

Deep CNN and twin support vector machine based model for detecting potholes in road network

Mohit Misra^{1,*}, Saptarshi Gupta¹ and Shailesh Tiwari²

¹Department of Electronics and Communication Engineering, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Ghaziabad, Uttar Pradesh 201204, India

²School of Computer Science Engineering & Technology, Bennett University, Greater Noida, Uttar Pradesh, 201310, India

*E-mail: misramohit8585@gmail.com

Received for publication April 14, 2025.

Abstract

Potholes are a persistent issue in road infrastructure that can be responsible for safety risks and economic burdens. For effective road maintenance and ensuring the safety of road user, it is essential to detect potholes quickly. It is noticed that traditional methods are time-consuming and labor-intensive task to repair the potholes. Recently, several machine learning (ML) techniques have been integrated to design an automated pothole detection model. These methods can detect anomalies that indicate the presence of potholes by examining the features and patterns of the road surface. Furthermore, these methods are integrated into existing road infrastructure and maintenance workflows, which enable proactive maintenance strategies and resource optimization. Hence, is the present study aims to explore the efficacy of the deep convolutional neural network (CNN) and twin support vector machine (TSVM) methods for accurate identification of the potholes in road infrastructure. The deep CNN method is applied to extract the relevant feature for the road image dataset, while the twin parametric support vector machine (TPSVM) method is employed for accurate detection of potholes. The performance of deep CNN and TPSVM combination is evaluated using several performance measures. The results indicate that deep CNN-TPSVM method achieves better results than existing models for detecting the potholes in road infrastructure.

Keywords

pothole detection, twin support vector machine, road infrastructure, deep convolutional network

1. Introduction

In transportation infrastructure, detecting and managing road potholes are critical concerns. Potholes, which are typically caused by road surface degradation over time, pose substantial risks to road users, resulting in accidents, vehicle damage, and increased maintenance expenses (Du et al., 2020; Singh et al., 2020). These issues are addressed through proactive efforts to ensure prompt discovery and correction. Traditional pothole detection systems are mostly based on physical inspections by road repair staff or residents reporting occurrences. However, these

methods are frequently time-consuming, inefficient, and subject to human mistakes (Pan et al., 2017; Zalama et al., 2013; Laurent et al., 2012; Cao et al. 2020). Furthermore, they may only cover part of the road network, resulting in unnoticed potholes and delayed repairs. In recent decades, the confluence of modern technologies, including machine learning (ML), computer vision, and the Internet of Things (IoT), has changed the pothole-detecting process (Dhiman & Klette, 2019; Hoang & Nguyen, 2018; Wu et al., 2020). The automated pothole systems integrated with sensors and cameras can detect and categorize potholes precisely and effectively, and

help authorities to prioritize repairs and improve road safety. Several strategies are used for real-time pothole detection and monitoring, including sensor-based systems, image processing algorithms, and data analytics techniques. Furthermore, the present study investigate the obstacles and alternatives related to the implementation of these technologies in various rural and urban areas. The aim of these cutting-edge methodologies and developing trends is to identify potholes on road and make global efforts to improve the dependability, sustainability, and safety of transportation infrastructure.

Asphalt roads are prevalent in modern transportation networks and act as vital conduits for automotive mobility. However, the structural integrity of these roads is constantly compromised by the appearance of potholes, which not only puts road safety but also involves enormous maintenance expenses for authorities. Hence, potholes on asphalt roads must be detected and addressed promptly to provide a smooth and safe journey for vehicles. Traditionally, the identification of potholes on asphalt roads has depended mainly on physical inspections, which are labor-intensive, time-consuming, and prone to error. Furthermore, because these checks are reactive in nature, potholes are frequently discovered after causing damage or an accident. To address these constraints, researchers and engineers have developed improved technology and novel approaches for automatic pothole identification on asphalt roads. Some prominent pothole detection approaches include 3D reconstruction and 3D laser-based scanning, vibration-based systems, and vision-based models. The benefits and drawbacks of these technologies include considerably greater performance at a higher cost (3D laser approaches), but dependability and accuracy are important challenges (vibration method). Consequently, real-time monitoring and recognition of potholes can considerably reduce the frequency of accidents and are realistic solutions to the problem. They also take time to build an accurate framework for mapping potholes throughout the road network. Thus, the goal of the present study is to identify potholes using 2D vision, and it focuses on road images to find potholes. However, detecting potholes is difficult because road images include a variety of shapes, sizes, shadows, and scales, as well as a complicated background. It has also been discovered that convolutional neural networks (CNNs) have been widely adopted for the accurate detection of potholes in road images, but there are several limitations, such as with high-resolution images, the training time of CNN increased, potholes are sometimes small in

size, discriminative features, and time-consuming. Hence, this study considers the two-fold approaches for accurate identification of potholes. The highlights of the present study are expressed below:

- An efficient pothole detection model based on the combination of the deep CNN and twin support vector machine (TSVM) is designed.
- Deep CNN model is utilized for the identification of the prominent features from the road images for the detection of potholes in road infrastructure.
- TSVM approach is considered for accurate identification of potholes in road infrastructure.
- The performance of the proposed pothole detection model is tested on a real-world road image dataset. This dataset comprises 11,150 images with binary classifications, i.e., pothole and no pothole. The 860 images belong to the pothole class, whereas the others belong to no pothole class.
- The simulation findings are assessed using the accuracy, precision, recall, F1-score, and AUC metrics. The accuracy of training and validation sets is also computed to investigate the overfitting issue.

The remainder of the present study is organized as follows: Section II discusses existing works on pothole identification and paving in road networks. Section III summarizes the proposed pothole detection model. Section IV summarizes the experimental findings of the proposed detection model. Section V concludes the study.

II. Related Works

This section discusses the related works reported in the literature on pothole detection in road infrastructure.

Tamagusko and Adelino presented a comparative analysis of different YOLO models for the accurate identification of potholes in asphalt roads. This study aimed to investigate the efficacy of computer vision methods for accurate detection of potholes. The results of this study revealed that YOLO version 4 achieves higher mean precision compared with other YOLO versions. This study also reported that YOLO version 5 also gets good results and has the potential to accurately identify potholes in road infrastructure.

Lopez et al. considered the deep neural network for the detection of potholes and speed bumps in roads. For accurate identification of potholes and speed bumps, this study considered the color

road images that are captured by the ZED camera mounted in front of the vehicles. The road images dataset comprises 714 images, and tagging is done manually either pothole or a bump. The deep neural network model is trained on 70% of the data, and the rest of the data is used to evaluate the performance of the model. The results stated that the deep neural network model obtains a 98.17% precision rate using 37 convolutional layers.

Real-time monitoring of the road network is a tedious task. To address this issue of road infrastructure, Xin et al. (2023) developed a crowdsourcing-based fusion approach for the identification of potholes, cracks, and other defects. The fusion approach combines the vehicle video data and accelerometer data for the detection of potholes and cracks. Further, a spatial clustering-based technique is also utilized for optimizing the potholes detection results. The results showed that the proposed model achieves higher accuracy compared with traditional methods.

