International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Research Article)

Check for updates

A hybrid approach to detect and classify pothole on Bangladeshi roads using deep learning

Shafi Ullah Adid *, Md. Emon and Taofica Amrine

Department of Computer Science and Engineering, Port City International University, Chittagong - 4000, Bangladesh.

International Journal of Science and Research Archive, 2024, 12(01), 1045–1053

Publication history: Received on 18 April 2024 revised on 24 May 2024; accepted on 27 May 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.12.1.0950

Abstract

Potholes are a major problem for Bangladesh, a country with a developing economy and infrastructure. Traditional pothole detection methods, often based on manual inspection, are insufficient for effectively managing the vast national road network. The primary focus of this study is on pothole detection on roads in Bangladesh, employing deep learning techniques. The pothole dataset, comprising images captured by us, has been curated, leading to the development of two distinct datasets: one for pothole classification, totaling 6000 images, and another for pothole detection, comprising 1300 images. To improve the diversity of the datasets, the preprocessing procedure includes motion blur reduction and optical distortion correction. Our methodology employs state-of-the-art deep learning models, including Convolutional Neural Network (CNN), ResNet101, ResNet50, VGG16, DenseNet201, DenseNet169, InceptionResNetV2, MobileNetV2, EfficientNetB0, and a combined CNN along with ResNet101 model. The CNN and ResNet101 combination exhibit an exceptional accuracy of 97.89%. YoloV7 and YoloV8 demonstrate promising accuracy levels in pothole detection tasks, achieving 0.709 and 0.959, respectively. The datasets and models developed in this study have significant implications for improving road safety in Bangladesh. The proposed hybrid of CNN and pretrained ResNet101 exhibits the best performance than all other classification models applied in this study and the YoloV8 performed better than YoloV7 in the pothole detection task.

Keyword: Pothole; Bangladeshi Road; Optical Distortion Correction; Convolutional Neural Network; YoloV8**.**

1. Introduction

Bangladesh, a nation with a growing economy and rapidly expanding infrastructure, faces a significant challenge: the persistent presence of potholes on its roads. These ubiquitous obstacles not only pose a safety hazard to motorists and pedestrians but also contribute to vehicle damage and economic losses. Potholes can cause drivers to lose control of their vehicles, leading to accidents, injuries, and even fatalities. This risk is particularly high for motorcyclists and cyclists. Driving through potholes can damage tires, rims, suspension systems, and other vehicle parts, leading to costly repairs. It forces drivers to navigate cautiously, significantly slowing down traffic flow and increasing travel time. Frequent repairs and patching due to potholes degrade the overall quality of the roads, leading to a shorter lifespan. Several factors contribute to the prevalence of potholes in Bangladesh: The high volume of vehicles, especially heavy trucks, puts excessive strain on the roads, accelerating wear and tear, Heavy rainfall weakens the road base and washes away soil, further accelerating pothole formation. While efforts are underway, potholes remain a significant challenge in Bangladesh.

Deep learning has demonstrated remarkable capabilities in image and video analysis, making them well-suited for automated pothole detection tasks. Pothole detection can be accomplished with accuracy and efficiency by utilizing deep learning algorithms, which are highly proficient in tasks related to pattern recognition and image processing. Deep learning models can be trained on vast amounts of data, allowing them to adapt to the diverse road conditions and

* Corresponding author: Shafi Ullah Adid

environments prevalent in Bangladesh. This includes variations in road types, weather patterns, and even the unique characteristics of potholes in the region. This adaptability ensures the effectiveness of the system across different road networks. Deep Learning can effectively identify potholes even in challenging conditions like varying lighting, shadows, and road textures, surpassing the limitations of simpler methods like thresholding or edge detection. The application of deep learning in pothole detection holds significant promise for enhancing road safety, supporting infrastructure development, and fostering innovation and collaboration.

