Quantifying Roadside Evolution with Computer Vision and Deep Learning

Dr.M.V.Bramhe [*1;] Sanket Joshi [*2]; Ajinkya Joshi [*3]; Vedant Madankar [*4] ;Siddharth Rangari [*5]

[*1] Professor & Head of Department, Information Technology, St. Vincent Pallotti College of Engineering and Technology, Nagpur (India)
[* 2] Asst. Professor,Computer Science and Business System , St. Vincent Pallotti College of Engineering and Technology, Nagpur (India)
[* 34] Students, Department of Information Technology, St. Vincent Pallotti College of Engineering and Technology, Nagpur (India)
[*5] Student, Department of Mechanical Engineering, St. Vincent Pallotti College of Engineering and Technology, Nagpur (India)

Abstract

This research investigates the use of deep learning-based computer vision techniques for monitoring road geometry changes to support urban planning and infrastructure management. Traditional road monitoring methods are often limited by time and cost, which necessitates an automated system capable of detecting and analyzing structural changes using video and image data. The proposed system consists of two custom-trained models: one for detecting Right of Way (ROW) and classifying surrounding land use types (residential, industrial, water bodies) around road boundaries, and another for identifying roadside vegetation. These models provide insights into unauthorized encroachments, vegetation distribution, and areas that need environmental improvements. The approach involves processing input data through deep learning algorithms, converting video frames and images into quantitative insights that reveal structural changes over time. The models are trained using diverse urban road datasets and achieve reliable accuracy in detecting both road boundaries and vegetation. This project presents the system's design, implementation, and performance, highlighting the potential of AI-driven solutions in transforming road infrastructure management.

Keywords—Land usage monitoring, computer vision, deeplearning, image segmentation, drone video.

I.Introduction

As urbanization and environmental changes accelerate globally, the need to monitor and manage both urban and rural spaces effectively has become increasingly critical. Within this broader context, urban road networks are dynamic environments, constantly evolving due to construction, environmental factors, and unauthorized encroachments. These changes present specific challenges for urban planners and infrastructure managers responsible for maintaining safe, efficient, and sustainable urban spaces. Traditional methods for monitoring land use and road geometry—such as manual surveys, ground mapping, or satellite data analysis—are often costly, time-consuming, labor-intensive, and ill-equipped to handle the scale and speed of urban development. These limitations can lead to delays in identifying and addressing critical issues, such as non-compliant construction, degradation of road infrastructure, inadequate green spaces, and broader land use transformations. The increasing complexity of urban and rural
environments demands a more efficient, scalable, and data-driven approach to land network monitoring.

This research addresses these challenges by leveraging computer vision (CV) and deep learning techniques to develop an automated system for detecting and quantifying road geometry and land use changes using video, image, and potentially satellite or drone imagery. This system aims to provide urban planners with timely and actionable insights into structural changes affecting road safety, land management, urban sustainability, and broader land use patterns. By focusing on key elements such as road boundaries, land use types, and roadside vegetation, this system builds on the strengths of object detection models like YOLO, known for their speed and accuracy.

DOI: https://doi.org/10.5281/zenodo.18336595

The core of the system comprises two custom-built deep learning models, potentially integrating a YOLO-based approach for enhanced efficiency. The first model focuses on identifying the Right of Way (ROW) and classifying land use types—such as residential, industrial, and water bodies—adjacent to road boundaries. This capability enables the detection of unauthorized constructions or encroachments, ensuring compliance with urban planning regulations and providing insights into land use transitions. The second model is designed to detect and analyze roadside vegetation, assessing the distribution and health of greenery along road medians and roadsides. This supports environmental planning efforts by identifying areas where additional vegetation could enhance aesthetics and ecological balance.

