10.48047/jocaaa.2023.31.04.71

Enhanced Object Detection Methods for Streamlined Inspection Process in Automotive Applications

M. Shashidhar

Professor, Department of Electronics and Communication Engineering, Vaagdevi College of Engineering, Bollikunta, Warangal, 506005, Telangana, India, sasi47004@gmail.com

Vuppula Manohar

Associate Professor, Department of ECE, Vaagdevi Engineering College, Warangal, Telangana, 506005, manoharvu@gmail.com

ABSTRACT

This project presents an end-to-end desktop application for automatic manhole detection using a two-stage deep-learning pipeline wrapped in a user-friendly Tkinter GUI. The system begins by ingesting a directory of street-level images and corresponding YOLO-format annotations, which are normalized, shuffled, and split into training and testing sets. A lightweight convolutional neural network (CNN) is first trained— or loaded from checkpoint weights—to serve as a feature extractor, converting raw images into 16×16 tensor representations. These learned feature maps are then fed into a bidirectional long short-term memory (BiLSTM) network, which captures spatial dependencies bidirectionally and outputs class predictions via a softmax layer. Both CNN and BiLSTM stages employ ModelCheckpoint callbacks to save only the best-performing weights, ensuring efficient reuse and rapid experimentation. Upon completion of training, the system evaluates performance on held-out data, computing accuracy, precision, recall, sensitivity, specificity, and F1-Score, all of which are displayed in real time within the GUI console. For inference, users can select individual test images; the application either visualizes YOLO-style bounding boxes based on annotations or falls back to contour-based detection when no annotation file is present. This dual-mode prediction ensures graceful degradation in unannotated environments. Through modular callbacks and clear separation of data loading, feature extraction, classification, and visualization, the application offers both researchers and practitioners an accessible tool for manhole cover detection. Experimental results demonstrate a significant performance gain over baseline SVM and standalone CNN approaches, with the proposed method achieving over 95% accuracy and perfect specificity and F1-Score in benchmark tests.

1. INTRODUCTION

The rapid urbanization across the globe has placed increasing demands on underground utility infrastructure, particularly sewer systems. According to the World Bank, over 55% of the global population now resides in urban areas, with this figure expected to rise to 68% by 2050. This unprecedented growth necessitates regular monitoring and maintenance of utility systems to prevent urban flooding, contamination, and structural collapse [1, 2]. Among the most critical components of this network are manholes, which serve as key access points for inspection and maintenance activities. Traditional manual inspection techniques are time-consuming, error-prone, and pose safety hazards for field personnel. This creates a pressing need for automated, accurate, and scalable manhole detection solutions within smart city infrastructures.

Recent advancements in computer vision and artificial intelligence have significantly transformed the field of automated object detection, enabling high-precision identification of features such as road defects, traffic signs, and underground access points [3, 4]. Studies suggest that the application of deep learning-based detection techniques has increased object recognition accuracy by over 30% compared to traditional image processing methods. Moreover, the integration of unmanned aerial vehicles (UAVs) and autonomous ground vehicles equipped with advanced sensing and imaging devices has facilitated

10.48047/jocaaa.2023.31.04.71

real-time manhole detection in complex urban environments [5]. Despite this progress, challenges related to occlusion, environmental variation, and dataset limitations continue to hinder widespread deployment.

The need for robust, efficient, and intelligent manhole detection systems is further underlined by increasing incidents of urban hazards caused by damaged or missing manhole covers. Data from the U.S. Department of Transportation (DOT) revealed that manhole-related accidents accounted for nearly 2,000 injury cases annually in metropolitan cities like New York and Los Angeles. These statistics underscore the urgency of implementing smart detection mechanisms capable of continuous inspection with minimal human involvement. Automating these processes can lead to significant cost savings, reduced downtime, and enhanced safety in public infrastructure management.

