Advancing Urban Infrastructure Safety: Modern Research in Deep Learning for Manhole Situation Supervision Through Drone Imaging and Geographic Information System Integration

Ayoub Oulahyane, Mohcine Kodad
MATSI Research Lab, ESTO, Mohammed First University, Oujda, Morocco

***Abstract*—This paper research introduces a cutting-edge approach to enhancing urban infrastructure safety through the integration of modern technologies. Leveraging state of the art deep learning techniques, specifically the recent object detection models, with a focus on YOLOv8, we propose a system for supervising and detecting manhole situations using drone imagery and GPS location data. Our experiments with object detection models demonstrate exceptional results, showcasing high accuracy and efficiency in the detection of manhole covers and potential hazards in real-time drone imagery. The best trained model is YOLOv8, which achieves a mAP@50 rate of 89% and a Precision rate of 95%, surpassing existing methods. By combining this visual information with precise GPS location data, our system offers a comprehensive solution for monitoring urban landscapes. The integration of YOLOv8 not only improves the efficiency of manhole detection but also contributes to proactive maintenance and risk mitigation in urban environments. This research represents also a significant step forward in leveraging modern research methodologies, and the outstanding results of our trained models underscore the effectiveness of Object detection models in addressing critical infrastructure challenges.**

Keywords*—*Urban infrastructure safety; object detection; Deep Learning (DL); UAV (Drones); Computer Vision (CV)

I. INTRODUCTION

In the contemporary landscape of urban development, ensuring the safety and integrity of critical infrastructure is a paramount concern. Among the myriad challenges faced by urban planners and maintenance authorities [1], [2], the efficient and accurate detection of manhole covers, and potential hazards stands out as a pivotal aspect of proactive risk management [3]. This research seeks to address this challenge head-on by embracing the convergence of advanced technologies, with a specific focus on the You Only Look Once (YOLO) object detection model, particularly the latest iteration, YOLOv8 [4]. The proliferation of unmanned aerial vehicles (UAVs) or drones [5], [6], coupled with the advancements in deep learning [7], has opened new avenues for real-time surveillance and analysis of urban landscapes. The ability to deploy drones for high-resolution imaging provides a dynamic and flexible solution for monitoring infrastructure elements that are typically challenging to inspect manually or through conventional means.

Concurrently, the integration of Global Positioning System (GPS) technology adds a layer of precision by providing accurate geospatial information [8]. At the heart of this research are the Object detection models, renowned for its state-of-the-art object detection capabilities. YOLOv8 excels in processing images swiftly while maintaining a high level of accuracy, making it an ideal candidate for real-time applications. By training the models to recognize manhole covers and potential hazards in diverse urban settings, we aim to harness the full potential of deep learning models to bolster the efficiency and efficacy of infrastructure monitoring [9]. The integration of these technologies holds the promise of transforming traditional approaches to urban infrastructure supervision. Rather than relying on periodic inspections or reactive measures, our proposed system aims to establish a proactive and intelligent framework. By fusing the power of object detection methods with drone imagery and GPS location data, we aspire to create a comprehensive solution that not only detects and supervises manhole situations (see Fig. 1), but also contributes to a deeper understanding of the evolving dynamics within urban environments.

Fig. 1. Samples of varied manhole situations.

As we embark on this exploration of technology-driven urban research, the subsequent sections will delve into the methodology employed, the experimental results obtained, and the broader implications of our findings. Through this interdisciplinary approach, we endeavor to contribute to the burgeoning field of intelligent infrastructure management, paving the way for safer, more resilient urban environments in the face of evolving challenges.

This paper will be structured as follows: Background in Section II will be followed by related work in Section III. Next, we will present our methodology in the Section IV, followed by the presentation of our results in Section V. Finally, we will conclude this paper in Section VI.

II. Background

In this section we will try to give a general vision of the deep learning models that we will use in our approach and related work, which will be divided into two parts, the first related to data measurement classification models, as for the second presents some models of image detection models.

A. Computer Vision

Computer vision is a field of artificial intelligence (AI) that enables machines to interpret and make decisions based on visual data. It seeks to teach computers how to gain a high-level understanding of digital images or videos, similar to the way humans interpret and understand visual information. This involves tasks such as image recognition, object detection, image segmentation, and more. Computer vision has diverse applications, ranging from facial recognition and autonomous vehicles to medical image analysis and industrial automation [10].

1) Enhancing urban infrastructure safety: The integration of computer vision with an edge approach can significantly contribute to enhancing urban infrastructure safety [12].

2) Computer vision edge approach: The term "edge" in computer vision often refers to processing data closer to the source of generation rather than relying on a centralized server or cloud. The edge approach involves deploying computer vision algorithms on devices like edge computing devices, cameras, or sensors, enabling real-time analysis and decision-making without the need for constant connectivity to a central server. This approach is particularly beneficial in scenarios where low latency and immediate responses are crucial [11].

3) Surveillance and monitoring: Computer vision can be used to analyze live camera feeds for suspicious activities, unauthorized access, or potential safety hazards. By deploying this capability at the edge, responses can be immediate, addressing security concerns in real-time [13].