We aim to develop a robust and scalable AI model that can automatically identify and localize potholes in images and videos. This system will not only help to enhance road safety by providing timely alerts to authorities but also optimize road maintenance efforts by pinpointing areas requiring repair. In addition, the study aims to investigate the feasibility and effectiveness of applying cutting-edge technologies in considering Bangladesh's infrastructure difficulties, with the ultimate objective of enhancing road quality, lowering accident rates, and maximizing resource distribution for road maintenance initiatives. The goal of this research is to support Bangladesh's efforts to build resilient and sustainable transportation systems, which will ultimately help the people and economy of the nation.

2. Literature Review

In 2020, Arjapure [1] et al. addresses the problem of manual classification and evaluation of potholes on roads, which is time-consuming and slows down the road maintenance process. They propose a method for automatic classification and detection of potholes using convolutional neural networks (CNNs) on road images. The method is implemented in Python using OpenCV library and tested on 722 raw images. The results show that pre-trained models InceptionResNetV2 and DenseNet201 can detect potholes with reasonable accuracy of 89.66%. In 2022, Asad [2] et al. explore the potential of various state-of-the-art deep learning models and object detection frameworks, including YOLOv1, YOLOv2, YOLOv3, YOLOv4, Tiny-YOLOv4, YOLOv5, and SSD-mobilenetv2, for pothole detection on a single-board computer (Raspberry Pi) using the AI kit (OAK-D) as an edge platform. The authors perform experiments on both an image dataset with potholes in diverse road conditions and illumination variations, as well as on real-time video captured through a moving vehicle. The results show that Tiny-YOLOv4, YOLOv4, and YOLOv5 achieve the highest mean average precision (mAP) of 80.04%, 85.48%, and 95%, respectively, on the image set, demonstrating the effectiveness of the proposed approach for pothole detection. In 2021, Dewangan [3] et al. proposes an improved prototype model for pothole detection and intelligent driving behavior in autonomous vehicles. Existing approaches have limitations in detecting potholes due to their non-uniform structure and dynamic road environments. The proposed model uses a convolutional neural network with a vision camera to detect potholes and validate its driving behavior in a prepared road environment. The experimental analysis shows that the model achieves high accuracy, sensitivity, and F-measure values of 99.02%, 99.03%, and 98.33%, respectively, which are comparable to state-of-the-art techniques. In 2020, Chen [4] et al. addresses the challenge of pothole detection in road images, which is crucial for pavement maintenance and rehabilitation. The authors propose a new method based on location-aware convolutional neural networks (CNNs) that focuses on discriminative regions in the road rather than the global context. The method comprises two subnetworks: a localization subnetwork that identifies candidate regions with high recall, and a part-based subnetwork that classifies the candidates. The approach achieves high precision (95.2%) and recall (92.0%) simultaneously and outperforms existing methods. The results demonstrate that accurate part localization significantly improves classification performance while maintaining high computational efficiency. In 2021, Patra [5] et al. describes the development of an end-to-end system called PotSpot for real-time detection, monitoring, and spatial mapping of potholes on roads. The system uses a Convolutional Neural Network (CNN) model to detect potholes and generates real-time pothole-marked maps using the Google Maps API. An Android application integrates both the pothole detection and mapping components. The proposed model is compared to six baseline methods, including artificial neural networks, support vector machines, and pre-trained CNN models. The results show that the proposed model achieves better accuracy (approximately 97.6%) and a higher Area Under the Curve (AUC) value (0.97) than the baseline methods. Overall, the PotSpot system provides an effective solution for real-time pothole detection and mapping, which can help improve road maintenance and safety. In 2020, Ping [6] et al. proposes an efficient pothole detection system using deep learning algorithms, specifically YOLO V3, SSD, HOG with SVM, and Faster R-CNN. The system is designed to detect potholes on roads automatically, addressing the problem of severe traffic accidents and reduced road efficiency caused by potholes. The authors train and test four models using a preprocessed dataset and compare their accuracy and loss. The results show that YOLO V3 performs best, providing faster and more reliable detection results. In 2021, Bibi [7] et al. discusses the importance of detecting road surface defects for safe and smooth traffic flow, particularly in the context of autonomous vehicles. The authors propose a novel mechanism for automatic detection of road anomalies using Edge AI and Vehicular Ad Hoc Network (VANET). They use a combination of sensors and Deep Neural Network (DNN) techniques to enable autonomous vehicles to perceive their surroundings and detect tracks and obstacles. The proposed mechanism utilizes road images captured via camera and deploys a trained model for road anomaly detection in a vehicle. In 2021, Arjapure [8] et al. proposes a deep learning technique, Mask Region-Based Convolutional Neural