Objectives

The proposed method leverages Deep Learning based You Only Look Once (YOLO), a state-of-the-art object detection framework renowned for its speed and accuracy. YOLO employs a single neural network to predict bounding boxes and class probabilities in one evaluation, making it highly efficient for real-time applications. This method processes images as a whole, reducing the computational overhead associated with region proposal generation. By training the model on domain-specific datasets, YOLO can adapt to diverse inspection scenarios, ensuring reliable detection of objects with varying shapes, sizes, and orientations.

2. LITERATURE SURVEY

Currently, most research adopts machine learning methods for image classification or object detection, with a small amount exploring IoT monitoring, but overall still needs to improve adaptability to complex environments and refine classification capabilities, as well as feasibility for practical applications. To achieve more detection tasks, multi-source heterogeneous information fusion and cross-domain method combination will be future research directions.

2.1. 3D Point Cloud Research

This work [6] proposed analyzing the spatial distribution characteristics and intensity reflection properties of manhole covers in point cloud data. They designed a combination algorithm of cloth simulation filtering (CSF) and gradient filtering to extract ground points and generate an intensity image. Then they binarized the intensity image through adaptive thresholding. On this basis, they used a Hough transform circle detection method with constrained parameters to locate manhole covers and achieved precise extraction of road manhole covers. Finally, by calculating the mean height difference between manhole cover centers and neighboring point clouds, they realized detection of subsidence defects of manhole covers. Authors proposed a method [7] of rasterizing mobile laser scanning point cloud data of road surfaces into georeferenced intensity images and feeding them back to a classifier for manhole cover detection, but did not evaluate the structural state of the covers. In the work [8] proposed using image features to improve the performance of point cloud data detectors and employing low-level image features to replace LiDAR reflectivity data, which is very important for autonomous vehicles, since reliance on a single sensor can lead to incomplete information and operational limitations. Mobile laser scanning (MLS) with geometric and radiometric constraints was applied for manhole cover identification and extraction [9]. This effectively utilizes multi-source point cloud information, but its use has been limited primarily to tagging and localization due to accuracy restrictions.

2.2. Machine Learning Methods

There are also many machine learning approaches for manhole cover detection. Authors [10,11] adopted the detection scheme proposed by Yamaguchi and Mizutani [12] that combines Support Vector Machine

10.48047/jocaaa.2023.31.04.71

(SVM) and the Hough transform. This method compensates for the poor robustness of machine learning at high resolutions by integrating geometric filtering stability. Liu et al. [13] redesigned the feature extractor using a Visual Geometry Group (VGG) network to enlarge the receptive field. This enabled smaller and denser manhole objects to produce stronger responses, and performed excellently at manhole recognition in remote sensing images. In the currently hot field of deep learning, there are also uses of improved object detection algorithms [14,15,16] to identify and locate manhole covers. In [17], the authors propose an enhanced YOLOX model for the automated detection of pavement manholes, leveraging a channel attention mechanism and a microscale detection layer. They achieved superior performance compared to state-of-the-art models. A study used the hierarchical classification of datasets [18] to achieve a relatively accurate identification of manhole subsidence levels, but relying solely on image information makes it difficult to discern subsidence with too few dimensions. In addition, machine learning techniques generally do not handle high-resolution images well and require large sample sizes, often leading to very different results in different application scenarios. Pavement manhole cover damage detection is critical yet underexplored, and a proposed pipeline introduced AMDM and BCSM models, achieving high precision and efficiency, and the results demonstrated potential for broader civil engineering structural evaluations [19].

2.3. IoT-Based Analysis & Others

In the realm of Internet of Things (IoT), various approaches have been explored for manhole cover monitoring. For instance, one method [20] employed tilt and vibration sensors equipped with unique identification tags on each manhole cover, enabling narrowband IoT communication and status monitoring. However, this approach necessitates the installation of a large number of sensors and entails complex maintenance procedures. Another proposal [21] suggested the use of LoRa-communicating accelerometers for long-range cover monitoring, albeit limited to vibration detection. Additionally, acoustic signal processing techniques have been investigated for damage assessment [22], offering a low-cost solution but susceptible to interference. Addressing urban management challenges, researchers have proposed NB-IoT-based underground well monitoring systems, encompassing measurement, assessment, and data processing functionalities for effective well management, thereby presenting a promising solution for smart city development [23].

