Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 4 (2024)

Pothole Detection Model based on Convolutional Neural Network and Metaheuristic Group Learning Algorithm

[1]Puneet Singh, [2]Mohit Sharma, [3]Ravi Gaba

[1]*Research Scholar, Department of Civil Engineering, BGIET, Sangrur, Punjab, India.*

[2]*Assistant Professor, Department of Civil Engineering, BGIET, Sangrur, Punjab, India*

[3]*AEE(OS), Punjab Pollution Control Board, Punjab, India*

Abstract: In this paper, we have presented a pothole detection model in order to enhance the road safety. In this model, convolutional neural network (CNN) and metaheuristic group learning algorithm (CNN) is employed for classify the pothole. The proposed model has four stages: reading the dataset, pre-processing the dataset, classification of potholes using machine learning, and performance evaluation. In the first stage, a standard dataset is read. In the second stage, the pre-processing of the dataset is done using the filter, enhancement, and segmentation methods to find the region of interest. Further, in this stage, a metaheuristic group learning algorithm is utilized to fine-tune the pre-processing method. Next, in the third stage, the classification of potholes is done using CNN algorithms. To accomplish this goal, the CNN algorithm is trained and tested using the dataset generated after the pre-processing method. Finally, the evaluation of the proposed model is done using various performance metrics such as accuracy, precision, recall, and F1-score. The evaluation of the proposed model is done on the standard dataset of potholes, which is available on the Kaggle database. This dataset contains smooth and pothole images. The result shows that the proposed model achieves better performance metrics than the existing method.

Keywords:CNN, Deep Learning, Detection, Machine Learning, Metaheuristic,Pothole, Segmentation.

1. Introduction

One significant structural road failure is a pothole [1]. It is created when transportation and water are present together [2]. Traffic destroys the road surface when water weakens the soil beneath it, removing pieces of the road surface. Road potholes pose a serious risk to vehicle conditions and traffic safety and are irritating [3]. For example, in the first two months of 2018, 11,706 complaints regarding potholes on the roads were made by vehicles, according to the Chicago Sun-Times [4]. *The Pothole Facts* estimates that poor road conditions are responsible for around one-third of the 33,000 traffic deaths that occur in the US each year. Therefore, it is very important to regularly check roads and fix potholes [5].

Currently, the primary method for detecting potholes in roads is still manual eye assessment [6]. Road potholes are often seen and reported by qualified inspectors and structural engineers. This procedure is risky, costly, and ineffective. For instance, pothole detection and repair cost New Zealand city governments millions of dollars in 2017 (Christchurch alone paid $525,000 USD) [7]. Furthermore, it has been estimated that San Diego, California, repairs around 30,000 potholes annually. The municipal road maintenance staff in San Diego could use some help finding potholes, so they asked locals to report them [8]. Additionally, inspectors and engineers' manual road pothole identification findings are usually subjective since they rely on their expertise and judgment [9]. Because of these factors, researchers have been working hard to create automated methods for evaluating the state of roads that can effectively, precisely, and completely locate, identify, and reconstruct potholes [10]. As a result, we have developed an automated road evaluation model in this study that can identify potholes. Machine learning and metaheuristic algorithms are considered in this study to achieve this objective. The main contribution of this paper is to enhance the accuracy of the pothole detection using the CNN algorithm. In order to accomplish this goal, the pre-processing of the dataset is done using the metaheuristic

GLA algorithm. Further, simulation evaluation is done on the standard dataset and evaluated using the four parameters, namely, accuracy, precision, recall, F-score.

The remaining structure of the paper as follows. Section 2 show the related work on pothole detection using the machine learning. Section 3 defines the methodology of this research. Section 4 explains the proposed road pothole detection model. Section 5 shows the simulation results on the standard dataset and evaluation using the various performance metrics. Finally, conclusion is drawn in Section 6.