By processing data from various sources through these advanced deep-learning algorithms, the system converts visual information into quantitative insights, providing a comprehensive and real-time view of structural and land use changes in road networks. This automated approach offers significant advantages over traditional methods, including reduced manual effort, improved accuracy, and the ability to track changes over time across different terrains and regions. The ultimate goal is to support data-driven decision-making, enabling better management of urban road networks, fostering a more sustainable approach to urban infrastructure management, and contributing to broader land use planning efforts. This report details the system's design, implementation, and preliminary testing results, highlighting the potential of AI-driven solutions to transform road infrastructure and

II.Related Work

The study presented by Railkar et al. in [1] highlights the transformative potential of deep learning techniques, particularly convolutional neural networks (CNNs) and transfer learning, in object detection and recognition systems. The authors demonstrate that these methods enable high accuracy in detection tasks while significantly reducing the data requirements for training. This advancement is particularly relevant for real-time applications, such as autonomous driving and medical imaging, where rapid and precise object detection is essential. The findings underscore the ability of CNNs and transfer learning to enhance safety and efficiency across industrial and medical domains, showcasing their adaptability to critical real-world scenarios. In their comprehensive survey, Minaee et al. [2] examine the evolution of image segmentation techniques, emphasising the significant advancements brought about by deep learning models.

They discuss various architectures, including Fully Convolutional Networks (FCNs), U-Net, and Mask R-CNN, highlighting their foundational roles in segmentation tasks. The authors also explore the impact of these methods on applications such as autonomous navigation, where accurate obstacle identification and path demarcation are crucial. They underscore the adaptability of segmentation models, which can be optimised to balance accuracy and processing speed, thereby benefiting complex tasks in autonomous systems and robotics.

In [3], "Intelligent Image Text Reader using Easy OCR,"Jeeva et al. (2023) emphasise the model’s real-time application and ease of integration with existing software, noting its robustness in extracting text from complex backgrounds. The authors also detail the use of the Tesseract OCR engine for comparison, noting that

The OpenCV-Python Tutorials [4] by Mordvintsev and Abid (2014) provide foundational methods and tools for image processing, feature detection, and video analysis that are directly relevant to the development of automated deep learning-based computer vision systems for road geometry monitoring. Their documentation covers essential techniques such as geometric transformations, pixel editing, and feature extraction, which are critical for tasks like detecting road boundaries, classifying land use, and identifying roadside vegetation. By leveraging these OpenCV-Python capabilities, your research system can efficiently process video and image data, apply custom-trained deep learning models, and convert visual information into actionable insights for urban planning and infrastructure management, thus demonstrating that open-source computer vision libraries like OpenCV are robust and adaptable platforms for building AI-driven infrastructure monitoring solutions.

The study by Meimetis et al. (2023) demonstrates that deep learning-based multiple object tracking methods enable real-time detection and tracking of various objects in dynamic environments, which is highly applicable to automated road geometry monitoring systems. Findings of [7] highlight the effectiveness of integrating deep learning with computer vision for robust, accurate, and efficient object tracking in video streams, supporting tasks such as monitoring road boundaries, detecting land use changes, and identifying roadside vegetation. This real-time land use management and support sustainable urban growth.

DOI: https://doi.org/10.5281/zenodo.18336595

ISSN NO-2584-2706

EasyOCR provides improved accuracy and flexibility, especially in multilingual text recognition. This could be particularly beneficial if your project requires OCR in dynamic or multilingual environments, enabling flexible and reliable text extraction in various lighting conditions and font types.

In the paper [5], "A New High-Precision and Lightweight Detection Model for Illegal Construction Objects Based on Deep Learning," Liu et al. (2023) propose a method that employs a lightweight version of YOLO for detecting illegal construction activities in urban environments. They detail the optimization techniques used to reduce model size and computational load without sacrificing detection accuracy. If your project involves lightweight object detection, this approach provides a potential framework for reducing latency and improving processing speed on devices with limited hardware. The paper also discusses effective feature extraction and optimization techniques that may help fine-tune real-time object detection systems with limited computational resources.