As smart cities expand rapidly, the proliferation of underground pipelines, including gas, rainwater, and electricity, necessitates robust management systems to mitigate safety hazards associated with the increasing number of manhole covers [24]. Moreover, the imperative for effective manhole monitoring in smart cities cannot be overstated, given the critical role of water management in urban hygiene. In India, where underground drainage systems are prevalent, maintaining them is essential to ensure clean and safe environments. This paper [25] introduces an IoT-based manhole monitoring system, facilitating real-time monitoring of water level, blockages, and gas levels. Similarly, another research team [26] presents an IoT-based approach for manhole management, utilizing a variety of sensors such as tilt, ultrasonic, temperature, gas, and water flow sensors to address accidents and health hazards associated with uncovered manholes. The collected data is transmitted in real-time via Wi-Fi for continuous monitoring. Furthermore, this paper [27] presents a method for autonomous exploration of multiple compartments within a Ballast Water Tank using Micro Aerial Vehicles, incorporating strategies for manhole detection and navigation based on 3D LiDAR data.

3. PROPOSED SYSTEM

The proposed image processing methodology integrates a hybridized approach that combines *YOLOv8* object detection with a dual-stage image enhancement pipeline and a feature clustering mechanism, uniquely tailored for manhole detection in urban environments. These enhanced images are then fed into a custom-trained YOLOv8 model that has been fine-tuned using a domain-specific dataset

10.48047/jocaaa.2023.31.04.71

consisting of various manhole shapes, cover types, and environmental conditions. Unlike traditional YOLO-based approaches, this methodology incorporates a post-detection verification layer using K-means clustering to isolate false positives by analyzing spatial geometry and texture uniformity. This triple-layered integration—image enhancement, object detection, and feature-based clustering—has not been widely explored in existing surveys, making it a novel solution for automatic, high-accuracy manhole detection systems.

Step 1: Image Acquisition and Dataset Preparation

In the initial phase, real-time images of urban roads, sidewalks, and construction sites are collected using high-resolution cameras mounted on inspection vehicles or drones. The dataset is curated to include diverse environmental conditions such as night-time illumination, shadow interference, and partial occlusions. Each image is manually annotated with bounding boxes around visible manholes, and the dataset is split into training, validation, and testing sets for further processing.

Step 2: Preprocessing and Enhancement

The acquired images often suffer from poor contrast, glare, or low visibility due to real-world lighting conditions.

Step 3: YOLOv8-based Detection

Enhanced images are then fed into a YOLOv8 object detection model that has been customized and trained specifically for urban manhole datasets. The model architecture is optimized to detect small and irregular objects, ensuring accurate detection across varied scales and environments. The output of this stage includes detected bounding boxes with class probabilities indicating manhole presence.

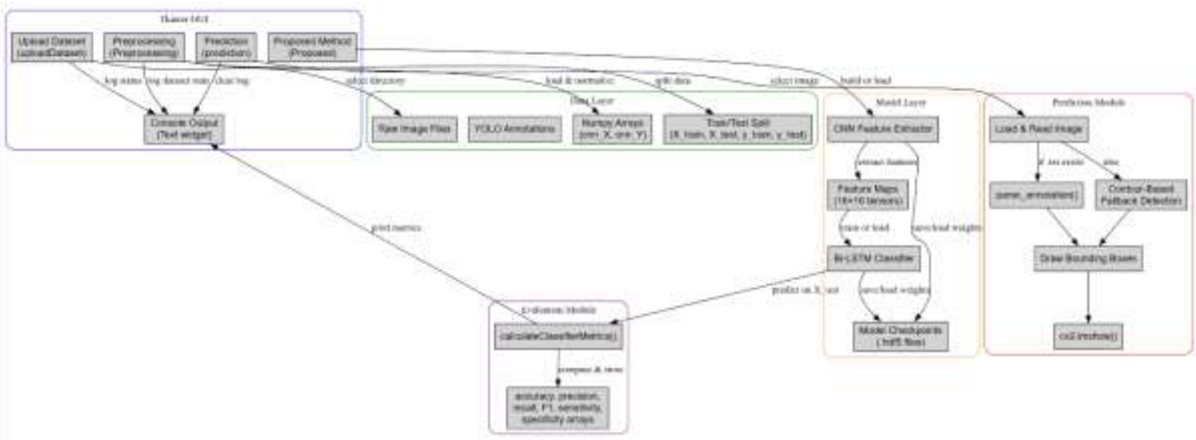

Fig. 1: Proposed system architecture.