By combining the capabilities of computer vision with an edge computing approach, urban areas can benefit from faster and more efficient responses to safety and infrastructure challenges, ultimately creating smarter and safer cities.

B. YOLO (You Only Look Once)

YOLO (You Only Look Once) is a popular object detection algorithm that is widely used in computer vision applications. The key idea behind YOLO is to divide the input image into a grid and, for each grid cell, predict bounding boxes and class probabilities.

This allows YOLO to simultaneously detect multiple objects in an image in real-time [14], [15]. Here are some key features and concepts associated with YOLO:

1) Real-time detection: YOLO is known for its efficiency, and it can perform object detection in real-time, making it suitable for applications like video analysis.

2) Single forward pass: YOLO performs object detection in a single forward pass through the neural network, as opposed to two-stage detectors, which involve region proposal networks and classification networks separately.

3) Bounding box prediction: For each grid cell, YOLO predicts bounding boxes along with confidence scores and class probabilities. This allows it to detect multiple objects of different classes in a single pass.

4) Anchor boxes: YOLO uses anchor boxes to improve the accuracy of bounding box predictions. These anchor boxes are pre-defined bounding box shapes, and the model learns to adjust these anchors during training.

5) Darknet: YOLO is typically implemented using the Darknet framework, which is an open-source neural network framework written in C and CUDA. Darknet supports YOLO and allows for training and using YOLO models.

C. ArcGIS (Geographic Information System)

ArcGIS, developed by Esri (Environmental Systems Research Institute), is a geographic information system (GIS) software suite. It is widely used for creating, managing, analyzing, and displaying spatial data. GIS is a technology that combines geography (maps) and data to provide valuable insights, enabling users to make informed decisions based on geographic information. ArcGIS provides a comprehensive platform for working with spatial data at various scales, from local to global. The suite includes a range of desktop, server, and web-based applications. Overall, ArcGIS is a powerful tool for spatial analysis and mapping across various industries, including environmental management, urban planning, public health, transportation, and more. It is widely used by professionals and organizations to understand, interpret, and visualize geographic patterns and relationships in their data [16], [17].

III. Related Work

This study integrates multiple technologies, including deep learning, object detection, UAV (Unmanned Aerial Vehicle) technology, and ArcGIS positioning, with a specific focus on manhole detection. There is a limited number of published papers that cite these technologies in conjunction. Existing studies in this field attempt to develop new datasets to enhance detection accuracy. Additionally, they strive to balance accuracy with computational efficiency. This balance is crucial for drone implementation, especially when utilizing parallel processing co-processors. The objective is to optimize the system for real-time applications while maintaining high detection accuracy, which is essential for efficient urban planning and infrastructure management.

In their research, Pang et al. developed a method for detecting road manhole covers using a stereo depth camera and the MGB-YOLO model, achieving a notable accuracy of 96.6%. This approach, which outperforms several existing models, is particularly efficient for deployment in in-vehicle devices, contributing significantly to urban infrastructure management and vehicular safety [18].

In their work, Andersen et al. address the challenge of drone navigation in dark, GPS-denied, and confined spaces, focusing on the high processing power required for maintaining detailed environmental maps. They note the particular difficulty in navigating narrow spaces where low-resolution voxel representations can impede trajectory planning. Inspired by the Inspectrone Project, which involves inspecting large marine vessels, the authors propose a deep learning model for detecting manholes using only depth images. This study aims to balance accuracy with computational efficiency, making it suitable for drone implementation on parallel processing co-processors. A key feature of their approach is the use of a temporal filter to enhance robustness and reduce false positives, requiring multiple detections within a timeframe to confirm the manhole's location. The effectiveness of their method, which is agnostic to scene texture, is demonstrated through successful drone flights through a standard-sized manhole on a marine vessel, showcasing a viable solution for manhole detection in challenging environments [19].

In their research, Timofte, Radu et al. focus on the challenge of accurately 3D localizing road fixtures, particularly manhole covers, across extensive road networks. They propose an innovative pipeline utilizing images captured by vans to detect, recognize, and localize manholes, a task complicated by issues like occlusions, varying illumination conditions, and significant viewpoint differences. Additionally, the diversity in manhole cover designs adds to the complexity. Their approach effectively combines 2D and 3D computer vision techniques to handle large volumes of image data, achieving notable performance. This study is distinguished as the first to report on manhole mapping using solely computer vision techniques and GPS, marking a significant advancement in the field of automated road surveying [20].

Despite the advancements, these papers also address challenges such as the difficulty in detecting manholes under certain conditions (e.g., poor lighting, obscured by objects, or in densely built-up areas) [21].

IV. Methodology

A. Proposed Method

The research paper introduces a comprehensive methodology employing drones and deep learning (DL) models for the efficient monitoring, detection, and precise localization of manholes. This approach is methodically structured into interconnected phases as shown in Fig. 2:

1) *Drones for surveillance*: Utilizing drones equipped with cameras and potentially other sensors, the methodology involves aerial surveillance to collect visual data of manholes. This step is pivotal for acquiring the necessary imagery for further analysis [22].

