Engineered Science
DOI: https://dx.doi.org/10.30919/es1626

Label-Efficient Self-Supervised Representation Learning-based Road Pothole Detection from Thermal Images

Deepika Vikas Agrawal,[1,*] Varun Gupta,[2] C Rama Krishna,[1] Muskaan Chopra[2] and Abhinav Puri[2]

Abstract

Potholes represent one of the leading causes of road accidents and vehicle damage globally, contributing to issues such as punctured tyres and wheel damage. Prompt and accurate detection of potholes through automated systems is crucial for ensuring road safety and infrastructure integrity. However, traditional pothole detection techniques that predominantly utilize supervised learning require extensive human-labelled datasets for training, which are both expensive and time-consuming. To address this challenge, we propose self-supervised representation learning (SSL) - based approach for pothole detection is reported and demands fewer labelled data. Our proposed approach involves two main phases: an initial pretext phase and a subsequent downstream fine-tuning phase. In the pretext phase, the model autonomously extracts robust and informative image representations from the unlabelled dataset of pothole images. In the downstream phase, the model is fine-tuned on a smaller, labelled dataset to specifically address thermal pothole detection. Extensive experimental evaluations were performed with varying proportions of labelled data (10%, 50%, and 100%), achieving accuracies of 95.71%, 98.29% and 99.2% respectively. The findings clearly demonstrate that the proposed SSL-based approach outperforms the conventional and existing supervised learning-based methods of pothole detection. Notably, the self-supervised learning-based model trained on 10% or 50% of the labelled data surpasses fully supervised models trained on the entire dataset.

Keywords: Potholes detection; Self-supervised learning; Supervised learning; Residual networks; Thermal imaging.
Received: 10 March 2025; Revised: 08 July 2025; Accepted: 08 July 2025
Article type: Research article.

1. Introduction

Road networks constitute the backbone of transportation infrastructure and significantly influence economic development and societal well-being. Ensuring their consistent maintenance and optimal condition remains a global challenge due to relentless traffic loads, environmental factors, and material degradation, all of which contribute to road surface deterioration. One prevalent appearance of this deterioration is the formation of potholes, hollow, surface-level cracks that pose considerable hazards if not promptly detected and repaired. Undetected potholes pose severe threats, including an elevated risk of traffic accidents, substantial vehicle damage, and increased fuel consumption. Annually, potholes contribute to thousands of vehicle damages and numerous traffic accidents worldwide. In India, despite prolonged lockdowns due to the COVID-19 pandemic in 2021, the National Crime Records Bureau (NCRB) reported a 17% rise in road fatalities, highlighting potholes as a critical factor influencing road safety.[1] Unfortunately, traditional methods for detecting potholes continue to be reliant on manual assessment methods and inspections, a labour-intensive and time-consuming process fraught with limitations and inefficiencies. Recent advancements in machine learning offer promising alternatives through automated pothole detection. Nonetheless, existing automated approaches predominantly depend on supervised learning techniques, which necessitate extensive automated datasets. Collecting high-quality, accurately labelled datasets is both resource-intensive and costly, posing significant practical barriers.

Numerous studies have attempted to address these challenges through supervised and deep learning-based pothole detection methods. Considerable research has been conducted in pothole classification, localization and segmentation. For instance, S. Gupta et al,[2] J. Liu *et al.*[3] and Yousaf et al[4] proposed a convolutional neural network (CNN)

[1]*Department of Computer Science and Engineering, National Institute of Technical Teachers Training and Research, Chandigarh, 160019, India*
[2]*Department of Computer Science and Engineering, Chandigarh College of Engineering and Technology, Chandigarh, 160019, India*
[*]*Email:* deepika.cse22@nitttrchd.ac.in (Deepika Vikas Agrawal)

Table 1: Pothole Detection Approaches.

Reference	Approach	Dataset	Accuracy	Remarks
Patra *et al.*[27]	Custom CNN (PotSpot)	3,424 real-world images	97.6%	Demonstrates high accuracy at scale
Gagliardi *et al.*[17]	Lightweight CNN	Diverse road conditions custom dataset	90-93%	Emphasizes efficiency suitable for resource-constrained deployments
Aparna *et al.*[18]	CNN (thermal imaging data)	Pothole Thermal Images	97.08%	Leveraging thermal data may enhance feature distinctiveness but requires specialized sensors.
Chu *et al.*[5]	Custom CNN	Pothole Images from Lahore roads	97%	High accuracy on a region-specific dataset potential for domain adaptation elsewhere
Yishun Li *et al.*[19]	Pretrained EfficientNet	Six pavement distress types	~86% (average)	Covers multiple distress classes, though average accuracy suggests further fine-tuning might be needed
N. Cinar *et al.*[20]	Pretrained DenseNet121	67 images (limited validation dataset)	Above 99%	Extremely high accuracy but limited by small data size caution on generalizability and computational cost
Ahmad *et al.*[21]	MobileNet v2, ResNet50, ResNet18	Aggregated datasets (online + real-world)	Up to 98%	Shows strong performance with varied data sources, emphasizes the benefit of model choice
Li *et al.*[22]	Multiple pre-trained CNN(Res Net-50 best)	Six road distress types	96.24%	ResNet-50 outperforms others in a multi-class scenario, reaffirming general strong performance.
Gupta *et al.*[23]	ResNet-based model	RDD2020 dataset	94.79%	Demonstrates consistent performance on a well-known benchmark (RDD2020)
A. Egaji *et al.*[24]	Random Forest (among ML classifiers)	Mobile devices & navigation maps data	94.44%	Random Forest outperforms other ML algorithms simpler than deep learning but is still effective
S. Bhatlawande *et al.*[9]	Classical ML classifiers	Custom Dataset	88%	Illustrates baseline approaches room for improvement using more advanced deep models
M. Asmadi *et al.*[25]	KNN-based method	200 real-world images	95.83%	Shows competitive performance with minimal data, yet KNN can be sensitive to dataset size and feature selection

based localization model for potholes while Chu et al [5] developed a localization technique for real-time deployment. Bucko et al[6] and Asad *et al.*[7] used YOLOv3 and YOLOv5, respectively, for pothole detection. Jakubec *et al.*[8] enhanced YOLOv8 with GAN-based augmentation, improving robustness at the cost of higher complexity. For segmentation, Bhatlawande *et al.*[9] utilized classical segmentation techniques, whereas Asmadi *et al.*[10] explored KNN-based models showing promising accuracy on limited data. Z. Feng *et al.*[11] and J. Fan *et al.*[12] used attention fusion network combining RGB and disparity images for pothole segmentation on Pothole-600. C. Liu *et al.*[13] adopted Faster R-CNN and FCN for segmenting pavement distresses, though their evaluation was limited to daytime images. U. K. *et al.*[14] proposed the Mask R-CNN model for dry/wet pothole segmentation, while V. Pereira *et al.*[15] and J. Guan *et al.*[16] used U-Net for semantic segmentation. As the present work explicitly targets pothole classification, this section provides a comprehensive method that predominantly utilizes customized convolutional neural networks (CNNs) or pre-trained network architectures for the classification of potholes.