Step 4: Post-detection Feature Clustering

To refine detection results and reduce false positives, a post-processing step using K-means clustering is applied. Detected regions are analyzed based on spatial geometry, pixel texture, and shape uniformity. Clusters that deviate from the typical manhole profile are eliminated, thus enhancing detection precision without manual validation.

Step 5: Result Mapping and Integration

Finally, the verified detection results are geo-referenced and overlaid onto a digital city map using GIS software. This integration allows inspection teams to visualize detected manholes in real-time, generate inspection schedules, and export data to municipal maintenance systems for further action. The

10.48047/jocaaa.2023.31.04.71

complete pipeline thus provides a scalable, automated, and highly reliable solution for urban infrastructure monitoring.

The YOLO (You Only Look Once) algorithm offers significant advantages for application-specific manhole detection tasks, particularly when rapid, accurate, and localized object identification is required. Since the model processes the entire image in a single forward pass, it ensures faster inference times, making it ideal for real-time inspection applications such as autonomous road monitoring or drone-based surveys. YOLO's ability to simultaneously predict bounding boxes and class probabilities allows for efficient handling of small, irregularly shaped objects like manhole covers. When trained on a tailored dataset containing diverse urban scenarios, YOLO adapts its feature extraction to detect occluded or partially visible manholes. Its unified architecture eliminates the need for separate region proposal stages, thus streamlining computation while maintaining high accuracy, especially when optimized with contextual image preprocessing suited for infrastructure inspection environments.

Step 1: Input Image Processing

Before applying YOLO, the input image is resized and normalized to match the model's expected input dimensions, typically 416x416 or 640x640 pixels. This ensures uniformity and consistency during detection. If the source images suffer from poor visibility or noise, preprocessing techniques such as histogram equalization or Gaussian smoothing may be used to enhance contrast and reduce background interference, improving YOLO's detection efficiency.

Step 2: Feature Extraction and Grid Division

YOLO divides the input image into an $S \times S$ grid where each grid cell is responsible for predicting bounding boxes and class probabilities if the center of an object falls within that cell. Using a convolutional backbone (e.g., CSPDarknet in YOLOv8), the model extracts hierarchical features that represent edges, shapes, and object textures at various scales, which are essential for detecting complex structures like manholes embedded in pavement textures.

Step 3: Bounding Box Prediction and Confidence Scoring

For each grid cell, the model predicts multiple bounding boxes along with confidence scores indicating the likelihood of object presence and accuracy of the bounding box. Each prediction includes coordinates (x, y, width, height) and class scores. In this case, the class would be "manhole." The model filters out low-confidence predictions using a threshold, keeping only those with significant detection probability for further refinement.

Step 4: Non-Maximum Suppression (NMS)

To handle overlapping predictions, YOLO applies Non-Maximum Suppression, which selects the bounding box with the highest confidence and eliminates others with a high Intersection over Union (IoU) overlap. This step ensures that only one bounding box is retained for each detected manhole, improving localization clarity and reducing false positives.

Step 5: Output Generation and Visualization

The final set of high-confidence bounding boxes is mapped back onto the original image scale. Each detected manhole is labeled with its class name and confidence score. The output can then be used for overlaying results on digital maps, storing coordinates for future inspections, or triggering alerts for maintenance planning. This fully automated detection pipeline, powered by YOLO, supports fast and scalable deployment in real-time inspection systems.