2) *Detection of manhole conditions*: The visual data gathered by drones are transmitted to a cloud-based framework. Here, a specialized deep learning model, already trained, scrutinizes the images. The primary task of this DL model is to identify both the presence and condition of manholes from the collected visuals [23].

3) *Precise localization of manholes*: Concurrent with condition detection, the system also focuses on accurately localizing the manholes. It leverages geographical data obtained from the drones to pinpoint the exact physical locations of the manholes, a crucial element for subsequent maintenance or monitoring operations [24].

4) *Informed decision making*: Following the successful detection and localization of manholes and the evaluation of their state, the system then classifies these findings. This classification is essential for determining the appropriate actions needed, such as cleaning, repairs, cover replacements, sediment removal, coatings, or comprehensive structural assessments [25].

This enhanced methodology signifies a significant advancement in the field, leveraging cutting-edge technology for urban infrastructure management.

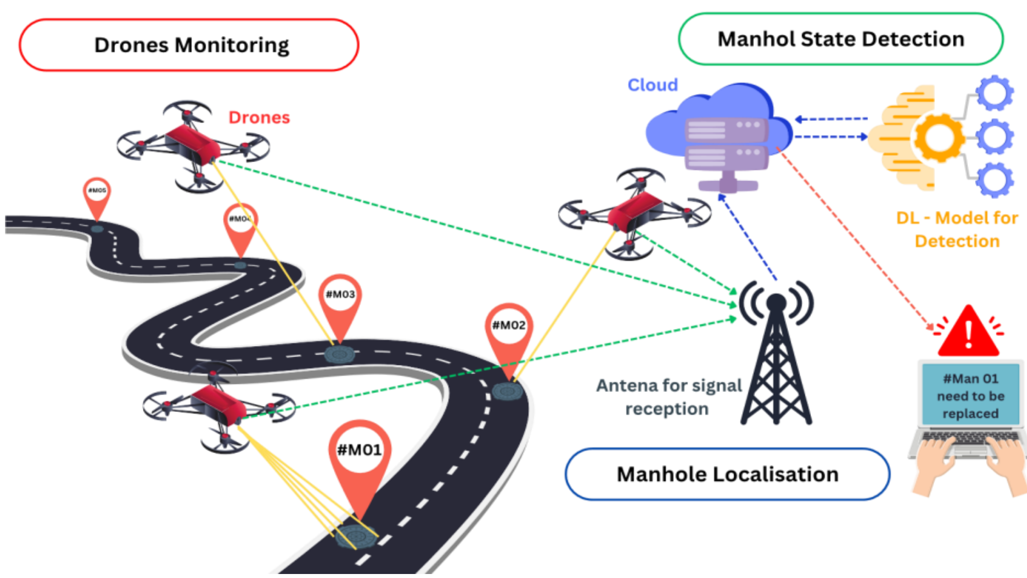

Fig. 2. Proposed monitoring, detection, and localization system for manholes.

B. Architecture for Training Process

The proposed methodology (see Fig. 3), delineates a sophisticated and integrated system for manhole identification and maintenance, utilizing a combination of modern technologies including unmanned aerial vehicles (UAVs), cloud computing, deep learning, and potentially geographic information systems (GIS). This system is structured into four critical steps:

1) Data collection and input: The initial phase involves compiling a comprehensive dataset of manhole images. These images are meticulously annotated, likely with bounding boxes or similar markers, to highlight the presence of manholes. This dataset forms the foundation of the entire process.

2) Model development and training: A deep learning model, though its specific neural network architecture is not detailed, is meticulously trained using the aforementioned dataset. This model is intricately designed to effectively recognize and pinpoint manholes from the visual data collected by the drones.

3) Object detection and output: Once the model is trained, it enters the object detection phase. In this stage, the model applies its learned patterns to new images captured by drones, successfully identifying manholes in these fresh visuals.

4) Analysis and decision making: The final stage involves a critical analysis of the model's output. This analysis is pivotal in making informed decisions regarding the maintenance and other necessary actions for the manholes detected.

This method represents a significant leap in infrastructure monitoring and maintenance, harnessing the power of advanced technologies to create an automated and intelligent system. This system not only increases efficiency but also potentially enhances the safety and reliability of urban infrastructure management.

C. Testing Sample

The chosen area for testing our manhole cover detection and monitoring method via drones is illustrated in Fig. 4, which displays a region with diverse urban characteristics. The area is demarcated by a series of waypoints forming a boundary within which the drone operations are to be conducted. On the left side, we have a simplified schematic from ArcGIS Maps, which provides a clear and uncluttered view of streets and key establishments like the "Coffee El Jabah," a "Pharmacy," and educational institutions like "El Ouafa School" and "Pythagoras Private School." This representation is beneficial for initial planning and coordination purposes.

The right side of the figure contrasts this with two views from Google - one from satellite imagery offering a detailed, real-world perspective of the area's layout and another from Google Maps, which includes street names and a blue overlay indicating the operational path of the drone. The satellite imagery provides a comprehensive view of the density and structure of buildings, roads, and vegetation, which is crucial for understanding potential obstacles and optimizing flight paths.

