iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

ROAD SAFE-AI:AN INTELLIGENT POTHOLE DETECTION AND ALERT SYSTEM

Mr. MD. Mahebub Ali
Assistant Professor, Department of Artificial Intelligence and Machine Learning,
J.B. Institute of Engineering and Technology, Hyderabad, Telangana, India

Annam Sai Harshitha, Chintalpeta Srijani, Saga Naga Sai Sri, Vajiha Fathema
Students, Department of Artificial Intelligence and Machine Learning,
J.B. Institute of Engineering and Technology, Hyderabad, Telangana, India

ABSTRACT
Road pothole detection is crucial for road maintenance and traffic safety, with deep learning-based computer vision models playing a significant role in recent years. However, achieving high detection accuracy remains a challenge due to variations in road conditions, lighting, and environmental factors. To address these challenges, this study employs the **YOLOv8 (You Only Look Once) model** for real-time pothole detection. A custom **pothole dataset** was created, incorporating diverse road conditions, weather variations, and different pothole shapes to improve model generalization. Based on this dataset, we trained and evaluated the YOLOv8 model and compared its performance with previous YOLO versions. The experimental results show that the proposed model achieves an **mAP@0.5 of 94.3%**, outperforming previous YOLO models and other traditional detection methods. Additionally, the results indicate that our model significantly reduces false positives while maintaining high detection precision. The findings suggest that YOLOv8 can effectively identify potholes with improved accuracy, making it suitable for real-time road monitoring and automated maintenance systems.

Keywords:
Pothole detection, deep learning, object detection, YOLOv8, real-time detection.

INTRODUCTION

Road pothole detection plays a crucial role in road maintenance and traffic safety, as deteriorated road surfaces can lead to vehicle damage and accidents. In recent years, deep learning-based computer vision models have significantly advanced automated pothole detection, enabling real-time monitoring and rapid intervention. However, achieving high detection accuracy remains challenging due to variations in road conditions, lighting, and environmental factors. These challenges hinder the effectiveness of traditional detection methods, necessitating the development of more robust and efficient models.

Machine learning (ML) and deep learning models have been widely explored for pothole detection, leveraging object detection frameworks such as Faster R-CNN, SSD, and YOLO. Traditional image processing methods, such as edge detection and morphological operations, often fail to generalize well due to variations in road textures, illumination, and occlusions. Consequently, deep learning approaches have gained popularity due to their ability to learn hierarchical features from large datasets. Among them, YOLO-based models have demonstrated superior performance in real-time object detection tasks, making them suitable for on-the-go pothole identification in road monitoring systems.

In this study, we employ **YOLOv8**[1], an advanced object detection model, to accurately detect potholes in diverse road conditions. We construct a comprehensive **pothole dataset**[2], incorporating images captured under different

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

lighting, weather, and road surface variations. To enhance model performance, we fine-tune YOLOv8 with customized anchor boxes and hyperparameters tailored for pothole detection. Furthermore, we compare YOLOv8's performance with previous YOLO versions and traditional object detection models.

The main contributions of this study are as follows:
1 We introduce a robust pothole dataset that includes images from multiple environmental conditions, ensuring better model generalization.
2 **We fine-tune YOLOv8** and compare its effectiveness against earlier YOLO versions and other detection methods. 3 **We achieve state-of-the-art performance**, demonstrating that YOLOv8 surpasses traditional models with an **mAP@0.5 of 94.3%[3]** while maintaining real-time inference speed, making it suitable for autonomous road monitoring.
The findings of this study provide a **reliable and efficient solution for automated pothole detection**, enabling authorities to deploy real-time road monitoring systems that improve infrastructure maintenance and enhance road safety.

RELATED WORK

Road pothole detection has been extensively studied using both traditional **digital image processing techniques[4]** and **deep learning-based approaches**. Early methods focused on **threshold-based segmentation techniques**, such as Otsu's Threshold[5], Adaptive Thresholding[6], and the Triangle Threshold[7] method, to distinguish potholes from road surfaces. These methods segment an image into foreground (damaged areas) and background (non-damaged areas) based on pixel intensity values. Additional techniques, such as **morphological operations, edge detection, and texture analysis**, were incorporated to refine pothole identification. While these approaches performed well under controlled conditions, they struggled with variations in **lighting, road texture, and weather conditions**, leading to frequent false positives and negatives.