2. Related Work

Abdelmalak et al. [11], The critical issue of pothole identification plays an important role in maintaining infrastructure integrity and road safety in the context of intelligent transportation systems. This research project uses a sophisticated and multidimensional strategy to carefully negotiate the complexities of automated pothole identification. The dataset consists of more than 300 carefully tagged images of roads with and without potholes. The suggested strategy demonstrates outstanding effectiveness in identifying road defects by using the powerful GoogLeNet for feature extraction and coordinating the optimization of XGBoost using the Al-Biruni Earth Radius Metaheuristic Algorithm. The results show how effective the techniques that were used were, with BER-XGBoost being the best performer with an accuracy score of 96.01%. This model is very accurate and shows a wide range of measures, such as F-score, specificity, sensitivity, positive predictive value, and negative predictive value. The usefulness of this technique is supported by rigorous statistical studies, such as the ANOVA and the Wilcoxon Signed Rank Test. In summary, this research not only provides useful information to the relevant area but also raises important questions about the implications of optimization techniques and the complex function of feature extraction in the context of automated pothole identification. By successfully bridging the gap between scientific advancements and practical applications, this study advances the ongoing development of intelligent systems, enhancing road safety and improving infrastructure management.**Anandhalli et al. [12],** The main topic of the study was comparing how well two deep learning models can find potholes. The Yolo model was trained on a set of images with labels, and a self-built basic sequential CNN model was trained on datasets without labels. Annotating images is a time-consuming process, and training annotated images demands both personnel and high-end resources, such as GPUs. In contrast, a self-built fundamental CNN model requires fewer resources and requires less time for training. The main objective of the article is to highlight the results of comparing YOLO and CNN in identifying potholes. The results demonstrate that CNN has a high accuracy rate of 98.51%.

3. Methodology

This section explains the dataset, CNN, metaheuristic GL algorithm, and evaluation parameters are employed in the proposed pothole detection model.

3.1 **Pothole Database**

In this research, we have considered the secondary database for evaluate the proposed roadside safety assessment model. This database is publicly available on the Kaggle website [13]. This database contains the normal and pothole images which are captured from the different angle of the road.

3.2 CNN

Convolutional neural networks are very effective in extracting local characteristics from images. The primary distinction between convolutional neural networks and classic neural networks is the usage of a partial connection network and the introduction of the idea of local receptive fields. To train the standard neural network for image feature learning, the picture of two-dimensional data is converted to one-dimensional data and placed in the network input layer. It is difficult for neural networks to learn the spatial features between pixels since this process degrades the image's spatial structural information. Furthermore, the neural network's training period is too lengthy due to its excessive number of parameters. Studies in the biological processing of visual information served as an inspiration for the local receptive field. Within the cerebral cortex, the neurons responsible for vision receive a signal that is a two-dimensional image. Based on this idea, convolutional neural networks are created. There is only one connection between each neuron in the network and the local neuron in the top layer. The values in the network are also kept low. Parameter sharing is another essential convolution neural network feature. To extract characteristics from images, convolution neural networks use convolution kernels. Although each convolution kernel may extract a feature from a local image, various local images may have identical statistical properties. The same characteristics may be learned by using a convolution kernel. This allows the global image to be convoluted by extracting the same features at various locations inside the global image using a convolution check. The network's training time is shortened, and its parameters are significantly decreased via the use of parameter sharing. As seen in Figure 1, the input, activation, convolution, full connection, pooling, and output layers constitute the current standard topology for convolution neural networks [14].

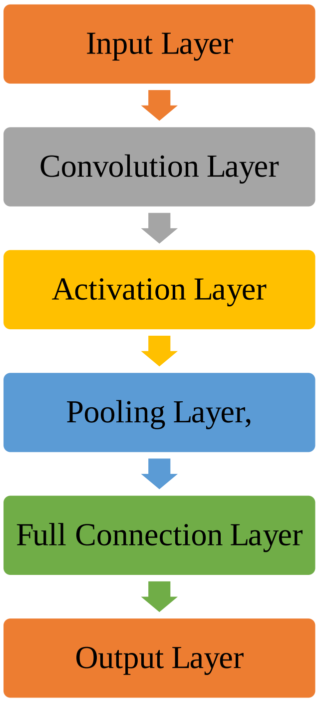


Figure 1. Layers of CNN [14]