Tahir and Jung (2023) utilized the distributed deep learning techniques for detecting the potholes on road networks. They revealed that the neural network models deployed on edge devices have compromised accuracy and precision rates with small datasets. This issue is addressed by the distributed model and data parallel techniques. Further, PyTorch and TensorFlow methods are employed for improving the detection rate. The aforementioned combination is implemented in Google cluster and Edge cluster environments. The results revealed that the proposed combination obtains superior results in terms of total loss and cross entropy.

Singh et al. (2023) considered the low accuracy issue of the pothole detection models due to low light conditions and presented an improved long short-term memory (ILSTM) model and one-dimensional local binary pattern (1D-LBP). In the proposed model, the road surface data are collected through accelerometer and gyroscope sensor data using a smartphone. The prominent features are extracted using the 1D-LBP technique, while ILSTM is utilized for detecting the potholes. The results revealed that the proposed model achieves a higher accuracy rate and less execution time compared with other existing methods.

Heo et al. (2023) presented an image-based pothole detection model that combines the multi-scale feature network and risk assessment. In the proposed model, the multi-scale feature network is designed using the SPFPN-YOLOv4 and CSPDarknet53-tiny. The spatial pyramid pooling is developed using SPFPN-YOLOv4, while the feature pyramid network is

obtained through CSPDarknet53-tiny. They claimed that SPFPN-YOLOv4 tiny improves the detection rate by an average of 2%–5% compared with other YOLO variants.

Potholes detection is one of the challenging tasks for driver assistance and autonomous vehicles. It is reported that current approaches frequently ignore water-filled and lighted potholes. Satti et al. (2024) considered these issues and developed a new approach based on a cascade classifier and a vision transformer. The task of the cascade classifier is to determine the pattern related to the road surface, while the vision transformer is utilized for the detection of potholes and analysis. The performance of the above combination is evaluated using four well-known pothole datasets and compared with a vast number of existing models. It is reported that the proposed combination achieves superior results in terms of precision, recall, and mAP.

Chougule and Barhatte (2023) developed a smart pothole detection model for the accurate identification of potholes in the road surface. In the proposed model, a camera is mounted in front of the vehicles for capturing the road images. Further, the deep learning models (i.e., YOLOv3 and YOLOv5) are utilized for detecting the potholes in images. The performance of the proposed model is assessed using the precision, recall, and average precision. The results stated that YOLOv5 obtains superior results than YOLOv3. The simulation results of YOLOv5 obtain 0.763, 0.548, and 0.635, respectively.

Zhao et al. (2024) presented a road damage detection model called MED-YOLOv8s and YOLOv8s for the detection of potholes and damages in the road surface. MobileNetv3 model is utilized for reducing the number of parameters and computations. Furthermore, an ultra-lightweight attention mechanism is adopted for improving the model's generalization performance. The results showed that the MED-YOLOv8s model obtains a 95.2% detection rate compared with YOLOv8s. It is also seen that model complexity is significantly reduced.

Lee et al. (2022) considered the accuracy issue of the pothole prediction models and developed a new prediction model based on the ML and deep learning models. This study also considered the independent features for the accurate prediction of potholes. Such features are minimum temperature, relative humidity, precipitation, and traffic volume. The findings of this study revealed that the mask-RCNN model provides superior results compared with other methods.

For ensuring the stability of vehicles and safety of drivers, Xu et al. (2023) presented a vision-IMU based

detection and ranging (VIDAR) model for detecting the cracks and potholes in road infrastructure. This model is the combination of the IMU to filter, mark, and frame for pothole detection. Recall and accuracy parameters are considered for evaluating the performance of the VIDAR model. The simulation results are also compared with traditional methods, and it is stated that the VIDAR model obtains higher precision and accuracy rates.

Vinodhini and Sidhaarth (2023) combined the CNN and transfer learning for accurate detection of potholes and cracks in the road surface. The aim of this combination is to assess the initial maintenance cost and address the repairs in the road surface. This study is also extended to send the message to drivers regarding the road surface. The efficacy of the above combination is evaluated using the accuracy parameter and compared with transfer learning-RNN and transfer learning-GAN. The results stated that the proposed combination obtains a 96% accuracy rate compared with the other two models.

III. Proposed Pothole Detection Model

This section discusses the proposed pothole detection model based on a deep CNN and TSVM. The proposed model contains four layers: (i) data layer, (ii) preprocessing layer, (iii) feature extraction layer, and (iv) detection layer. These layers are discussed in detail in next sections. Figure 1 illustrates the proposed pothole detection model based on a deep CNN and TSVM.

a. Data layer

This layer is responsible for collecting the data from the asphalt road, either in the context of pothole or no pothole. The real-world road images are captured using the smart camera in daytime. These images are labeled either pothole road images or no pothole road images. The final dataset contains 650 images with latitude and longitude information. A local folder is created on the system to store these images.

b. Data preprocessing layer

This subsection accomplishes the different preprocessing tasks associated with the road images. A Python environment is chosen to perform the preprocessing tasks. Initially, the images are resized

using the morphological operations. The image resizing is done by using the affine transformation. The size of the initial images is 512×512 , and it can be reduced to 64×64 using the affine transformation. It is also stated that most of the asphalt images have black color, but the present study considers the grayscale images for pothole detection. Hence, all road images are converted into grayscale images in which pothole is highlighted with black color and the rest of the area is highlighted with white color. Furthermore, the canny function is adopted to compute the edges of the pothole, and the region of interest is determined using Kapoor entropy. ROI is considered to extract the relevant features in the feature extraction phase. In the proposed model, the morphological operations are considered in the data-preprocessing layer, and the aim of these operations is to determine the shape or structure of features in an image. These operations are applied on the grayscale images of potholes to enhance or extract structural components. In pothole detection, erosion, dilation, opening, and closing are used. Erosion removes small white noises and shrinks the boundaries of foreground regions, and also helps eliminate irrelevant features. On the contrary, dilation can enlarge the white regions (e.g., pothole regions after thresholding), which helps in connecting broken parts of the detected pothole contour. Opening is useful for removing small objects or noise while preserving the shape of larger objects, making it ideal for cleaning up preprocessed images. Closing helps in filling small holes or gaps within the pothole region, which improves contour integrity. These operations are particularly useful before edge detection or contour extraction as they enhance the structural consistency of potholes in the image, facilitating accurate feature localization. Affine transformations are a set of geometric transformations that preserve lines and parallelism in pothole images. It includes a diverse set of operations like scaling, rotation, translation, shearing, and reflection. In pothole detection, scaling is especially important for resizing images to a uniform resolution while maintaining aspect ratios, which ensures consistent input dimensions for deep learning models. Translation is used to align features spatially, which is helpful when merging datasets with varying coordinate origins. Rotation helps the model become invariant to different orientations of potholes on the road, thereby improving generalizability. Additionally, shearing can be used for data augmentation by slightly distorting the images to simulate different camera angles and perspectives. These transformations not only standardize the data