1.1 Customized CNN-based methods

Patra *et al.* proposed PotSpot, an IoT-cloud-integrated system employing a customized CNN designed for real-time pothole detection. Evaluated on a substantial dataset comprising 3,424 real-world images. PotSpot achieved an impressive classification accuracy of 97.6%. A. Gagliardi *et al.*[17] introduced a lightweight CNN model aimed at categorizing diverse road conditions, including potholes and cracks, yielding accuracies between 90% and 93%. Aparna *et al.*[18] developed a CNN-based methodology leveraging thermal imaging data for pothole detection, reporting an accuracy of 97.08%. Similarly, Chu *et al.*[5] designed and trained a CNN model on images collected from Lahore roads, successfully detecting potholes with a commendable accuracy of 97%.

1.2 Pretrained networks-based methods

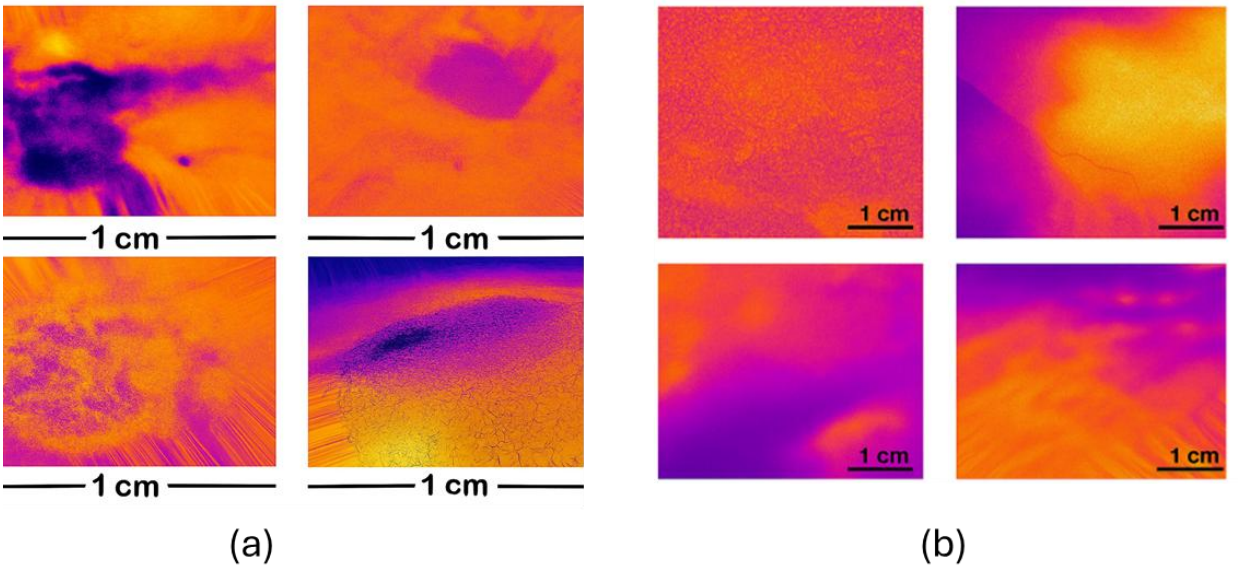

Fig. 1: (a) Pothole thermal images, (b) Non-pothole thermal images.

Yishun Li *et al.*[19] employed pre-trained Efficient Net models to identify six pavement distress types, including severe potholes, achieving an average classification accuracy of approximately 86%. N. Cinar *et al.*[20] applied a pre-trained DenseNet121 model to pothole detection tasks and achieved an accuracy above 99%. However, this result was derived from a limited validation dataset of only 67 images, and the computational requirements of DenseNet121 present practical limitations. Ahmad *et al.*[21] assessed various pre-trained architectures, including MobileNet v2, ResNet50, and ResNet18, for classifying pavement damage. Their analysis, conducted on datasets aggregated from online repositories and actual road images in Pakistan, demonstrated an accuracy of up to 98%. Li *et al.*[22] investigated the performance of multiple pre-trained CNN models in classifying six road distress types, concluding that ResNet-50 achieved superior results with an accuracy of 96.24% in both binary and multi-class scenarios. Gupta *et al.*[23] reported strong classification performance, achieving an accuracy of 94.79% using a ResNet-based model trained on the Road Damage 2020 (RDD2020) dataset.

1.3 Alternative machine learning methods

Additional research has examined non-deep learning methodologies for pothole classification. A. Egaji *et al.*[24] conducted a comparative analysis of five different machine learning classifiers, leveraging datasets collected from mobile devices and vehicle navigation maps, with the random forest model exhibiting the high accuracy of 94.44%. S. Bhatlawande *et al.*[9] explored various classical machine learning classification algorithms, achieving an accuracy of 88%, thereby highlighting opportunities for methodological improvements. M. Asmadi *et al.*[25] proposed KNN-based pothole classification on the dataset of 200 images on real-world images and got an accuracy of 95.83%.

Table 2: Details of the dataset.

Dataset	Total Images	Training set	Validation set	No. of classes
Thermal images	4904	3923	981	2

1.4 Unsupervised pothole detection method

Xinxiang *et al.*[26] proposed an unsupervised pothole detection method leveraging mobile laser scanning point cloud data, utilizing directed distances and density clustering for rapid localization and denoising. Negative skewness and skewness coefficients determined potholes accurately. Experimental validation confirmed robust performance across various pothole geometries.

A comparative overview of recent pothole classification

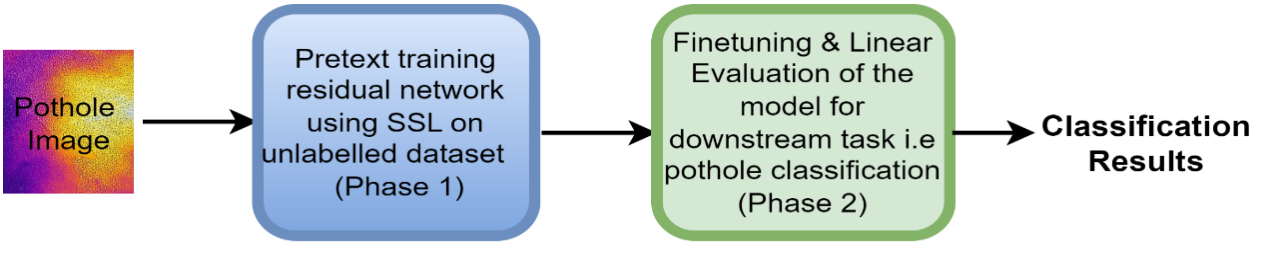

Fig. 2: Overall SSL-based Approach for pothole detection.