- **Input Layer:** The convolutional neural network's input layer is its first layer. The main input for convolutional neural networks for processing images is either the image's pixel value matrix or a local image block. Convolutional neural networks may receive single-channel or multi-channel image matrices as input.
- **Convolution Layer:** The most significant component of a CNN is convolution layer whose primary function is convolution. In the convolution layer, each node only gets a small piece of the network matrix from the previous layer. This is different from the full connection layer. The pieces are usually 3 x 3 or 5 x 5. The convolution layer is made up of various characteristic matrices that are generated by convolution. The eigenmatrix is generated by moving several convolution kernels over the input matrix and inner-producting them. By following these steps, we can get a description of the image's features in a certain area.
- **Activate Layer:** Each convolution layer element is activated using the activation function, resulting in a nonlinear mapping process between the network's input and output that keeps the matrix's size constant.
- **Pool Layer:** The down sampling method in the pool layer extracts the most significant features from the local feature matrix. This further abstracts the reduction in dimension features of the feature matrix and further decreases the number of factors in the final full connection layer network. This thereby lowers the network's overall parameters, lowers the model's complexity, lowers the likelihood of overfitting issues during training, and enhances the model's resilience. Additionally, the model's computation speed is increased. Typically, the pool layer's operation involves calculating the area matrix's mean, maximum, and random values. The mean pooling technique eliminates the information about the image's structure yet successfully reduces the influence

of image noise. Maximum pooling is popular because it may successfully preserve the image's structural information while lowering the convolution error.

- **Full Connection Layer:** There is a fully linked layer that is made up of deep-layer feature transformation. This is done after multiple convolutional and pooling layers have extracted features.
- **Output Layer:** To provide the classification result, the classifier maps the high-level characteristics in the complete connection layer to the input image's category probabilities. Convolutional neural networks are trained under supervision. Using the back propagation technique, the advanced line optimizes the network properties after propagating forward to the network.

3.3 Metaheuristic GL Algorithm: The GLA exhibits how managers and group leaders influence their members' abilities, as the name indicates. This method differs significantly from previous algorithms in that it splits the population into many equal groups. Then, it chooses a subset of people as group leaders (based on fitness), and then identifies the manager, or the person with the highest overall fitness. The following are the GLA algorithm's key features [15]:

• The initial population is generated at random.

• The most fit person within the population assumes the role of manager for every individual.

• Several groups are represented by the whole population (four groups were considered for this study).

• The leader of every group is the best member (the one with the greatest level of fitness within the group).

• Each group's leader has an impact on its fellow members.

• Group leaders and other people are impacted by the manager.

• Mutations are used to randomly alter an individual's basic structure.

3.4 Evaluation Parameters

This section presents the performance evaluation parameters are utilized to evaluate the proposed roadside assessment model. Initially, in this section, confusion matrix is explained which the most preferred matrix is to evaluate the machine learning model. Further, based on this matrix, various other parameters, namely, accuracy, precision, recall, and F-score is measured [16-18].

- One special tabular representation of a supervised machine learning algorithm's performance is the confusion matrix. Every row indicates the number of cases of an actual class, and every column indicates the number of instances of the predicted class. The confusion matrix is used to generate most classification metrics, which are based on the number of false positives (FP), false negatives (FN), true positives (TP), and true negatives (TN).

	Predicted Value		
Actual Value		Yes	No
	Yes	TP	FN
	No	FP	TN

Figure 2 Confusion Matrix

- **Accuracy:** Accuracy is the most accurate metric for performance and may be simply defined as the ratio of accurately predicted observations to total observations. The solution is found using Equation (1).

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \qquad (1)$$

- Recall: The classifier's ability to predict positive samples is measured by its recall, also known as its true positive rate. It is computed as follows.

Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 4 (2024)

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

- Precision: It is usually computed by dividing the total number of properly labelled images for the class by the number of correctly labelled images.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

- F1-Score: The F1 Score is calculated by taking the harmonic mean of recall and accuracy. The higher the model's F1 Score, the better its performance. The following formula is used to get the F1-score:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

4. Proposed Road Pothole Detection Model

This part talks about the suggested roadside safety assessment model. It is meant to make finding potholes more accurate by using the CNN and metaheuristic algorithm. The flowchart of the proposed roadside safety assessment model is shown in Figure 3.Initially, we downloaded the dataset from the Kaggle website. Next, we convert the color images to grayscale and resize them. Next, we pre-process the image using the median filter, adaptive histogram method, and segmentation method to prepare it for the CNN algorithm. In addition, the segmentation method employs the metaheuristic group learning algorithm to determine the optimal threshold value. Next, we split the dataset images into a 70:30 ratio. The CNN algorithm trains on 70% of the dataset images, while it uses 30% for testing. The CNN algorithm processes these images and detects the pothole or not pothole in the road. Finally, we evaluate the proposed model using the accuracy, precision, recall, and F-score parameters.

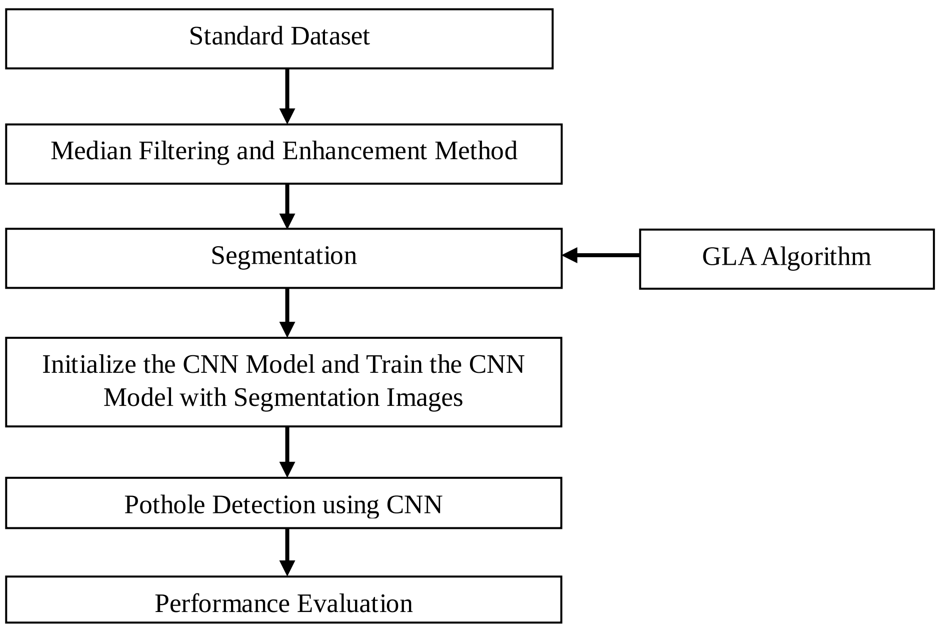


Figure 3. Proposed Roadside Safety Assessment Model

5. Simulation Results

This section presents the simulation results of the proposed pothole detection model and comparative analysis with the existing methods. MATLAB 2018a software is used for simulation purposes and the configuration of the system is i7 processor, 8GB RAM, 64-bit window operating system. Table 1 shows the parameters are initialized for simulate the proposed pothole detection model.In this table, the parameter values of the metaheuristic algorithm and CNN algorithm is shown.

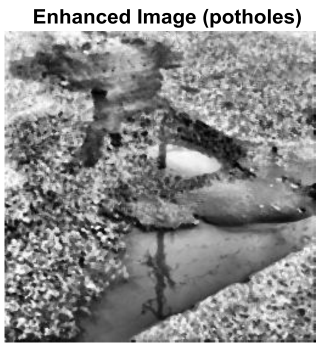
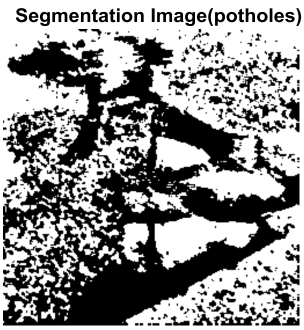
Table 1. Parameter Values of the Proposed Pothole Detection Model for Simulation Purposes

Parameter	**Value**
Population	10
Iteration	100
Total Number of Groups	2
Total Number of Members	5

5.3 Simulation Results

Table 2 shows the visual quality of the proposed roadside safety assessment model. Initially, the input dataset image is shown. After that, pre-processing images are generated after filtering, enhancement, and segmentation is shown.