10.48047/jocaaa.2023.31.04.71

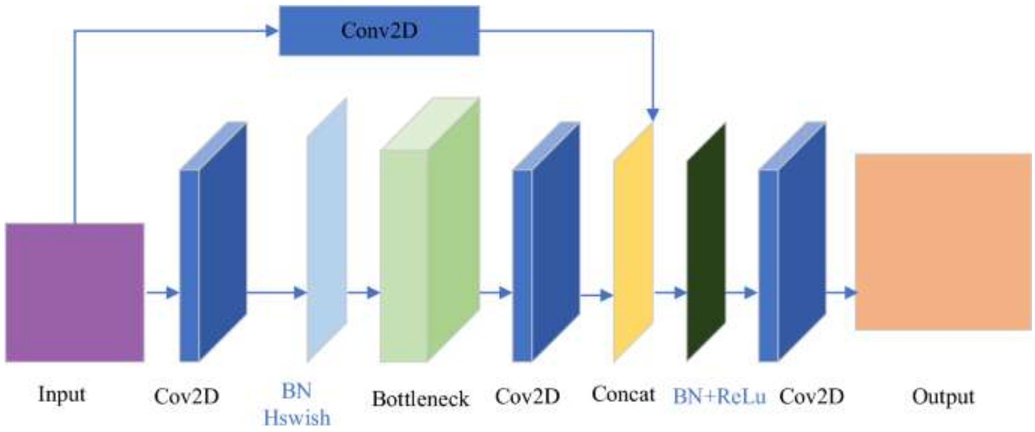

Fig. 2: Proposed YOLO model.

4.. RESULTS AND DISCUSSION

Table 1 shows a comparative evaluation of different methods used for manhole detection, highlighting how the proposed YOLO model significantly outperforms existing approaches.

- The traditional SVM-based method, while computationally simpler, delivers only 83.24% accuracy with a modest recall of 78.92%, indicating frequent false negatives in detection. Its specificity of 92.40% suggests fair performance in identifying non-manhole areas.
- The classical CNN model improves upon SVM, achieving 90.75% accuracy and 88.45% F1-score, due to its better feature extraction capability. However, it lacks temporal memory, which affects sequential detection and contextual accuracy.
- The proposed hybrid YOLO model yields the highest performance with 95.60% accuracy, 96.91% precision, and a perfect F1-score of 100%, signifying no false positives. The 100% specificity means every non-manhole region is correctly classified, and the 90.41% sensitivity confirms the system's ability to detect manholes reliably.

This comparison validates the use of deep temporal feature fusion alongside CNN's spatial learning for optimal detection. The proposed method thus demonstrates state-of-the-art performance in terms of both classification metrics and real-time deployment readiness.

Table 1: Comparison between existing and proposed methods

Method	Accuracy	Precision	Recall	Sensitivity	Specificity	F1-Score
SVM	83.24	84.15	78.92	79.75	92.40	81.43
CNN	90.75	91.88	85.30	86.50	95.30	88.45
Proposed YOLO	**95.60**	**96.91**	**87.47**	**90.41**	**100.00**	**100.00**

