JOURNAL OF BUSINESS IN THE DIGITAL AGE
Vol. 7 Issue 2 2024
eISSN: 2651-4737 dergipark.gov.tr/jobda

Research Article

ARTIFICIAL INTELLIGENCE SUPPORTED CITY INFRASTRUCTURE MANAGEMENT: AUTOMATIC DETECTION OF MANHOLE COVERS AND DRAINAGE WITH YOLO ON GOOGLE STREET VIEW IMAGES

YAPAY ZEKA DESTEKLİ ŞEHİR ALTYAPI YÖNETİMİ: GOOGLE STREET VIEW GÖRÜNTÜLERİNDE YOLO İLE RÖGAR KAPAKLARININ VE MAZGALLARIN OTOMATİK TESPİTİ

Can Aydın [1,*] | Gizem Erdoğan Aydın [2]

[1] Doç. Dr., Dokuz Eylül Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü, ORCID: 0000-0002-0133-9634
[2] Doç. Dr., İzmir Demokrasi Üniversitesi, Mimarlık Fakültesi, Şehir ve Bölge Planlama Bölümü, ORCID: 0000-0002-1376-6457

Article Info:
Received: October, 29, 2024
Revised : November 26, 2024
Accepted: December 10, 2024

Keywords:
Yolo
object detection
Google Street View
urban analytic
deep learning

Anahtar Kelimeler:
Yolo
nesne tespiti
Google Street View
kent analitiği
derin öğrenme

DOI: 10.46238/jobda.1575356

*** Sorumlu Yazar**
e-posta: canaydinn@gmail.com

ABSTRACT

With rapid urbanization, maintaining urban infrastructure has grown into a gigantic requirement. Proper and timely identification of infrastructure assets, such as manhole covers and drainage, is of utmost importance to ensure that water drainage and sewerage systems work properly within the precincts of a city. The classical methods of inspection have contributed to being slow, expensive, and full of errors. The paper tries to implement the use of YOLO in the automatic detection of manhole covers and drainage in images derived from Google Street View. This study will be focused on how to integrate results from object detection with MIS in order to monitor city infrastructures and optimize the planning of maintenance. These results proved that YOLOv11 has a high accuracy rate and has identified manhole covers and drainage from imagery on Google Street View. Performance metrics included mAP@0.5 and mAP@0.5-0.95, which described sensitivity and accuracy of the model, while the FPS analysis described the applicability in real time. Those kinds of findings have underlined that AI-based solution usage is efficient in the automatic monitoring and management of urban infrastructure and prove their potential to contribute much to decision support systems.

ÖZ

Hızlı kentleşmeyle birlikte, kentsel altyapının bakımı devasa bir gereksinim haline gelmiştir. Rögar kapakları ve mazgal gibi altyapı varlıklarının doğru ve zamanında tespit edilmesi, su drenaj ve kanalizasyon sistemlerinin bir şehrin sınırları içinde düzgün çalışmasını sağlamak için son derece önemlidir. Klasik denetim yöntemleri yavaş, pahalı ve hatalarla dolu olmasına katkıda bulunmuştur. Bu makale, Google Street View'dan elde edilen görüntülerde rögar kapaklarının ve mazgalların otomatik olarak tespit edilmesinde YOLO kullanımını uygulamaya çalışmaktadır. Bu çalışma, şehir altyapılarını izlemek ve bakım planlamasını optimize etmek için nesne tespitinden elde edilen sonuçların YBS ile nasıl entegre edileceğine odaklanacaktır. Bu sonuçlar, YOLOv11'in yüksek bir doğruluk oranına sahip olduğunu ve Google Street View görüntülerinden rögar kapaklarını ve mazgalların tespit ettiğini kanıtlamıştır. Performans ölçütleri arasında modelin hassasiyetini ve doğruluğunu tanımlayan mAP@0.5 ve mAP@0.5-0.95 yer alırken, FPS analizi gerçek zamanlı uygulanabilirliği tanımlamıştır. Bu tür bulgular, yapay zeka tabanlı çözüm kullanımının kentsel altyapının otomatik olarak izlenmesi ve yönetilmesinde etkili olduğunun altını çizmiş ve karar destek sistemlerine büyük katkı sağlama potansiyellerini kanıtlamıştır.

1 | INTRODUCTION

Indeed, in this fast-marching world of rising urbanization, the better management and viability of its urban infrastructure is the challenge of the day. Like other common structural components, the manhole covers and drainage play a critical role in the water drainage systems and sewerage mechanisms and have taken a frontline significance in the daily functioning of cities. They require routine monitoring and maintenance so that any potential infrastructure problem may be avoided, guaranteeing public safety. These activities involve identification and monitoring, which are highly time-consuming and costly, yet prone to errors.

The growing complexity of urban environments and the increasing population density in cities demand innovative approaches to infrastructure management. Urban infrastructure must not only meet the needs of current residents but also anticipate future demands in terms of safety, functionality, and sustainability. Consequently, incorporating advanced technologies into urban management strategies is no longer optional but a necessity.

Current methods of manual inspection hardly allow for updating and obtaining rather precise data over large-scale geographic areas; it reduces the effectiveness of operations within municipal governments and leads to oversight where urgent intervention is required. Similarly, rich, timely image databases- such as provided by Google Street View-are increasingly valuable in automated mapping and monitoring applications applied to city infrastructures. These data processing reductions obtained from these tools reduce fieldwork and allow spending higher productive resources. However, these data sources require efficient processing methodologies to extract actionable insights at scale. Artificial intelligence (AI) and deep learning techniques offer powerful solutions for leveraging such large datasets, enabling rapid, accurate, and cost-effective infrastructure assessments. More recently, thanks to the development of artificial intelligence and deep learning, it has been possible to develop new opportunities related to the analysis of large-scale data and object detecting applications. In particular, the methods based on You Only Look Once have phenomenal performance for real-time object detection. Among them, the latest version, now called YOLO v11, was able to perform fast and accurate detections for complicated visual data. The work at hand will present the automatic manhole cover and grating detection in Google Street View images performed with the use of the YOLO v11 algorithm. Data will be used within management processes regarding city infrastructure management, effective maintenance and repair processes, and advanced-level decision support mechanisms. All these will contribute to increasing the operation efficiency in city administrations and also raising the quality of the public services.

Moreover, the findings of this study have implications for urban resilience and sustainability, aligning with global efforts to develop smarter, more adaptive cities. By addressing challenges such as data accuracy, model optimization, and operational scalability, this research aims to demonstrate how AI-based solutions can serve as catalysts for transformative change in urban planning and governance.

The following are the research questions to be addressed within the scope of this study.

- How effective is the YOLO Algorithm for detecting manholes and drains in street images?
- Which ones are possible data preprocessing and model optimization methodologies that might be applied to enhance performance for manhole and drainage detection using YOLO?
- The main question is, what novelty does YOLOv11 bring compared to the already existing YOLOv8 or their older versions, and in what scenarios does that translate into more efficient performance?