Network (Mask RCNN), to detect and segment potholes on roads. The dataset used consists of 291 images collected manually on local roads and highways in Mumbai, and the annotations were created using the VGG Image Annotator tool. The proposed method detects potholes as a region of interest and computes their area based on the generated mask. The computed area is then compared with the actual measured area, showing an overall accuracy of 90% with ± 10% deviation. The results are promising, and the information extracted can be used for cost estimation in road repair management. In 2021, Bhavya [9] et al. proposes a deep learning-based model for automating the process of pothole detection on roads. The model uses a dataset of images of roads and classifies them as either plain roads or roads with potholes. The authors claim that the model can accurately detect potholes and replace the need for manual inspections, which require a lot of manpower and time. The model is intended to assist the government in maintaining roads more efficiently and effectively, reducing the need for manual labor and improving road conditions. In 2021, Rohitaa [10] et al. proposes an automatic pothole identification system using deep learning algorithms to address the challenges of manual road assessment. The system uses a dataset to train and test three deep learning models: CNN, Mask RCNN, and YOLOv3. The performance of the models is compared using evaluation metrics. The system also includes hardware components for reporting potholes and warning drivers of their presence. The proposed system aims to improve road maintenance and reduce the likelihood of accidents caused by potholes.

3. Methodology

The proposed approaches will be covered in detail in this section. Figure 1 illustrates the overall methodology of our proposed pothole classification and detection system.

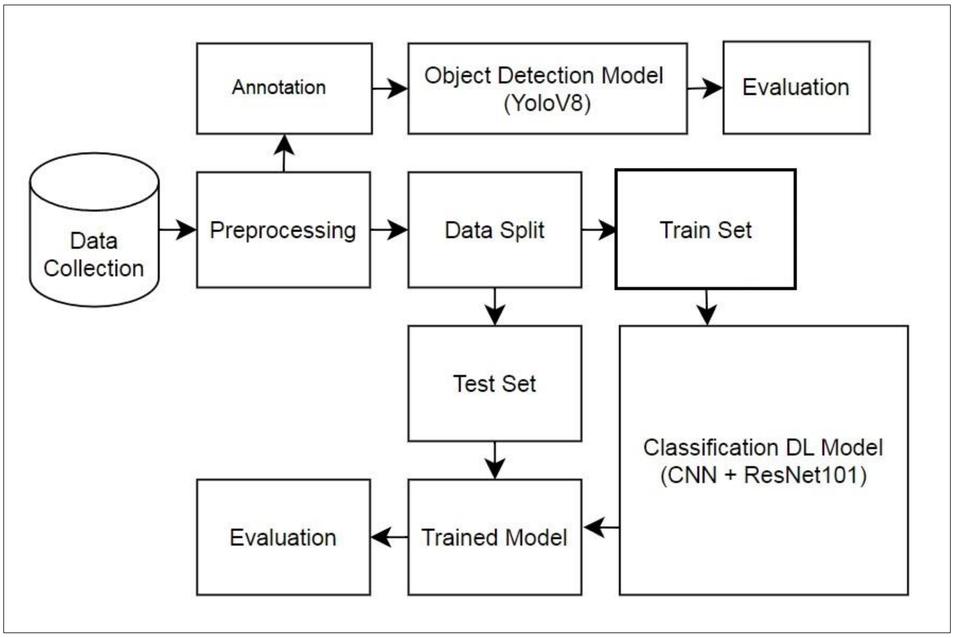

Figure 1 Proposed Methodology of Pothole Detection and Classification.