We suggest to studying this area that has eleven manhole covers within a one-kilometer range, signifying the target objects for the drone's detection system. The area is chosen for its typical urban features and the presence of manhole covers that need monitoring, and close to our laboratory making it an ideal test bed for validating the effectiveness of the drone-based surveillance system in actual road settings. The dual representation of the area through different mapping services aids in cross-verifying details and planning the drone's flight more accurately.

D. Decisions Related to Manhole Situations

Certainly, there are more specific examples of decisions related to manhole situations, including actions like cleaning and replacement:

- Cleaning Procedures: Implement regular cleaning schedules to remove debris, sediment, and blockages from manholes, ensuring optimal functionality.
- Repairs and Patching: Promptly address minor damages through patching or localized repairs to prevent further deterioration.
- Manhole Cover Replacement: Evaluate and replace worn-out or damaged manhole covers to ensure the safety of pedestrians and motorists.
- Sediment Removal: Implement strategies for the systematic removal of sediment buildup within manholes to maintain proper drainage and prevent blockages.
- Coating and Sealing: Apply protective coatings or sealants to manhole surfaces to enhance durability and resistance to environmental factors.
- Structural Assessment: Conduct thorough structural assessments to identify weak- nesses or defects, making informed decisions on repairs or replacements.

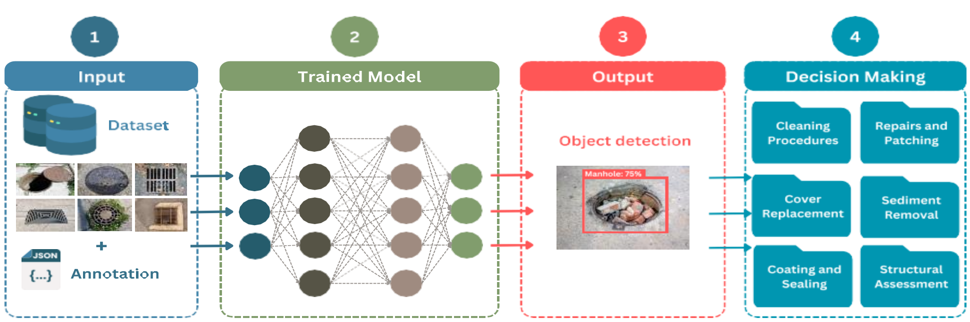

Fig. 3. Approaches of our experimental studies.