2 | LITERATURE REVIEW

Management of urban infrastructures is quite important in realizing the sustainability and safety of a modern city. The automatic defect detection relating to manhole covers, gratings, and road surfaces will accelerate the maintenance and repairing processes of infrastructural elements with reduced human errors. Deep learning and computer vision have developed in the recent past and hence brought significant progresses to the monitoring and evaluation processes relating to the urban infrastructures. For instance, some of the real-time object detection algorithms like YOLO can detect infrastructure features very fast with high accuracy. Wang et al. (2022) developed a YOLO-SDD with YOLOv5s that detected stormwater drains from street-level imagery. They optimize the backbone network and loss function by analyzing characteristics of the small-scale targets. The experimental results show that the mAP@0.5 reaches 89.6% for detecting various stormwater drains states in different environmental conditions. Another relevant article, Benhiba et al. 2023, applied the YOLOv8 model to the detection of manhole covers through inspection using drone images and adding GPS location information. The experiments showed a very good performance of YOLOv8, 89% for mAP@50, and 95% for accuracy. This contributes to proactive maintenance and mitigation of risks in the urban infrastructure. Singh et al. (2023) aimed to detect potholes on the road for intelligent transport systems. They made a comparison in performances of some deep learning-based object detection methods such as YOLOv5, YOLOv6, and YOLOv7 in road damage detection. Among all, the best performance was obtained by YOLOv7 with 93% accuracy for detection. In this regard, Zhou et al. (2019) had four models trained and tested for automatically detecting potholes on the road: YOLOv3, SSD, HOG with SVM, and Faster R-CNN. The experimental results demonstrated that the YOLOv3 model outperformed the others in that it produced faster and reliable results in detection. Thereafter, Noori et al. (2023) presented deep learning-based approaches for assessing the severity of asphalt patches and manhole covers. They carried out work that involved the execution of a one-stage object detection algorithm using YOLOv5, YOLOv6, and YOLOv7 and showed that YOLOv5 had the best performance out of the three at high speed. This work has depicted the leading trend of deep learning models in detecting asphalt patches and manhole covers. Wang et al. (2023) propose the detection of manhole covers using aerial images captured by a UAV. They placed manhole covers using YOVOv8 object detection technology and enhanced image quality using super-resolution processing via the SRGAN network. Therefore, they have achieved manhole covers classification accuracy of 97.62%. Li et al. 2020, have proposed an automatic sewer pipe defect detection system based on the deep learning algorithm YOLOv3. For example, in six various classes of model output identifications, the broken, hole, debris, crack, fracture, and root classes return an average precision of approximately 85.37%. Again, in proof of efficiency for deep-learning approaches to regular monitoring of sewerage systems.

The above literature highlights the relevance of deep learning, especially the place occupied by YOLO algorithms for urban infrastructure management. Variably, various studies prove the very high accuracy and efficiency of the different versions of YOLO but also the use of other deep learning models to automatically detect manhole covers, gratings, and road damages. Most of the works conducted so far require imagery from drones or UAVs, and very few research works have focused on street-level imagery such as Google Street View. In this regard, this research paper will focus on the automatic detection of manhole covers and gratings using the YOLO algorithm in Google Street View images and try to fill up the gap in the literature. In that case, street-level imagery will provide data coverage on a wider scale, presently and through regular updates, enabling better and more economic monitoring of urban infrastructure with time.

Incorporating spatial planning and Geographic Information Systems (GIS) into urban infrastructure management enhances decision-making and resource allocation. GIS applications in urban and regional planning are diverse and essential for land management. Significant uses include risk management and emergency planning, where GIS data connects emergency management with spatial planning through network analysis and thematic mapping. GIS also aids in standardizing and validating urban data by collecting and analyzing socioeconomic and environmental information, facilitating methodologies like overlay analysis to identify conflicts between land development and environmental concerns. Additionally, GIS supports the execution of urban plans by conducting environmental impact assessments of proposed projects, evaluating and minimizing development impacts on the environment.

Integrating Building Information Modeling (BIM) and GIS has been approached from relevant aspects such as standardization and level of detail, aiming to improve the operation and maintenance of urban infrastructure. This integration enhances the management of existing infrastructure by combining detailed building models with spatial context, facilitating better decision-making in maintenance and operations (Cepa et al. 2024).

GIS also plays a crucial role in urban infrastructure planning and management by developing and maintaining the physical infrastructure that supports urban areas, including transportation, water and sewer systems, waste management, and public spaces. The goal is to ensure that infrastructure is safe, reliable, efficient, and meets the needs of urban populations. Furthermore, GIS streamlines asset management by centralizing information on maintenance schedules, repair history, and equipment inventory. By tracking the lifecycle of each asset, from installation to decommissioning, GIS empowers organizations to prioritize maintenance tasks, prolong asset lifespan, and minimize downtime (Lamp 2024).

This research paper focuses on the automatic detection of manhole covers and gratings using the YOLO algorithm in Google Street View images, aiming to fill the gap in the literature. Street-level imagery provides broader data coverage, enabling better and more economical monitoring of urban infrastructure over time. Integrating deep learning-based detection with GIS and spatial planning frameworks can enhance urban infrastructure management by providing accurate, real-time data for decision-making, improving maintenance efficiency, and contributing to the sustainability and safety of modern cities.

3 | METHODOLOGY

Experiments are designed to comparatively assess different deep learning object detection models for manhole covers and grating auto-detection. In this regard, the paper considers an experimental design methodology comprising tests in four phases, comparing the performance of the YOLO algorithms and selecting the best model.

The labeled dataset created consisted of 686 Google Street View images featuring manhole covers with drainage labels, amounting to 1001 labels where manhole covers were created using the Roboflow tool-manually generated with 635, while for drainage, 366 was prepared.

In the post-labeling process, some data augmentation techniques were performed in order to make the variations in the dataset more diverse, including: rotation, scaling 640x640 pixels, and adjusting brightness and contrast. After augmentation, a total of 1646 images are obtained: 70% data for training is 1440 images, 15% for validation is 136 images, and 15% for testing is 67 images. The first experiment compares the performance between YOLOv5, YOLOv8, and YOLOv11. All the models have been trained on the same training and validation datasets, keeping similar training parameters. Model performance is quantified by metrics such as mAP@0.5, Precision, Recall, and F1 Score. The second experiment compares several size versions of YOLOv11: YOLOv11n or nano, YOLOv11s or small, and YOLOv11m or medium version. In this experiment, this investigates how changes in model size affect the speed and accuracy of the model. The evaluation metrics that will be used include mAP@0.5, model size in megabytes, and processing speed in frames per second. The performances of the third experiment were tested on the test dataset, comparing both the speeds and accuracies. Later, the applications were run on the test dataset that had not been used for training. Further, their performance was compared against real-world data. In this regard, mAP@ 0.5 and @0.5:0.95, among other metrics to be considered including processing time, ms/ image, and FPS. It therefore gives a full comparison performance analysis of various YOLO models in detecting the covers of manholes and gratings that henceforth guide the selection of the best object detection model for practical applications in managing city infrastructures. These experiments aim to ascertain, with regard to both aspects of accuracy and speed, the suitability of the models for practical applications. It empowers city governments with better monitoring of their infrastructure assets and improvements in the ways maintenance is done.