3.1. Dataset Collection

As no public dataset of potholes in Bangladeshi roads was available, we undertook the creation of our own dataset tailored to our research objectives. The dataset creation process involved two distinct stages to support the development of both the classification and object detection models.

For the training of the classification model, we curated a dataset comprising 6000 images. These images were captured using mobile phone cameras and encompassed various road conditions within Chittagong City, Bangladesh. Care was taken to capture images in diverse lighting conditions to ensure the robustness and generalizability of the classification model. In addition to the classification dataset, we also assembled a separate dataset consisting of 1300 images with annotation specifically designed for training the object detection model. Similar to the classification dataset, these images were acquired using mobile phone cameras and were representative of the road conditions prevalent in Chittagong City. Table 1 shows the data distribution of classification dataset.

Table 1 Data Distribution for Classification Task.

Class	Image Count	Total
Pothole	3000	6000
No Pothole	3000	

3.2. Image Pre-processing

Preparing image data for deep learning tasks, including our research, requires a vital step called image data pre-processing. This phase entails several operations to improve the quality and relevance of the images in the dataset.

3.2.1. Image Resizing

Every image in the object detection dataset has been stretched to 512 x 512, and each one in the classification dataset has been resized to 224 x 224. It will speed up the processing process and significantly decrease the model's performance.

3.2.2. Data Augmentation

Data augmentation has been applied to introduce variations in the dataset. This process involves making adjustments to the existing photos, which may include flipping, rotating, enlarging, or altering the brightness. By diversifying the dataset, augmentation helps prevent overfitting and enhances the model's ability to generalize to various situations.

3.2.3. Annotation

We marked or labeled specific areas or features within the images to indicate the presence of potholes. This annotation process helped us train the deep learning object detection models to detect potholes accurately.

3.2.4. Rescale

We adjusted the size of the images to a uniform scale. This resizing ensured consistency in the dimensions of the images, which is important for training the models effectively and efficiently.

3.2.5. Denoising

We removed any unwanted or distracting noise from the images. This could include blurriness or graininess caused by factors such as poor lighting conditions or camera imperfections [8].

3.2.6. Gray scaling

We converted the images from color to grayscale. This simplified the images by representing them in shades of gray instead of full color [13].

3.2.7. Contrast Enhancement

We adjusted the contrast of the images to make the differences between light and dark areas more pronounced. This enhancement improved the visibility of details within the images [14].

3.3. Dataset Splitting

We have experimented with the ResNet101 model on datasets with varying split ratios in order to determine the optimal split ratio. The accuracy score for each split ratio in this experiment is displayed in Table 2. Since the 70:30 split shows the best accuracy in this testing, we have chosen it.

Table 2 Accuracy Scores of Each Split of Dataset.

Ratio (Train: Test)	70:30	80:20	90:10
Accuracy	**95.11**	94.00	94.83

3.4. Deep Learning Model for Pothole Classification

3.4.1. Convolutional Neural Network – CNN

A Convolutional Neural Network (CNN) [11] is a deep learning model specifically designed for processing structured grid data, such as images. CNNs are composed of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The key innovation of CNNs lies in the convolutional layers, which consist of filters that convolve over input images to extract local features. These filters learn to detect patterns such as edges, textures, and shapes, capturing hierarchical representations of the input data. Pooling layers are then used to down sample the feature maps, reducing computational complexity and increasing translation invariance. Their ability to automatically learn hierarchical representations of visual data makes them highly effective for tasks requiring understanding and interpretation of complex visual information.

3.4.2. Residual Network 101 – ResNet101

ResNet101 [12] is a deep convolutional neural network architecture that is widely used for various computer vision tasks, including image classification. It is part of the ResNet (Residual Network) family of models developed by Microsoft Research. The "101" in ResNet101 refers to the depth of the network, specifically indicating that it has 101 layers. One of the key innovations introduced by ResNet is the concept of residual connections, which enable the training of very deep networks by addressing the problem of vanishing gradients. In this study, ResNet101 is employed to analyze road images from Bangladeshi roads, aiming to identify and classify potholes accurately.