Table 2 Visual Quality Analysis of the Proposed Model

Original Image	Pre-Processed Image
	
Enhanced Image	Segmentation Image
	

Further, Figure 4 shows the confusion matrix is evaluated for the proposed pothole detection model, which represents whether the CNN model correctly detects the pothole or not. In this table, the total number of TP cases is 131, FP cases are 0, FN cases are 2, and TN cases are 64, which are detected by the CNN model for the dataset used for evaluation purposes.

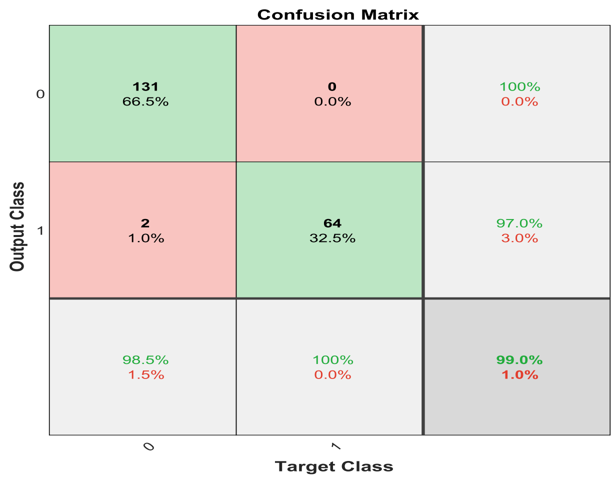


Figure 4 Confusion Matrix for the Proposed Pothole Detection Model

Table 3 shows the evaluation parameters, namely, accuracy, sensitivity, specificity, and F-score are evaluated for the proposed pothole detection model to detect two classes (pothole and non-pothole) of the dataset. The result shows that the proposed model achieves high value for both classes such as 0.98985 accuracy for both classes.

Table 1 Performance Evaluation of the Proposed Pothole Detection Model

N Classes	Accuracy	Sensitivity	Specificity	F-Score
1	0.98985	1	0.98496	0.98462
2	0.98985	0.98496	1	0.99242

Figure 5 shows the accuracy of the different classes of the dataset. The result shows that the proposed pothole detection model achieves the same accuracy for both classes.

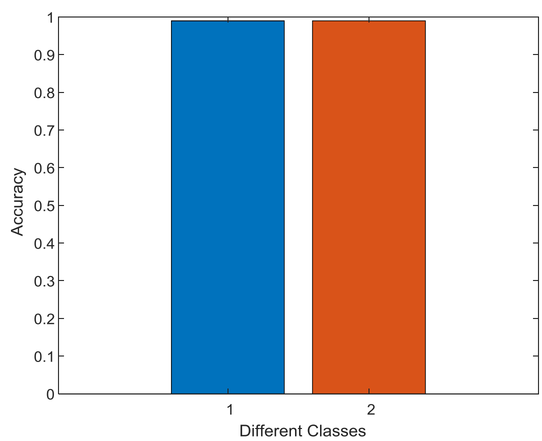


Figure 5Comparative Analysis of Accuracy Parameter for Different Classes of the Dataset

Figure 6 shows the sensitivity of the different classes of the dataset. The result shows that the proposed pothole detection model achieves 1 value for Class 1 and 0.98496 value for class 2.