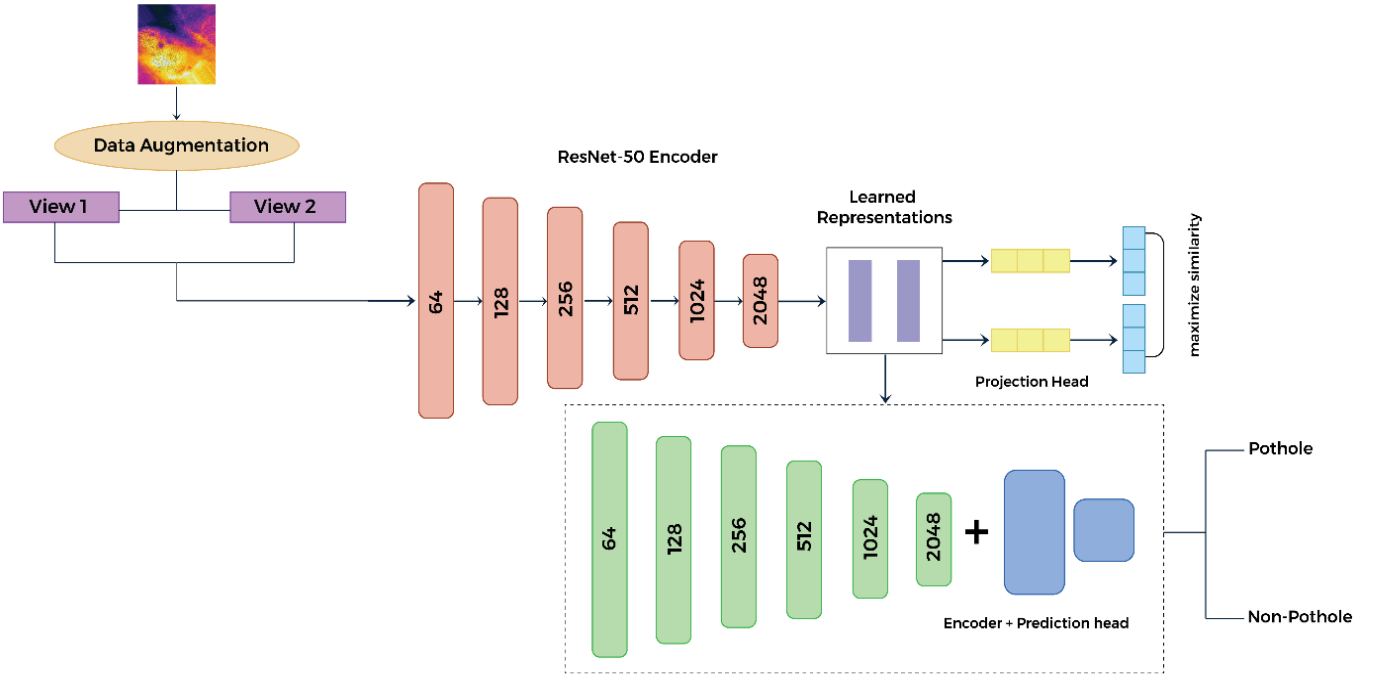

Fig. 3: Proposed SSL-based architecture for pothole detection.

approaches is presented in Table 1. Among CNN-based methods, PotSpot[27] and the thermal imaging model by Aparna *et al.*[20] show strong performance (97%+), with the latter leveraging thermal imaging for improved robustness. Lightweight CNNs[17] emphasize deployment efficiency, though with a slight trade-off in accuracy. Pretrained models such as DenseNet121[20] and MobileNet v2[21] achieve over 98%, though DenseNet shows limitations in generalizability due to a small dataset. Notably, Random Forest [24] and KNN-based [25] methods also yield competitive accuracy (94–95%) with simpler architectures, but deep learning methods generally provide more consistent results. This discussion underscores the practical merit of integrating self-supervised learning and thermal imaging, as pursued in the proposed approach, to balance performance and label efficiency. Table 1 details the parametric evaluation of the pothole detection approaches:

1.5 Gap analysis

From the above discussions, we have identified the following limitations in the existing methods:

1. Existing methods are based on the supervised learning approach and require a vast amount of labelled data for training. However, a lot of time and financial costs are involved in annotating data.
2. Most existing methods for pothole detection are trained on RGB images that are taken only in daytime and good weather conditions, potentially limiting their applicability to nighttime and extreme weather conditions.
3. There is a need to improve the performance of existing supervised learning-based pothole detection methods.

In summary, the majority of existing pothole detection methodologies heavily rely on supervised learning paradigms, necessitating extensively annotated datasets which impose substantial resource demands in terms of time and financial costs. To overcome these limitations, it is imperative to develop methods capable of achieving robust pothole detection performance with minimal annotated data. Addressing this critical research gap, the proposed methodology introduces a self-supervised learning-based approach that significantly reduces the reliance on labelled data. Unlike conventional supervised methods, self-

Table 3: Hyperparameter values for the pretext task.

Hyperparameter	Value
Image resizes	224 x 224
p-value of Vertical and Horizontal Flip	0.5
Rotation	Degree - (-90, 90)
p-value of Color jitter	0.8
p-value of Grayscale	0.2
p-value of Gaussian Blur	0.5, Kernel size – 9
Batch size	256
Optimizer	Stochastic Gradient Descent
Learning rate	0.0005
Weight decay	0.0005
Temperature value (NT-XENT)	0.5

Table 4: Hyperparameter values for Phase 2.

Hyperparameter	Value
Image resize	255x255
Center crop	224x224
Momentum	0.9
Batch size for 100% dataset	256
Batch size for 10% dataset	64
Optimizer	LARS
Learning rate	0.005
Weight decay	10^{-6}

supervised learning harnesses inherent data properties as pseudo-labels, significantly reducing reliance on extensive labelled data. The self-supervised methodology comprised two essential phases. The initial phase termed the pretext task, involves training models on pseudo-labels derived from unlabelled datasets to learn rich and generalized image representations effectively.[28] Subsequently, in the downstream phase, these learned representations are employed to train classifiers specifically targeting pothole detection, thus drastically reducing the volume of labelled images required for effective training [29]. Furthermore, the integration of thermal infrared imaging enhances detection robustness under challenging environmental conditions such as fog, rain, and nighttime, thus broadening the applicability and reliability of pothole detection systems.

The primary contributions of this research are succinctly outlined as follows:

- The proposed self-supervised approach substantially reduces the dependency on large-scale annotated datasets for effective pothole detection.
- The proposed approach demonstrates superior performance compared to traditional and state-of-the-art supervised learning-based pothole detection methods reported in the literature.

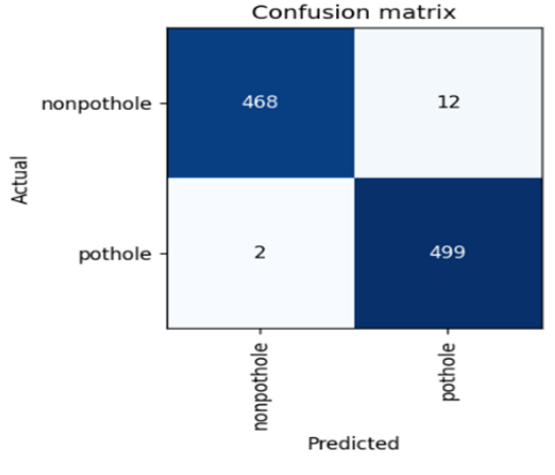

Fig. 4: Confusion matrix for 4.2.1

- Remarkably, even when the proposed self-supervised model utilizes significantly fewer labelled samples (e.g. 10% or 50% of the full dataset), it consistently outperforms supervised models trained on the complete dataset.
- This research uniquely leverages thermal imaging for pothole detection, ensuring robust and reliable performance under various adverse environmental conditions, including rainy, foggy, and nighttime scenarios.