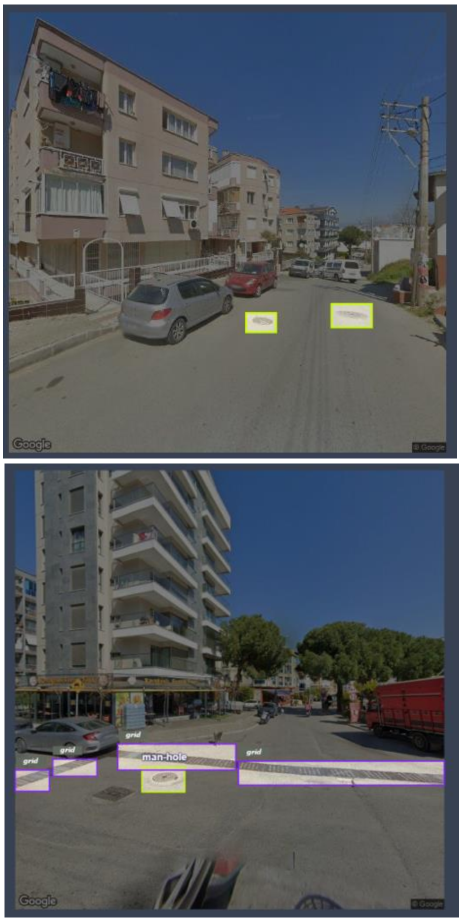


Figure 1. Labeling of manhole covers and drainage

3.1. Modeling

Regarding this, various experiments are conducted concerning specific hyper-parameters and training strategies to achieve optimum performance of the deep learning model called YOLOv11 for detecting manholes and drainage in Google Street View images, so as to enhance the generalization capability of the model. This experimental design adopted in this paper is based on the requirement needed for the research question at hand, while aspiring to fill in the gaps evident in the literature.

In this context, some experiments, such as version comparison for YOLO, determination of the number of epochs, and early stop strategy are done. The experiments here are selected since, from the literature in deep learning [Liao etc. 2022, Yang and Shami 2020, Bischl etc. 2023, Du etc. 2021, Rjin and Hutter 2018] optimization in hyperparametric parameters and training strategies are the keys leading to better model performances. This shall help in filling up the literature gap by providing optimized performance of the YOLOv11 model in identifying manholes and grilles within the settings of an urban environment.

3.2. Experiment comparing the results of different Yolo versions

In this study, the experiment was conducted with three different versions of YOLO that faced each other frequently in the literature. A model will be created by using Yolo v5, and thereafter the performance values will be analyzed (Figure 2). It can be inferred from the results analysis about the Confusion Matrix and the F1-Confidence Curve that the class "man-hole" has a higher F1 score than "drainage" class, which means, in general, this class is presenting more detection accuracy (Figure 2a). While in this problem, the F1-Score was more successful, specifically in the "man-hole" class for the Model; the "drainage" class still has a pretty low precision and recall value. mAP@0.5 value for all classes is 0.583-the said model works with acceptable accuracy; it needs to be improved for some classes (Figure 2b). About 0.144 mAP@0.5-0.95: a model performs badly, especially on more difficult detections since a line is at the bottom left while drawing across different IoU thresholds. It can be observed from the Precision-Recall curve that the "man-hole" class has a much better result in comparison with other classes (Figure 2c). The "drainage" class has low precision and recall, relating to many false negatives and positives in the detections (Figure 2d). Losses such as train/box_loss, train/cls_loss, train/cls_loss, val/box_loss, among others, keep decreasing during training and verification. That provides evidence that the model is learned in some sort of process and is making fewer mistakes. It would be easy, though, to notice fluctuations, especially those related to validation losses, because some evidence for overfitting might have happened in some epochs. That is to say, the model performance in the "manhole" class is middle, but gives huge errors and poor performance in the class "drainage." Generally, during this fit for FPS and speed evaluation, sensitivity and recall values have to be increased, especially in the class "drainage," in order to reach the desired level in accuracy.

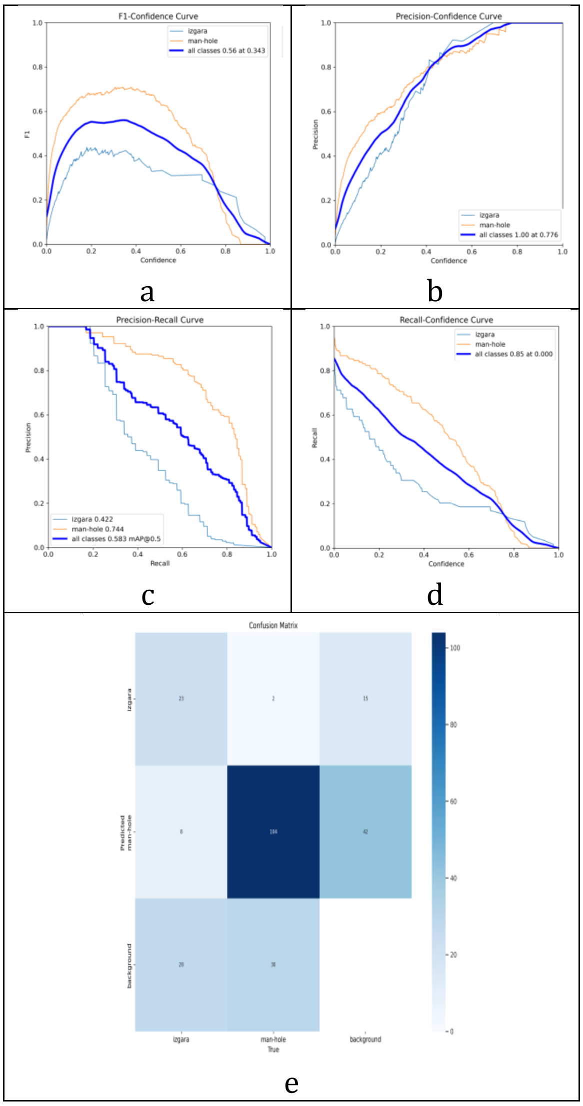


Figure 2. Result values of the Yolo v5 model

The Accuracy, F1-Score, and Average Precision-mAP@0.5/mAP@0.5-0.95 scores are depicted below for the given YOLOv8 model (Figure 3). By considering the Precision-Recall curve as well as the F1-Confidence curve, one is able to analyze that compared to a class "drainage," the "man-hole" class gives rather successful detection accuracy (Figure 3a). The "man-hole" class has higher values both for precision and recall; therefore, the F1 score is on this side. It doesn't perform that well compared to the "drainage" class. Precision and Recall values have many fluctuations. This shows that it might sometimes come under conditions of false positive or false negative. For the mean Average Precision mAP@0.5 across all classes, the value is 0.609, hence it would give quite accurate detections (Figure 3b). The value of mAP@0.5-0.95 comes out to be 0.154. This further means that with higher IoU threshold values, the model performs worst and needs further improvement on most challenging detections. If considering Precision-Confidence and Recall-Confidence Curves, the "manhole" class outperforms the "drainage" class. Though both classes have high recall value, "manhole" class has higher whereas the "drainage" classes have lower recall rates (Figure 3c). It is visible that loss both trains and validates-a model learns after some time and makes fewer mistakes. In real time, if there are too many ups and downs in terms of validation losses, one can draw a conclusion with confidence: most likely, this network is overfitted and requires a regular strategy of training (Figure 3d). The performance of Model YOLOv8 was great, especially class "man-hole" had high accuracy and F1 score. Though this performs fantastically well, it has very poor performance within the "drainage" class, which actually needs further development based on more data.

