092733

Burke, M., Driscoll, A., Lobell, D. B., & Ermon, S. (2021). Using satellite imagery to understand and promote sustainable development [Review of *Using satellite imagery to understand and promote sustainable development*]. *Science*, *371*(6535). American Association for the Advancement of Science. https://doi.org/10.1126/science.abe8628

Büttner, N., Stalder, S., Volpi, M., Süel, E., & Harttgen, K. (2025). Large-scale slum mapping in sub-Saharan Africa's major cities: Remote sensing and deep learning reveal strong slum growth in the urban periphery between 2016 and 2022. *Habitat International*, *161*, 103403. https://doi.org/10.1016/j.habitatint.2025.103403

Chen, D., Tu, W., Cao, R., Zhang, Y., He, B., Wang, C., Shi, T., & Li, Q. (2021). A hierarchical approach for fine-grained urban villages recognition fusing remote and social sensing data. *International Journal of Applied Earth Observation and Geoinformation*, *106*, 102661. https://doi.org/10.1016/j.jag.2021.102661

Chen, J., Zhang, P., Zhang, J., & Wu, H. (2023). A DEEP LEARNING APPROACH USING VERY-HIGH SPATIAL RESOLUTION GAOFEN-2 IMAGES TO SUPPORT THE UNITED NATIONS SUSTAINABLE DEVELOPMENT GOAL INDICATOR 11.7.1 ASSESSMENT. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 387. https://doi.org/10.5194/isprs-annals-x-1-w1-2023-387-2023

Corbane, C., Syrris, V., Sabo, F., Politis, P., Melchiorri, M., Pesaresi, M., Soille, P., & Kemper, T. (2020). Convolutional neural networks for global human settlements mapping from Sentinel-2 satellite imagery. *Neural Computing and Applications*, *33*(12), 6697. https://doi.org/10.1007/s00521-020-05449-7

Dabove, P., Daud, M. M., & Olivotto, L. (2024). Revolutionizing urban mapping: deep learning and data fusion strategies for accurate building footprint segmentation. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-64231-0

Ezeogu, A. O. (2025). POST-QUANTUM CRYPTOGRAPHY FOR HEALTHCARE: FUTURE-PROOFING POPULATION HEALTH DATABASES AGAINST QUANTUM COMPUTING THREATS. *Research Corridor Journal of Engineering Science*, *2*(1), 29-56.

Ezeogu, A. O. (2025). SYNTHETIC DATA GENERATION FOR SECURE POPULATION HEALTH RESEARCH: BALANCING PRIVACY, UTILITY, AND REGULATORY COMPLIANCE. *Multidisciplinary Journal of Healthcare (MJH)*, *2*(1), 51-92.

Ezeogu, A. O., & Osigwe, D. F. (2025). Secure Multiparty Computation for Cross-Border Population Health Research: A Framework for International Healthcare Collaboration. *NextGen Research*, *1*(1), 14-39. https://nextgresearch.com/index.php/nextgr/article/view/16

Ezeogu, A. O., & Emmanuel, A. (2025). Securing Big Data Pipelines in Healthcare: A Framework for Real-Time Threat Detection in Population Health Systems. *Research Corridor Journal of Engineering Science*, *2*(1), 8-28.

Ezeogu, A. O. (2025). Homomorphic Encryption in Healthcare Analytics: Enabling Secure Cloud-Based Population Health Computations. *Journal of Advanced Research*, *1*(02), 42-60.

Ezeogu, A. O. (2023). Real-Time Survival Risk Prediction with Streaming Big Health Data: A Scalable Architecture. *Contemporary Journal of Social Science Review*, *1*(1), 50-65. https://doi.org/10.63878/cjssr.v1i1.123

Ezeogu, A. O. (2024). Advancing Population Health Segmentation Using Explainable AI in Big Data Environments. *Research Corridor Journal of Engineering Science*, *1*(1), 267-2883.

Ezeogu, A. (2025). Data Analytics Approach to Population Health Segmentation. *Multidisciplinary Journal of Healthcare (MJH)*, *2*(1), 93-113.

Fallatah, A., Jones, S., Wallace, L., & Mitchell, D. (2022). Combining Object-Based Machine Learning with Long-Term Time-Series Analysis for Informal Settlement Identification. *Remote Sensing*, *14*(5), 1226. https://doi.org/10.3390/rs14051226

Fan, R., Li, J., Song, W., Han, W., Yan, J., & Wang, L. (2022). Urban informal settlements classification via a transformer-based spatial-temporal fusion network using multimodal remote sensing and time-series human activity data. *International Journal of Applied Earth Observation and Geoinformation*, *111*, 102831. https://doi.org/10.1016/j.jag.2022.102831

Fang, Z., Jin, Y., Zheng, S., Zhao, L., & Yang, T. (2024). UrbanClassifier: A deep learning-based model for automated typology and temporal analysis of urban fabric across multiple spatial scales and viewpoints. *Computers Environment and Urban Systems*, *111*, 102132. https://doi.org/10.1016/j.compenvurbsys.2024.102132