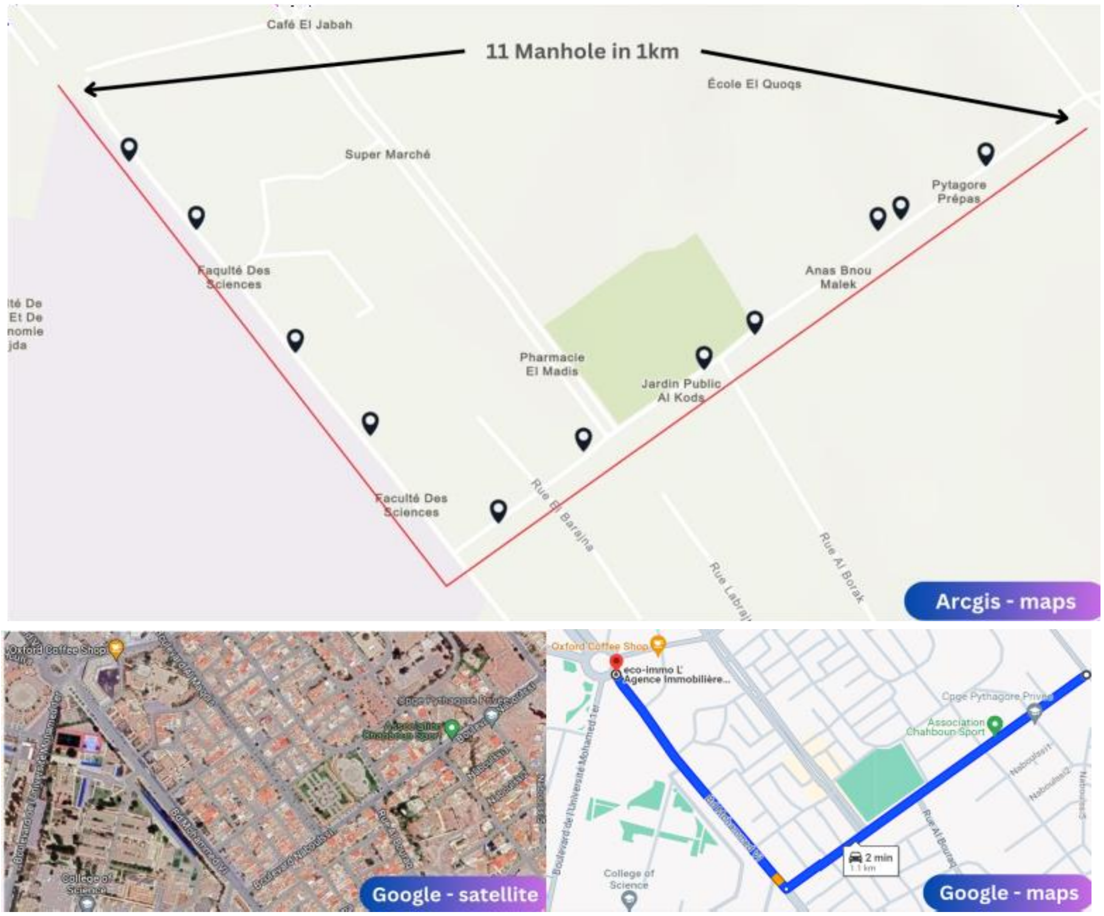

Fig. 4. Region designated for analyzing our architectures in actual road settings.

- Odor Control Measures: Introduce measures such as deodorizing agents or ventilation systems to address unpleasant odors associated with manhole situations.
- Emergency Pumping: Establish protocols for emergency pumping in situations where water accumulates rapidly, preventing potential flooding and infrastructure damage.
- Rehabilitation Programs: Develop rehabilitation plans for aging manholes, including strategies for structural reinforcement and longevity extension.
- Upgraded Materials: Consider using advanced, durable materials for manhole construction and covers to enhance longevity and reduce maintenance needs.

These decisions encompass a range of actions aimed at addressing specific issues within manhole situations, from routine maintenance to emergency response and infra- structure upgrades.

V. Results and Discussions

A. Hardware and Software Characteristics

To assemble a comprehensive dataset of manhole covers from web sources for effective labeling, we employ a variety of online databases and repositories, including platforms such as Kaggle and other internet resources that provide copyright-free imagery. Following the collection phase, we utilize a suite of labeling tools for image annotation, such as LabelImg and VGG Image Annotator (VIA). These tools facilitate the precise placement of bounding boxes around each manhole cover within the images. Each image must undergoes a meticulous review process to verify the accuracy of the annotations. This ensures the integrity of the dataset, which is crucial for the subsequent training of machine learning models.

B. Implementation Setup

For our implementation, we have used TensorFlow and PyTorch, two open-source data analysis and deep learning software library, on a high-performance computing system (HPC) equipped with the following hardware specifications:

- Two Intel Gold 6148 (2.4 GHz/20 cores) processors.
- Two NVIDIA Tesla V100 graphics cards, each with 32GB of RAM.

C. Evaluation Metrics

Table I summarizes the metrics, their application in the context of computer vision and object detection, and their respective formulas [26], [27].

TABLE I. ENHANCED AND REFINED METRICS FOR MODEL EVALUATION ANALYSIS

Metric	Explanation	Mathematical Representation
IoU	Measures the overlap between two bounding boxes, used in evaluating the accuracy of object detection models.	IoU = Area of Overlap / Area of Union
mAP	Average of the Average Precision (AP) across all classes and/or IoU thresholds, used in object detection.	mAP = (1/N) * Σ(AP_i) from i=1 to N
Precision	Ratio of correctly predicted positive observations to the total predicted positives.	Precision = TP / (TP + FP)
Recall	Ratio of correctly predicted positive observations to all observations in the actual class.	Recall = TP / (TP + FN)
F1-Score	Harmonic mean of Precision and Recall, used for balancing the two and helpful in case of uneven class distribution.	F1-Score = 2 * (Precision * Re- call) / (Precision + Recall)

D. Evaluating the Results

Table II, presents the performance results of various deep learning models for the task of object detection, specifically for detecting manhole covers. The models listed are YOLOv8, GroundingDINO, DETR, Faster R-CNN, MobileNet SSD v2, and Detectron2. They are evaluated on several metrics, which include Intersection over Union (IoU), mean Average Precision (mAP) at IoU thresholds of 0.5 and 0.75, Inference time per image, Precision, Recall, and F1-Score.

YOLOv8 outperforms the other models in almost all the metrics, with an IoU of 94%, mAP@0.5 of 87.74%, and mAP@0.75 of 89.44%. Its precision, recall, and F1- Score are all equal at 95%. These numbers indicate a highly accurate and reliable model for the specified detection task. On the other side, the inference time per image is 60 frames per second for YOLOv8 which outperforms other models. The inference time per image is an important factor in real-world applications where processing speed can be crucial.

The other models show varying degrees of success. GroundingDINO and MobileNet SSD v2 demonstrate moderate performance, with GroundingDINO achieving a higher mAP@0.5 but lower F1-Score compared to MobileNet SSD v2. DETR and Faster RCNN have comparable performance, with DETR having a slightly better IoU and mAP@0.75, suggesting better localization and confidence in detections at stricter thresholds.

The graph depicts the training progress of various deep learning models for the object detection of manhole covers, with a focus on the mean Average Precision (mAP) at an Intersection over Union (IoU) threshold of 0.5. This metric, mAP@0.5, is a standard performance measure in object detection that combines both precision and recall to evaluate the quality of the predictions, specifically at an IoU threshold of 50%.