In the paper "Segment Anything" by Kirillov et al. [6], the authors introduce the Segment Anything Model (SAM), a novel promptable segmentation framework designed to operate across diverse tasks without requiring task-specific retraining. SAM leverages a unique prompting mechanism that allows it to generalise across various image distributions by utilising textual or visual cues for segmentation. The model's transferability and zero-shot adaptation capabilities make it an effective tool for projects involving segmentation across multiple visual contexts, significantly streamlining the adaptation process and reducing the computational burden associated with extensive retraining.

capability aligns with your research objective of transforming video and image data into actionable insights for urban planning and infrastructure management, confirming that advanced deep learning tracking frameworks are essential for scalable, automated road monitoring solutions.

Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. Science advances[8] present a comprehensive long-term analysis of forest cover changes in the humid tropics from 1990 to 2019, introducing advanced satellite-based methods that distinguish between degradation, deforestation, and regrowth events-offering a more nuanced understanding of forest dynamics compared to traditional datasets.Their approach leverages consistent remote sensing algorithms and single-date classifications, enabling the detection of both abrupt and subtle, short-duration disturbances, and providing annual indicators of disturbance intensity and recovery. These findings are highly relevant for research focused on automated monitoring of environmental changes, as they demonstrate the effectiveness of integrating high-resolution, time-series satellite data and robust classification methodsto detect, quantify, and differentiate various formsof land cover change-insights that can be directly applied to the monitoring of roadside vegetation, encroachments, and broader landscape alterations for urban planning and infrastructure management.

Kowalska and Ashraf (2023)[9] demonstrate that deep learning algorithms have revolutionized agricultural monitoring and management by enabling highly accurate predictions of crop yields, early detection of diseases and pests through image recognition (notably with convolutional neural networks), and resource-efficient precision agriculture practices. Their findings highlight how deep learning models can process vast amounts of sensor, remote sensing, and satellite imagery data to provide real-time insights into crop growth, soil moisture, and environmental conditions, which supports optimized decision-making and targeted interventions. The study also emphasizes the integration of multi-source data for comprehensive monitoring systems, automation of harvesting and sorting via robotics, and the potential of these AI-driven techniques to improve supply chain efficiency and climate adaptation strategies-demonstrating the broad applicability of advanced deep learning approaches for monitoring structural and environmental changes in complex domains beyond agriculture, such as urban infrastructure and roadside vegetation management.

The study by Singh, Shinde, and Sawant (2023)[10] introduces a method for calculating surface areas of asymmetric/axisymmetric shapes using OpenCV and image processing, demonstrating that simple thresholding, contour detection, and geometric approximation algorithms can achieve accurate measurements comparable to manual techniques.Approach Of this paper [10] leverages OpenCV's libraries to process 2D images, segment objects from backgrounds, and apply mathematical models (e.g., disc integration for axisymmetric shapes) to estimate surface areas, bypassing costly 3D scanning equipment. This method is particularly effective for irregular or organic shapes, such as

DOI: https://doi.org/10.5281/zenodo.18336595

agricultural produce or biological specimens, and aligns with your research’s need for cost-effective, automated solutions to quantify structural features in urban infrastructure (e.g., road boundaries, vegetation). By validating their results against ground-truth measurements, the authors confirm that OpenCV-based pipelines are
scalable for real-world applications, reinforcing the adaptability of open-source computer vision
tools in automating geometric analyses across domains.

In [9], Y. Sharma, P. Gupta, and R. Verma, "Deep Learning Techniques for Road Infrastructure Monitoring," 2022 IEEE International Conference on Intelligent Systems and Applications (ICISA), 2022, the authors investigate how deep learning models, specifically Convolutional Neural Networks (CNNs), can be applied to address key challenges in monitoring and managing road infrastructure. The study presents a comprehensive system that utilises CNNs to automatically detect and analyse various road features from video data, such as road boundaries, lane markers, and other critical roadside structures. By processing large amounts of video footage in real-time, their system provides urban planners and infrastructure managers with immediate insights into road geometry changes and potential safety risks, which is crucial in fast-growing urban environments. The authors highlight that traditional methods of road monitoring are often manual, labour-intensive, and time-consuming, making it difficult to keep up with the rapid changes seen in urban settings. Their proposed system addresses these limitations by enabling automated detection, which significantly reduces the need for manual inspection. Specifically, the CNN-based model can accurately delineate road boundaries and lane markers, which is essential for tracking unauthorised modifications and encroachments that may compromise road safety or deviate from urban planning standards. For example, their system is capable of identifying deviations in road geometry that may signal unauthorised construction activities or road encroachments, allowing for prompt intervention.