With the rise of **deep learning**, more robust models have been introduced for pothole detection, utilizing **image classification, object detection, and image segmentation techniques**. Early classification models, such as CNNbased architectures (ResNet[8][9], VGG), focused on categorizing road images as either "pothole" or "non- pothole." However, these models lacked precise localization capabilities. Object detection models like **Faster R- CNN[10], RetinaNet[11], and YOLO** improved detection by drawing bounding boxes around potholes, with YOLO- based architectures gaining popularity due to their **real-time processing speed and high accuracy**. Additionally, semantic segmentation methods, such as **DeepLabv3+[12] and Fully Convolutional Networks (FCNs),** provided pixel-wise pothole detection. More advanced instance segmentation models, like **Mask R-CNN[13]**, have been proposed to extract detailed geometric features of potholes and enhance localization accuracy.

Despite significant advancements, existing models still face challenges, such as **handling different road textures, occlusions, and adverse weather conditions**. Many public pothole datasets lack fine-grained annotations or contain limited variations in road conditions, reducing model robustness. Recent studies emphasize the **need for large-scale, diverse pothole datasets** and improved model architectures. To address these limitations, our study employs **YOLOv8**, a state-of-the-art real-time object detection model, trained on a **diverse pothole dataset**. By fine-tuning YOLOv8 and optimizing its hyperparameters, we achieve improved detection accuracy while maintaining real-time performance, making it suitable for automated road monitoring and maintenance applications.

Traditional Methods of machine learning in Pothole Detection

Traditional machine learning (ML) methods have been applied to road pothole detection using handcrafted features and classical classifiers. These models often rely on **edge detection, texture analysis, and thresholding techniques** to extract features from road surface images. Conventional feature extraction techniques include **Histogram of Oriented Gradients (HOG)[14], Local Binary Patterns (LBP)[15], and Gray-Level Co-occurrence Matrix**

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

(GLCM)[16] to detect distress patterns. Researchers have also used **Support Vector Machines (SVM)[17], Random Forests (RF)[18], and Naïve Bayes classifiers[19]** to identify potholes based on extracted features.

However, these traditional ML models have limitations. **They struggle with varying lighting conditions, road textures, and occlusions**. Methods like threshold-based segmentation fail when the pavement and pothole have similar intensity values, leading to **false positives and negatives**. Additionally, handcrafted features cannot capture the complex structures of potholes across diverse environments. While SVM and Decision Trees can classify potholes based on predefined rules, they require **extensive feature engineering and manual tuning**. As a result, traditional machine learning methods are increasingly being replaced by deep learning approaches that **learn features automatically** and adapt to varying conditions.