The training process is shown in Fig. 5, over several epochs, which represent full iterations over the entire dataset. As the epochs increase, we generally expect the model to improve in its detection capabilities as it learns from the data. YOLOv8 shows a rapid and steady improvement, achieving a high mAP@0.5 early on and maintaining that lead throughout the training process. This indicates that YOLOv8 is learning effectively and can generalize well from the training data to detect manhole covers with high precision and recall.

TABLE II. OBTAINED RESULTS FOR THE IMPLEMENTED MODELS

DEEP LEARNING MODEL	IoU	mAP	mAP	Inference	Precision	Recall	F1-Score
	%	@0.5 %	@0.75 %	s/Image	%	%	%
YOLOv8	94.00	87.74	89.44	60	95.03	95.02	95.02
GroundingDINO	73.00	82.74	81.12	9	78.74	74.21	76.41
DETR	89.00	79.56	81.17	11	83.95	72.58	77.85
Faster R-CNN	80.00	80.68	84.44	15	71.20	76.31	73.67
MobileNet SSD v2	78.00	75.19	81.82	45	83.34	77.27	80.19
Detectron2	0.00	81.07	67.74	39	-	-	-

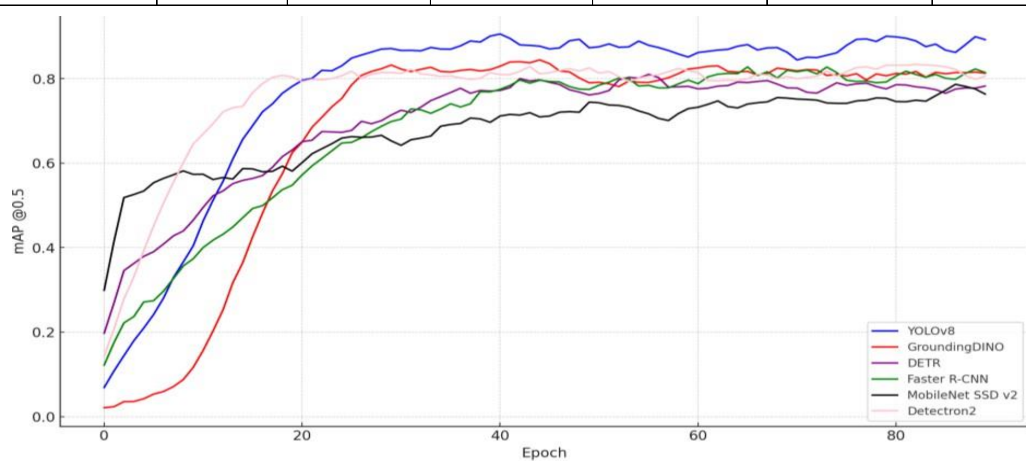

Fig. 5. Simulation of mAP@0.5 training progress over 100 epochs.

Other models, such as GroundingDINO, DETR, Faster R-CNN, and MobileNet SSD v2, also show improvement over time but with different learning curves. GroundingDINO and Faster R-CNN, for instance, demonstrate a more gradual improvement. DETR and MobileNet SSD v2 have similar trajectories, with MobileNet SSD v2 starting off stronger but DETR overtaking it by the end. These variations in learning curves can be due to differences in model architectures, learning rates, data augmentation, and other hyperparameters that affect how quickly and effectively a model learns.

Detectron2, however, appears to have a different trend. It starts with poor performance and takes a longer time to begin improving. Once it does start to improve, it shows a more gradual and less stable increase in mAP@0.5, with fluctuations that suggest the model may not be learning consistently or is struggling with the dataset. This could be indicative of issues such as overfitting, underfitting, or inadequate training data, which may require further investigation and adjustment of the training process. Overall, the graph is an essential tool for understanding the learning dynamics of each model and for diagnosing potential issues in the training process.

Fig. 6 provided showcases the results of a machine learning object detection algorithm, specifically YOLOv8, as it attempts to identify manhole covers in various settings. The image demonstrates instances, where the YOLOv8 model has successfully identified manhole covers with high confidence scores, as indicated by the numbers next to the word "Manhole" within the red bounding boxes. These scores represent the model's confidence in its predictions, with 1.0 being the highest, signifying 100% confidence.

On the other hand, Fig. 7 illustrates scenarios where the YOLOv8 model has incorrectly detected manhole covers or assigned lower confidence scores to its predictions. These false positives or less certain detections can occur due to a variety of factors such as occlusions, varying lighting conditions, unusual manhole cover designs, or similarities between the manhole covers and other objects in the environment.

YOLOv8, like other machine learning models, is not infallible and can sometimes fail to make accurate predictions. This is often due to the limitations in the training data or the inherent challenges in interpreting complex and dynamic real-world scenes. These misclassifications and uncertainties in object detection models highlight the need for continuous improvement and training with diverse datasets to enhance the model's accuracy and reliability in various conditions.

