Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Smart Pothole Detection and Geospatial Visualization using Deep Learning and ArcGIS

Suchitra Patil[1], Nilkamal More[1], Abhijeet Pasi[1], Bhakti Palkar[1], Manisha Galphade[2], Nishant More[3]
K.J. Somaiya School of engineering, Somaiya Vidyavihar University, Vidyavihar, Mumbai, India[1]
School of Computing, MIT Art, Design and Technology, Pune,India[2], Endicott College of International Studies, Woosong University, Daejeon,South Korea[3]

ARTICLE INFO

Received: 29 Dec 2024

Revised: 17 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

This paper presents a complete system that combines artificial intelligence and geospatial analysis for pothole detection and demonstration on roads. Overcoming the inefficiencies of manual inspection of delay, time-consuming, and susceptible to human error, the system combines deep learning algorithms and ArcGIS to effectively detect potholes. Deep networks like VGG16, ResNet50, and Mask R-CNN are used in the model to enable accurate pothole detection on road surfaces. Moreover, integration with ArcGIS enables spatial mapping for better maintenance planning, while web interface with GPS support enables citizen engagement through real-time pothole reporting. The system performed quite well with an accuracy of 93% and F1 score of 95.86%. The Mask R-CNN model achieved 73% precision much higher than baseline models like YOLOv3. Results vindicate the system's potential as a viable, scalable, and community-based device for road monitoring and infrastructure maintenance.

Keywords: Deep learning ArcGIS, Pothole detection, VGG, Resnet, RCNN, Object Detection

INTRODUCTION

Road networks play important roles in enabling transportation and communication between nations. As much as there is dependence on road travel, there is a necessity to ensure safety and reliability of road infrastructure. Potholes—sudden depressions or hollows on road surfaces caused mainly by water infiltration, structural weakening, or abandonment—are among the principal causes of traffic accidents. Road defects can cause damage to vehicles, traffic congestion, and, in extreme cases, injury or fatality. In cities like Mumbai, for example, municipal records show an enormous number of road-defect accidents, particularly during the monsoon season. The traditional technique of detecting such defects, namely manual inspection, is not just time-consuming and labor-intensive but also subject to variable results because of human fallibility and subjective decision. In addition, the methods used are typically reactive, designed to sweep the road after it has taken a lot of damage, resulting in delayed maintenance and excessive maintenance costs. In lieu of such weaknesses, the research here puts forth a smart and independent pothole detection system that integrates deep learning models and geospatial mapping protocols. Convolutional neural networks like VGG16, ResNet50, and Mask R-CNN are employed by the system to identify objects in images, and their results are merged with ArcGIS for spatial representations of where the potholes have been detected. A complementary component—a web-based interface with GPS capabilities—allows for public participation, enabling users to report potholes directly. This integrated approach promises not only more efficient monitoring but also proactive and participatory road maintenance.

The system that is suggested in this paper aims to ana lyze multiple data sources like satellite imagery and street-level photos in order to detect potholes precisely. We evaluate our system on a set of street-level images and effectively detect potholes with good accuracy. It also saves costs in the long term by preventing expensive road repairs after neglecting minor repairs.The blending of ArcGIS with deep learning enables holistic road infrastructure analysis to be more precise and increase pothole detection accuracy. Our sys tem has, therefore, been concluded as able to limit the destructive power of potholes against road infrastructure by being able to timely

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

detect and perform maintenance. Application of ArcGIS and deep learning technology, and also the pothole reporting website, makes the system cost-efficient, effective, and accessible to the public.

OBJECTIVES

Urban road networks are constantly exposed to wear and tear due to high vehicle density, adverse weather conditions, and aging infrastructure. Timely detection and management of road damage, especially potholes, is essential for ensuring road safety and reducing long-term maintenance costs. Traditional approaches to damage identification are inadequate for modern smart city needs due to their inefficiency and lack of scalability. Leveraging artificial intelligence in combination with geospatial technologies presents a viable and innovative alternative.