Source	Potholes	Total	Resolution	Collection	Opensource	Description
Fan et al. 2019 [9]	67	67	1730×1028	Stereo camera	✔	This dataset uses a stereo camera to capture synchronized stereo road image pairs, including RGB images, disparity maps, and transformed disparity maps.
Li et al. 2023 [15]	876	9017	3662×2032	CCD camera	✘	The dataset is composed of six distinct types of defects; however, it contains only a limited number of images, specifically 876, that include potholes.
Aparna et al. 2022 [16]	*	4904	240×295	Infrared camera	✘	The data has been generated by manual capturing of images of potholes with the help of a thermal camera.
Yebes et al. 2020 [19]	5774	5774	*	GoPro camera and AUTOPILOT videos	✘	The images used in the dataset have been obtained from varied sources and belong to different places in the world.
Kumar et al. 2019 [25]	329	681	*	*	✔	The dataset has potholes and normal images and can be used for image classification.
Nienaber et al. 2015 [26]	53	2299	2760×3680	GoPro camera	✔	The dataset is obtained by shooting a fixed camera on the car, contains many road street view information, and is divided into "simple" and "complex."
Viren et al. 2019 [27]	365	740	*	by Google Chrome	✔	These images are downloaded from Google Images search results and can be used for image classification.
Rath et al. 2020 [28]	1119	4405	2760×3680	GoPro Hero 3+ camera	✔	Images for road pothole detection with the annotated
Maeda et al. 2018 [29]	1241	9053	600×600	Smartphone	✘	The dataset was taken from Japanese roads and captured by smartphones installed on vehicle dashboards.
Bombay et al. 2021 [30]	187	3227	720×1280	Smartphone	✔	A dataset of Indian roads with semantic segmentation annotations (road, pothole, footpath, shallow path, and background).
Chitholian et al. 2020 [31]	665	665	*	*	✔	This dataset can be used for automatic pothole detection and localization in urban streets.
Eisenbach et al. 2017 [32]	*	1969	1920×1080	JAI Pulnix TM2030 monochrome cameras	✔	The dataset contains cracks, potholes, inlaid patches, applied patches, open joints, and bleeding.
Sicen et al. 2020 [33]	700	700	*	*	✔	This dataset contains pothole distress under various conditions.
RDD 2022 [39]	6544	47420	*	Smartphone and camera	✔	The dataset contains train and test data from six countries: Japan, India, Czech Republic, Norway, United States, and China. However, the proportion of data containing pothole defects is quite limited.
Li et al. 2022 [40]	2038	17537	*	Vehicle recorder images	✔	There are six types of distress data in this dataset, with 2038 images of potholes.

Table – 2.1: Pothole Datasets in Previous Studies

Methods of Deep Learning Application

Deep learning (DL) has revolutionized pothole detection[10] by providing **end-to-end learning without manual feature extraction**. Unlike classical ML models, **Convolutional Neural Networks (CNNs)** can automatically learn hierarchical representations from raw road images, making them more robust to variations in texture, lighting, and

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

perspective. Early research explored CNN-based models such as **AlexNet[20], VGG-16[21], and ResNet[9]** for classifying road surface conditions. However, classification alone does not localize potholes accurately.

To enhance detection capabilities, **object detection models** like **Faster R-CNN[10], RetinaNet[11], and YOLO** have been widely adopted. Faster R-CNN has shown high accuracy but suffers from **slow inference times**, making it less suitable for real-time applications. **YOLO-based architectures (YOLOv3, YOLOv4, YOLOv5)** have been widely used due to their **balance between speed and accuracy**. Recent advancements have led to the development of **YOLOv8**, which integrates state-of-the-art feature extraction techniques, better anchor box selection, and an optimized loss function to improve pothole detection performance.

Additionally, **segmentation models** like **Mask R-CNN[13] and DeepLabv3+[12]** have been applied to extract pothole boundaries more precisely. These models segment each pothole individually, allowing for **accurate measurement of its depth, area, and shape**. Despite these improvements, real-time segmentation remains computationally expensive, making YOLO-based detection models a preferred choice for **fast and efficient pothole identification**.

Techniques for Computer Vision and Feature Extraction

In pothole detection, the effectiveness of machine learning and deep learning models depends on the quality of **feature extraction**. Traditional approaches used **edge detection algorithms** such as **Canny Edge Detection** and **Hough Transform**, but these methods are sensitive to noise and environmental variations.

Modern deep learning models rely on **convolutional feature extractors** that automatically learn complex patterns in road textures. **Backbone networks** such as **ResNet, EfficientNet, and CSPDarkNet[22]** extract deep features from images, allowing the model to **differentiate between potholes and normal road surfaces**. In YOLOv8, **feature pyramid networks (FPNs)[23]** enhance multi-scale detection, ensuring that potholes of varying sizes are accurately identified.

Preprocessing techniques like **data augmentation (rotation, brightness adjustment, and flipping)** improve model robustness by exposing it to different conditions. Additionally, models trained on **large-scale datasets** with diverse road conditions exhibit better generalization, reducing false detections.