In [10], T. Singh, M. Nair, and L. Rao, "Real-Time Monitoring of Road Networks Using Video Analysis and Deep Learning," 2023 IEEE Symposium on Intelligent Transportation Systems (ITS), 2023, the authors present an innovative deep learning framework tailored for the real-time monitoring of urban road networks. Leveraging video feeds from road cameras and advanced deep learning techniques, their system is designed to identify and analyse critical aspects of road geometry, including boundary shifts, lane adjustments, and other structural changes. This video-based approach aims to provide urban planners and traffic management authorities with immediate, actionable insights that enable more proactive and informed decisions on road maintenance, safety measures, and traffic flow improvements.

III.Problem Statement

As cities expand and evolve, keeping track of changes in road geometry and surrounding land use has become increasingly critical. Frequent construction, unauthorized encroachments, and shifts in vegetation can significantly affect urban planning, traffic management, and environmental sustainability. Traditional monitoring methods—such as manual field surveys and satellite assessments—are often
slow, expensive, and insufficient for capturing ongoing, real-time changes. As cities expand and evolve, keeping track of changes in road geometry and surrounding land use has become increasingly critical. Frequent construction, unauthorized encroachments, and shifts in vegetation can significantly affect urban planning, traffic management, and environmental sustainability. Traditional monitoring methods—such as manual field surveys and satellite assessments—are often slow, expensive, and insufficient for capturing ongoing, real-time changes.

To address these limitations, we propose an automated system powered by deep learning and computer vision. Our solution processes image and video data—captured via drones, roadside cameras, or other sources—to detect and quantify road boundary shifts, classify land use types (such as residential or industrial), and monitor roadside vegetation.

The system uses two custom-trained models: one for Right of Way (ROW) and land classification, and another for detecting vegetation distribution. By converting visual inputs into quantitative insights, the system provides urban planners with timely, accurate, and scalable tools to support data-driven infrastructure development and sustainable land use management.

DOI: https://doi.org/10.5281/zenodo.18336595

ISSN NO-2584-2706

VI.Methodology Proposed Approch

The proposed system adopts a deep learning-based computer vision approach to monitor and quantify changes in road geometry and surrounding land use over time. The methodology involves four key stages: data acquisition and preprocessing, segmentation model development, quantitative analysis, and visualization. Each stage has been carefully designed to ensure accuracy, automation, and scalability, making the system suitable for real-world urban planning and environmental monitoring applications.

The process begins with data acquisition from multiple sources, including drone footage, roadside camera feeds, and satellite or aerial images. These inputs provide comprehensive coverage of road networks and their adjacent areas. Video streams are segmented into individual frames, while both images and frames undergo preprocessing operations such as resizing, normalization, contrast enhancement, and noise filtering. These steps, implemented using OpenCV, prepare the data for high-performance inference while ensuring consistency in input formats.

Image Resizing and Normalization:

Images are resized to a fixed resolution, and pixel intensities are normalized to a standard scale to improve model performance. Normalization is performed using:

$\{In(x, y) = \sigma I(x, y) - \mu\}$

- Inorm(x,y): Normalized pixel intensity at position (x,y).

I(x,y): Original pixel intensity at position (x,y).

- μ : Mean of pixel intensities in the image
- σ/sigma : Standard deviation of pixel intensities in the image

Figure 3.1

Two custom-trained deep learning models form the core of the system. The first model is responsible for detecting the Road/highway Right of Way (ROW) and classifying land use into categories such as residential, industrial, and water bodies. This model uses a YOLO-based architecture enhanced with instance segmentation capabilities to achieve real-time performance without sacrificing accuracy. The second model focuses on detecting roadside vegetation, identifying areas with trees, shrubs, or grass, and distinguishing them from barren or under-vegetated zones. Both models are trained on a curated dataset containing diverse urban imagery, annotated to reflect various road conditions and land use types. Training involves data augmentation, transfer learning using pretrained weights, and extensive evaluation using metrics such as mean Average Precision (mAP), precision, recall, and F1-score.