In light of this, the primary aims of the present research are as follows:

1. To develop an automated pothole detection system using advanced deep learning techniques, specifically convolutional neural networks like VGG16, ResNet50, and Mask R-CNN, to classify and detect potholes with high accuracy and efficiency.

2. To incorporate geospatial analysis through ArcGIS for visualizing and mapping pothole locations across large road networks, supporting data-driven decision-making for infrastructure maintenance.

3. To implement a web-based reporting platform equipped with GPS functionality that enables citizens to report potholes in real-time, thereby enhancing public involvement and streamlining communication with municipal authorities.

LITERATURE SURVEY

Studies have extensively researched automated pothole detection using deep learning techniques with the purpose of substituting manual inspection methods that prove inefficient. The authors in [1] developed a modified version of the VGG16 network by removing select convolutional layers and adding dilated convolutions to reduce computational complexity without sacrificing accuracy. A comparison of the proposed model with YOLOv5 (ResNet101), Faster R-CNN (ResNet50 with FPN), MobileNetV2, and InceptionV3 architectures showed that it obtained a precision score of 91.9% when evaluated against a custom dataset of road images acquired from windshield-mounted smartphones. The study in [2] investigated the detection performance of SSD-TensorFlow and YOLOv3 (Darknet53) and YOLOv4 (CSPDarknet53) on pothole data that included images from Lebanese roads together with online sources. The research used more than 2,000 annotated potholes in over 1,000 images and achieved a detection precision of 88%.

A hybrid system that employed DBSCAN clustering and Z-score analysis with the Xception network served as the detection method in [3]. The analysis of vehicle dashboard accelerometer data combined with public image collections formed the basis of the dataset. The method produced results with a precision of 93.44% and used blockchain-enabled API storage to maintain transparency. Ukhwah et al. [4] created an automated system which combined deep learning models with ArcGIS for identifying potholes through Mask R-CNN and VGG16 and ResNet32 integration. The system included a public reporting portal together with accelerometer and image data functionality. This framework provided blockchain-based storage for its data while achieving the same precision rate of 93.44% and ensuring both robustness and public accessibility.

The research in [5] focused on performing binary classification of pothole and non-pothole images which were obtained from dashboard-mounted cameras in Timor Leste. A deep CNN received pre-processed images to achieve 95.2% accuracy which was recommended for improvement by incorporating extra sensors to measure severity. Research presented in [6] used Mask R-CNN for both detecting pothole shapes and calculating their surface area. The 460-image dataset processed by the model yielded a mean average precision (map) of 0.84 which demonstrated its potential for both detection and measurement capabilities. The authors suggested that building a larger training dataset along with testing different models would help enhance the results.

A real-time detection framework presented in [7] combined convolutional neural networks with recurrent layers to improve temporal tracking capabilities. The final model achieved 94.62% accuracy after processing smartphone-

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

acquired images from a custom dataset that outperformed various baseline models in comparative testing. Javed et al. described in [8] a three-stage detection system which consisted of a Region Proposal Network followed by a CNN classifier and ended with a final pothole localization layer. The R-CNN model achieved 94.5% accuracy through training with 1000 images while performing better than VGG16, InceptionV3 and ResNet50. The two-stage system described in [9] first utilized traditional image processing to detect potholes followed by machine learning classifiers for severity evaluation. The presented model demonstrated its ability to make precise classifications during testing on Indian road images and showed potential for real-time implementation.

Egaji et al. [10] established a real-time convolutional neural network system that used images from road surfaces for its operation. The system proved successful in differentiating between damaged and undamaged road conditions and showed potential as a tool for on-site infrastructure evaluation. The researchers in [11] trained YOLOv3 to detect potholes on Indian road datasets at high speed. The model delivered fast processing and reliable accuracy but needed further improvement through the integration of contextual road texture information. The research in [12] focused on statistical pothole characteristics of Indian roads to develop maintenance strategies while [13] presented a smartphone-based detection system which used accelerometer data for crowdsourced reporting and real-time alerts.