3.4.3. Hybrid of CNN and ResNet101

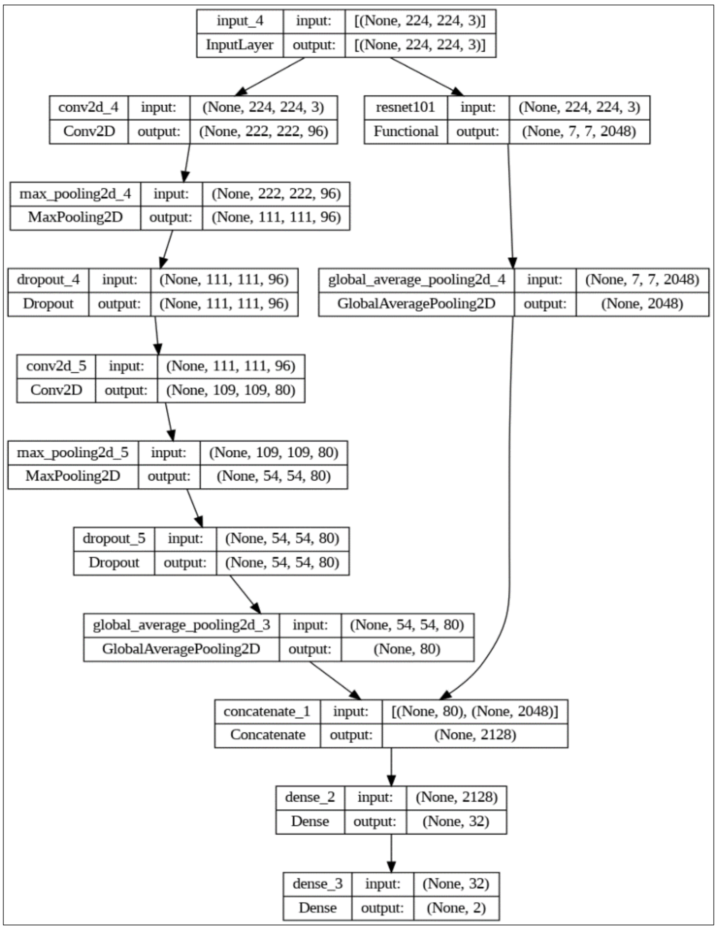

Figure 2 Architecture of Proposed Hybrid Model.

We have innovatively fused the power of CNN and ResNet101 models to create a hybrid architecture. CNNs are renowned for their ability to automatically learn hierarchical features from images, while ResNet101, an extension of the ResNet architecture, introduces skip connections to address the challenges associated with training very deep networks. By combining these two architectures, we aim to harness the feature extraction capabilities of CNNs and the gradient flow optimization of ResNet101. In such a hybrid model, the initial layers consist of standard convolutional and pooling layers from a CNN, allowing the network to capture low-level features effectively. The subsequent layers could incorporate residual blocks inspired by ResNet101, incorporating skip connections to facilitate the flow of gradients during training.

The Figure 2 shows that the hybrid model combines a convolutional neural network with a pre-trained ResNet101 model. The CNN consists of convolutional layers with max-pooling and dropout, followed by a global average pooling layer. The ResNet101 model is initialized with weights from ImageNet and includes a global average pooling layer. The outputs of both the custom CNN and the ResNet101 model are concatenated, creating a fused feature representation. The combined features are then passed through dense layers for further processing, with the final output layer having two units and a sigmoid activation function, indicating a binary classification task. The model is compiled using the Adam optimizer with a low learning rate, binary cross-entropy loss, and accuracy as the evaluation metric.