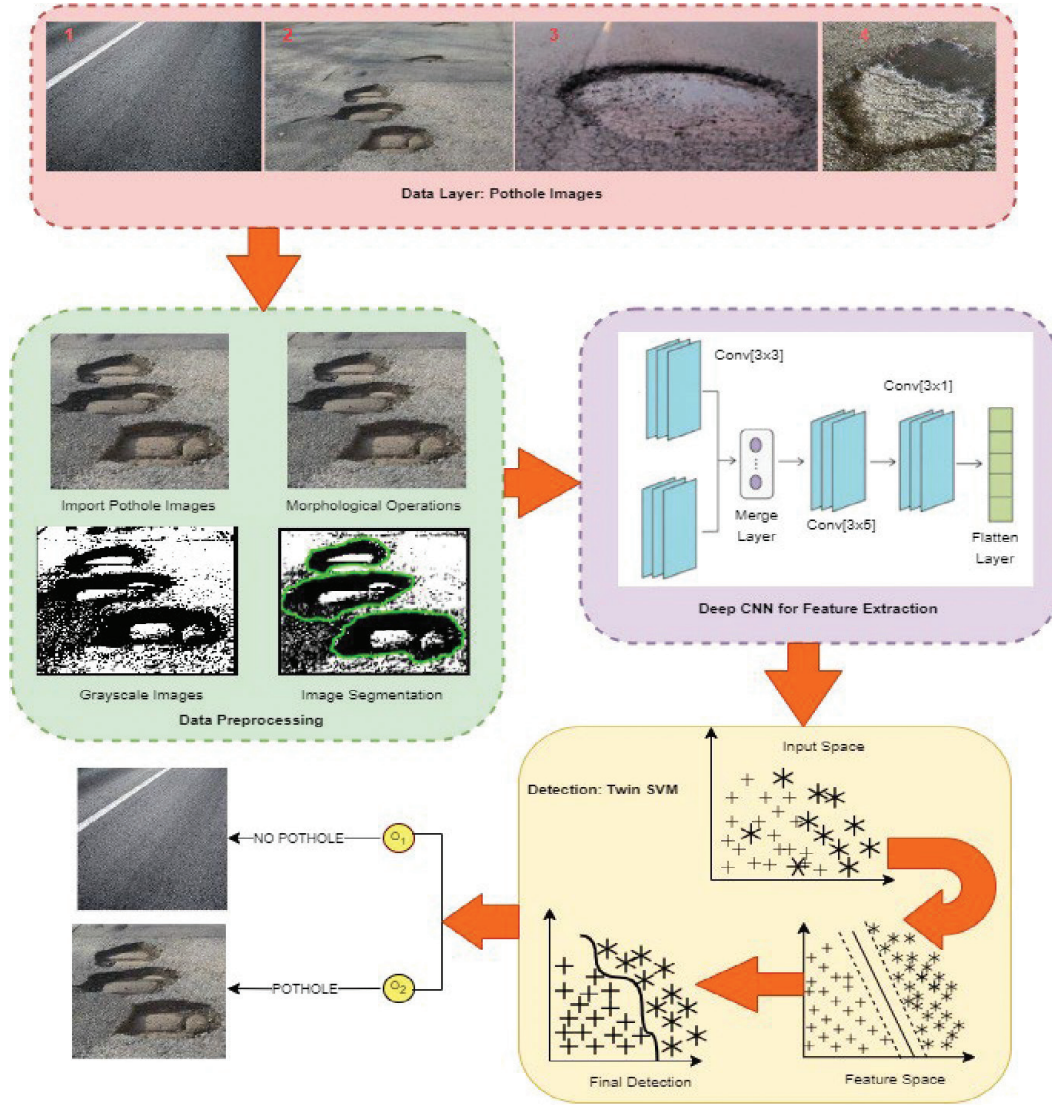


Figure 1: Illustrates the proposed pothole detection model based on deep CNN and TSVM. CNN, convolutional neural network; TSVM, twin support vector machine.

but also increase dataset diversity and enhance model robustness.

Kapoor entropy is a generalized form of entropy that measures the uncertainty or randomness within an image. It can consider the pixel intensities and also the underlying structure and spatial distributions for measuring the uncertainty. Unlike the other traditional entropy measures, such as Shannon entropy, Kapoor entropy also provides a more quantification of information content by introducing a parametric component that adjusts sensitivity to variations in pixel intensities. In our approach, Kapoor entropy is employed to dynamically assess the information in different sub-regions of the preprocessed image. The regions with higher Kapoor entropy values are considered to

have a higher degree of variability. It can be described as follows:

$$H(th_1, th_2, \dots, th_n) = H_0 + H_1 + \dots + H_n$$

$$H_0 = -\sum_{j=0}^{th_1-1} \frac{p_j}{w_0} \ln \frac{p_j}{w_0}, w_0 = \sum_{j=0}^{j=0} p_j$$

$$H_1 = -\sum_{j=th_1}^{th_2-1} \frac{p_j}{w_1} \ln \frac{p_j}{w_1}, w_1 = \sum_{j=th_1}^{th_2-1} p_j$$

$$H_n = -\sum_{j=th_n}^{L-1} \frac{p_j}{w_n} \ln \frac{p_j}{w_n}, w_n = \sum_{j=th_n}^{L-1} p_j$$

$$f_{\text{Kapoor}}(th_1, th_2, \dots, th_n) = \arg\max \{H(th_1, th_2, \dots, th_n)\} \quad (1)$$

c. Feature extraction using deep CNN

In this layer, a deep CNN model is considered to capture the relevant features from the road image dataset. To determine the relevant feature, this layer considers the segmented greyscale images as input. It is observed that the deep CNN model avoids the manual generation of the high-order features from the road image dataset. This model generates efficient and high-order features from the low-order feature inputs. Further, a feature extractor is designed with the help of the deep CNN model for extracting the effective features. This process is illustrated in Figure 2. By inspiring the wide popularity and applicability of the deep CNN model (Nogales & Benalcázar, 2023; Deepak & Ameer, 2022b; Khater & Gamel, 2023; Gagliardi et al., 2023), a feature extractor model is designed based on the small multiple convolutional kernels that are stacked to each other and can generate an extended receptive field. It is also

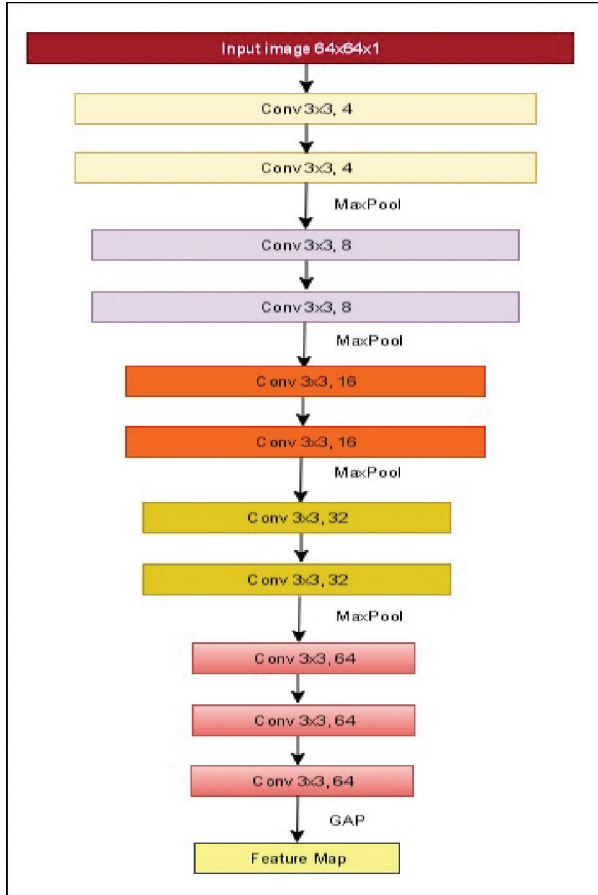


Figure 2: Proposed deep CNN-based feature extractor model. CNN, convolutional neural network.