While [14] was unrelated to pothole detection, it addressed institutional strategies for internationalization—highlighting planning and community involvement, which interestingly parallels the civic engagement theme in smart infrastructure systems. In [15], a computer vision-based technique for analyzing pavement images was described, successfully demonstrating the use of digital image processing for pothole identification and maintenance efficiency. Lastly, [16] described the use of YOLO for pothole detection in Indian roads, confirming its value for fast, accurate object detection. Despite the advances demonstrated across these studies, most focus solely on visual detection without incorporating geospatial intelligence tools like GIS. Very few offer a combined solution that allows for real-time detection, spatial mapping, and public reporting in a unified platform. Additionally, many systems lack comprehensive performance evaluations using multiple metrics, such as F1 score or AUC. Addressing these gaps, the present study introduces a multi-component framework that utilizes high-performing CNNs [18], integrates ArcGIS for geospatial visualization, and includes a GPS-enabled reporting interface to actively involve citizens in infrastructure upkeep.

METHODS

A. Dataset

The dataset used for this project is available and comprises 8484 RGB images of various road structures, including pot- holes and roadside barriers. The algorithm approach considered for implementation involves using a deep learning pretrained VGG-16 algorithm and image processing for pothole detection. Additionally, ArcGIS mapping will be used to visualize the detected potholes over a specific area. In case the data set is not available, the project will create a data set of 100 captured images with a minimum of six attributes required for the analysis. The work aims to detect potholes accurately, which will help in minimizing the accidents and damage to vehicles on the roads.

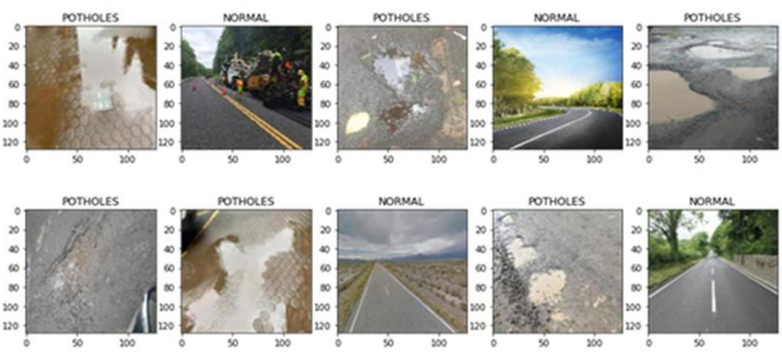

Figure 1: Pothole Dataset Sample Images.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

A. Python Implementation

1. Preparing Training Data

To begin with, we collected a dataset of road images, where each image was labeled as 0(normal roads) or (containing potholes). We did Image preprocessing to improve image quality, eliminate noise, and extract significant features for further analysis. Some image preprocessing techniques include:

(a) Resizing: This is performed to resize the image to a specific resolution or size, so that havin all the images are of standardize size makes it easier to process them.

(b) Cropping: This involves removing unwanted parts of an image. Cropping can help to remove irrelevant information from an image and focus on the relevant parts.

(c) Augmentation: This step involved applying trans- formations like rotation, scaling and flipping to generate new images from existing ones.

2. Training Classification Model

The preprocessed dataset to train a sequential model for classification. In a sequential model layers are stacked sequentially, with each layer feeding into the next. Sequential models also work well for tasks where there is a clear input-output mapping, such as image classification. The layers in a sequential model process the input data in a feedforward manner, passing the output of each layer as input to the following layer. This allows the model to learn hierarchical representations of the input data, where each layer learns to identify increasingly complex features. To learn hierarchical representations of the input images, a sequential model with multiple convolutional and max pooling layers is build here. This output from the layers was flattened and then fed into fully connected layers with a dropout layer to prevent overfitting. Finally, the model outputs a probability distribution over the two classes using a softmax activation function. During training of model an Adam optimizer was used, and binary cross-entropy loss was also utilised as the objective function. After this classification model was trained, it was used to classify new images of roads into two categories: those containing potholes and those that do not.