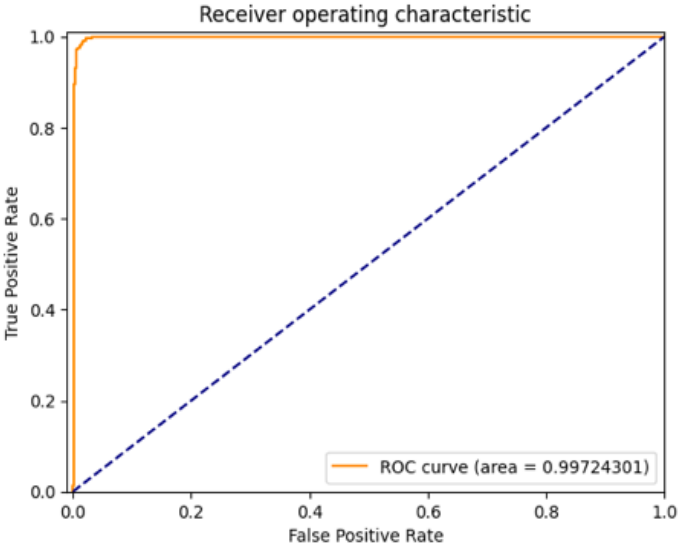

Fig. 5: ROC graph obtained for 4.2.1

The organization of this manuscript is as follows: Section 2 presents the experimental methodology, Section 3 describes the model architecture, and Section 4 discusses the results. Finally, Section 5 summarizes the research findings and presents conclusion and avenues for future exploration.

2. Experimental methodology

The following section outlines the dataset used for training and validation, followed by the proposed self-supervised learning-based approach and architectural details for pothole detection using thermal images.

2.1 Dataset description

The images utilized in this study were sourced from Panjab University's Design Innovation Center dataset. This dataset comprises manually captured thermal images of potholes from multiple locations in Chandigarh, India, recorded under diverse environmental conditions, including rain, fog, and nighttime. The dataset has been previously employed in existing literature.[18] Images were obtained using a thermographic camera that captures infrared radiation to create images analogous to conventional cameras that rely on visible light. The primary advantage of thermal imaging technology is its robustness under various atmospheric conditions, effectively detecting potholes during nighttime, wet road conditions, and foggy environments. Table 2

Table 5: SSL Fine-tuning results while initializing model with pre-trained weights during Phase 1.

Model training description				
			\	
			100% dataset (Batch Size: 256)	
	Accuracy	Precision	Recall	F1-Score
	99.2	99.3	99.5	99.1
ResNet50 was initialized with pre-trained (ImageNet) weights during phase 1 and fine-tuned (10 warm epochs followed by whole network training) during phase 2			50% dataset (Batch Size: 256)	
	Accuracy	Precision	Recall	F1-Score
	98.29	98.12	98.3	98.1
			10% dataset (Batch Size: 64)	
	Accuracy	Precision	Recall	F1-Score
	95.71	95.4	95.3	95.5

provides detailed information regarding the training and validation datasets. Illustrative examples of pothole and non-pothole thermal images from the dataset are shown in Fig. 1.

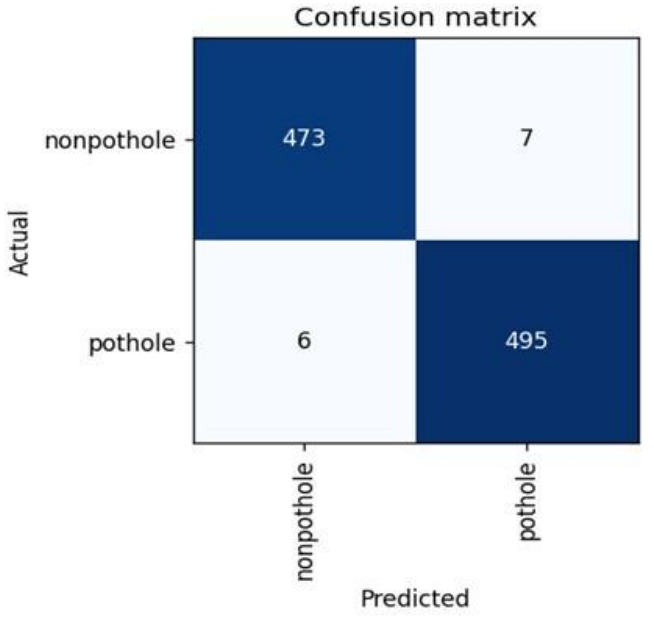

Fig. 6: Confusion matrix for 4.2.2

The dataset is partitioned into two distinct subsets: a training set comprising approximately 3923 images and a validation/ test set containing around 981 images. Model performance is evaluated using cross-validation based on the validation set.

The selected dataset from Panjab University is especially suitable due to its diversity in environmental conditions, including nighttime, foggy, and rainy scenarios. Thermal imaging ensures consistent visibility regardless of lighting and weather, directly addressing critical limitations faced by RGB-based pothole detection systems.

2.2 Proposed approach

Fig. 2 illustrates the proposed self-supervised learning approach for pothole detection using thermal imagery. The methodology is divided into two principal phases: the pretext task and the downstream task. During the pretext task, various data augmentation techniques- including resizing, random horizontal flips, random vertical flips, random jittering, random grayscale transformation, Gaussian blur, and affine transformations–are applied to generate distinct views from each other original image. Data augmentation during the pretext task phase is crucial for generating varied views of the same image, thus training the model to recognize inherent similarities despite visual transformations. This helps the model to extract generalized and robust features, improving its performance in real-world scenarios. Two augmented views generated from the same image constitute a positive pair, while views originating from different images form negative pairs. Contrastive learning is employed to extract meaningful image embeddings, which subsequently serve as input for the downstream classification task.

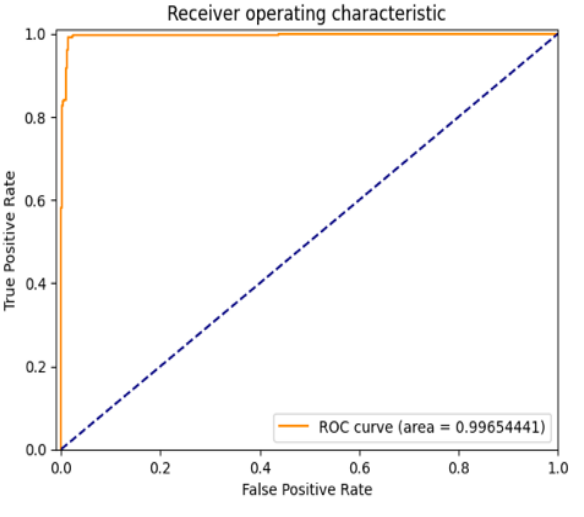

Fig. 7: ROC Curve for 4.2.2

Table 6: SSL Fine-tuning results without initializing model with pre-trained weights during Phase 1.

Model training description	100% dataset (Batch Size: 256)			
	Accuracy	Precision	Recall	F1-Score
	98.78	98.12	98.6	98.34
	50% dataset (Batch Size: 256)			
ResNet50 was not initialized with pre-trained (ImageNet) weights during phase 1 and fine-tuned (10 warm epochs followed by whole network training) during phase 2	Accuracy	Precision	Recall	F1-Score
	98.12	97.87	98.1	98.04
	10% dataset (Batch Size: 64)			
	Accuracy	Precision	Recall	F1-Score
	94.99	95.1	94.34	94.98

In the downstream task, residual networks are initialized using the embeddings obtained from the pretext phase. This strategy significantly reduces the number of labelled images required for network training since the model has already learned critical visual representations during the pretext phase.