assumed that such a network can increase the depth of the model and also non-linear transformations as compared to large convolutional kernels. It is also noticed that fewer parameters are required to learn the model. In our proposed feature extractor, 6 convolutional blocks are designed by combining the 13 convolutional layers. Further, the convolutional layers in a convolutional block produce the same quantity of feature maps and size of the output feature map, except last convolutional block. The last block consists of three cascading convolutional layers, while the rest of the blocks contain two stacked convolutional layers. These convolutional layers consider the 3×3 continuous convolutional kernel using stride 1. The feature maps for this work are defined as 8×8 , 16×16 , 32×32 , and 64×64 . Further, a max-pooling operation is applied with a 2×2 pixel window with stride 2. The task of this operation is to reduce the size of the feature map but preserve the information. The pooling operation also restricts the parameters as well as the computation of the proposed model. Finally, an average pooling operation is implemented to obtain the features using the input image size 64×64 . When comparing the average pooling operation to the conventional fully connected layer, the average pooling is less prone to overfitting and does not require any additional parameters. Finally, the last layer of the last convolutional block considers the softmax function, and the task of this function is to map the extracted feature with the probabilities of each feature vector. The feature extractor model is trained using the multi-class logarithmic loss function, and it can provide the discriminative features. In the training phase, the model employs the structural entropy graphs in the average pooling layer to generate feature characteristics. Consequently, the present study builds a trained feature extractor model that can take the input from the deep CNN model and the output from the average-pooling layer, respectively. The logarithmic loss function is defined using Eq. (2).

$$\text{Loss} = -\frac{1}{N} \sum_{m=1}^N \sum_{n=1}^M X_{m,n} \log_2 p_{m,n} \quad (2)$$

d. Detection layer: Twin parametric support vector machine (TPSVM)

The aim of the detection layer is to identify the potholes. For the identification of the potholes, this layer considers the relevant features extracted by the previous layer. Further, the potholes are detected using the TPSVM classifiers. The original TPSVM framework constructs two non-parallel hyperplanes by solving

two smaller quadratic programming problems and each involving fewer samples. While this enhances computational efficiency, it introduces challenges related to model generalization and margin robustness. In the context of road condition detection, we analyzed the influence of reduced training data size on classifier stability and accuracy. This implies that when TPSVM is efficient, a minimum subset threshold is necessary to maintain structural risk minimization and margin fidelity. It is noticed that the TPSVM classifier provides better state-of-the-art classification results for imbalanced dataset and is widely used in the literature for classification tasks. TPSVM is an extension of the TSVM that computes the separate hyperplane using the two non-parallel parametric hyperplanes. It can be achieved by using the two small-sized support vector machines (SVMs), which are described below.

$$f_a(x) = (w_a \times x) + b_a = 0; \text{ and } f_c(x) = (w_c \times x) + b_c = 0 \quad (3)$$

Assume, a binary classification problem that can divide the data items d_1 and d_2 such as data item d_1 belongs to class a, while data item d_2 belong to class c in the given n -dimensional space (S). Suppose the matrix (M) in $M^{d_1 \times n}$ denotes the data items belong to class a, the matrix (N) in $N^{d_2 \times n}$ denotes the data items belong to class c. Now, the TPSVM that allocates the data in two classes must follow the Eqs (4) and (5).

$$(w_a \times x) + b_a \geq 0 \text{ such that } i = 1, 2, 3, \dots, d_1 \quad (4)$$

$$(w_c \times x) + b_c \geq 0 \text{ such that } i = 1, 2, 3, \dots, d_2 \quad (5)$$

Further, the pair of the parametric margin of the hyperplane is obtained using the following constraint-based optimization problem, which is mentioned in Eqs (6) and (7).

$$\begin{aligned} \min_{w_a, b_a, \xi} & \left(\frac{1}{2} \|w_a\|^2 + \frac{p_1}{d_2} e_2^T (Nw_a + e_2 b_a) + \frac{v_1}{d_1} e_1^T \xi \right) \\ \text{s.t.} & (Mw_a + e_2 b_a \geq 0 - \xi) \text{ such that } \xi > 0 \quad e_1 \end{aligned} \quad (6)$$

and

$$\begin{aligned} \min_{w_c, b_c, \eta} & \left(\frac{1}{2} \|w_c\|^2 + \frac{p_2}{d_1} e_1^T (Nw_c + e_1 b_c) + \frac{v_2}{d_2} e_2^T \eta \right) \\ \text{s.t.} & (Mw_c + e_1 b_c \geq 0 - \eta) \text{ such that } \eta > 0 \quad e_2 \end{aligned} \quad (7)$$

In Eqs (5) and (6), ξ and η can be described as the slack variables. The variables $p_1, p_2 > 0$; $v_1, v_2 > 0$

are defined as regularization parameters that are utilized to compute the penalty weight. e_1 and e_2 are defined as the vectors of appropriate dimension. Finally, Eq. (7) is used to allocate the data object (o) that belongs to either class a or b.

$$\arg \min_{k=1,2} \frac{|w_k \cdot x + b_k|}{\|w_k\|} \quad (8)$$

Apart from above, the regularization parameters in TPSVM serve key roles such as p_1 and p_2 penalize misclassified points and control the tradeoff between margin maximization and training error for each hyperplane. While, v_1 and v_2 scale the contributions of each class to prevent class imbalance from biasing the optimization process.

IV. Experimental Results and Discussion

The simulation results of the present study are discussed in this section. The benchmark road pothole dataset is considered to evaluate the performance of the proposed model. Several well-known parameters, such as accuracy, precision, recall, and F1-score, are considered to determine the efficacy of the proposed model. The simulation results of the proposed model are also compared with several well-known classifiers/models. The proposed model is implemented in MATLAB environment with Windows operating system, 16 GB RAM, and a Core i7 processor. The details of the performance parameters are discussed in subsection "Performance parameter," while the subsection "Results and discussion" discusses the simulation results of the proposed model and other existing classifiers/models.

a. Performance parameter

This subsection describes the performance parameters that are used in the present study. It is noticed that the research community widely adopts these parameters for evaluating the performance of the pothole detection models. The description of these parameters are summarized as

Accuracy rate is described as the correctly classified images with respect to the total number of images can be defined as the number of images. It can be evaluated using Eq. (9).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \quad (9)$$

In Eq. (7), TP is true positive, TN is true negative, FP is false positive, and FN is false negative. TP and TN describe the correctly classified images. Whereas incorrectly classified images are described by FP and FN.