3. Building Object Detection Model

For the task identifying potholes in the images of paved roads using object detection the VGG16 architecture was employed. VGG16 is a popular deep neural network structure for image classification tasks. It consists of 16 layers, including several convolutional and very few max pooling layers, followed by fully connected layers. The network employs 3x3 filters for the convolution operation that allows more filters with less number of parameters. The VGG16 model derives the features of the input image and utilizes these features to suggest regions of interest (RoIs) in the image that can potentially hold objects. These RoIs are further processed through extra layers to classify and smooth the object boundaries. The model is trained on huge image datasets such as ImageNet and obtains the best performance on many image classification tasks. We trained this model on a road image dataset, and it learned to detect the position and size of potholes in the image. We applied transfer learning to fine-tune the pre-trained VGG16 model on our road image dataset.

B. GIS Implementation

1. Preparing Training Data

We used the tool “Label Objects for Deep Learning” to create a labeled imagery dataset, that can be used to train a deep learning model for our imagery workflows. Using this tool, we were able to actively identify and label objects within an image, and later we exported the result into training data which consist of image chips, labels, and statistics that we required for the model training process. In the event that we already had a pre-existing labeled vector or raster data, we could have chosen to use the Geoprocessing tool that is “Export Training Data For Deep Learning” which could have helped us to generate the training data required for the next step. However in our scenario, we took map inputs by identifying locations with potholes and labeling them on the TIFF images. Later, we utilized Export Training Data, which uses a remote sensing image to transform labelled vector into deep

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

learning training datasets. The output of this process was a folder of image chips and a folder of metadata files. Image chips are typically 256 pixels in row and 256 pixels in column with each chip containing one or more objects.

2. Training Deep Learning Model

We used the 'Train Deep Learning Model' tool to train our model for object detection in imagery workflows. To do this, we had to install the deep learning frameworks for ArcGIS. The tool we used to produce the input training data was called "Export Training Data For Deep Learning", which consisted of images and labels folders. Our main aim was to fine-tune an existing trained model, and we tested two different model types: MASKRCNN and YOLOV3. During the training process, we had several arguments to configure, including Batch Size, Epochs, Chip Size, and Learning Rate. One critical parameter that we focused on was the backbone model, which determines the preconfigured neural network that would be used as the architecture for training the new deep learning model as a part of Transfer Learning. We specifically prioritized three models: RESNET50, RESNET101, and VGG16. RESNET50 and RESNET101 are residual networks trained on the Imagenet Dataset, consisting of over 1 million images and are 50 and 101 layers deep, respectively. On the other hand, VGG16 is a convolutional neural network trained on the Imagenet Dataset to classify images into 1,000 object categories and has 16 layers deep.

3. Model Inferencing

The Esri Model Definition file, which is composed of a file in JSON format (.emd), a JSON string, and a deep learning model package (.dlpk), was utilized. When we employed this tool on the server, we were able to utilize a JSON string as it allowed us to paste the string instead of uploading the .emd file. In addition, we had to ensure that the .dlpk file was stored locally. The tool mentioned in the scenario can process input imagery in map space as well as pixel space.Map space is used to denote images in a coordinate system relative to a map, whereas pixel space refers to unprocessed image space without rotation or distortions. We were able to improve the performance of the tool by increasing the batch size. We had to remember that increasing the batch size would use more memory. Also, if the input value given was not a perfect square, we needed to make the batch size equal to the highest possible square value. For example, if the input value was 6, the batch size would be made equal to 4. Finally, we also made use of other parameters, such as boxes, padding, and threshold, to fine-tune the pothole detection process and ensure that our deep learning model was well-equipped to handle the task at hand.

RESULTS

This section presents the evaluation outcomes of the proposed deep learning models used for pothole detection. The models were trained on a curated dataset of road images and assessed based on various performance metrics including accuracy, F1 score, loss, and area under the ROC curve (AUC).

Sequential Classification Model Performance

A deep learning-based sequential model was developed and trained on labeled road images, where roads were classified as either containing potholes or not. During training, the model achieved an accuracy of 95%, while the test accuracy reached 93%. The associated test loss was approximately 0.70, indicating reliable generalization. To further validate the classification performance, a confusion matrix was generated using a test set of 480 images. The model successfully classified 450 images correctly, and the resulting ROC curve yielded a score of 89%, demonstrating robust discriminatory ability (Figure 2).