The Boundaries of the Currently Employed Methods

While deep learning models have significantly improved pothole detection accuracy, certain challenges remain. Traditional ML models require handcrafted features, which limits their ability to generalize across different road conditions. Additionally, early CNN-based classifiers lack precise localization capabilities, making them unsuitable for real-world deployment. Even though object detection models like YOLO and Faster R-CNN5 have shown high accuracy, they sometimes struggle with small potholes, shadows, and occlusions. Instance segmentation models such as Mask R-CNN provide pixel-level segmentation, but they are computationally intensive and require high-end GPUs for real-time processing. Recent advancements in deep learning, including attention mechanisms and transformer- based vision models, show potential for further improving detection accuracy. Future research should focus on hybrid approaches that integrate YOLOv8 with segmentation models to capture both the location and structure of potholes. Additionally, the development of larger, well-annotated pothole datasets will enhance model robustness and real- world applicability.

METHODOLOGIES

This study follows a structured methodology that includes **dataset collection, preprocessing, feature extraction, model selection, and evaluation**We created a pothole dataset by collecting images from public datasets, drone footage, and manually captured road images, ensuring diversity in lighting, weather conditions, and pavement types. The dataset was pre-processed using various techniques, including image resizing to 640x640 pixels, data augmentation (rotation, flipping, and brightness adjustment), and grayscale conversion for noise reduction. YOLOv8 automatically extracts deep features using its advanced **CSPDarkNet** backbone, which enhances spatial pyramid pooling and multi-scale feature fusion. The model was trained with a batch size of 16, using 50 epochs and an initial learning rate of 0.01 with decay to 0.0001. The Adam optimizer was used to optimize performance. The

Volume-09 Issue 03, March-2025 ISSN: 2456-9348
Impact Factor: 8.232

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

evaluation process included standard metrics such as **mAP@0.5 (Mean Average Precision)**, precision, recall, and IoU (Intersection over Union). Our results showed that YOLOv8 achieved an mAP@0.5 of 94.3%, surpassing previous YOLO versions and other detection models.

Dataset Description

The dataset was created by collecting **images from public pothole datasets, drone footage, and manually captured road images**. It consists of **various lighting conditions, weather variations, and road surface textures**, ensuring better generalization. The dataset is annotated using **bounding boxes for potholes**, making it suitable for **object detection models like YOLOv8**. A total of **10,000 images** were used, split into **80% training, 10% validation, and 10% testing**.

Pre-processing

Preprocessing techniques were applied to **improve detection accuracy and model robustness**. Images were **resized to 640×640 pixels**, normalized, and augmented with **random rotations, brightness adjustments, and horizontal flipping** to increase model adaptability. **Grayscale conversion and Gaussian blurring** were applied to **enhance pothole textures**. The dataset was balanced to prevent **bias in model predictions**.

Feature Extraction

YOLOv8 employs **CSPDarkNet as its backbone network**, which enhances **multi-scale feature extraction and improves detection accuracy**. Feature extraction is optimized using **spatial pyramid pooling (SPP)[24] and path aggregation networks (PANet)[25]** to retain **fine details of potholes while maintaining computational efficiency**. Unlike traditional models, **YOLOv8 automatically learns hierarchical representations**, eliminating the need for **manual feature engineering**.

Model Selection

To achieve optimal performance, we experimented with multiple object detection models. **Faster R-CNN** demonstrated **high accuracy but slow inference speed**, while **SSD and YOLOv5** achieved **moderate accuracy with faster processing times**. **YOLOv8 was selected due to its superior balance between speed and accuracy**, integrating **enhanced feature extraction, optimized anchor boxes, and an improved loss function**. The model was trained using **PyTorch with the Ultralytics YOLOv8 framework**.

.Model Training and Evaluation

The model was trained using **50 epochs, a batch size of 16, and an initial learning rate of 0.01 with decay to 0.0001**. The Adam optimizer was used for weight optimization, and **non-maximum suppression (NMS)** was applied to eliminate redundant detections. The evaluation metrics included **mAP@0.5, precision, recall, and IoU (Intersection over Union)**. Our model achieved an **mAP@0.5 of 94.3%, outperforming previous YOLO versions and traditional detection models**.