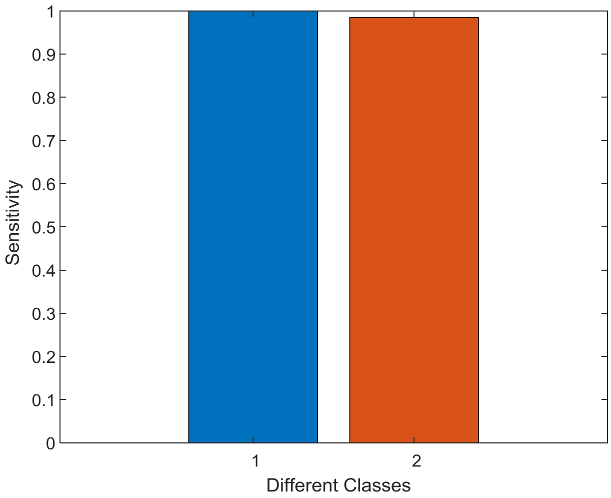


Figure 6 Comparative Analysis of Sensitivity Parameter for Different Classes of the Dataset

Figure 7 shows the specificity of the different classes of the dataset. The result shows that the proposed pothole detection model achieves 0.98496 value for Class 1 and 1 value for class 2.

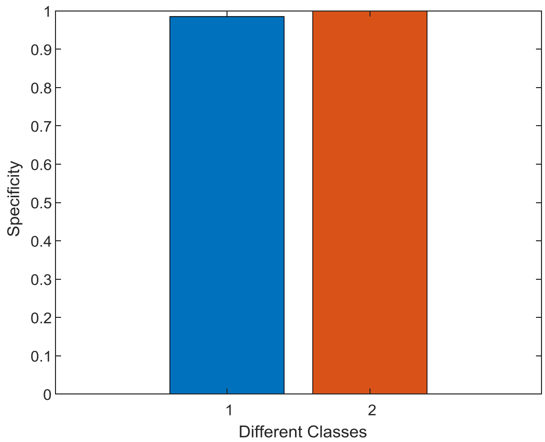


Figure 7 Comparative Analysis of Specificity Parameter for Different Classes of the Dataset

Figure 8 shows the F-Score of the different classes of the dataset. The result shows that the proposed pothole detection model achieves 0.98462 value for Class 1 and 0.99242 value for class 2.

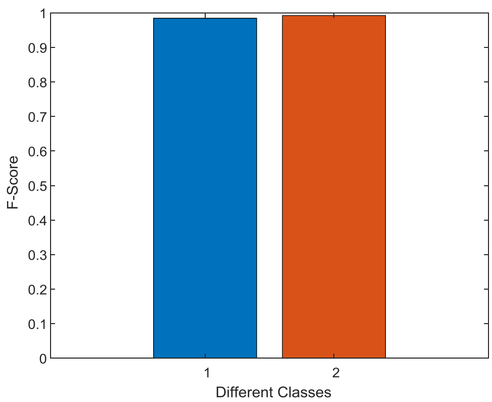


Figure 8 Comparative Analysis of F-Score Parameter for Different Classes of the Dataset

5.4 Comparative Analysis

Table 4 shows the comparative analysis of the proposed pothole detection model with the existing models based on XGBoost and its hybrid approaches with metaheuristic algorithms, namely, FOA, POS, and BER [11]. The result shows that the proposed model outperforms the existing models in terms of accuracy, sensitivity, specificity, and F-score.

Table 4 Comparative Analysis

Algorithm	Accuracy	Sensitivity	Specificity	F-Score
XGBoost	0.9211	0.9434	0.8807	0.939
FOA-XGBoost	0.9381	0.9524	0.9174	0.9479
PSO-XGBoost	0.947	0.9677	0.9174	0.9554
BER-XGBoost	0.9601	0.9908	0.9174	0.9665
Proposed Pothole Detection Model	0.98985	0.99248	0.99248	0.98852

6. **Conclusion**

In this research, we have presented a roadside safety assessment model that classifies the potholes. We use the CNN algorithm for classification purposes to achieve this goal. Additionally, we pre-process the pothole images using the filtering, enhancement, and segmentation methods. Moreover, we optimize the segmentation method using the metaheuristic GLA algorithm during the pre-processing stage. The simulation evaluation is done on the secondary dataset of potholes, which is available on the Kaggle website. This dataset contains two types of images, such as with and without pothole images. The evaluation is done using the performance parameters: accuracy, sensitivity, specificity, and F-score. The proposed roadside safety assessment model outperforms others and achieves an accuracy of 0.98985, sensitivity of 0.99248, specificity of 0.99248, and F-score of 0.98852. In the future, we will enhance the performance of the CNN by fine-tune the hyperparameters using the metaheuristic algorithm. Further, we will consider the real-time dataset for checking the robustness of the proposed model.