E. Discussions

For the improved and expanded implementation, the drone system operates by transmitting real-time images of detected manholes via an RTMP server. These images are accompanied by precise geolocation data, including the exact address and GPS coordinates. Once this information is relayed to the central monitoring system, operators can assess the condition of the manhole and determine the necessary course of action. This decision-making process involves evaluating the status of the manhole, such as its current state, potential hazards, and maintenance requirements. The information, along with the operator's decision, is then systematically cataloged in a structured database. An example of such data organization can be seen in Table III, which illustrates the format and type of data stored.

Fig. 6. Examples of good detection of manholes cover using Yolov8.

Fig. 7. Examples of wrongly detection of manholes cover using Yolov8.

TABLE III. DATA USED FOR THE IMPLEMENTATION

ID	City	Address	Latitude	Longitude	Description	Status	PicURL	Action
1	OUJDA	Boulevard Nabloussi	34.655204	-1.892503	Manhole in the middle of the road.	1	/Manhole/#man01.jpg	none
2	OUJDA	Boulevard Nabloussi	34.655902	-1.891422	Manhole in the left side of the road.	0	/Manhole/#man02.jpg	Cleaning Procedures
6	OUJDA	Bd Mohammed VI	34.654240	-1.898296	Manhole in the middle of the road.	1	/Manhole/#man06.jpg	none
11	OUJDA	Bd Mohammed VI	34.655922	-1.899748	Manhole in the right side of the road.	0	/Manhole/#man011.jpg	Repairs and patching

As the database grows with more entries, it becomes a rich source of information for training machine learning algorithms. By analyzing the accumulated data, these algorithms can learn to recognize patterns and anomalies associated with different manhole conditions. Over time, with sufficient training and refinement, the system can evolve to autonomously identify issues and suggest or even initiate appropriate actions without human intervention. This advancement in autonomous decision-making not only enhances efficiency but also reduces the response time in addressing urban infrastructure issues, thereby contributing to a safer and more effectively managed city environment.

VI. CONCLUSION

In conclusion, this research has successfully demonstrated the effectiveness of employing the YOLOv8 object detection model for the real-time supervision and detection of manhole situations using drone imagery and GPS integration. Our experiments showcased the model's exceptional precision and efficiency, underscoring its potential as a robust solution for proactive urban infrastructure monitoring. The integration of YOLOv8 proved instrumental in surpassing traditional methods, offering a swift and accurate means of identifying manhole covers and potential hazards across diverse urban landscapes. The synergy between deep learning, drone technology, and GPS data not only enhanced the speed of detection but also provided a comprehensive understanding of the spatial dynamics inherent in complex urban environments. The integration of advanced deep learning models, coupled with real-world performance metrics, establishes a robust foundation for the practical implementation of our proposed system in urban environments.

As a perspective, we will continue to find more models and to make another experiment to strengthen our research in this field.

REFERENCES

[1] U. G. Sandström, P. Angelstam, and A. Khakee, ‘Urban comprehensive planning – identifying barriers for the maintenance of functional habitat networks’, Landsc. Urban Plan., vol. 75, no. 1, pp. 43–57, Feb. 2006, doi: 10.1016/j.landurbplan.2004.11.016.

[2] A. Yani, M. L. Fadhillah, and R. W. Saputra, ‘Supervision of the Implementation of Road Maintenance in the Department of Public Works and Urban Planning’, East Asian J. Multidiscip. Res., vol. 2, no. 2, Art. no. 2, Feb. 2023, doi: 10.55927/eajmr.v2i2.3102.

[3] G. Jia, G. Han, H. Rao, and L. Shu, ‘Edge Computing-Based Intelligent Manhole Cover Management System for Smart Cities’, IEEE Internet Things J., vol. 5, no. 3, pp. 1648–1656, Jun. 2018, doi: 10.1109/JIOT.2017.2786349.

[4] T. Zhou, L. Zheng, Y. Peng, and R. Jiang, ‘A Survey of Research on Crowd Abnormal Behavior Detection Algorithm Based on YOLO Network’, 2022 2nd Int. Conf. Consum. Electron. Comput. Eng. ICCECE, pp. 783–786, Jan. 2022, doi: 10.1109/ICCECE54139.2022.9712684.

[5] N. M. Noor, A. Abdullah, and M. Hashim, ‘Remote sensing UAV/drones and its applications for urban areas: a review’, IOP Conf. Ser. Earth Environ. Sci., vol. 169, no. 1, p. 012003, Jun. 2018, doi: 10.1088/1755-1315/169/1/012003.

[6] B. Fan, Y. Li, R. Zhang, and Q. Fu, ‘Review on the Technological Development and Application of UAV Systems’, Chin. J. Electron., vol. 29, no. 2, pp. 199–207, 2020, doi: 10.1049/cje.2019.12.006.

[7] L. Alzubaidi et al., ‘Review of deep learning: concepts, CNN architectures, challenges, applications, future directions’, J. Big Data 2021 81, vol. 8, no. 1, pp. 1– 74, Mar. 2021, doi: 10.1186/S40537-021-00444-8.

[8] S. Kumar and K. B. Moore, ‘The Evolution of Global Positioning System (GPS) Technology’, J. Sci. Educ. Technol., vol. 11, no. 1, pp. 59–80, Mar. 2002, doi: 10.1023/A:1013999415003.