Recall parameter gives the number of correctly classified images with respect to all positive labeled images. It is evaluated using Eq. (10).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (10)$$

Precision gives correctly classified images with respect to all correctly classified images (TP and FP). It is determined using Eq. (11).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100 \quad (11)$$

F1-score can be described as a balanced parameter that can be used to evaluate the efficacy of the model, and it is computed using precision and recall. Eq. (12) is used to compute the F1-score.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (12)$$

b. Results and discussion

The experimental results of the proposed deep CNN-TPSVM model and other existing techniques are discussed in this section. Several well-known parameters, such as accuracy, precision, recall, and F1-score, are considered for evaluating the efficiency of the proposed deep CNN-TPSVM model. Before evaluating the simulation parameters, the confusion matrix is also computed for each technique. Before proceeding the experimentation, the pothole image dataset is divided into a training set and a validation set. Further, 70% of the pothole images dataset is considered for the training set, while the rest of the images dataset is considered for the validation set. The present study also computes the accuracy and loss rates of the training and validation sets. AUC parameter is also considered to evaluate the efficacy of the proposed deep CNN-TPSVM model along with the rest of the techniques. Figure 3 depicts the confusion matrix of the proposed deep CNN-TPSVM model. It is also mentioned that the pothole dataset consists of 1,150 images, in which 860 images are labeled as pothole images and the rest of the images are labeled as no pothole images. Figure 3 illustrated that the proposed

Confusion Matrix Deep CNN-SVM					
Actual Label	841		19		
	16		274		
	Pot	hole	No	Pot	hole
Predicated Label					

Figure 3: Depicts the confusion matrix of proposed deep CNN-TPSVM model. CNN, convolutional neural network; TPSVM, twin parametric support vector machine.

deep CNN-TPSVM model accurately identifies 841 pothole images out of 860 images. While the proposed model also identifies 16 images as pothole images, but in actual these images can be labeled as no pothole images. Figure 4 depicts the confusion matrix of the DBN, Inceptionv3, Vgg16, VGG19, ANN, and SVM techniques. It is observed that the DBN model identifies the pothole images more accurately among these models.

The experimental results of the proposed deep CNN-TPSVM model are illustrated in Table 1. The results are evaluated based on accuracy, F1-score, recall, and precision parameters. The results are also compared with popular techniques existing in the literature such as ANN, VM, VGG16, VGG19, Inceptionv3, and DBN. By analyzing the results, it is found that the proposed deep CNN-TPSVM model achieves a higher accuracy rate (96.12%) compared with other existing techniques. It is also seen that ANN exhibits a lower accuracy rate (79.73%) among all techniques. By analyzing the variants of the CNN models (VGG16, Vgg19, InceptionV3, and DBN), it is reported that the DBN model obtains higher accuracy (93.04%) among these variants. The proposed deep CNN-TPSVM model also obtains good results in terms of precision and recall parameters compared with other techniques. The proposed model has higher recall and precision rates (98.13% and 97.79%). Whereas, recall and precision rates of the DBN model are 94.30% and 96.31%, respectively. It is also stated that the precision and recall rates (85.82% and 87.32%) of the ANN technique are lower among all. Moreover, the F1-score parameter is also considers to validate the performance of the proposed deep CNN-TPSVM model. This parameter can be acted as a balance parameter that considers the only positive instance of the dataset. Hence, this parameter is evaluated using



Figure 4: Depicts the confusion matrix of other existing models.

Table 1: Depicts the simulation results of the proposed deep CNN-TPSVM model and other existing models

Technique	Accuracy (%)	Recall (%)	Precision (%)	F1-score (%)
ANN	79.73	87.32	85.82	86.56
SVM	82.60	88.83	88.01	90.31
VGG16	84.43	90.58	88.82	90.99
VGG19	87.65	91.97	91.55	93.87
InceptionV3	90.06	92.55	94.87	95.69
DBN model	93.04	94.30	96.31	96.92
Proposed deep CNN-TPSVM Model	96.12	98.13	97.79	97.96

CNN, convolutional neural network; SVM, support vector machine; TPSVM, twin parametric support vector machine.

the TP, TN, and FP. Hence, this parameter specifies the goodness of the model and is more significant than the accuracy parameter. F1-score parameter results are also summarized in Table 1. It is seen that the proposed deep CNN-TPSVM model gets a higher F1-score rate (97.96%) than other techniques. On the contrary, ANN obtains a lower F1-score rate (86.56%) among all techniques. It is also stated that the proposed deep CNN-TPSVM model obtains better F1-score rate than an accuracy rate for the detection of potholes. Hence, it is said that the proposed model is a more significant and effective mode for detecting the potholes in road network. Figure 5 presents the comparative results of all techniques based on the accuracy, precision, recall, and F1-score parameters. The results illustrate the efficiency of the proposed deep CNN-TPSVM model in terms of accuracy, recall, precision, and F1-score. It is also observed that the F1-score of the proposed CNN-TPSVM model is better than its accuracy rate. Hence, it is said that the proposed deep CNN-TPSVM model is more prominent than other models.

The present study also considers accuracy and loss rates of the training and validation sets. The accuracy rate of the training and validation sets, along with epoch, is depicted in Figure 6. The accuracy rate of both sets is represented using a different color scheme. It is found that the accuracy rate of validation set is higher than the training set. The accuracy rate of the validation set is 0.958, whereas the accuracy rate of the training set is 0.921. The loss rate of the training and validation sets is depicted in Figure 7. It is seen that the loss rate of the validation set is lower than the training set. The loss rate of the validation set is 0.042, whereas the loss rate of the training set is 0.079. The accuracy and loss rates of the training and validation sets are used to understand the overfitting issue of the model. So, it is found that the loss rate of the validation set is lower than the training set. Hence, the proposed deep CNN-TPSVM model cannot over fit the data, and the overfitting issue of the binary classification problem is successfully handled by the proposed model. In the present study, the AUC parameter is also computed for the proposed