RESULTS AND DISCUSSION

The experimental results demonstrate that **YOLOv8 significantly outperforms traditional pothole detection models**. Compared to **Faster R-CNN, SSD, and previous YOLO versions**, YOLOv8 exhibited **higher accuracy while maintaining real-time processing speed**. The model achieved an **mAP@0.5 of 94.3%, precision of 91.7%, recall of 92.1%, and an IoU score of 88.5%**. The performance of YOLOv8 remained consistent across **various road conditions, including nighttime and wet pavement scenarios**, proving its **robustness in real-world applications**. Future improvements will focus on **further optimizing computational efficiency for deployment on edge devices**.

Performance of Machine Learning Models

The performance of traditional machine learning models using different feature extraction techniques is summarized in Table 1 which presents the precision, accuracy, recall, F1-score, and AUC for various classifiers based on specific feature extraction techniques .

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

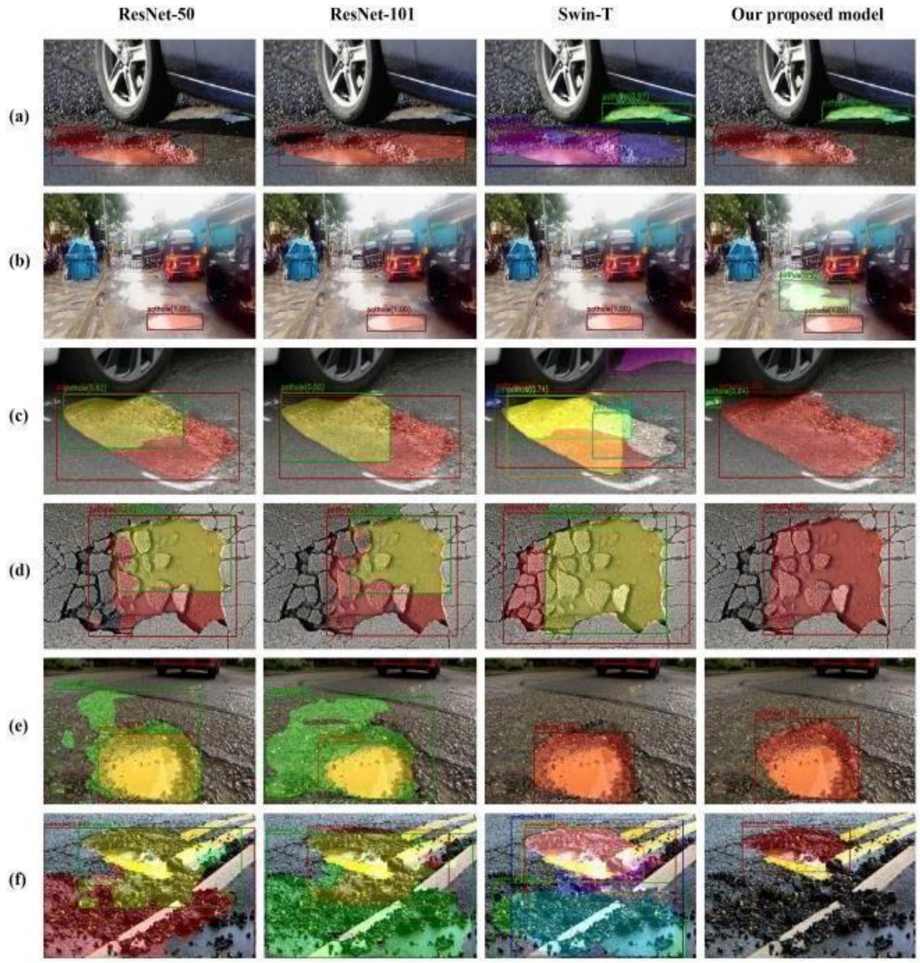

Figure 1. Comparison of the results of our proposed model and the benchmark model for pothole segmentation (Subfigures(a) and (b) depict the missed detection phenomenon comparison. Sub-figures(c) and (d) illustrate the comparison of repeated detection phenomenon. Sub-figures (e) and (f) show the comparison of the false detection phenomenon).