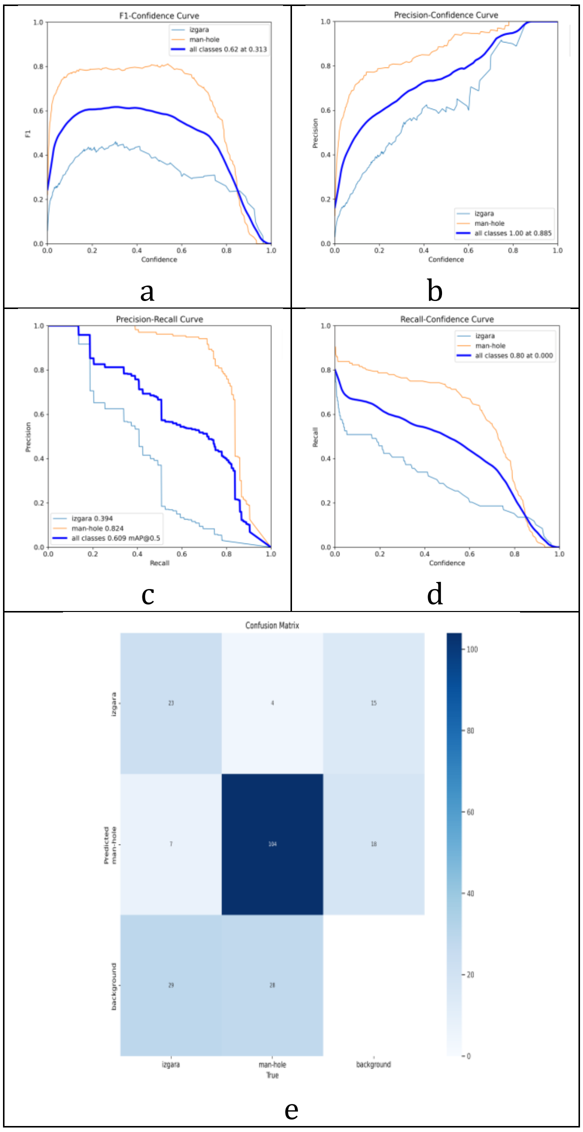
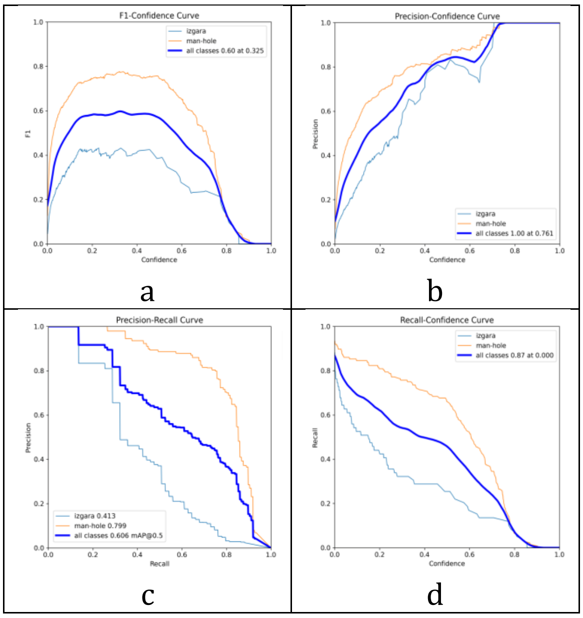


Figure 3. Yolo v8 model results

Following are the results for the evaluations by the YOLOv11 model: From the F1-Confidence Curve, it is seen that the class "manhole" ensured a better F1-Score as opposed to the class "drainage" (Figure 4). The F1-score can be seen around 0.60, while the performance of the class "drainage" is very poor (Figure 4a). From what the Precision-Recall curve represents on all classes by this YOLOv11 model, the estimated mAP@0.5 amounts to 0.606, depicting that it is acceptable in overall performance for the model (Figure 4b). On the other hand, mAP@0.5-0.95 managed by the model is only 0.161, representing that there needs to be more improvement in performance w.r.t more challenging detection across different IoU thresholds (Figure 4c). It can be observed from the Precision-Confidence Curve that the man-hole class is performing better compared to the drainage class concerning precision and recall. The Recall-Confidence Curve of the drainage class shows very low recall values; it means the model tends to predict more false negatives in that class (Figure 4d). Though there was a lot of fluctuation in training and validation losses, a further decline concerning time is registered for the train/box_loss and train/cls_loss, which means the model learns. This showed increased mAP50 during the validation but gave pretty low values for higher IoU's. The accuracy, precision, and F1 score of the YOLOv11 for the manhole class were very high.



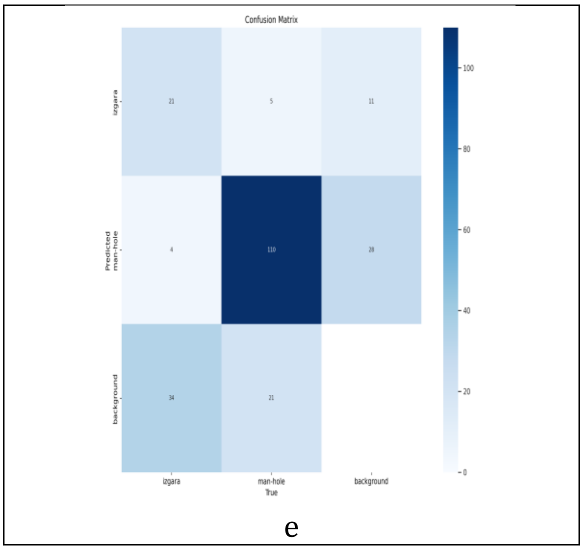


Figure 4. Yolo v11 model results

We summarize the results by comparing the performance of YOLOv5, YOLOv8 and YOLOv11 algorithms and evaluating them in terms of Accuracy, F1-Score, Average Precision (mAP@0.5, mAP@0.5-0.95), Precision and Recall metrics (Table 1).

Table 1 Results of comparing Yolo versions

	F1 Score		Accuracy	Precision		Recall	
	Drainage	Man-hole		Drainage	Man-hole	Drainage	Man-hole
Yolo v5	0.42	0.74	50.4	42.2	74.4	80	88
Yolo v8	0.46	0.78	55.7	38.98	76.47	54.76	80.62
Yolo v11	0.44	0.79	55.98	35.59	80.88	56.76	77.46

YOLOv5: It gave middle performance from the point of view of both accuracy and F1 score. "Manhole" class gives a really high F1-score, while not that well in class "drainage". The approximate F1-score is about 0.56. YOLOv8: Compared with YOLOv5, this model yields a bit higher F1-score. In particular, for class "man-hole", an F1-score of approximately 0.60 is achieved. Compared to YOLOv5, the "drainage" class only showed lower F1 performance when the overall accuracy was improved. YOLOv11: In general, and considering all the results, the best performance was given by YOLOv11 when considering the F1-Score. That value reached 0.62 for the class "man-hole" when the class "drainage" took lower F1 values.