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

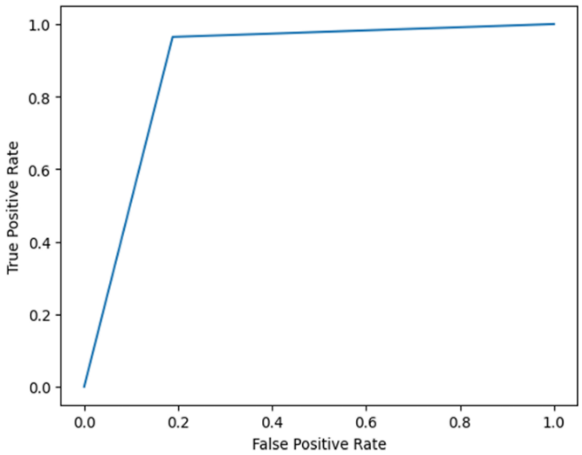

Figure 2: ROC Curve for training accuracy of pothole detection sequential model.

Epoch and Batch Size Optimization

To determine optimal model performance, experiments were conducted with varying epochs and batch sizes. A comparative analysis is provided in Table 1. The model trained for 200 epochs with a batch size of 12 exhibited the best results, achieving:

- Test Accuracy: 93%
- F1 Score: 95.86%
- Test Loss: 47.05%
- AUC Score: 85.51%

This confirms that extended training contributes positively to the model's ability to generalize.

Table 1: Comparision of metrics for each epoch batch.

Metrics	**30 Epochs**	**100 Epochs**	**200 Epochs**
Batch Size	4	12	12
Learning Rate	0.0005	0.0001	0.0001
Conv2d Layers	3	5	5
Test Accuracy	89%	91%	93%
F1 Score	90.14%	92.60%	95.86%
Test Loss	47.67%	51.01%	47.05%
AUC score	80.01%	80.43%	85.51%

Object Detection Using VGG16 and Mask R-CNN

The object detection capability of the system was evaluated using VGG16 and Mask R-CNN architectures. VGG16-based detection effectively identified pothole regions in high-resolution road images (Figure 3), while the Mask R-CNN model—utilizing ResNet50 as its backbone—delivered higher precision in spatial localization (Figure 4).

Comparison of Deep Learning Models

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Table 2 presents a performance comparison of different models tested in this study. Among them, the Mask R-CNN model with ResNet50 yielded the highest precision of 73%, outperforming both ResNet101 and YOLOv3 (Darknet53), the latter of which showed only 16% precision, highlighting its limitations in this specific context. The pothole detected using Vgg16 can be represented visually through the Figure 3

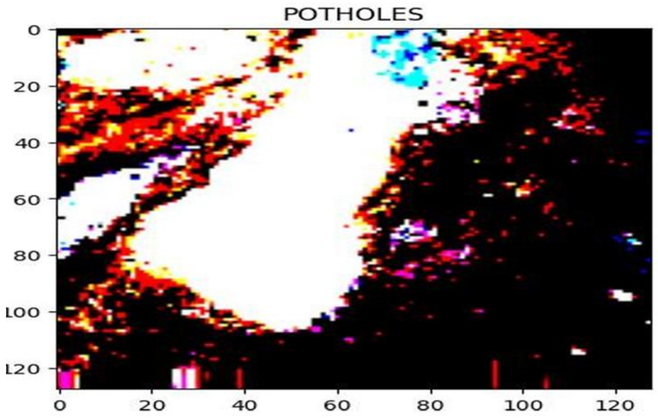

Figure 3: Pothole Detected using VGG16.

The pothole detected using MaskRCNN model having Resnet-50 as the backbone model can be represented visually through the Figure 4.