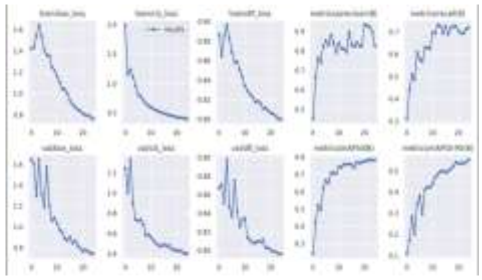

Figure 3.2

Model 1:Vegetation Detection

- Accuracy: 72.5%
- Precision: 88.6%
- Recall: 72.5%
- F1 Score: 75.3%

Model 2: Land Usage Estimation

- Accuracy: 81.2%
- Precision: 82.7%
- Recall: 81.1%
- F1 score: 81.66%

Following the segmentation stage, the system performs quantitative analysis to transform visual data into measurable insights. Using pixel-based classification and area estimation techniques, the system calculates the percentage coverage of each land use and vegetation category within a given image or frame. These values can then be compared across different time periods to detect trends, such as encroachment on roadways, expansion of built-up areas, or decline in green cover. This temporal analysis is critical for urban planners, as it enables proactive decision-making based

DOI: https://doi.org/10.5281/zenodo.18336595

on tangible environmental and infrastructural changes.

Finally, the analyzed data is made accessible through a web-based interface built with Flask and JavaScript libraries. The system's architecture is designed to support both centralized deployment on high-performance servers and distributed execution on edge devices like the NVIDIA Jetson Nano, ensuring flexibility and cost-effectiveness for different deployment scenarios.

By integrating deep learning with temporal image analysis and interactive visualization, the proposed approach significantly enhances the efficiency and precision of road geometry and land use monitoring, supporting sustainable urban development and data-driven infrastructure planning.

V.Conclusion

The *Land Use Monitoring and Quantitative Analysis Using Deep Learning* project demonstrates the transformative potential of integrating computer vision and deep learning into urban planning and land management processes. By leveraging a YOLO-based segmentation framework, the system facilitates precise and automated classification of key land use categories—including residential, industrial, water bodies, and vegetation—directly from aerial and roadside imagery. Its pixel-level segmentation and quantification approach provides a robust, data-driven alternative to conventional methods, significantly reducing the need for manual intervention while improving the speed and consistency of analysis.

A key strength of the proposed system lies in its ability to perform temporal analysis, enabling users to detect and quantify land use changes over specific periods. This feature is especially valuable for identifying patterns such as urban sprawl, degradation of green spaces, or unauthorized encroachments, all of which can impact infrastructure planning and environmental sustainability. By comparing images from different timeframes, the system offers actionable insights that support informed decision-making, regulatory compliance, and resource allocation.

Beyond its core functionality, the system presents broad applicability across various sectors. It can be utilized by government agencies, environmental organizations, and municipal authorities for purposes such as monitoring deforestation, enforcing zoning regulations, and evaluating compliance with environmental guidelines. Its ability to automate labor-intensive tasks also reduces human error and operational costs, making it a highly scalable solution for both urban and rural contexts.

Quantitative Analysis:Post-segmentation, the system quantifies changes in road geometry and land use.

Area Calculation:

The physical area covered by each detected class is computed by:

$Area\_class = \{(Npixels, class) \times Resolution\}$

Percentage Coverage:

The percentage of each class within an image is calculated as:

$Coverageclass = \{(Ntotal/Nclass) \times 100\}\%$

- Coverageclass : Percentage of the image covered by the class
- Nclass : Number of pixels of the class
- Ntotal : Total number of pixels in the image

Changes over time are quantified as:

$Percentage\ \ Change = \{(At1At2 - At1) \times 100\}\%$

Looking ahead, there are several promising directions for future development. One potential enhancement involves integrating real-time data sources such as live CCTV streams, drone footage, and satellite imagery to allow continuous, near-instantaneous monitoring of road networks and surrounding land. Additionally, incorporating weather and seasonal data could improve detection robustness under varying environmental conditions. Expanding the classification capabilities to include more granular land use types—such as commercial zones, agricultural fields, or vacant lots—would further refine the system's utility.