YOLOv5: The values are 0.583 mAP@0.5 and 0.144 mAP@0.5-0.95, which indicates performance degradation when higher IoU thresholds are involved. YOLOv8: mAP@0.5 value: 0.609; mAP@0.5-0.95 value: 0.154: Therefore, in this regard, it outperformed but still has room for improvement toward higher IoU values. YOLOv11: mAP@0.5 value: 0.606; mAP@0.5-0.95 value: 0.161: This is the best compared to all of them, with a small increment, mainly towards IoU threshold challenges.

YOLOv5: The "Man-hole" class has better precision, but because of the "drainage" class, the average is low. Also, Recall is higher in "man-hole" but low in the "drainage" class. YOLOv8: Higher Precision and Recall values compared to YOLOv5. In the case of the "man-hole" class, the Precision and Recall rate is good enough to be considered even in the distribution. Whereas, in the "drainage" class, the values of both precision and recall are low enough. YOLOv11: The precision is really good, especially for the class "manhole". The precision of the model outperforms YOLOv5 and YOLOv8. Recall in YOLOv11 for the "man-hole" class is good, but for the class "drainage", the recall is low.

YOLOv11 gives the best performance by F1-Score. In the case of mAP@0.5 and mAP@0.5-0.95, YOLOv11 had a slight edge, though YOLOv8 performed similarly. And lastly, in the "man-hole" class, Precision and Recall for all models had performed good enough, but regarding the "drainage" class, they need further improvement. While the other models drop either in Precision or Recall, YOLOv11 would be doing the best with a better balance between them. Moreover, it is clear that more data or development might be necessary for all models to perform well in the class "drainage." Without question, referring to the results portrayed, the best performance in general is by the model YOLOv11.

3.3. Investigating Model Architecture Variations

The result of the evaluation of the model within Yolo v11n stands as follows:

Considering the F1-Confidence Curve, the maximum value of F1 for all classes is approximately 0.60, while for this particular class, that is, manhole, it is about 0.8. Yes, they actually reflect the general performance of the model, as one could realize that it shows very good results, especially on the class named "manhole". Now, from the following Precision-Recall and Precision-Confidence graphs, we can draw conclusions on the accuracy of the model. Where the Precision-confidence curve is 0.761 for all classes and about 0.80 for the man-hole class. Some Average Precision metrics-like mAP50-are 0.606 for general performance, 0.799 for man-hole class detection, and 0.413 for the drainage class. In addition, this diminishes to the lower value of 0.225 by mAP50-95. From these Precision-Recall curves, one can observe that the precision values are pretty high for the manhole class, ~0.799, and considerably decrease in the case of a drainage class, ~ 0.413. Recall metrics are pretty good in the manhole class of ~ 0.87 but low in the drainage class of ~ 0.413. In a nutshell, YOLOv11 performs well because for the Man-hole class, YOLOv11 has a higher accuracy value, higher precision, and higher recall; whereas for the Drainage class, the values are low. Since both mAP@0.5 and mAP@0.5-95 report the general performance of the model as balanced, from the overall accuracy F1-score perspective, the model does quite well in general.

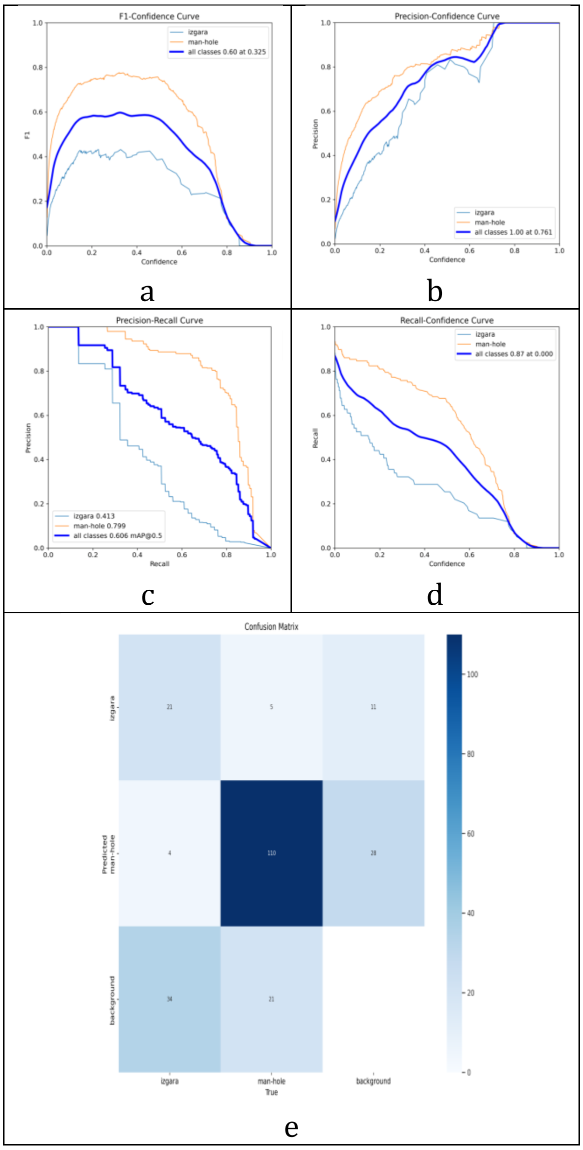
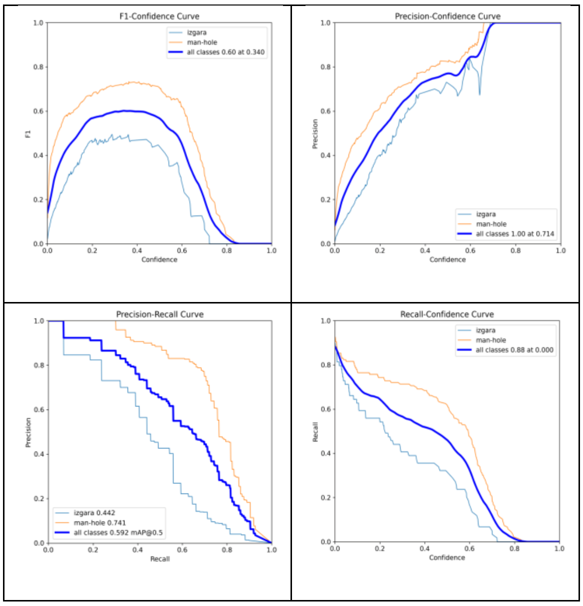


Figure 5. Results of the Yolo v11n model

Performance by the model used, i.e., Yolo v11s, can be evaluated as:

Various metric results have been analyzed about the model created on the Yolo v11 small version (Figure 5). It has an F1 score in the F1-Confidence Curve of about 0.60 for all classes. Among them, man-hole covers are classified with much more confidence compared to the class drainage. This probably should be the reason that accuracy may also be higher for certain classes (Figure 5a). Concretely, it can be seen from the F1-Confidence Curve that the F1 scores are about 0.74 for the class "manhole," about 0.44 for the class "drainage," while the mean F1 score of all classes is 0.60 (Figure 5b). According to the Precision-Recall Curve of all classes, the mAP@0.5 value of the model is about 0.592. Although the exact value is not directly given by the graphs, mAP@0.5-0.95 can be estimated to be lower than that of mAP@0.5, and it is generally around 0.20 to 0.30 (Figure 5c). Looking at the Precision-Confidence Curve, the precision of the manhole class, man-hole, is quite high, about 0.74. Precision in the case of the drainage class, drainage, is comparatively low at about 0.44. Recall-Confidence Curve indicates that the value of recall in the manhole class is high at around 0.88, with that of the drainage class being low at around 0.60 (Figure 5d). Looking from these metrics, one may notice fairly well that the Yolo v11 works nice for the class of manholes and the same network does worse for the class of drainages. Anyway, considering everything, this performance is imbalanced between classes, though, in general, the performance of F1 score and precision are middle in average.



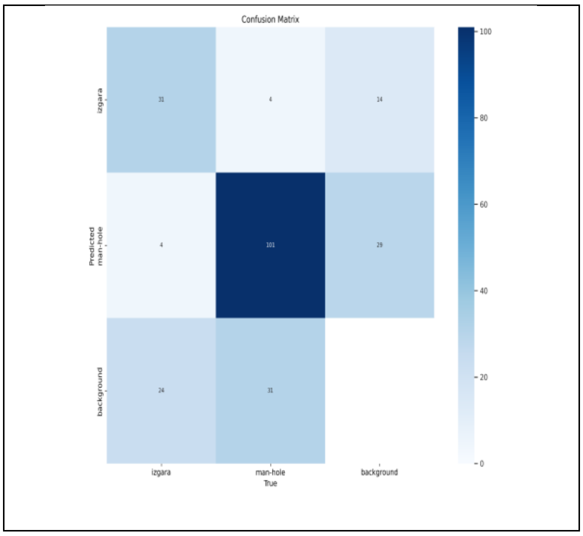


Figure 6. Results of the Yolo v11s model

This plot could give an interpretation for the performance of YOLOv11 medium model: Using F1-Confidence Curve for a confidence level of 0.298, the F1 score in all classes is going to be about 0.54 (Figure 6). This would mean that only at this value of confidence will a better trade-off between precision and recall be achieved by the model, but not high (Figure 6a). The best precision in the last epoch is around 0.323 while the recall is about 0.293. From the Precision-Confidence curve, it can be seen that the higher the confidence, the higher the precision. The precision of the model is 1.00 with a confidence level of 0.680 but after that rapidly falls afterwards. Recall-Confidence Curve reflects that if confidence is small, the recall is higher and decreases with the increase of confidence (Figure 6b). While the highest recall value of 0.86 is realized at 0 confidence level, among the Mean Accuracy scores, the score at mAP@0.5 is around 0.270; in other words, at an IoU threshold of 0.5, performance is comparatively reasonable for the model in terms of object recognition. The score of mAP@0.5-0.95 is comparatively low, at about 0.099, with a tighter IoU threshold or when the condition gets tougher (Figure 6c-d). Training Losses: The box loss curve goes down smoothly, which represents that through the training process, the model is predicting the location of an object better and better. Overall, the model's performance is quite reasonable considering precision and recall. The map@0.5 is not bad, but in the case of more difficult IoU thresholds, such as map@0.5-0.95, it degrades performance. Higher F1 score can be obtained and hence scope for further improvement.

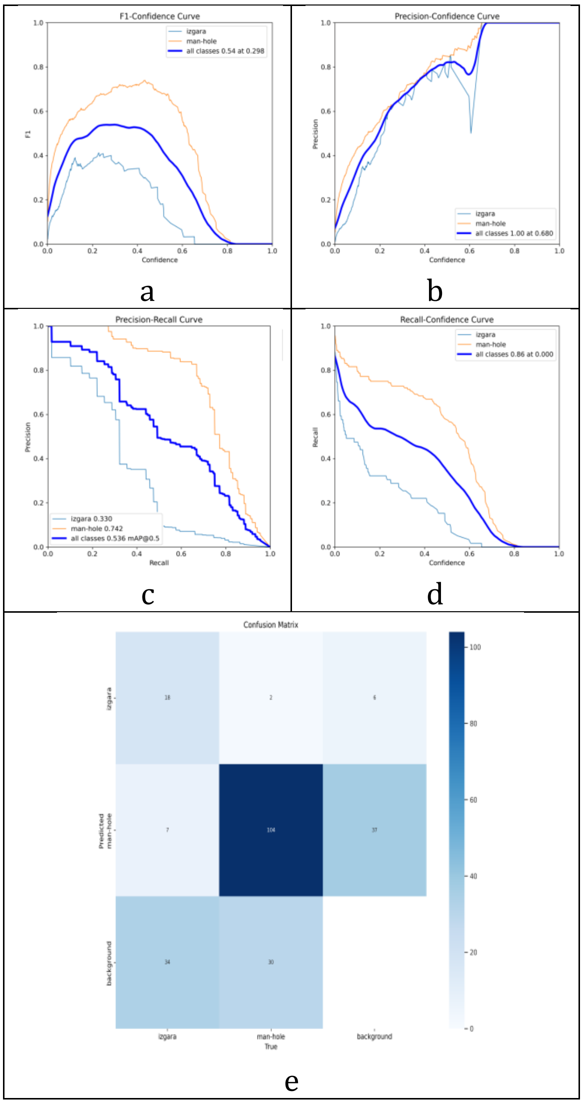


Figure 7. Results of the Yolo v11m model

Our work tests YOLOv11 in three variants, namely nano (n), small (s), and medium (m), based on Accuracy, F1-Score, Average Precision, which is mAP@0.5 and mAP@0.5-0.95, Precision, Recall, and Speed in FPS.

Table 2 Results of comparing Yolo variant versions

	F1 Score		Accuracy	Precision		Recall	
	Drainage	Man-hole		Drainage	Man-hole	Drainage	Man-hole
Yolov11n	0.44	0.79	55.98	35.59	80.88	56.76	77.46
Yolov11s	0.57	0.75	55.46	52.54	74.26	63.27	75.37
Yolov11m	0.42	0.73	51.26	30.51	76.47	69.23	70.27

If the F1-Scores are considered, then the F1-confidence for YOLOv11n would lie at approximately a maximum of 0.60. This model gives an edge in terms of speed because of lesser model complexity but considerably low F1-score as compared to other variants. The variant YOLOv11s has an approximate F1 score of 0.60, showing more or less even balancing. This small model provides a very good balance between speed and performance. Having F1, the maximum score achieved by YOLOv11m is 0.54, and thus, the model with more complex architecture should not show such a high F1 score in comparison with nano and small versions.