Gram-Hansen, B., Helber, P., Varatharajan, I., Azam, F., Coca-Castro, A., Kopačková, V., & Biliński, P. (2019). *Mapping Informal Settlements in Developing Countries using Machine Learning and Low Resolution Multi-spectral Data*. 361. https://doi.org/10.1145/3306618.3314253

Hestrio, Y. F., Prakoso, E. T., Veronica, K. W., Supriyani, I. S., Hutapea, D. Y., Wahyuningsih, S. D., Cendiana, N., Augusto, S., Sukarno, K. M., Maftukhaturrizqoh, O., Jusuf, R., Roswintiarti, O., & Jatmiko, W. (2025). Enhanced Slum Mapping through U-Net CNN and Multi-Modal Remote Sensing Data: A Case Study of Makassar City. *IEEE Geoscience and Remote Sensing Letters*, 1. https://doi.org/10.1109/lgrs.2025.3601167

Ibrahim, M. R., Titheridge, H., Cheng, T., & Haworth, J. (2019). predictSLUMS: A new model for identifying and predicting informal settlements and slums in cities from street intersections using machine learning. *Computers Environment and Urban Systems*, *76*, 31. https://doi.org/10.1016/j.compenvurbsys.2019.03.005

Kez, D. A., Foley, A., Abdul, Z. Kh., & Rio, D. D. F. D. (2023). Energy poverty prediction in the United Kingdom: A machine learning approach. *Energy Policy*, *184*, 113909. https://doi.org/10.1016/j.enpol.2023.113909

Khan, M., Hanan, A., Kenzhebay, M., Gazzea, M., & Arghandeh, R. (2024). Transformer-based land use and land cover classification with explainability using satellite imagery. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-67186-4

Li, D., Sun, L., Yu, Y., & Tian, P. (2025). *Geospatial micro-estimates of slum populations in 129 Global South countries using machine learning and public data*. https://doi.org/10.5194/essd-2025-260

Li, F., Yiğitcanlar, T., Nepal, M. P., Nguyen, K., & Dur, F. (2023). Machine learning and remote sensing integration for leveraging urban sustainability: A review and framework [Review of *Machine learning and remote sensing integration for leveraging urban sustainability: A review and framework*]. *Sustainable Cities and Society*, *96*, 104653. Elsevier BV. https://doi.org/10.1016/j.scs.2023.104653

Liu, R., Kuffer, M., & Persello, C. (2019). The Temporal Dynamics of Slums Employing a CNN-Based Change Detection Approach. *Remote Sensing*, *11*(23), 2844. https://doi.org/10.3390/rs11232844

Lu, W., Zhuang, D., Peng, F., Feng, Z., & Yang, Y. (2024). A Geoscience-Aware Network (GASlumNet) Combining UNet and ConvNeXt for Slum Mapping. *Remote Sensing*, *16*(2), 260. https://doi.org/10.3390/rs16020260

Lumban-Gaol, Y., Rizaldy, A., & Murtiyoso, A. (2023). COMPARISON OF DEEP LEARNING ARCHITECTURES FOR THE SEMANTIC SEGMENTATION OF SLUM AREAS FROM SATELLITE IMAGES. ˜ *The œ international Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences/International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1439. https://doi.org/10.5194/isprs-archives-xlviii-1-w2-2023-1439-2023

Mahabir, R., Croitoru, A., Crooks, A., Agouris, P., & Stefanidis, A. (2018). A Critical Review of High and Very High-Resolution Remote Sensing Approaches for Detecting and Mapping Slums: Trends, Challenges and Emerging Opportunities [Review of *A Critical Review of High and Very High-Resolution Remote Sensing Approaches for Detecting and Mapping Slums: Trends, Challenges and Emerging Opportunities*]. *Urban Science*, *2*(1), 8. Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/urbansci2010008

Marasinghe, R., Yiğitcanlar, T., Mayere, S., Washington, T. L., & Limb, M. (2024). Towards Responsible Urban Geospatial AI: Insights From the White and Grey Literatures. *Journal of Geovisualization and Spatial Analysis*, *8*(2). https://doi.org/10.1007/s41651-024-00184-2

Moukheiber, D., Restrepo, D., Cajas, S. A., Arbeláez, M. P., Celi, L. A., Kuo, K.-T., López, D. M., Moukheiber, L., Moukheiber, M., Moukheiber, S., Osorio-Valencia, J. S., Purkayastha, S., Paddo, A. R., Wu, C., & Kuo, P. (2024). A multimodal framework for extraction and fusion of satellite images and public health data. *Scientific Data*, *11*(1). https://doi.org/10.1038/s41597-024-03366-1

Neal, I., Seth, S., Watmough, G. R., & Diallo, M. S. (2022). Census-independent population estimation using representation learning. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-08935-1

Neupane, B., Aryal, J., & Rajabifard, A. (2022). BUILDING FOOTPRINT SEGMENTATION USING TRANSFER LEARNING: A CASE STUDY OF THE CITY OF MELBOURNE. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 173. https://doi.org/10.5194/isprs-annals-x-4-w3-2022-173-2022