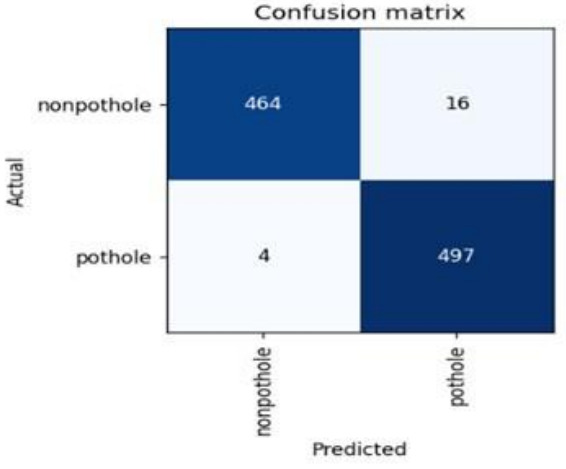

Fig. 8: Confusion matrix for 4.3.1

2.2.1 Proposed architecture

In this study, we propose a self-supervised learning (SSL) based approach for pothole detection, utilizing contrastive learning through SimCLR (simple framework for contrastive learning of visual representations)[30] to learn robust image embeddings. The proposed architecture begins with extensive data augmentation techniques applied to thermal images, generating multiple augmented views. In the initial phase (pretext task), we employ a pre-trained ResNet50 backbone along with a projection head composed of three fully connected layers. Contrastive learning facilitates embedding extraction by pulling representations from positive pairs closer and pushing representations from negative pairs further apart.[31] In the subsequent phase (downstream task), the learned embeddings from the pretext phase initialize the residual network. A prediction head consisting of two fully connected layers replaces the projection head, enabling the classification of thermal images into potholes and non-potholes. This process significantly reduces dependency on extensive labelled data for training. Fig. 3 illustrates the detailed SSL-based architecture for pothole detection.

The similarity between augmented image pairs is computed using cosine similarity between their respective image embeddings, as shown in Eq. (1).

$$\text{similarity}(x_i, x_j) = \text{cosine similarity}(z_i, z_j) \quad [30] \tag{1}$$

where x_i and x_j are augmented views of images and z_i and

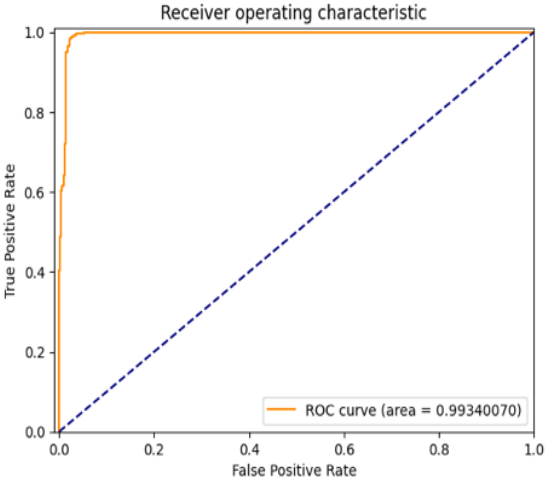

Fig. 9: ROC graph for 4.3.1

Table 7: SSL linear evaluation Results while initializing the model with pre-trained weights during Phase 1.

Model training description	100% dataset (Batch Size: 256)			
	Accuracy	Precision	Recall	F1-Score
	97.96	97.8	97.47	97.77
	50% dataset (Batch Size: 256)			
ResNet50 was initialized with pre-trained (ImageNet) weights during phase 1 and linear evaluated (the encoder part is frozen, and only the fully connected layers are trained) during phase 2	Accuracy	Precision	Recall	F1-Score
	96.34	96.4	96.47	96.39
	10% dataset (Batch Size: 64)			
	Accuracy	Precision	Recall	F1-Score
	93.99	93.8	94.1	93.6

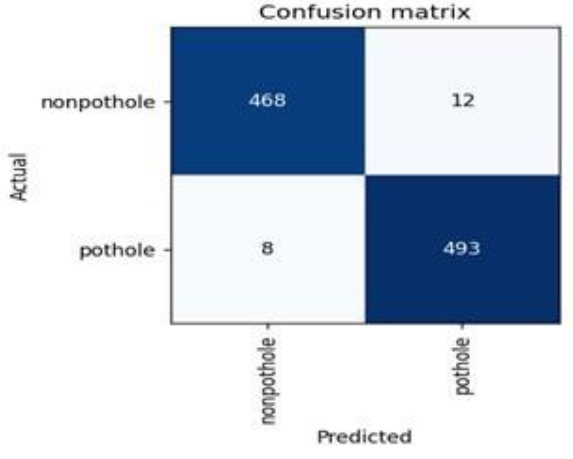

Fig. 10: Confusion matrix for 4.3.2

z_j are the encoded image representations. The cosine similarity between two image embeddings is calculated as given in Eq. (2).

$$S_{ij} = \frac{z_i^T z_j}{\|z_i\|\|z_j\|} \quad [30] \qquad (2)$$

The softmax function is applied to get the probability of two similar images. Softmax functions are generalized logistic functions that can be used to classify multiple types of data, as given in Eq. (3)

$$\text{softmax} = e^{\text{similarity(pair 1)}} / \text{Sum over all possible pairs}$$

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad [32] \qquad (3)$$

The loss for pairs of similar images is calculated as negative of the log of the softmax function as given in Eq 4.

$$l(i,j) = -log \frac{\exp(sim(z_i z_j)/\tau)}{\sum_{k=1}^{2n} 1_{k \neq i} \exp(sim(z_i z_k)/\tau)} \quad [30] \qquad (4)$$

This loss function is known as 'Normalized Temperature-scaled cross-entropy (NT-Xent)',[33] a modified version of cross-entropy loss. Here, τ is the adjustable temperature parameter for scaling the terms in the contrastive loss function. We have used a value of 0.5 for the temperature parameter (τ) in the proposed work.

3. Results and discussions

The proposed method was implemented using the open-source Fastai and Pytorch libraries and trained on Nvidia Quadro RTX 8000 GPU (48 GB). Following pretext training on unlabelled images, two evaluation strategies were conducted in the downstream phase: fine-tuning (entire model training) and linear evaluation (encoder frozen, fully connected layers trained).

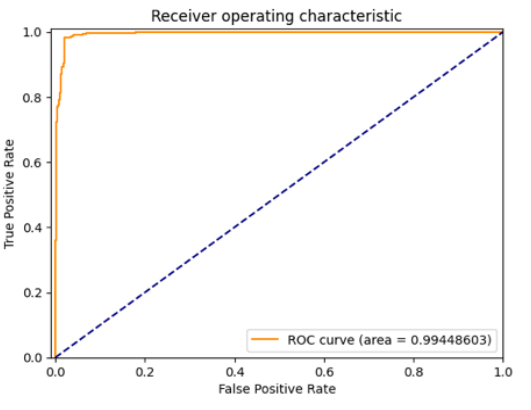

Fig. 11: ROC graph obtained for 4.3.2

Table 8: SSL Linear evaluation results without initializing model with pre-trained weights during Phase 1.