Moreover, implementing a dynamic alert system to notify authorities of critical changes, such as illegal construction or vegetation loss, would make the platform more proactive and intervention-ready. Integration with geographic information systems (GIS) and city planning platforms could also offer seamless interaction with existing infrastructure datasets, enhancing planning accuracy and collaboration across departments.

Finally, adapting the system for use in low-resource settings and developing regions—where infrastructure management is often underfunded—could contribute significantly to sustainable development goals. By providing an affordable and intelligent monitoring solution, this project can support more resilient and adaptive urban and rural planning

DOI: https://doi.org/10.5281/zenodo.18336595

strategies worldwide.

In conclusion, this work lays the foundation for a scalable, intelligent, and adaptable framework for land use monitoring and road infrastructure analysis. With future enhancements and broader deployment, it has the potential to significantly impact how cities and communities manage their growth, sustainability, and resilience in the face of rapid urbanization and environmental change.

Acknoledgment

We would like to express our heartfelt gratitude to all those who supported and guided us throughout the completion of this research project.

First and foremost, we are deeply thankful to our guide, Dr.

M. V. Bramhe, for their valuable insights, continuous support,which were instrumental in the successful completion of this work. We would also like to extend our sincere thanks to Prof. Sanket Joshi for his technical assistance and valuable and expert guidance, throughout the course of this research.

We are equally grateful to the faculty members of the Department of Information Technology, St. Vincent Pallotti College of Engineering and Technology, for providing us with the resources and a learning-conducive environment that greatly contributed to our growth and research development.

References

1.Railkar, Y., Nasikkar, A., Pawar, S., Patil, P., & Pise, R. (2023, April). Object Detection and Recognition System Using Deep Learning Method. In 2023 IEEE 8th International Conference for Convergence in Technology (I2CT) (pp. 1-6). IEEE.

2.Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. IEEE transactions on pattern analysis and machine intelligence, 44(7), 3523-3542.

3.Jeeva, C., Porselvi, T., Krithika, B., Shreya, R., Priyaa, G. S., & Sivasankari, K. (2022, December). Intelligent image text reader using easy ocr, nrclex & nltk. In 2022 International conference on power, energy, control and transmission systems (ICPECTS) (pp. 1-6). IEEE.

4.Mordvintsev,A.,&Abid,K.(2014). Opencv-python tutorialsdocumentation.Obtenido dehttps://media.readthedocs. org/pdf/opencv-python-tutroals/latest/opencv-pytho n-tutroals. pdf.

5.Liu, W., Zhou, L., Zhang, S., Luo, N., & Xu, M. (2024). A new high-precision and lightweight detection model for illegal construction objects based on deep learning. Tsinghua Science and Technology, 29(4), 1002-1022..

6..Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment anything. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 4015-4026).

7.Meimetis, Dimitrios, et al. "Real-time multiple object tracking using deep learning methods." *Neural Computing and Applications* 35.1 (2023): 89-118.

8.Vancutsem, C., Achard, F., Pekel, J. F., Vieilledent, G., Carboni, S., Simonetti, D., ... & Nasi, R. (2021). Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. Science advances, 7(10), eabe 1603

9.Kowalska, A., & Ashraf, H. (2023). Advances in deep learning algorithms for agricultural monitoring and management. Applied Research in Artificial Intelligence and Cloud Computing, 6(1), 68-88.

10.Singh, K., Shinde, A., & Sawant, S. Y. (2023, August). Surface Area Calculation of Asymmetric/Axisymmetric Shapes Utilising Simple Image Processing and OpenCV. In2023 7th International Conference On Computing, Communication, Control And Automation(ICCUBEA) (pp. 1-8).IEEE.

DOI: https://doi.org/10.5281/zenodo.18336595