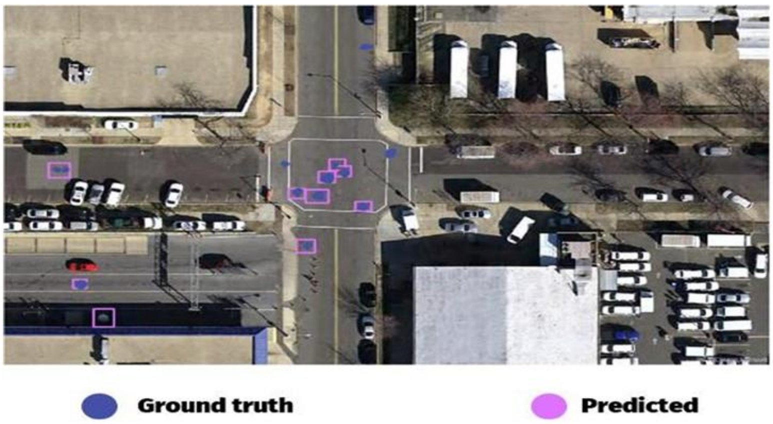

Figure 4: Pothole Deteced using ArcGIS.

Table 2: Comparision of metrics for different models used.

Metrics	Darknet 53	Resnet 50	Resnet 101
Batch Size	64	**16**	16
Chip Size	224	**224**	224
Model Type	YoloV3	**MaskRCNN**	MaskRCNN
Epochs	50	**50**	50
Training Loss	13227.5	**1.03**	0.73
Precision	16%	**73%**	41%

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The training and validation loss for the above best model MaskRCNN with Resnet 50 is shown in the Figure. 5

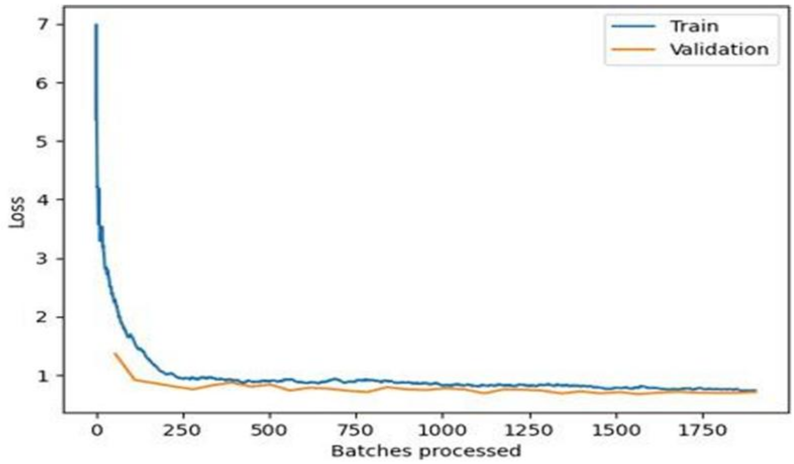

Figure 5: Training and Validation Loss Trend

DISCUSSION

The results obtained in this study demonstrate significant improvement over traditional manual pothole inspection systems. Our best-performing deep learning model (Mask R-CNN with ResNet50 backbone) achieved a precision of 73%, which aligns with several recent studies that implemented similar architectures.

For example, Arjapure & Kalbande in [6] reported a mean average precision (mAP) of 0.84 using Mask R-CNN for pothole area detection, showing comparable effectiveness to our ResNet-based implementation. Similarly, Javed et al. [8] demonstrated 94.5% accuracy using an R-CNN approach, indicating that further fine-tuning of region proposal stages could enhance our model's precision.

In contrast, models such as YOLOv3 with Darknet53 underperformed in our tests, yielding a precision of only 16%, despite other works like [11] reporting better results on Indian road datasets. This discrepancy likely stems from dataset differences and image resolution variations during training.

Notably, the integration of ArcGIS in our system adds a unique geospatial dimension absent in many prior studies. This not only improves visualization and interpretation but also offers municipal authorities a practical interface for real-time infrastructure monitoring.

By combining GIS and deep learning, our model achieves a balance between accuracy (~93%), generalization, and practical usability. Compared to other systems in literature that focus solely on image-based models or sensor-based detection, our approach offers a holistic and scalable solution with both technical and civic engagement benefits.