3.5. Deep Learning Model for Pothole Detection

3.5.1. You Only Look Once version 8 – Yolov8

YOLOv8 is the latest star in the YOLO family of real-time object detection models. It pushes the boundaries of both speed and accuracy, making it a top choice for various computer vision tasks. YOLOv8 is user-friendly, adapting to different hardware and offering a Python package and command line interface for easy implementation. YOLOv8 employs a deep neural network architecture, making it adept at detecting and classifying objects with high efficiency. Its versatility and performance have led to its widespread adoption in various applications, including surveillance, autonomous vehicles, and general object recognition tasks.

4. Results and discussion

Table 3 shows the classification report of each model that we applied on our study. The results showcase the effectiveness of deep learning models, which exhibit high precision, recall, and F1-Score. Notably, the CNN+ResNet101 combination achieves exceptional performance with a precision of 0.98 across all metrics, while YoloV8 demonstrates strong results with a precision of 0.959 in the case of the detection scenario.

Table 3 Model Performance Summary**.**

Task	Model	Precision	Recall	F1-Score
Pothole Classification	CNN	0.97	0.97	0.94
	ResNet101	0.95	0.95	0.95
	ResNet50	0.96	0.96	0.96
	VGG16	0.91	0.91	0.91
	DenseNet201	0.95	0.95	0.95
	DenseNet169	0.89	0.89	0.89
	InceptionResNetV2	0.94	0.94	0.94
	MobileNetV2	0.88	0.88	0.88
	EfficientNetB0	0.91	0.91	0.91
	CNN+ResNet101	0.98	0.98	0.98
Pothole Detection	YoloV7	0.709	0.709	0.709
	YoloV8	0.959	0.959	0.959

Figure 3 visually encapsulates the hybrid of CNN and ResNet101 model's performance through a confusion matrix, delineating the accurate predictions and misclassifications for Pothole Detection and Classification no_pothole and pothole.

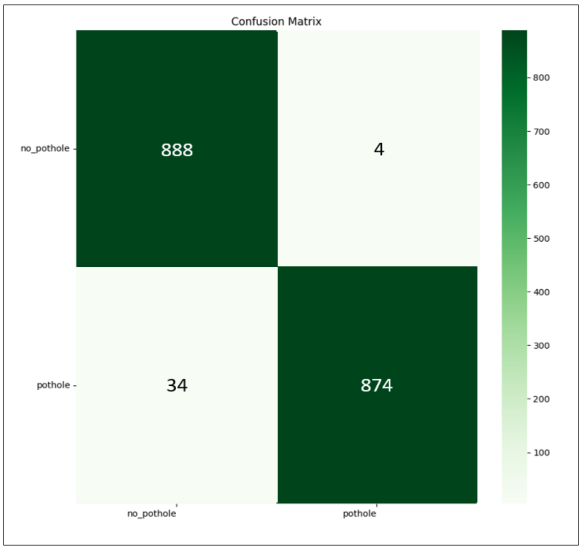

Figure 3 Confusion Matrix of Hybrid CNN and ResNet101 Model.

The confusion matrix in Figure 4 illustrates the performance of the YoloV8 model by showing the accurate predictions and misclassifications for both the background and the pothole in the task of Pothole Detection.

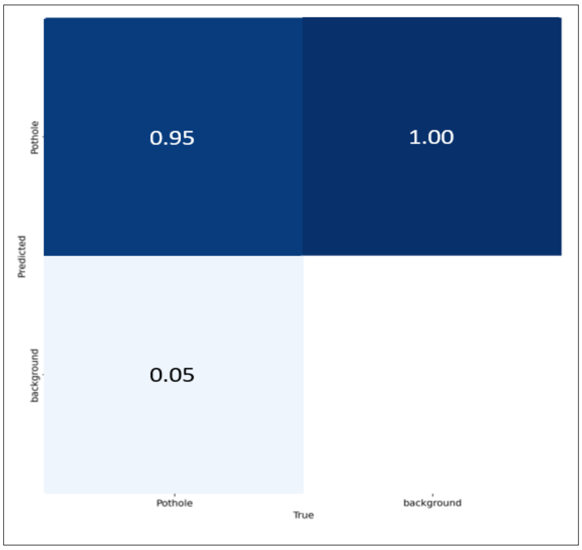

Figure 4 Confusion Matrix of YoloV8 Model.