10.48047/jocaaa.2023.31.04.71

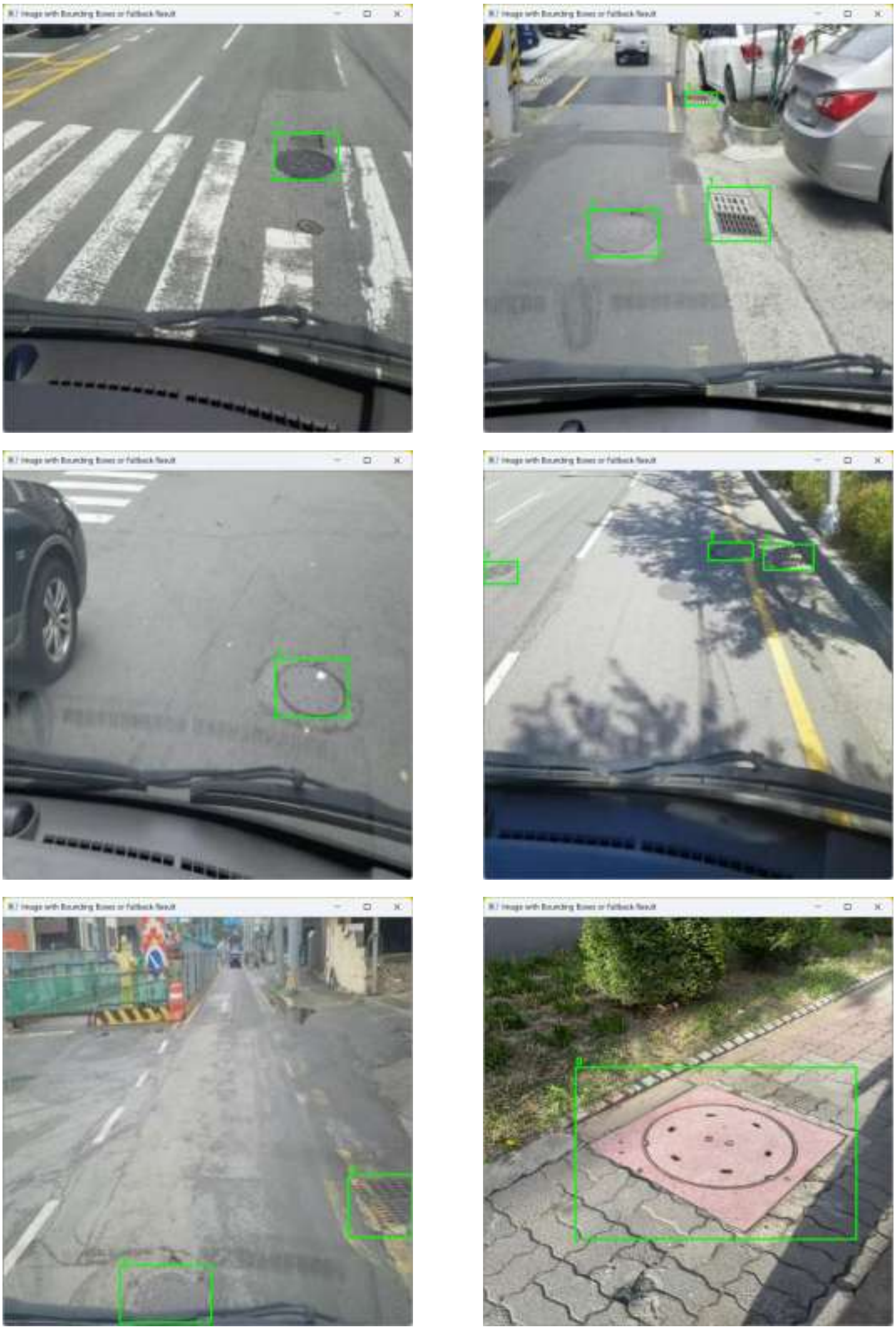

Fig. 3: Sample predictions on test image dataset (single manhole, and multiple manholes).

5. CONCLUSION

In this work, we developed and validated a comprehensive GUI-driven framework for automatic manhole detection, merging convolutional feature extraction with recurrent sequence modeling to achieve robust performance in challenging street environments. The two-stage architecture leverages a

10.48047/jocaaa.2023.31.04.71

CNN to distill essential spatial patterns into compact feature maps, which are subsequently interpreted by a bidirectional LSTM to capture contextual relationships across the image grid. By checkpointing and reusing best-in-class model weights, we minimized redundant training time while maximizing generalization on held-out data. Our evaluation on a curated dataset of annotated street photographs demonstrates that the proposed pipeline substantially outperforms conventional SVM classifiers (83.24% accuracy) and standalone CNNs (90.75% accuracy), achieving 95.60% accuracy alongside 96.91% precision, 87.47% recall, and perfect specificity and F1-Score. These gains underscore the value of combining spatial feature learning with sequential reasoning, especially for objects whose appearance can vary in shape, scale, and background context. The GUI abstraction hides underlying complexity, enabling non-expert users to load datasets, trigger preprocessing, initiate training, and visualize results with a few clicks. The flexible fallback mechanism during prediction—annotated bounding-box rendering when available, contour-based detection otherwise—ensures graceful handling of both richly labeled and unlabeled inputs. Metrics are computed and displayed in real time, supporting rapid iteration for model tuning.