REFERENCES

[1] K. R. Ahmed, "Smart Pothole Detection Using Deep Learning Based on Dilated Convolution," Sensors, 2021, 8406.
[2] A. Al Shaghouri, R. Alkhatib, and S. Berjaoui, "Real-Time Pothole Detection Using Deep Learning," ArXiv abs/2107.06356, 2021.
[3] V. B. Nikam, M. Marathe, O. Hetavkar, D. Bakhai, P. Mehta, N. More, and A. Bambole, "Gamification of Pothole Reporting and Road Assessment using a Reward Based system Supported by the Blockchain Technology," Proc. ACM Sigspatial, 2022, 10 pages.
[4] E. N. Ukhwah, E. M. Yuniarno, and Y. K. Suprapto, "Asphalt Pavement Pothole Detection using Deep Learning method based on YOLO Neural Network," 2019 Int. Seminar on Intelligent Technology and Its Applications (ISITIA), pp. 35-40.

Journal of Information Systems Engineering and Management
2025, 10(38s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[5] V. Pereira, S. Tamura, S. Hayamizu, and H. Fukai, “A Deep Learning-Based Approach for Road Pothole Detection in Timor Leste,” 2018 IEEE Int. Conf. on Service Operations and Logistics, and Informatics (SOLI), pp. 279-284.
[6] S. Arjapure and D. R. Kalbande, “Deep Learning Model for Pothole Detection and Area Computation,” 2021 Int. Conf. on Communication Information and Computing Technology (ICCICT), pp. 1-6.
[7] P. Bhavya, C. Sharmila, Y. S. Sadhvi, C. M. L. Prasanna, and V. Ganesan, “Pothole Detection Using Deep Learning,” in *Smart Technologies in Data Science and Communication*, vol. 210, Springer, Singapore, 2021.
[8] A. Javed et al., “Pothole Detection System Using Region-Based Convolutional Neural Network,” 2021 IEEE 4th Int. Conf. on Computer and Communication Engineering Technology (CCET), pp. 6-11.
[9] A. Shah, G. Sharma, and L. Bhargava, “Smart Implementation of Computer Vision and Machine Learning for Pothole Detection,” 2021 11th Int. Conf. on Cloud Computing, Data Science & Engineering (Confluence), pp. 65-69.
[10] O. A. Egaji, G. Evans, M. G. Griffiths, and G. Islas, “Real-time machine learning-based approach for pothole detection,” *Expert Systems with Applications*, vol. 184, 2021, 115562.
[11] J. Dharneeshkar, “Deep Learning based Detection of potholes in Indian roads using YOLO,” 2020 Int. Conf. on Inventive Computation Technologies (ICICT), pp. 381–385.
[12] V. L. Solanke, D. D. Patil, A. S. Patkar, and G. S. Kale, “Analysis of existing road surface on the basis of pothole characteristics,” *Global Journal of Research in Engineering*, vol. 19, pp. 1–5, 2019.
[13] A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo, “Real-time pothole detection using android smartphones with accelerometers,” Proc. 2011 Int. Conf. on Distributed Computing in Sensor Systems (DCOSS), Barcelona, Spain, Jun. 2011.
[14] A. M. Legreid, “Potholes and Strategies on the Road to Campus Internationalization,” *International Research and Review: Journal of Phi Beta Delta Honor Society for International Scholars*, vol. 6, no. 1, 2016.
[15] C. Koch and I. Brilakis, “Pothole detection in asphalt pavement images,” *Advanced Engineering Informatics*, vol. 25, no. 3, pp. 507–515, 2011.
[16] J. Dharneeshkar, S. Aniruthan, R. Karthika, and L. Parameswaran, “Deep Learning based Detection of potholes in Indian roads using YOLO,” 2020 Int. Conf. on Inventive Computation Technologies (ICICT), pp. 381-385.
[17] M. J. De Smith, M. F. Goodchild, and P. Longley, *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*, 2018.
[18] S. Weidman, *Deep Learning From Scratch: Building with Python from First Principles*, O’Reilly, 2020.