Nigar, A., Li, Y., Baloch, M. Y. J., Alrefaei, A. F., & Almutairi, M. H. (2024). Comparison of machine and deep learning algorithms using Google Earth Engine and Python for land classifications. *Frontiers in Environmental Science*, *12*. https://doi.org/10.3389/fenvs.2024.1378443

Nuaimi, M. A. A., Salem, E. H., Ibrahim, H. F. S., Ali, A. A. A. A., Abdulmuttalib, H., Biswas, R., Hà, L. H., Banerjee, S., & Mittal, A. S. (2024). Deep Learning-Based Assessment of Urban and Vegetation Changes Using High-Resolution Khalifasat Satellite Imagery over Dubai. *ISPRS Annals of the Photogrammetry, Remote Sensing* *and Spatial Information Sciences*, 137. https://doi.org/10.5194/isprs-annals-x-4-w4-2024-137-2024

Park, S., Lee, S., Kim, J., Lee, J.-G., Cha, M., Yang, J., & Ahn, D. (2025). *Learning Multidimensional Urban Poverty Representation with Satellite Imagery*. https://doi.org/10.48550/ARXIV.2509.04958

Persello, C., & Stein, A. (2017). Deep Fully Convolutional Networks for the Detection of Informal Settlements in VHR Images. *IEEE Geoscience and Remote Sensing Letters*, *14*(12), 2325. https://doi.org/10.1109/lgrs.2017.2763738

Pettersson, M., Kakooei, M., Ortheden, J., Johansson, F., & Daoud, A. (2023). *Time Series of Satellite Imagery Improve Deep Learning Estimates of Neighborhood-Level Poverty in Africa*. 6165. https://doi.org/10.24963/ijcai.2023/684

Raj, A., Mitra, A., & Sinha, M. (2024a). *Deep Learning for Slum Mapping in Remote Sensing Images: A Meta-analysis and Review*. https://doi.org/10.48550/ARXIV.2406.08031

Raj, A., Mitra, A., & Sinha, M. (2024b). Deep Learning for Slum Mapping in Remote Sensing Images: A Meta-analysis and Review. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2406.08031

Russo, L., Sorriso, A., Ullo, S. L., & Gamba, P. (2025). A Deep Learning Architecture for Land Cover Mapping Using Spatio-Temporal Sentinel-1 Features. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1. https://doi.org/10.1109/jstars.2025.3557687

Sabbata, S. D., Ballatore, A., Miller, H. J., Sieber, R., Tyukin, I., & Yeboah, G. (2023). GeoAI in urban analytics. *International Journal of Geographical Information Science*, *37*(12), 2455. https://doi.org/10.1080/13658816.2023.2279978

Sabir, A., Jameela, M., & Malik, A. W. (2023). LAND COVERAGE ANALYSIS OF PAKISTAN USING SATELLITE IMAGERY. ˜ *The œinternational Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences/International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 187. https://doi.org/10.5194/isprs-archives-xlviii-1-w2-2023-187-2023

Sakamoto, K., Jerzak, C., & Daoud, A. (2024). Planetary Causal Inference: Implications for the Geography of Poverty. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2406.02584

Salem, M., & Tsurusaki, N. (2023). Empowering Sustainability: Innovative Deep Learning and Remote Sensing Solutions for Urban Monitoring and Sustainable Development. *Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES)*, *9*, 391. https://doi.org/10.5109/7158006

Sierra, S., Ramo, R., Padilla, M., & Cobo, A. (2025). Optimizing deep neural networks for high-resolution land cover classification through data augmentation. *Environmental Monitoring and Assessment*, *197*(4). https://doi.org/10.1007/s10661-025-13870-5

Silva, J. P. da, Rodrigues, J. F., & Albuquerque, J. P. de. (2025). On the power of CNNs to detect slums in Brazil. *Computers Environment and Urban Systems*, *121*, 102306. https://doi.org/10.1016/j.compenvurbsys.2025.102306

Stark, T., Wurm, M., Debray, H., Zhu, X. X., & Taubenböck, H. (2024). *Uncertainty Aware Slum Mapping in 55 Heterogeneous Cities*. https://doi.org/10.2139/ssrn.4893813

Stephen, A. J., Juba, O. O., Ezeogu, A. O., & Oluwafunmise, F. (2025). AI-Based fall prevention and monitoring systems for aged adults in residential care facilities. *International Journal of Innovative Science and Research Technology*, 2371-2379. https://doi.org/10.38124/ijisrt/25may1548

Vitale, A., & Lamonaca, F. (2025). Advancing Built-Up Area Monitoring Through Multi-Temporal Satellite Data Fusion and Machine Learning-Based Geospatial Analysis. *Remote Sensing*, *17*(11), 1830. https://doi.org/10.3390/rs17111830

Wurm, M., Stark, T., Zhu, X. X., Weigand, M., & Taubenböck, H. (2019). Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, *150*, 59. https://doi.org/10.1016/j.isprsjprs.2019.02.006