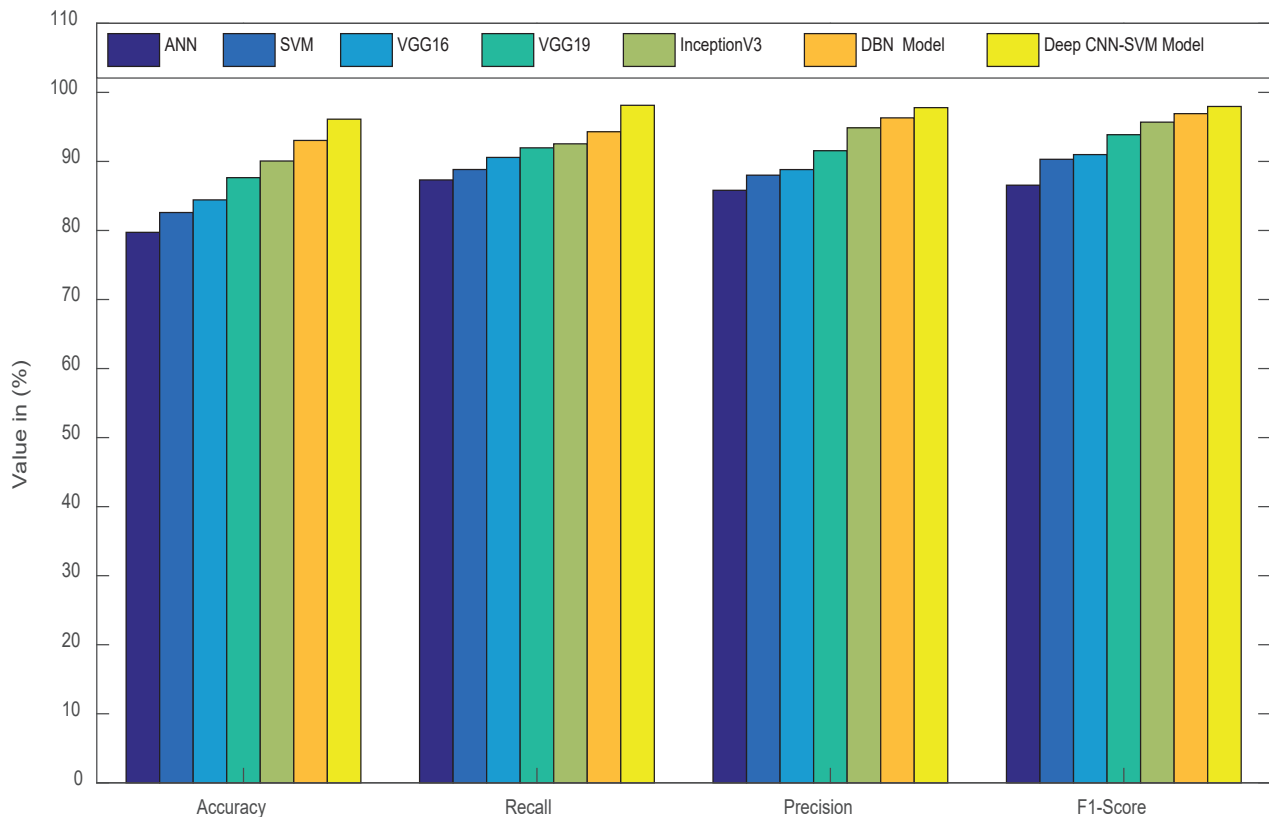


Figure 5: Comparative analysis of the results using proposed deep CNN-TPSVM model and existing models based on different performance parameters. CNN, convolutional neural network; TPSVM, twin parametric support vector machine.

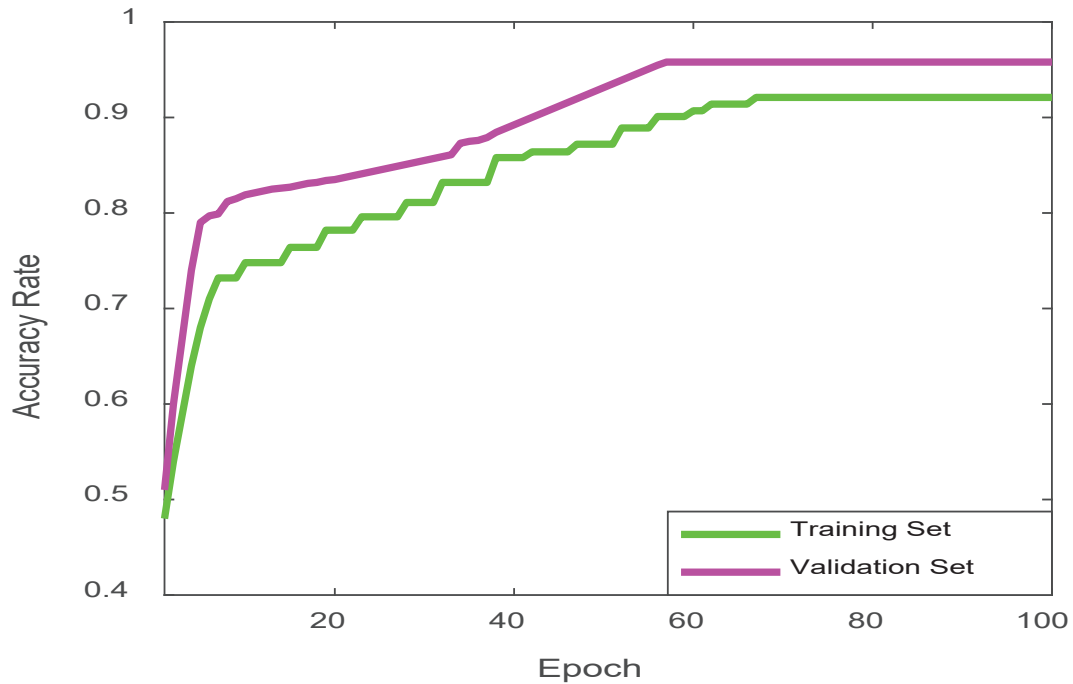


Figure 6: Demonstrates the accuracy rate of the proposed deep CNN-TPSVM model based on training and validation sets. CNN, convolutional neural network; TPSVM, twin parametric support vector machine.

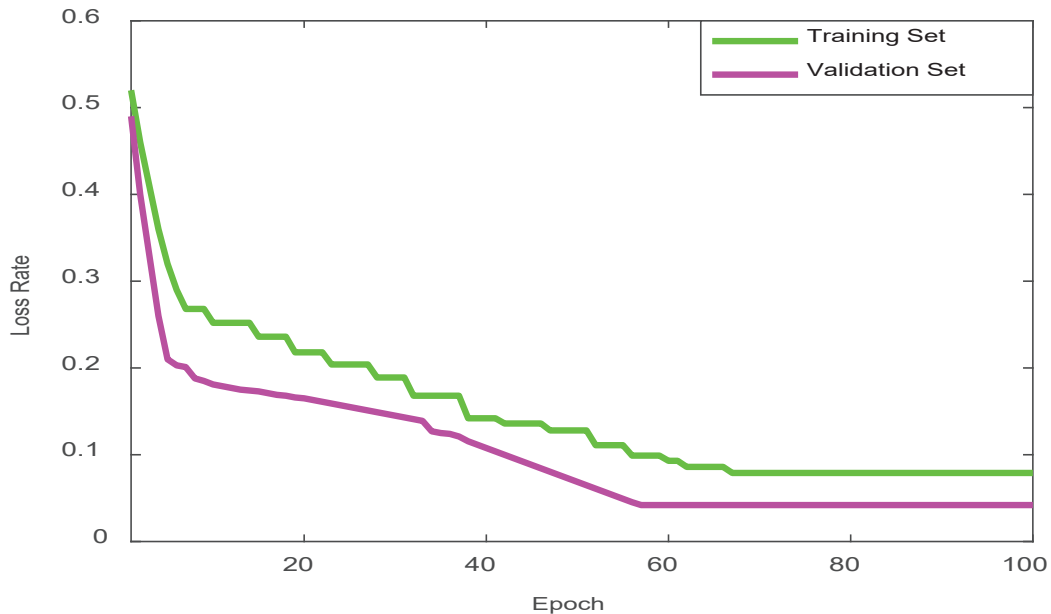


Figure 7: Demonstrates the loss rate of the proposed deep CNN-TPSVM model based on training and validation sets. CNN, convolutional neural network; TPSVM, twin parametric support vector machine.