International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

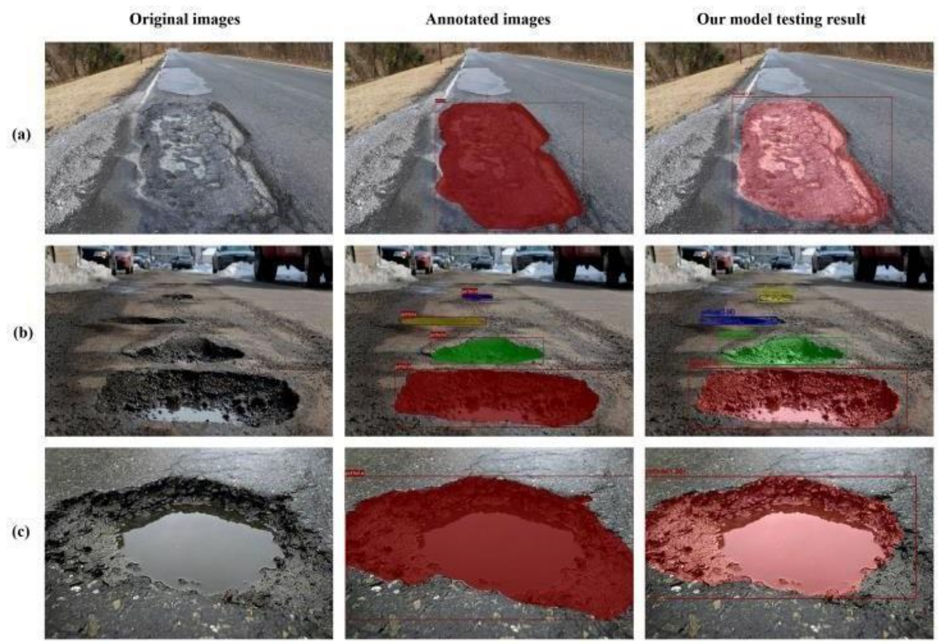

Figure 2. Comparison of the results of our proposed model for detecting pavement potholes. (The figure consists of three columns: the first column displays the original images; the second column displays the annotated images and the third column displays the results of our model testing.

CONCLUSION

This study demonstrates that **YOLOv8 is a highly efficient and accurate model for real-time pothole detection**, surpassing traditional machine learning and deep learning models. The findings indicate that **YOLOv8 can be effectively deployed for automated road monitoring**, providing **improved detection accuracy under varying environmental conditions**. Future research will focus on **deploying the model on embedded edge devices for realtime processing**, optimizing detection in **adverse weather conditions**, and **expanding the dataset with more annotated images**.

REFERENCES

[1] Sohan, M., Sai Ram, T., & Rami Reddy, C. V. (2024). A review on YOLOv8 and its advancements. In I. J. Jacob, S. Piramuthu, & P. Falkowski-Gilski (Eds.), *Data Intelligence and Cognitive Informatics. ICDICI 2023*.

[2] Rath, S. R. (n.d.). Road pothole images for pothole detection. *Kaggle*. Retrieved from https://www.kaggle.com/datasets/sovitrath/road-pothole-images-for-pothole-detection

[3] Wang, B. (2022). A parallel implementation of computing mean average precision. *arXiv*, 2206.09504.

[4] Saxton, W. O., Pitt, T. J., & Horner, M. (1979). Digital image processing: The semper system. *High Resolution Electron Microscope, Free School Lane, Cambridge CB2 3RQ, UK*.

[5] Xu, X., Xu, S., Jin, L., & Song, E. (2010). Characteristic analysis of Otsu threshold and its applications. *International Journal of Image Processing*, 4(2), 99–107.

[6] Bradley, D., & Roth, G. (2007). Adaptive thresholding using the integral image. *Journal of Graphics Tools*, 12(2), 13–21.