Model Training Description				
	100% dataset (Batch Size: 256)			
	Accuracy	Precision	Recall	F1-Score
	96.16	95.9	95.54	96.1
ResNet50 was not initialized with pre-trained (ImageNet) weights during phase 1 and linear evaluated (the encoder part is frozen, and only the fully connected layers are trained) during phase 2	50% dataset (Batch Size: 256)			
	Accuracy	Precision	Recall	F1-Score
	93.36	93.01	93.2	93.4
	10% dataset (Batch Size: 64)			
	Accuracy	Precision	Recall	F1-Score
	93.00	92.45	92.91	92.4

To validate the robustness of the proposed approach under conditions of limited annotated data, three subsets (10%, 50% and 100%) were evaluated. Batch sizes of 256 for 50% and 100% subsets and 64 for the 10% subset were used due to smaller sample availability.

This section first describes hyperparameter settings (Section 4.1), followed by analyses of results from fine-tuning (Section 4.2) and linear evaluation (Section 4.3). Sections 4.4 and 4.5 compare the proposed approach with supervised learning methods and existing state-of-the-art approaches.

3.1 Hyperparameter values and other experimental settings

In this section, we discuss hyperparameters and other parameter settings during phase 1 (pretext task) and phase 2 (fine-tuning and linear evaluation of the model for the downstream task, i.e., classification).

3.1.1 Hyperparameter settings for Phase 1 (pretext task)

Hyperparameters for the pretext task are detailed in Table 3. Images were resized to 224 x 224 pixels. Horizontal and vertical flips had probabilities of 0.5, with additional augmentations including rotation (-90 degrees to 90 degrees), color jitter (probability 0.8), grayscale transformation (probability 0.2), and Gaussian blur (probability 0.5, kernel size 9). Training employed a stochastic gradient descent (SGD) optimizer with a batch size of 256, learning rate of 0.0005, weight decay of 0.0005, and cosine schedular without warm restarts. A temperature of 0.5 was used for NT-Xent loss.

3.1.2 Hyperparameter settings for Phase 2 (fine-tuning)

The model trained on the pretext task is used as initialization for training during phase 2 for the downstream task, i.e., classification. The encoder part of the model is frozen for the first ten epochs of the training (warmup), and only fully connected layers are trained**.** Then, we train the whole network during the rest of the epochs. The hyperparameter values used during phase 2 (fine-tuning) are given in Table 4. We have used layer-wise adaptive rate scaling (LARS) to fine-tune the model. It is an optimization method for large batches, in combination with Nesterov momentum,[34]setting the momentum to 0.9 to optimize the loss. We train a model with a batch size of 256 for 50% and 100 % of the data and a batch size of 64 for 10% of the data as less data is left. A cosine scheduler without warm restart is used with a learning rate of 0.005 and a weight decay of 10^{-6}.

3.1.3 Hyperparameter settings for Phase 2 (linear evaluation)

After pretext training, we needed a way to evaluate the quality of the representations learned by the self-supervised learning model during phase 1. A standard method is to use a linear evaluation of the model. In this evaluation, the encoder part is frozen, and only fully connected layers are trained. The hyperparameter values used during phase 2 (linear evaluation) are given in Table 4. We have taken the dataset in different percentages of 10, 50, and 100 to validate the proposed method. We have taken a batch size of 256 for 100% and 50% of the dataset and 64 for 10% of the dataset as fewer data is left. We have also used the LARS optimizer here.

3.2 Results obtained during Phase 2 (fine-tuning)

This section discusses the pothole detection results obtained from the proposed approach for fine-tuning (10 warm epochs where only fully connected layers are trained, followed by full network training) during phase 2. These results are subdivided into two parts: fine-tuning results with pre-trained weights and

Table 9: Comparison of the fine-tuning results.

Fine-tuning with/without weight initialization	Approach	Accuracy (on 100% data)	Accuracy (on 50% data)	Accuracy (on 10% data)
Fine-tuning while initializing Resnet50 with pre-trained weights	Proposed SSL-based approach	99.2	98.29	95.71
	Supervised learning approach	96.83	95.9	89.99
Fine-tuning without initializing Resnet50 with pre-trained weights	Proposed SSL-based approach	98.78	98.12	94.99
	Supervised learning approach	96.02	94.72	86.00

fine-tuning results without ImageNet weights. The first type of results was obtained when we initialized ResNet50 with ImageNet weights (ImageNet weights) during the pretext task phase 1 and fine-tuned it during phase 2. In contrast, second-type results were obtained when we fine-tuned ResNet50 during phase 2 without initializing it with ImageNet weights during phase 1.

3.2.1 Fine-tuning results with pre-trained weights

We have initialized ResNet50 with ImageNet weights during the pretext task phase for this experiment. Table 5 shows the fine-tuning results for different percentages of the pothole dataset. On 100% of the data with a batch size of 256, the proposed method achieves an accuracy of 99.2% and other metrics above 99%. For 50% of the data with a batch size of 256, the proposed method achieves an accuracy of 98.29%. Even for 10% of the data with a batch size of 64, the proposed method achieves an accuracy of 95%.

We can easily observe from the results given in Table 5 that the proposed method can classify potholes with good performance even when trained on a small proportion (i.e., 10% dataset) of the labelled data. Its classification performance does not drop much when trained on half of the labelled data. Even for 10% of labelled data, it achieves good classification results. Thus, the proposed approach has drastically reduced the requirement of annotated data for training the model for pothole detection. Fig. 4 shows the confusion matrix obtained on the validation set of the dataset. It is clear from this matrix that most cases are either true positives or true negatives. Fig. 5 shows the ROC curve obtained for the classification of pothole images, which is quite satisfactory.

3.2.2 Fine-tuning results without pre-trained weights

This section presents the fine-tuning results obtained during phase 2 when the ResNet50 model was not initialized with ImageNet weights during phase 1. Table 6 presents the fine-tuning results obtained using a self-supervised learning (SSL) based pothole detection approach for different proportions of the dataset. With a batch size of 256, the proposed method achieves over 98% on 100% and 50% datasets. On the 10% dataset, using a batch size of 64, as less data is left, it maintains strong performance with all metrics above 98%.

Results in Table 6 show that the proposed method performs well even when the model is trained on a small proportion of the dataset without initializing it with ImageNet weights during phase 1. The comparison of results presented in Table 5 & 6 indicates that although pre-initializing models with ImageNet weights during the pretext phase help obtain little better classification results, it does not impact the results to a greater extent.

Fig. 6 shows that the confusion matrix obtained on the validation set of images is quite satisfactory. Fig. 7 shows the ROC curves obtained for the binary classification of the pothole images dataset, which is quite near the perfect one.