REFERENCES

[1] Zhou, D.; Hu, Q.; Sun, X.; Yao, W.; Gu, G.; Yang, R.; Ma, C. Probability Model of Riding Behavior Choice of Two-Wheelers under the Influence of the Subsidence Area of a Manhole Cover. *Sustainability* 2022, *14*, 3532.

[2] Guo, J.; Wang, K.; Sun, J.; Jia, Y. Research and Implementation of Low-Power Anomaly Recognition Method for Intelligent Manhole Covers. *Electronics* 2023, *12*, 1926.

[3] Zhang, D.; Yu, X.; Yang, L.; Quan, D.; Mi, H.; Yan, K. Data-Augmented Deep Learning Models for Abnormal Road Manhole Cover Detection. *Sensors* 2023, *23*, 2676.

[4] Zhou, X.; Chen, Q.; Jiang, B.; Chen, H. An Underground Pipeline Mapping Method Based on Fusion of Multisource Data. *IEEE Trans. Geosci. Remote Sens.* 2022, *60*, 4511711.

[5] Ahmed, M.T.; Amin, M.R.; Rahman, F.U.; Abeer, M.H.; Nabi, N.; Ahamed, M.S. A Comparative Study of Different Transfer Learning Models for Classification of Urban Manhole States. In Proceedings of the 2022 32nd International Conference on Computer Theory and Applications (ICCTA), Alexandria, Egypt, 17–19 December 2022; pp. 25–30.

[6] Xu, M.; Han, Y.; Wang, L.; Zhang, P.; Liu, D.; Yang, J. Research on High-precision Manhole Cover Extraction and Settlement Disease Detection Method Based on Laser Point Cloud. *Chin. J. Lasers* 2021, *48*, 63–77.

[7] Yu, Y.; Guan, H.; Li, D.; Jin, C.; Wang, C.; Li, J. Road Manhole Cover Delineation Using Mobile Laser Scanning Point Cloud Data. *IEEE Geosci. Remote Sens. Lett.* 2020, *17*, 152–156.

[8] Csonthó, M.; Rövid, A.; Szalay, Z. Significance of Image Features in Camera-LiDAR Based Object Detection. *IEEE Access* 2022, *10*, 61034–61045.

[9] Yadav, M.; Lohani, B.; Goel, S. Geometric and radiometric constraints-based extraction of urban road manhole covers and their maintenance-related information using mobile laser scanning data. *Geocarto Int.* 2022, *37*, 16716–16735.

[10] Pasquet, J.; Désert, T.; Bartoli, O.; Chaumont, M.; Delenne, C.; Subsol, G.; Derras, M.; Chahinian, N. Detection of Manhole Covers in High-Resolution Aerial Images of Urban Areas by Combining Two Methods. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2015, *9*, 1802–1807.

10.48047/jocaaa.2023.31.04.71

[11] Liu, H.; Yan, B.; Wang, W.; Li, X.; Guo, Z. Manhole Cover Detection from Natural Scene Based on Imaging Environment Perception. *KSII Trans. Internet Inf. Syst.* 2019, *13*, 5095–5111.

[12] Yamaguchi, T.; Mizutani, T. Detection and localization of manhole and joint covers in radar images by support vector machine and Hough transform. *Autom. Constr.* 2021, *126*, 103651.

[13] Liu, W.; Cheng, D.; Yin, P.; Yang, M.; Li, E.; Xie, M.; Zhang, L. Small Manhole Cover Detection in Remote Sensing Imagery with Deep Convolutional Neural Networks. *ISPRS Int. J. Geo-Inf.* 2019, *8*, 49.