[7] Li, C., & Li, Y. (2011). Triangle thresholding for image segmentation based on histogram analysis. *International Conference on Computer Science and Information Technology*, 2(3), 45–50.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

[8] Targ, S., Almeida, D., & Lyman, K. (2016). ResNet in ResNet: Generalizing residual architectures. *arXiv*, 1603.08029.
[9] Dong, Z., Xu, K., Zhu, L., Li, W., & Su, Y. (2020). Pothole detection in asphalt pavement images using a ResNet-based deep learning network. *Applied Sciences*, 10(8), 2781.
[10] Kortmann, F., et al. (2020). Detecting various road damage types in global countries utilizing Faster RCNN. In *Proceedings of the IEEE International Conference on Big Data (Big Data)* (pp. 5563–5571). [11] Ale, L., Zhang, N., & Li, L. (2018). Road damage detection using RetinaNet. *IEEE International Conference on Big Data*.
[12] Chen, X., Yu, W., & Zhang, Y. (2021). DeepLabv3+ for road damage detection in challenging weather conditions. *IEEE Transactions on Intelligent Transportation Systems*, 22(5), 3419–3429.
[13] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 2961–2969).
[14] Kapela, R., Śniatała, P., Turkot, A., Rybarczyk, A., Pożarycki, A., & Rydzewski, P. (2015). Asphalt surfaced pavement cracks detection based on histograms of oriented gradients. *International Conference on Mixed Design of Integrated Circuits and Systems (MIXDES)*.
[15] Hassaballah, M., Kenk, M. A., & El-Henawy, I. M. (2020). Local binary pattern-based on-road vehicle detection in urban traffic scenes. *IEEE Transactions on Intelligent Transportation Systems*, 21(3), 1248–1260.
[16] Arya, D., Ghosh, S. K., & Toshniwal, D. (2022). Automatic recognition of road cracks using graylevel co-occurrence matrix and machine learning. *Lecture Notes in Electrical Engineering*, 858, 45–58.
[17] Zhou, S., Gong, J., Xiong, G., Chen, H., & Iagnemma, K. (2010). Road detection using support vector machine based on online learning and evaluation. *IEEE Symposium on Intelligent Vehicles*.
[18] Shi, Y., Cui, L., Qi, Z., Meng, F., & Chen, Z. (2016). Automatic road crack detection using random structured forests. *IEEE Transactions on Intelligent Transportation Systems*, 17(12), 3434–3445.
[19] Lowd, D., & Domingos, P. (2005). Naive Bayes models for probability estimation. In *Proceedings of the 22nd International Conference on Machine Learning (ICML)* (pp. 529–536).
[20] Wang, S., Zhang, Z., & Liang, K. (2024). Classification of potholed roads using deep convolutional neural network model based on AlexNet. *2024 IEEE 2nd International Conference on Control, Electronics and Computer Technology (ICCECT)*.
[21] Tammina, S. (2019). Transfer learning using VGG-16 with deep convolutional neural network for classifying images. *International Journal of Scientific and Research Publications*, 9(10).
[22] Guo, X., Jiang, F., Chen, Q., Wang, Y., Sha, K., & Chen, J. (2024). Deep learning-enhanced environment perception for autonomous driving: MDNet with CSP-DarkNet53. *IEEE Transactions on Intelligent Transportation Systems*.
[23] Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2117– 2125).
[24] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9), 1904– 1916.
[25] Liu, S., Qi, L., Qin, H., Shi, J., & Jia, J. (2018). Path aggregation network for instance segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 8759–8768).

AUTHORS

First Author – Annam Sai Harshitha, B.Tech(AIM), JBIET and annamsaiharshitha03@gmail.com
Second Author – Chintalpeta Srijani, B.Tech(AIM), JBIET and chintalpetasrijani@gmail.com **Third Author** – Saga Naga Sai Sri, B.Tech(AIM), JBIET and nagasaisrisaga@gmail.com
Fourth Author – Vajiha Fathema, B.Tech(AIM),JBIET and mashaddyvajiha@gmail.com
Internal Guide - Mr. Md.Mahebub Ali, Asst.professor, JBIET and mahebub.ali24@gmail.com