References

[1] Ma, N., Fan, J., Wang, W., Wu, J., Jiang, Y., Xie, L. and Fan, R., 2022. Computer vision for road imaging and pothole detection: a state-of-the-art review of systems and algorithms. *Transportation safety and Environment*, *4*(4), p.tdac026.

[2] Miller JS, Bellinger WY et al. Distress identification manual forthe long-term pavement performance program. United States.Federal Highway Administration. Office of Infrastructure Researchand Development, 2003.

[3] Fan R, Ozgunalp U, Hosking B et al. Pothole detection basedon disparity transformation and road surface modeling. *IEEETransactions on Image Processing* 2019;29:897–908.

[4] Heaton A. Potholes and more potholes: Is it just us?March 2018.URL: shorturl.at/mLN56

[5] Fan R, et al. Rethinking Road Surface 3-D Reconstruction and Pothole Detection: From Perspective Transformation to Disparity Map Segmentation. *IEEE Transactions on Cybernetics* 2022;52:(7) 5799–5808 doi:10.1109/TCYB.2021.3060461

[6] Fan R, Liu M. Road damage detection based on unsupervised disparity map segmentation. *IEEE Transactions on IntelligentTransportation Systems* 2019;21(11):4906–4911.

[7] Guildford J. Christchurch the pothole capital of New Zealand. February 2018. URL: shorturl.at/ayDP5

[8] Devine R. City of San Diego asking residents to report potholes. January 2017. URL: shorturl.at/gnLPV

[9] Koch C, Georgieva K, Kasireddy V et al. A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Advanced Engineering Informatics* 2015;29(2):196–210.

[10] Fan R, et al. Long-Awaited Next-Generation Road Damage Detection and Localization System is Finally Here. In: *2021 29th European Signal Processing Conference (EUSIPCO)*, 2021. p. 641–645. IEEE.

[11] Abdelmalak, M.E.S., Gaber, K.S., Ahmed, M.A., OubeBlika, N., Zaki, A.M. and Eid, M.M., 2023. BER-XGBoost: Pothole Detection based on Feature Extraction and Optimized XGBoost using BER Metaheuristic Algorithm. *Full Length Article*, *6*(2), pp.46-6.

[12] Anandhalli, M., Tanuja, A., Baligar, V.P. and Baligar, P., 2022. Indian pothole detection based on CNN and anchor-based deep learning method. *International Journal of Information Technology*, *14*(7), pp.3343-3353.

[13] *Pothole detection* 2020. https://www.kaggle.com/datasets/andrewmvd/pothole-detection.

[14] Wang, P., Fan, E., & Wang, P. (2021). Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognition Letters*, *141*, 61–67. https://doi.org/10.1016/j.patrec.2020.07.042

[15] Rahman, C.M., 2023. Group learning algorithm: a new metaheuristic algorithm. *Neural Computing and Applications*, *35*(19), pp.14013-14028.

[16] Basavaraju, A., Du, J., Zhou, F. and Ji, J., 2019. A machine learning approach to road surface anomaly assessment using smartphone sensors. *IEEE Sensors Journal*, *20*(5), pp.2635-2647.

[17] Dube, D., Sonawane, S., Nawale, S., Jangam, S. and Niyajali, J., Advancing Road Safety through Pothole Detection System: A Deep Learning Approach.

[18] Ferjani, I. and Alsaif, S.A., 2022. How to get best predictions for road monitoring using machine learning techniques. *PeerJ Computer Science*, *8*, p.e941.