3.3 Results obtained during phase 2 (linear evaluation)

This section discusses the results obtained for linear evaluation (the encoder part of the model is frozen, and only the fully connected layers are trained) of the model during phase 2 using the proposed SSL-based pothole detection approach. These results are subdivided into linear evaluation results with pre-trained weights and linear evaluation results without pre-trained weights. The first type of results was obtained when we initialized ResNet50 with ImageNet weights during the pretext task phase 1 and performed the linear evaluation of the model during phase 2. In comparison, second-type results were obtained when we performed the linear evaluation of the ResNet50 during phase 2 without initializing it with ImageNet weights during phase 1.

3.3.1 Linear evaluation results with pre-trained weights

We initialized ResNet50 with ImageNet weights during the pretext task phase and then performed the linear evaluation of

Table 10: Comparison of the linear evaluation results.

Linear evaluation with/without weight initialization	Approach	Accuracy (on 100% data)	Accuracy (on 50% data)	Accuracy (on 10% data)
Linear evaluation while initializing Resnet50 with pre-trained weights	Proposed SSL-based approach	97.96	96.34	93.99
	Supervised learning approach	95.84	94.32	85.00
Linear evaluation without initializing Resnet50 with pre-trained weights	Proposed SSL-based approach	96.16	93.36	93.00
	Supervised learning approach	93.5	88.23	77.99

the model during phase 2 of this experiment. The linear evaluation results obtained with the proposed SSL-based approach for pothole classification are shown in Table 7. The proposed method achieves over 97% across all performance metrics on the 100 % dataset. On the 50% dataset, it maintains strong performance with all metrics exceeding 96%, while on the 10% dataset, it achieves over 93% across all metrics.

From Table 7, we can easily observe that even without training the encoder part network for classification, we achieve values of all classification metrics above 97% (for 100% of the dataset). Even for 10% of the data, we achieve more than 93% value of all metrics. Fig. 8 depicts the confusion matrix obtained, illustrating that most cases are either true positives or true negatives. Similarly, Fig. 9 shows the ROC curves obtained for classifying the pothole images dataset, demonstrating satisfactory performance across all obtained curves.

3.3.2 Linear evaluation results without pre-trained weights

For this experiment, we did not initialize the Resnet50 with ImageNet weights during phase 1 and performed the linear evaluation of the model during phase 2. The linear evaluation results using the proposed SSL-based pothole classification results are shown in Table 8. The proposed method achieves over 95% in all performance metrics on the 100 % dataset. In comparison, the proposed approach achieves above 92% across all performance metrics on both the 50% dataset and the 10% dataset.

Results in Table 8 show that the proposed method performs well even when the model is not initialized with ImageNet weights during phase 1 and just fully connected layers are trained during phase 2. Moreover, the proposed approach obtains good classification results even when trained on a small proportion of the dataset. The comparison of results presented in Table 7 and 8 indicates that the pre-initializing model with ImageNet weights during the pretext phase does not impact its performance much. However, there is a marginal improvement in the results with pre-trained weights. Fig. 10 and Fig. 11 show the confusion matrix and ROC curves obtained for the binary classification of pothole images, which are quite satisfactory.

3.4 Comparison with the supervised learning approach

This section provides a comprehensive comparison between the proposed self-supervised learning (SSL) methodology and the conventional supervised learning approach for the pothole detection task. We first compare performance during fine-tuning (initially training only fully connected layers for 10 epochs, followed by complete network training) and then during linear evaluation (training only fully connected layers with a frozen encoder backbone).

3.4.1 Comparison of fine-tuning results

We compare the performance of two approaches (SSL and supervised learning) for fine-tuning the model in two ways, i.e., when ResNet50 is initialized with pre-trained weights and without initializing ResNet50 with pre-trained weights. For the first comparison, in the SSL-based approach, ResNet50 is initialized with pre-trained weights before pretext training during phase 1. In the supervised learning-based approach, ResNet50 is also initialized with pre-trained weights before training for pothole classification. In contrast, for the second comparison, ResNet50 is used without pre-trained weight initialization for both approaches. As shown in Table 9, for fine-tuning with the pre-trained weights, the proposed SSL-

Table 11: Comparative results for pothole classification using thermal images.

Authors	Dataset	Approach	Backbone	% of Data Used	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Sathya *et al.*[35]	Pothole Thermal images	Supervised learning	CNN - MAO	100	99.10	96.56	97.23	94.00
Pathmanaban *et al.*[36]	Pothole Thermal images		Custom CNN	100	98.00	98.00	98.00	96.2
Aparna *et al.*[18]	Pothole Thermal images		ResNet - 50	100	97.08	-	-	-
Proposed work	Pothole Thermal images	Proposed Self-supervised learning (SSL)	ResNet -50	100	99.20	99.3	99.5	99.1
				50	98.29	98.12	98.3	98.1
				10	95.71	95.4	95.3	95.5

Based approach outperforms the supervised learning approach by almost 3% for 50% and100% of the data and by about 6% when both trained on 10% of the data for pothole classification. For fine-tuning with the pre-trained weights, even when trained on half of the data (50%), the proposed approach outperforms the supervised approach being trained on the full dataset (100%) by almost 2%. Further, for fine-tuning with the pre-trained weights, the proposed approach trained using just ten percent of the data obtains comparable results to the supervised approach trained on the full and half datasets. This comparison shows the effectiveness of the proposed approach.

For fine-tuning without pre-trained weights, the proposed approach surpasses the supervised approach by about 3% for 50% and 100% of the data and 9% for 10% of the data for pothole classification. For fine-tuning without pre-trained weights, even when trained on half of the data, the proposed approach outperforms the supervised approach being trained on the full dataset by more than 2%. Moreover, for fine-tuning without using the pre-trained weights, the proposed approach trained on only 10 percent of the data obtains better results than the supervised approach trained on half of the dataset. This comparison establishes the superiority of the proposed approach over the traditional supervised learning approach for pothole detection.

3.4.2 Comparison of linear evaluation results

This section compares the performance of SSL and supervised learning approaches for linear evaluation of the model (the encoder part of the model is frozen, and only the fully connected layers are trained) on similar lines as done earlier in the case of fine-tuning comparison, i.e., initializing ResNet50 with pre-trained weights and without initializing ResNet50 with pre-trained weights. For the first comparison, in the SSL-based approach, ResNet50 is initialized with pre-trained weights before pretext training during phase 1. In the supervised learning-based approach, ResNet50 is also initialized with pre-trained weights before training for pothole classification. In contrast, for the second comparison, ResNet50 is used without pre-trained weight initialization for both approaches.

Table 10 compares the results of the two approaches for linear evaluation with and without initializing ResNet50 with pre-trained weights for pothole classification. For linear evaluation with the pre-trained weights, the proposed approach outperforms the supervised approach by more than 2% for 50% and 100% of the data and by about 9% for 10% of the data for pothole classification. For fine-tuning with the pre-trained weights, even when trained on half of the data, the proposed approach outperforms the supervised approach being trained on the full dataset by almost 1%. Further, for fine-tuning with the pre-trained weights, the proposed approach trained using only ten percent of the data obtains comparable results to the supervised approach trained on the full and half datasets. This comparison demonstrated the effectiveness of the proposed approach. In the linear evaluation phase without pre-trained weights, the proposed approach surpasses the supervised approach by about 3% for 100% of the data and more than 5% for 50% and 10% of the data for pothole classification. For fine-tuning without pre-trained weights, the proposed approach trained on only ten percent of the data obtains comparable results to the supervised approach trained on full data. Further, it

outperforms the supervised approach trained on half of the dataset by 5%. This comparison shows that the proposed approach outclasses the traditional supervised learning approach for pothole detection.