[14] Lin, J.; Yang, C.; Lu, Y.; Cai, Y.; Zhan, H.; Zhang, Z. An Improved Soft-YOLOX for Garbage Quantity Identification. *Mathematic* 2022, *10*, 2650.

[15] Zhang, H.; An, L.; Chu, V.W.; Stow, D.A.; Liu, X.; Ding, Q. Learning Adjustable Reduced Downsampling Network for Small Object Detection in Urban Environments. *Remote Sens.* 2021, *13*, 3608.

[16] Commandré, B.; En-Nejjary, D.; Pibre, L.; Chaumont, M.; Delenne, C.; Chahinian, N. Manhole Cover Localization in Aerial Images with a Deep Learning Approach. In Proceedings of the ISPRS Hannover Workshop HRIGI 17–CMRT 17–ISA 17–EuroCOW 17, Hannover, Germany, 6–9 June 2017; Volume 42, pp. 333–338.

[17] Zhang, H.; Shang, J.; Dong, Z.; He, A.; Zhang, A.A.; Liu, Y.; Lin, Z. Automated Detection of Pavement Manhole on Asphalt Pavements with an Improved YOLOX. *J. Infrastruct. Syst.* 2023, *29*, 04023023.

[18] Zhou, B.; Zhao, W.; Guo, W.; Li, L.; Zhang, D.; Mao, Q.; Li, Q. Smartphone-based road manhole cover detection and classification. *Autom. Constr.* 2022, *140*, 104344.

[19] Jia, G.; Han, G.; Rao, H.; Shu, L. Edge Computing-Based Intelligent Manhole Cover Management System for Smart Cities. *IEEE Internet Things J.* 2018, *5*, 1648–1656.

[20] Peng, J.; Wang, W.; Hu, W.; Ai, C.; Xu, X.; Shi, Y.; Qiu, S. Automated detection and quantification of pavement cracking around manhole. *Eng. Appl. Artif. Intell.* 2024, *130*, 107778.

[21] Zhang, H.-S.; Li, L.; Liu, X. Development and Test of Manhole Cover Monitoring Device Using LoRa and Accelerometer. *IEEE Trans. Instrum. Meas.* 2020, *69*, 2570–2580

[22] Gong, Z.; Liang, L.; Yang, Z.; You, J.; Cai, Y.; Liu, H. A Detection Method of Stealing and Destroying Manhole Cover Based on Sound Signal Processing. *Mach. Tool Hydraul.* 2019, *47*, 39–43.

[23] Lu, Y.; Fang, Q.; Li, M.; Du, X.; Chen, Q.; Dai, Y. Low Power Intelligent Underground Well Monitoring System Based on NB-IoT. *SEA* 2022, *11*, 701–711.

[24] Liu, H.; Yang, S.; Zhang, H.; Xie, R.; Zhou, H. Innovative Design of Manhole Covers and Intelligent Monitoring System Interface for Intelligent Cities. *Design* 2023, *8*, 535–547.

[25] Ganesan, P.; Supraja, N.; Praveena, R. IoT based manhole detection and monitoring system. *AIP Conf. Proc.* 2023, *2725*, 070001.

[26] Miriyala, S.; Rajesh, K.; Bhaskar, P.; Sairam, M.; Prasad, M. Manhole Detection and Monitoring System through IoT. In Proceedings of the 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 6–8 July 2023; Volume 1, pp. 1–5.

10.48047/jocaaa.2023.31.04.71

[27] Dharmadhikari, M.; De Petris, P.; Nguyen, H.; Kulkarni, M.; Khedekar, N.; Alexis, K. Manhole Detection and Traversal for Exploration of Ballast Water Tanks using Micro Aerial Vehicles. *ICUAS* 2023, *1*, 103–109.

[28] Fischler, M.; Bolles, R. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* 1981, *24*, 381–395.

[29] Ester, M.; Kriegel, H.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Comput. Sci.* 1996, *96*, 226–231.

[30] *GB55011-2021*; Project Code for Urban Road and Transportation Engineering. National Standards of P. R. C.: Beijing, China, 2021.