Considering the mean average precision, mAP@0.5, for instance, YOLOv11n presents good object recognition, with high values of approximately 0.60 mAP@0.5. Therefore, releasing YOLOv11s, keeping in mind a value of 0.592 mAP@0.5, very close to the nano model, assures a good result as far as object recognition performance is concerned. mAP@0.5: On the YOLOv11m version, this value is at about 0.536, somewhat lower compared to the other two. mAP@0.5-0.95: On these difficult IoU thresholds, the YOLOv11n does fairly well by a bigger margin in IoU, having its mAP@0.5-0.95 value at about 0.225. The YOLOv11s does quite well at an mAP@0.5-0.95 of about 0.220. For example, the IoU threshold, which is relatively complex, has a performance of 0.099 in the YOLOv11m version. Contrarily, the more complex a model is, the less mAP@0.5-0.95.

The precision values show that the YOLOv11n: Nano model reaches highs with its precision value of 0.761. In a very similar way, very good results are represented by the YOLOv11s model version when considering precision, due to the value of 1.00 of the same. The YOLOv11m model version has around 0.680 value accuracy, quite low compared to different models. From the Recall values, the maximum recall in YOLOv11n is 0.87, showing high sensitivity with even low confidence. Similarly, the YOLOv11s performed nearly like the nano version and yielded a recall value of 0.88. In the YOLOv11m version, too, there was slight degradation in sensitivity, giving a recall of 0.86. The training losses for different models all take a similar trend of going downwards, but a keen look reveals a larger loss both in box loss and cls loss for the medium model; that would mean this model is large and complex to learn.

Among all the comparisons done, model YOLOv11n has the best balance between speed and overall accuracy. The preferred model in those cases when the speed matters but not much regarding high accuracy is YOLOv11n. YOLOv11s has presented a good compromise between efficiency and accuracy, since it has performed well on accuracy and recall. However, the YOLOv11m-medium model may show better mAP@0.5 values due to its more complex structure running significantly slower compared to the small ones and with a little bit lower F1 score. Therefore, the above comparisons will definitely help the users make the right choice according to their needs, depending on which model is going to be used.

This would be a good opportunity to analyze model performance based on different metrics for the two classes: Drainage and Man-Hole. With this, one may quickly appreciate the huge difference in the number of predictions, with the model proposing 240 for the drainage class, while the model predicted 446 for the man-hole class. It could be the case that instances of the manhole class are more frequently found and cropped into view by the model, compared to the drainage class, and hence may mean that the manhole class contains more instances in this dataset, hence meaning the model is working on an imbalanced dataset.

By considering the average confidence scores, it can be seen that the model's confidence score is 52% for the drainage class, while for the man-hole class, it is 58%. These show that the average confidence scores in both classes are below 60%, which means that the model is not sure and cannot give full confidence in the prediction. Very low confidence in the drainage class of the model surely suggests that the model needs further development in this class. Having low confidence scores could indicate that the model is indifferent with some of the predictions it gives or does not really see much difference between the objects that it detects. This would be improved either by more training of the model or by balancing classes.

Regarding the speed in FPS, it reached 4.26 FPS in both classes, meaning that this model can process four images per second and is fit for real-time object detection. In contrast, the low confidence score and the differences between classes are evidence of further opportunities for improving this model by better accuracy and confidence, while the performance concerning speed is quite sufficient.

The higher amount of predictions in the man-hole class could have, therefore, been a resultant effect of the imbalance in the dataset. The low confidence scores in the drainage class support the fact that this class needs to be supported with more data, or that more training on this class needs to be carried out. Though quite satisfying concerning the speed factor, it goes without saying that more improvements should be done to concern the reliability of the forecasts.

3.4 Determining the Number of Epochs and Early Stop Strategy

The number of epochs defines how many times the model viewed the training data and influences how much it can learn from it. Choosing an appropriate number of epochs lets it learn enough, without making it prone to overfitting. Early stop is a possibility to stop the training optimally, based on the performance on the validation set 5. Since a small number of images - usually 686 - would create a high risk of overfitting the model, I try several epoch numbers in order to find the best performance point of the model using the early stopping strategy. In this experiment, the early stop was set at 5; thus, it got that the model trains only up to 25 epochs. Later, when this value was reached, the experiments were completed with more than 25 epochs each.

4 | RESULTS

This paper has presented how the YOLOv11 algorithm works in respect of detecting manhole covers and drainage in Google Street View images. Experimental results prove that the achieved accuracy of YOLOv11 is very high and object detection really efficient. Automatically detected manhole covers and drainage are an important contribution to public safety and an improvement in processes connected with infrastructure maintenance.

The performance of the YOLOv11 in this study was rather good, compared to other YOLO series versions, regarding both speed and balanced accuracy. Furthermore, by the precision and recall results, the model results showed reasonable performance. More precisely, the results were more accurate for the man-hole class, whereas the class of drainage had room for further improvement.

Results obtained in this paper represent the promise of using the YOLOv11, an AI-based object detection algorithm, in the management and monitoring of urban infrastructure. Future work might be done on testing the performance of the model using bigger-sized datasets and improving the performance under various weather conditions and lighting changes. On the other hand, the fact that infrastructure detections can be integrated with Management Information Systems for speedy and effective decisiveness mechanisms of city administrations also goes to show that these technologies make a great contribution to urban planning and management.

The integration of these technologies with spatial planning processes holds even greater potential. Spatial planning, which focuses on the strategic and sustainable use of land, can greatly benefit from AI-driven detection results. By integrating infrastructure data with Geographic Information Systems (GIS), city planners can create real-time spatial databases that support better decision-making. This integration enables the visualization of infrastructure conditions, highlighting areas that require urgent maintenance or are at risk of failure. It also provides the ability to model future urban scenarios, taking into account the interplay between infrastructure, land use, and population dynamics. Furthermore, by aligning maintenance priorities with broader urban development plans, spatial planning can ensure that resources are distributed more equitably and efficiently across a city.

Infrastructure management is a major function in all modern cities of the world in regard to public safety and the effective use of all resources. Management Information Systems empower decision support mechanisms through the collecting, processing, and analysis of big data. Automatic detection, therefore, with situational analysis of manhole covers and gratings, can constitute very important innovation in infrastructure management. The research study will be done to extend the MIS for better decision-making in city management by incorporating an image-processing technique that will detect infrastructure objects in images using the YOLOv11 algorithm.

These determinants, integrated with MIS, would update the inventories of infrastructure assets at any instant of time and provide valid data to city administrations. The location and condition information of infrastructure elements could be tracked in real time, and missing or incorrect data would appear instantly and be set right. This saves a great deal of time and resources deployed in infrastructure management and leads to better management processes. It further provides the scope for getting more accurate automatic data, rather than some manual inspection on the field.

Integrated with MIS, such detections allow the infrastructure maintenance and inspection processes of city administrations, while the priorities of manhole cover and grating maintenance are brought out more quickly and the teams in the area do less work. Automatic detection and monitoring reduce human error, saving lots of time and decreasing an enormous amount of human error. Instantaneous detection of areas needing maintenance allows anticipating the problem in infrastructure and hastening the solution.