deep CNN-TPSVM model and other techniques. The main reason for the selection of the AUC parameter is the nature of the pothole dataset, as the pothole

dataset contains imbalanced data, i.e., 860 images of pothole and 260 images of no pothole classes. In case of imbalanced dataset, the AUC parameter is

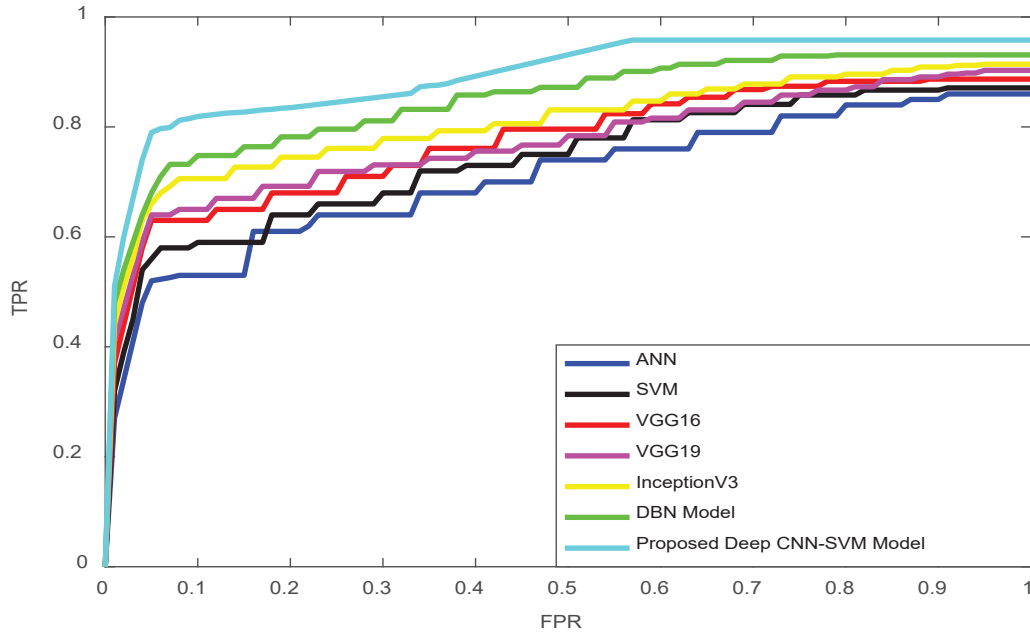


Figure 8: Depicts the AUC of proposed deep CNN-TPSVM model and other existing techniques. CNN, convolutional neural network; TPSVM, twin parametric support vector machine.

one of the significant parameters for evaluating the efficiency of the model. It can be computed by TP and FP rates. The results of the AUC parameter are illustrated in Figure 8, and it is stated that the proposed deep CNN-TPSVM model obtains a better AUC rate compared with other techniques.

V. Conclusion

The detection of potholes in the road network is a persistent challenge task including safety risks and economic burdens. But, it is a crucial task to ensure effective maintenance and also ensure the safety of road users. The present study advocates the combination of the deep CNN and TPSVM for accurate detection of potholes. In the proposed combination, the deep CNN method is utilized to extract the relevant features from the road image dataset, whereas, TPSVM technique is considered for detecting the potholes. The effectiveness of the deep CNN-TPSVM is explored based on the well-known performance measures and compared with existing pothole detection models. The results showed that the aforementioned combination reveals superior results compared with existing models up to 5%. The performance parameters such as AUC, accuracy, and loss rates of training and validation sets also demonstrate the effectiveness of the proposed deep CNN-TPSVM model in detecting potholes in the road network.

Finally, it can be concluded that the combination of deep CNN and TPSVM is a promising approach for automated pothole detection that can offer an efficient and accurate solution for addressing maintenance issues in the road network.

Acknowledgment

The authors would like to express their gratitude to SRM Institute of Science and Technology. The research was conducted independently without any external funding or institutional support.

Conflict of Interest

The authors declare no conflict of interests.

References

- Buza, E., Omanovic, S., & Huseinovic, A. (2013, October). Pothole Detection with Image Processing and Spectral Clustering. In proceedings of the 2nd International Conference on Information Technology and Computer Networks (Vol. 810, P. 4853).
- Cao, M., Tran, Q., Nguyen, N., & Chang, K. (2020). Survey on performance of deep learning models for

detecting road damages using multiple dashcam image resources. *Advanced Engineering Informatics*, 46, 101182. DOI: 10.1016/j.aei.2020.101182

Chougule, S., & Barhatte, A. (2023). Smart Pothole Detection System using Deep Learning Algorithms. *International Journal of Intelligent Transportation Systems Research*, 21(3), 483–492. DOI: 10.1007/s13177-023-00363-3

Deepak, S., & Ameer, P. (2022b). Brain tumor categorization from imbalanced MRI dataset using weighted loss and deep feature fusion. *Neurocomputing*, 520, 94–102. DOI: 10.1016/j.neucom.2022.11.039

Dhiman, A., & Klette, R. (2019). Pothole detection using computer vision and learning. *IEEE Transactions on Intelligent Transportation Systems*, 21(8), 3536–3550. DOI: 10.1109/tits.2019.2931297

Du, R., Qiu, G., Gao, K., Hu, L., & Liu, L. (2020). Abnormal road surface recognition based on smartphone acceleration sensor. *Sensors*, 20(2), 451. DOI: 10.3390/s20020451

Eriksson, J., Girod, L., Hull, B., Newton, R., Madden, S., & Balakrishnan, H. (2008). The pothole patrol. DOI: 10.1145/1378600.1378605

Fan, R., Ai, X., & Dahnoun, N. (2018). Road surface 3D reconstruction based on dense subpixel disparity map estimation. *IEEE Transactions on Image Processing*, 27(6), 3025–3035. DOI: 10.1109/tip.2018.2808770

Gagliardi, V., Giammorcaro, B., Francesco, B., & Sansonetti, G. (2023). Deep neural networks for asphalt pavement distress detection and condition assessment. In *Earth Resources and Environmental Remote Sensing/GIS Applications XIV*, 35. DOI: 10.1117/12.2688512

Gopalakrishnan, K., Khaitan, S. K., Choudhary, A., & Agrawal, A. (2017). Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials*, 157, 322–330. DOI: 10.1016/j.conbuildmat.2017.09.110

Heo, D., Choi, J., Kim, S., Tak, T., & Zhang, S. (2023). Image-Based pothole detection using Multi-Scale feature network and risk assessment. *Electronics*, 12(4), 826. DOI: 10.3390/electronics12040826

Hoang, N., & Nguyen, Q. (2018). A novel method for asphalt pavement crack classification based on image processing and machine learning. *Engineering With Computers*, 35(2), 487–498. DOI: 10.1007/s00366-018-0611-9

Khater, H. A., & Gamel, S. A. (2023). Early diagnosis of respiratory system diseases (RSD) using deep convolutional neural networks. *Journal of Ambient*

Intelligence and Humanized Computing, 14(9), 12273–12283. DOI: 10.1007/s12652-023-04659-w

Koch, C., & Brilakis, I. (2011). Pothole detection in asphalt pavement images. *Advanced Engineering Informatics*, 25(3), 507–515. DOI: 10.1016/j.aei.2011.01.002