[9] D. Amaxilatis, O. Akrivopoulos, G. Mylonas, and I. Chatzigiannakis, ‘An IoTBased Solution for Monitoring a Fleet of Educational Buildings Focusing on Energy Efficiency’, Sensors, vol. 17, no. 10, Art. no. 10, Oct. 2017, doi: 10.3390/s17102296.

[10] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, ‘Computer vision algorithms and hardware implementations: A survey’, Integration, vol. 69, pp. 309–320, Nov. 2019, doi: 10.1016/j.vlsi.2019.07.005.

[11] M. Leo, P. Carcagnì, P. L. Mazzeo, P. Spagnolo, D. Cazzato, and C. Distante, ‘Analysis of Facial Information for Healthcare Applications: A Survey on Computer Vision-Based Approaches’, Information, vol. 11, no. 3, Art. no. 3, Mar. 2020, doi: 10.3390/info11030128.

[12] X. Zhao et al., ‘Urban infrastructure safety system based on mobile crowdsensing’, Int. J. Disaster Risk Reduct., vol. 27, pp. 427–438, Mar. 2018, doi: 10.1016/j.ijdrr.2017.11.004.

[13] C. Calba et al., ‘Surveillance systems evaluation: a systematic review of the existing approaches’, BMC Public Health, vol. 15, no. 1, p. 448, May 2015, doi: 10.1186/s12889-015-1791-5.

[14] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, ‘Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection’, Sustain. Cities Soc., vol. 65, p. 102600, Feb. 2021, doi: 10.1016/J.SCS.2020.102600.

[15] D. Qi, W. Tan, Q. Yao, and J. Liu, ‘YOLO5Face: Why Reinventing a Face Detector’, May 2021, [Online]. Available: https://arxiv.org/abs/2105.12931v3.

[16] I. Lindsay and N. N. Kong, ‘Using the ArcGIS Collector Mobile App for Settlement Survey Data Collection in Armenia’, Adv. Archaeol. Pract., vol. 8, no. 4, pp. 322–336, Nov. 2020, doi: 10.1017/aap.2020.26.

[17] S. Phantuwongraj, P. Chenrai, and T. Assawincharoenkij, ‘Pilot Study Using ArcGIS Online to Enhance Students’ Learning Experience in Fieldwork’, Geosciences, vol. 11, no. 9, Art. no. 9, Sep. 2021, doi: 10.3390/geosciences11090357.

[18] D. Pang, Z. Guan, T. Luo, W. Su, and R. Dou, ‘Real-time detection of road manhole covers with a deep learning model’, Sci. Rep. 2023 131, vol. 13, no. 1, pp. 1– 14, Sep. 2023, doi: 10.1038/s41598-023-43173-z.

[19] R. E. Andersen, M. Zajaczkowski, H. Jaiswal, J. Xu, W. Fan, and E. Boukas, ‘Depth-based Deep Learning for Manhole Detection in UAV Navigation’, IST 2022 - IEEE Int. Conf. Imaging Syst. Tech. Proc., 2022, doi: 10.1109/IST55454.2022.9827720.

[20] R. Timofte and L. V. Gool, ‘Multi-view manhole detection, recognition, and 3D localisation’, Proc. IEEE Int. Conf. Comput. Vis., pp. 188–195, 2011, doi: 10.1109/ICCVW.2011.6130242.

[21] ‘Sensors | Free Full-Text | Survey and Performance Analysis of Deep Learning Based Object Detection in Challenging Environments’.

Accessed: Dec. 28, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/21/15/5116.

[22] A. V. Savkin and H. Huang, 'Asymptotically Optimal Deployment of Drones for Surveillance and Monitoring', Sensors, vol. 19, no. 9, Art. no. 9, Jan. 2019, doi: 10.3390/s19092068.

[23] V. Vishnani, A. Adhya, C. Bajpai, P. Chimurkar, and K. Khandagle, 'Manhole Detection using Image Processing on Google Street View imagery', in 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Aug. 2020, pp. 684–688. doi: 10.1109/ICSSIT48917.2020.9214219.

[24] B. Commandre, D. En-Nejjary, L. Pibre, M. Chaumont, C. Delenne, and N. Chahinian, 'Manhole Cover Localization in Aerial Images with a Deep Learning Approach', in ISPRS Hannover Workshop: HRIGI 17 – CMRT 17 – ISA 17 – EuroCOW 17, Hannover, Germany, Jun. 2017, pp. 333–338. doi: 10.5194/isprs-archives-XLII-1-W1-333-2017.

[25] 'Using large-scale experiments and machine learning to discover theories of human decision-making | Science'. Accessed: Dec. 28, 2023.[Online].Available:https://www.science.org/doi/full/10.1126/science.abe2629.

[26] I. Buzhinsky, A. Nerinovsky, and S. Tripakis, 'Metrics and Methods for Robustness Evaluation of Neural Networks With Generative Models'. 2020.

[27] 'Metrics to Evaluate your Machine Learning Algorithm' [Online]. Avaible https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e3823.