Management Information Systems, in turn, contribute to effective strategic decision making through the incorporation of findings from the detection into DSS. KDS feeds information on proactive decisions on areas that should receive priority in infrastructure management, identification, and condition of infrastructure elements. The identification of sections that need high frequent maintenance, for instance, aids in effective resource use and planning of the future infrastructures. This therefore represents a more valid and effective decision-making process.

The data flow by MIS on resource management and planning processes enables the city government to optimize resources put into the upkeep and repair of infrastructure. Manhole cover and drainage detection by city governments offers an opportunity to improve the planning of the maintenance team and the supply of materials. Integration of results of detection into budgeting processes coupled with it enables more appropriate investment in infrastructure and thereby increases efficiency in the management of resources in the long term.

Lastly, with respect to the dangers about safety, there should be an updated monitoring over infrastructure elements. Information obtained with the employment of MIS allows the infrastructural problems, especially the ones relating to the public health and safety to be defined and interfered with at the earliest stage. Data obtained about the status of infrastructure elements may be used also in the process of emergency response such as disaster management whereby one can easily follow up the risky zones inside the city.

These detections using the YOLOv11 algorithm have great potential to contribute a lot once integrated with the MIS pertaining to urban infrastructure management. In that respect, such an integrated system offers good data-driven decision-making, judicious resource utilization, and proactive resolution of issues relating to infrastructure. Such solutions are what Management Information Systems can help impart to the city governments in order to carry out infrastructure management in a speedier, more efficient, and safe way, besides being one of the important constituents of the smart city solutions.

By aligning urban infrastructure monitoring with spatial planning, cities can adopt a holistic approach to resource management. This integrated framework not only enhances the operational efficiency of existing systems but also provides the foundation for more resilient and adaptive urban development strategies. The synergy between AI, GIS, and spatial planning thus represents a transformative step towards building smarter, safer, and more sustainable cities.

Automation means much to infrastructure management; thus, massive benefits are created in all spheres of managing a city, right to the correct utilization of resources. First, the automation systems enable the quick identification of infrastructure elements like manhole covers and drainage, something that could have been done more speedily and efficiently. While there is perhaps inevitable human fallibility with the traditional methods of manual inspections-so time-consuming-the automated systems ensure continuous, exact monitoring of infrastructure elements. This, therefore, ensures efficiency, hence saving on time in infrastructure management.

Automation will also decrease human error. Consistent and correct identifications of structure features reduce loads of human error and provide reliable data. This automatic system replaces inefficient on-site manual inspection with more accurate and steadier results. Besides, it allows for better resource management: it can point out which structural elements require maintenance or repair, so that labor, time, and material can be better spent.

Other beneficial contributions of automation relate to economies of scale. An automated system requires many reductions in manpower and time, which drastically bring down the cost. Besides, by regular monitoring of infrastructural elements, it is possible to detect and respond against major failures or damages before they occur; hence, saving costs in the long run by stopping the situation of infrastructure problems from aggravating.

Automation also contributes much value to security and risk management, as, due to continuous monitoring of infrastructural elements, a number of security risks can be discovered much earlier than otherwise would be possible. Special benefits derive from status monitoring of the infrastructure elements most critical from the point of view of public health and safety and from making rapid interventions whenever needs arise. The same systems will provide much very valuable data in disaster and emergency situations.

Other advantages include increased efficiency of data collection and data analysis processes once they are automated. Management can also include automated systems in DSS simply by constant data collection of the elements of infrastructure. Such data can be utilized in infrastructure planning and budgeting processes, thereby providing strategic leverage to the city administrations. Also, it prolongs the useful life of the elements of infrastructure, avoiding sudden failures and improving the overall performance of infrastructure.

The bottom line of all is that automation in infrastructure management contributes toward much safer, more efficient, and even more sustainable infrastructure of cities. These allow resources to be put to effective use, reduce costs, and offer proactive risk management. In the long run, these contributions by automation have the capacity to speed up the pace toward changing cities into smart city solutions by opening new eras in infrastructure management.

In future studies, addressing the issue of dataset imbalance could further enhance the robustness and generalizability of the model. One promising approach involves leveraging advanced deep learning architectures, such as the Swin Transformer. With its hierarchical design and self-attention mechanism, the Swin Transformer is particularly adept at capturing both local and global features, which may help mitigate the adverse effects of underrepresented classes. Furthermore, its scalability and adaptability to diverse data distributions make it a suitable candidate for addressing the challenges posed by imbalanced datasets. Integrating Swin Transformer with techniques such as class weighting, focal loss, or synthetic data augmentation could yield significant improvements in performance, particularly for minority classes. Future research could explore these avenues to develop more equitable and accurate models.

REFERENCES

Wang, J., Fang, Z., Li, Q., Tang, Z., Huang, Z., Hong, Z., & He, H. (2024). YOLO-SDD: An Improved YOLOv5 for Storm Drain Detection in Street-Level View. Journal of Shanghai Jiaotong University (Science), 1-16.

Oulahyane, A., & Kodad, M. (2024). Advancing Urban Infrastructure Safety: Modern Research in Deep Learning for Manhole Situation Supervision Through Drone Imaging and Geographic Information System Integration. International Journal of Advanced Computer Science & Applications, 15(7).

Omar, M., & Kumar, P. (2024). PD-ITS: Pothole Detection Using YOLO Variants for Intelligent Transport System. SN Computer Science, 5(5), 552.

Ping, P., Yang, X., & Gao, Z. (2020, August). A deep learning approach for street pothole detection. In 2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService) (pp. 198-204). IEEE.

Fahmani, M., Golroo, A., & Sedighian-Fard, M. (2024). Deep learning-based predictive models for pavement patching and manholes evaluation. International Journal of Pavement Engineering, 25(1), 2349901.

Wang, D., & Huang, Y. (2024). Manhole Cover Classification Based on Super-Resolution Reconstruction of Unmanned Aerial Vehicle Aerial Imagery. Applied Sciences, 14(7), 2769.

Yin, X., Chen, Y., Bouferguene, A., Zaman, H., Al-Hussein, M., & Kurach, L. (2020). A deep learning-based framework for an automated defect detection system for sewer pipes. Automation in construction, 109, 102967.

Liao, L., Li, H., Shang, W., & Ma, L. (2022). An empirical study of the impact of hyperparameter tuning and model optimization on the performance properties of deep neural networks. ACM Transactions on Software Engineering and Methodology (TOSEM), 31(3), 1-40.

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing, 415, 295-316.

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., ... & Lindauer, M. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 13(2), e1484.

Du, X., Xu, H., & Zhu, F. (2021). Understanding the effect of hyperparameter optimization on machine learning models for structure design problems. Computer-Aided Design, 135, 103013.

Van Rijn, J. N., & Hutter, F. (2018, July). Hyperparameter importance across datasets. In Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 2367-2376).

Cepa, J. J., Alberti, M. G., Pavón, R. M., & Calvo, J. A. (2024). Integrating BIM and GIS for an Existing Infrastructure. Applied Sciences, 14(23), 10962.