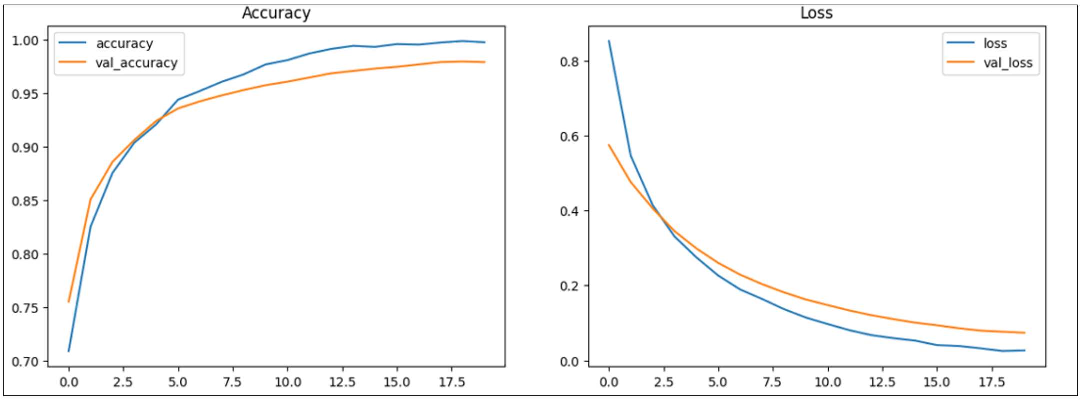

Figure 5 Accuracy vs Loss Curve for Training-validation of Proposed Hybrid Model.

Figure 5 presents a dynamic illustration of the hybrid CNN and ResNet101's training and validation accuracy, coupled with its loss curve, evolving across 20 epochs. In addition to being an essential instrument for assessing the model's generalization abilities and performance during training, this visual plot offers a comprehensive representation of the model's learning dynamics and convergence patterns.

An active illustration of the YoloV8's precision, recall accuracy curve is shown in Figure 6. The graphical plot provides a comprehensive representation of the model's learning dynamics and convergence patterns, and it is an essential aid for assessing the model's generalization abilities and performance throughout training.

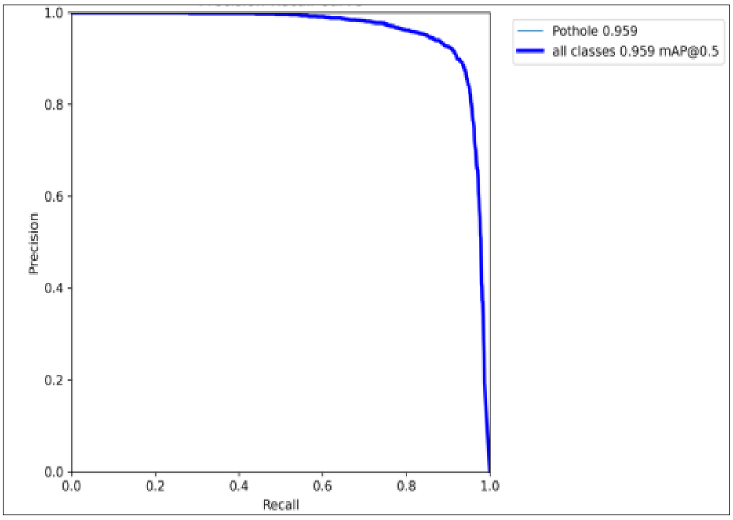

Figure 6 Precision vs Recall Curve of Proposed YoloV8 Model.

In Figure 7, we can see that, YOLOv8 primarily found a pothole on the right side of the road in a sample image. With great assurance, the model created a bounding box around it, demonstrating high identification certainty. Furthermore, there is writing in the same region that suggests a possible pothole detection with a confidence level of 0.55.

Figure 7 Sample Output of our Pothole Detection System by YoloV8 Model.