Laurent, J., Hébert, J. F., Lefebvre, D., & Savard, Y. (2012). Using 3D laser profiling sensors for the automated measurement of road surface conditions. In *Rilem bookseries* (pp. 157–167). DOI: 10.1007/978-94-007-4566-7_16

Lee, S., Le, T. H. M., & Kim, Y. (2022). Prediction and detection of potholes in urban roads: Machine learning and deep learning based image segmentation approaches. *Developments in the Built Environment*, 13, 100109. DOI: 10.1016/j.dibe.2022.100109

Li, S., Yuan, C., Liu, D., & Cai, H. (2016). Integrated processing of image and GPR data for automated pothole detection. *Journal of Computing in Civil Engineering*, 30(6). DOI: 10.1061/(asce)cp.1943-5487.0000582

Moosaei, H., Ganaie, M., Hladík, M., & Tanveer, M. (2022). Inverse free reduced universum twin support vector machine for imbalanced data classification. *Neural Networks*, 157, 125–135. DOI: 10.1016/j.neunet.2022.10.003

Nogales, R. E., & Benalcázar, M. E. (2023). Hand Gesture Recognition Using Automatic Feature Extraction and Deep Learning Algorithms with Memory. *Big Data and Cognitive Computing*, 7(2), 102. DOI: 10.3390/bdcc7020102

Pan, Y., Zhang, X., Tian, J., Jin, X., Luo, L., & Yang, K. (2017). Mapping asphalt pavement aging and condition using multiple endmember spectral mixture analysis in Beijing, China. *Journal of Applied Remote Sensing*, 11(1), 016003. DOI: 10.1117/1.jrs.11.016003

Peralta-López, J., Morales-Viscaya, J., Lázaro-Mata, D., Villaseñor-Aguilar, M., Prado-Olivarez, J., Pérez-Pinal, F., Padilla-Medina, J., Martínez-Nolasco, J., & Barranco-Gutiérrez, A. (2023). Speed Bump and Pothole Detection Using Deep Neural Network with Images Captured through ZED Camera. *Applied Sciences*, 13(14), 8349. DOI: 10.3390/app13148349

Pereira, V., Tamura, S., Hayamizu, S., & Fukai, H. (2018). Pereira, V., Tamura, S., Hayamizu, S., & Fukai, H. (2018, July). A deep learning-based approach for road pothole detection in timor leste. In *2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)* (pp. 279–284). IEEE. In *IEEE International Conference on Service Operations and Logistics, and Informatics*. DOI: 10.1109/soli.2018.8476795

- Satti, S. K., Rajareddy, G. N. V., Mishra, K., & Gandomi, A. H. (2024). Potholes and traffic signs detection by classifier with vision transformers. *Scientific Reports*, 14(1). DOI: 10.1038/s41598-024-52426-4
- Singh, P., Bansal, A., & Kumar, S. (2020). Performance Analysis of various Information Platforms for recognizing the quality of Indian Roads. 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 63–76. DOI: 10.1109/confluence47617.2020.9057829
- Singh, P., Kamal, A. E., Bansal, A., & Kumar, S. (2023). Road pothole detection from smartphone sensor data using improved LSTM. *Multimedia Tools and Applications*, 83(9), 26009–26030. DOI: 10.1007/s11042-023-16177-0
- Tahir, H., & Jung, E. (2023). Comparative study on distributed lightweight deep learning models for road pothole detection. *Sensors*, 23(9), 4347. DOI: 10.3390/s23094347
- Tamagusko, T., & Ferreira, A. (2023). Optimizing Pothole Detection in Pavements: A Comparative Analysis of Deep Learning Models. *Optimizing Pothole Detection in Pavements: A Comparative Analysis of Deep Learning Models*, 57, 11. DOI: 10.3390/engproc2023036011
- Tsai, Y., & Chatterjee, A. (2017). Pothole detection and classification using 3D technology and watershed method. *Journal of Computing in Civil Engineering*, 32(2). DOI: 10.1061/(asce)cp.1943-5487.0000726
- Vinodhini, K. A., & Sidhaarth, K. R. A. (2023). Pothole detection in bituminous road using CNN with transfer learning. *Measurement Sensors*, 31, 100940. DOI: 10.1016/j.measen.2023.100940
- Wang, H., Chen, C., Cheng, D., Lin, C., & Lo, C. (2015). A Real-Time pothole detection approach for intelligent transportation system. *Mathematical Problems in Engineering*, 2015, 1–7. DOI: 10.1155/2015/869627
- Wang, H., Zhu, J., & Feng, F. (2023). Elastic net twin support vector machine and its safe screening rules. *Information Sciences*, 635, 99–125. DOI: 10.1016/j.ins.2023.03.131
- Wu, C., Wang, Z., Hu, S., Lepine, J., Na, X., Ainalis, D., & Stettler, M. (2020). An Automated Machine-Learning approach for road pothole detection using smartphone sensor data. *Sensors*, 20(19), 5564. DOI: 10.3390/s20195564
- Xin, H., Ye, Y., Na, X., Hu, H., Wang, G., Wu, C., & Hu, S. (2023). Sustainable road pothole Detection: a crowdsourcing based Multi-Sensors fusion approach. *Sustainability*, 15(8), 6610. DOI: 10.3390/su15086610
- Xu, Y., Sun, T., Ding, S., Yu, J., Kong, X., Ni, J., & Shi, S. (2023). VIDAR-Based Road-Surface-Pothole-Detection Method. *Sensors*, 23(17), 7468. DOI: 10.3390/s23177468
- Xue, G., Zhu, H., Hu, Z., Yu, J., Zhu, Y., & Luo, Y. (2016). Pothole in the Dark: Perceiving Pothole Profiles with Participatory Urban Vehicles. *IEEE Transactions on Mobile Computing*, 16(5), 1408–1419. DOI: 10.1109/tmc.2016.2597839
- Yu, G., Ma, J., & Xie, C. (2022). Hessian scatter regularized twin support vector machine for semi-supervised classification. *Engineering Applications of Artificial Intelligence*, 119, 105751. DOI: 10.1016/j.engappai.2022.105751
- Zalama, E., Gómez-García-Bermejo, J., Medina, R., & Llamas, J. (2013). Road crack detection using visual features extracted by Gabor filters. *Computer-Aided Civil and Infrastructure Engineering*, 29(5), 342–358. DOI: 10.1111/mice.12042
- Zhang, A., Wang, K. C. P., & Ai, C. (2017). 3D shadow modeling for detection of descended patterns on 3D pavement surface. *Journal of Computing in Civil Engineering*, 31(4). DOI: 10.1061/(asce)cp.1943-5487.0000661
- Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016). Road crack detection using deep convolutional neural network. 2022 IEEE International Conference on Image Processing (ICIP), 3708–3712. DOI: 10.1109/icip.2016.7533052
- Zhao, M., Su, Y., Wang, J., Liu, X., Wang, K., Liu, Z., Liu, M., & Guo, Z. (2024). MED-YOLOv8s: a new real-time road crack, pothole, and patch detection model. *Journal of Real-Time Image Processing*, 21(2). DOI: 10.1007/s11554-023-01405-5