Table 4 presents a comparative analysis of our proposed approaches alongside Arjapure et al.'s [1] method for pothole detection using an 838-image dataset. The results showcase the effectiveness of our models in outperforming the benchmark, exemplified by Arjapure et al.'s accuracy rates for the corresponding models.

5. Conclusion

In conclusion, our proposed models demonstrate robust performance in pothole detection across diverse datasets, consistently achieving accuracy rates exceeding 90% for various architectures. These models not only showcase their efficacy in accurate classification but also hold the potential to significantly contribute to road safety. By promptly identifying and addressing road imperfections, our system plays a vital role in mitigating accidents and promoting a safer travel environment for all road users. We believe the high accuracy observed in our experiments reflects the reliability and practical applicability of our pothole detection system. In the future, we have planned to deploy our model for real time pothole detection.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Arjapure S, Kalbande DR. Road pothole detection using deep learning classifiers. International Journal of Recent Technology and Engineering (IJRTE). 2020;8(6):2-3.

[2] Asad, M. H., Khaliq, S., Yousaf, M. H., Ullah, M. O., & Ahmad, A. Pothole Detection Using Deep Learning: A Real-Time and AI-on-the-Edge Perspective. Advances in Civil Engineering. 2022.

[3] Dewangan, D. K., & Sahu, S. P. PotNet: Pothole detection for autonomous vehicle system using convolutional neural network. Electronics Letters. 2021;57(2): 53-56.

[4] Chen, H., Yao, M., & Gu, Q. Pothole detection using location-aware convolutional neural networks. International Journal of Machine Learning and Cybernetics. 2020;11(4):899-911.

[5] Patra, S., Middya, A. I., & Roy, S. PotSpot: Participatory sensing-based monitoring system for pothole detection using deep learning. Multimedia Tools and Applications. 2021;80: 25171-25195.

[6] Ping, P., Yang, X., & Gao, Z. A deep learning approach for street pothole detection. In: Big Data Computing Service and Applications (BigDataService). 6th ed. IEEE; 2020. p. 198-204.

[7] Bibi, R., Saeed, Y., Zeb, A., Ghazal, T. M., Rahman, T., Said, R. A., ... & Khan, M. A. Edge AI-based automated detection and classification of road anomalies in VANET using deep learning. Computational intelligence and neuroscience. 2021;1-16.

[8] Arjapure, S., & Kalbande, D. R. Deep learning model for pothole detection and area computation. In: International Conference on Communication information and Computing Technology (ICCICT). IEEE;2021. p. 1-6.

[9] Bhavya, P., Sharmila, C., Sai Sadhvi, Y., Prasanna, C. M., & Ganesan, V. Pothole detection using deep learning. In: Smart Technologies in Data Science and Communication: Proceedings of SMART-DSC. Singapore: Springer; 2021. p. 233-243.

[10] Rohitaa, R., Shreya, S., & Amutha, R. Intelligent Deep Learning based Pothole Detection and Reporting System. In: International Conference on Electrical, Computer and Communication Technologies (ICECCT). 4th ed. IEEE;2021. p.1-5.

[11] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. Gradient-based learning applied to document recognition. Proceedings of the IEEE. IEEE;1998; 86(11):2278-2324.

[12] He, K., Zhang, X., Ren, S., & Sun, J. Deep residual learning for image recognition. 2015; arXiv preprint arXiv:1512.03385.

[13] Donoho, D. L., & Johnstone, J. M. An influential paper on image denoising using wavelets. Ideal adaptation for wavelet shrinkage based on James-Stein estimation. 1994.

[14] Gonzalez, R. C., & Woods, R. E. While grayscale conversion itself is a simple process, here's a reference explaining color spaces and conversion. In: Digital image processing. 3rd ed. Pearson Prentice Hall. 2008.

[15] Pizer, S. M., Amburn, E. P., Austin, J. D., Greer, R. O., Jennings, B. T., Kim, Y., ... & Zucker, M. A. A paper on histogram equalization, a common contrast enhancement technique. Survey of multiscale image segmentation methods. 2003.