3.5 Comparison with existing approaches

We identified very few studies addressing pothole detection using thermal images. Table 11 provides a comparative analysis of the proposed self-supervised learning (SSL)-based method with existing supervised learning methods by Sathya *et al.*,[35] Pathmanaban *et al.*,[36] and Aparna *et al.*[18] using thermal images for pothole classification. These studies primarily rely on fully labelled datasets and supervised training strategies. In contrast, our SSL-based method demonstrates superior performance even with limited labelled data. When leveraging the ResNet50 backbone trained on 100% of the dataset, the proposed SSL based approach achieves an accuracy of 99.20%, surpassing Aparna *et al.*'s[18] supervised ResNet50-based approach by over 2%, Sathya *et al.*[35] CNN-MAO method by 0.1%, and Pathmanaban *et al.*[36] approach by 1.2%. Even when utilizing only 50% of the labelled data, our proposed method attains an accuracy of 98.29%, still exceeding the performance of Aparna *et al.*'s[18] results with 100% data by more than 1% and closely competing with Sathya *et al.*[35] and Pathmanaban *et al.*[36]. Remarkably, with only 10% of the labelled data, the proposed SSL approach maintains competitive performance, achieving an accuracy of 95.71%, which remains significantly high compared to fully supervised methods.

The higher precision, recall, and F1-score further emphasize the robustness of our method in detecting potholes, especially when label availiabiity is constrained. This demonstrates a significant advancement over prior studies, which require full supervision and are less practical for real-world deployement where annotated thermal data is limited or expensive to obtain.

The comparative evaluation underscores the effectiveness and robustness of our proposed SSL-based method, demonstrating consistently higher accuracy than existing supervised methods[20,35] and[36] - even when trained on significantly smaller subsets (10% or 50%) of labelled data. This highlights both the label efficiency and practical deployment potential of the proposed method compared to prior thermal image-based pothole detection approaches.

4. Conclusion

Automatic pothole detection plays a vital role in road infrastructure maintenance, directly influencing road safety and maintenance efficiency. In this research, we introduced a self-supervised learning (SSL) based approach for pothole detection utilizing thermal imagery. The proposed methodology comprises two primary phases: an initial self-supervised pretext phase that leverages unlabelled data to effectively learn robust image representations, followed by a downstream phase where the model is fine-tuned using significantly fewer annotated images. Extensive comparative analyses demonstrate that our SSL-based approach significantly outperforms traditional supervised learning methods and existing state-of-the-art approaches. Remarkably, our method achieves superior or comparable classification accuracy even when trained on only 10% of the labelled dataset, clearly highlighting its effectiveness and efficiency in data-constrained scenarios. Thus, the proposed SSL-based approach substantially reduces the reliance on extensively labelled datasets, presenting a highly practical solution for real-world deployment. Future work will focus on extending this approach to include semantic segmentation of potholes, further leveraging the robust visual representations acquired during the self-supervised pretext task.

Acknowledgments

Not applicable.

Conflict of Interest

The authors declare that they have no conflict of interest.

Supporting information

Not applicable.

CRediT Statement

Deepika Vikas Agrawal: Writing, Original draft, Investigation, Formal analysis, Validation. **Varun Gupta:** Conceptualization, Supervision, Writing, review & editing. **C. Rama Krishna:** Supervision, Writing, review & editing. **Muskaan Chopra:** Software. **Abhinav Puri:** Software.

References

[1] Express News Service, "NCRB report: Deaths in road accidents up by 17%," The Indian Express (P) Ltd, delhi, Aug. 30, 2022, https://indianexpress.com/article/india/ncrb-report-deaths-in-road-accidents-up-by-17-pc-8119743/

[2] S. Gupta, P. Sharma, D. Sharma, V. Gupta, N. Sambyal, Detection and localization of potholes in thermal images using deep neural networks, *Multimedia Tools and Applications*, 2020, **79**, 26265-26284, doi: 10.1007/s11042-020-09293-8.

[3] IEEE 6th advanced information technology, electronic and automation control conference, *2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. October 3-5, 2022, Beijing, China. IEEE, 2022, 1-2, doi: 10.1109/IAEAC54830.2022.9929564.

[4] T. Lee, Y. Yoon, C. Chun, S. Ryu, CNN-based road-surface crack detection model that responds to brightness changes, *Electronics*, 2021, **10**, 1402, doi: 10.3390/electronics10121402.
[5] H.-H. Chu, M. Rizwan Saeed, J. Rashid, M. Tahir Mehmood, I. Ahmad, R. Sohail Iqbal, G. Ali, Deep learning method to detect the road cracks and potholes for smart cities, *Computers, Materials & Continua*, 2023, **75**, 1863-1881, doi: 10.32604/cmc.2023.035287.
[6] B. Bučko, E. Lieskovská, K. Zábovská, M. Zábovský, Computer vision based pothole detection under challenging conditions, *Sensors*, 2022, **22**, 8878, doi: 10.3390/s22228878.
[7] M. H. Asad, S. Khaliq, M. H. Yousaf, M. O. Ullah, A. Ahmad, Pothole detection using deep learning: a real-time and AI-on-the-edge perspective, *Advances in Civil Engineering*, 2022, **2022**, 9221211, doi: 10.1155/2022/9221211.
[8] M. Jakubec, E. Lieskovska, B. Bucko, K. Zabovska, Pothole detection in adverse weather: leveraging synthetic images and attention-based object detection methods, *Multimedia Tools and Applications*, 2024, **83**, 86955-86982, doi: 10.1007/s11042-024-19723-6.
[9] S. Bhatlawande, A. Deshpande, S. Deshpande, S. Shilaskar, Proactive detection of pothole and walkable path for safe mobility of visually challenged, *3rd International Conference for Emerging Technology (INCET)*. May 27-29, 2022, Belgaum, India. IEEE, 2022, 1-5, doi: 10.1109/INCET54531.2022.9824637.
[10] M. Aiman Dani Asmadi, S. Zainuddin, H. Mohd Nasir, I. Syafiza Md Isa, N. Emileen Abd Rashid, I. Pasya, Doppler radar-based pothole sensing using spectral features in k-nearest neighbors, *Bulletin of Electrical Engineering and Informatics*, 2025, **14**, 276-286, doi: 10.11591/eei.v14i1.8398.
[11] H. Yu, Y. Deng, F. Guo, Real-time pavement surface crack detection based on lightweight semantic segmentation model, *Transportation Geotechnics*, 2024, **48**, 101335, doi: 10.1016/j.trgeo.2024.101335.
[12] J. Fan, M. J. Bocus, B. Hosking, R. Wu, Y. Liu, S. Vityazev, R. Fan, Multi-Scale Feature Fusion: Learning Better Semantic Segmentation For Road Pothole Detection, *IEEE International Conference on Autonomous Systems (ICAS)*. August 11-13, 2021. Montreal, QC, Canada. IEEE, 2021, pp. 1-5, doi: 10.1109/icas49788.2021.9551165.
[13] Y. Zhang, C. Liu, Real-time pavement damage detection with damage shape adaptation, *IEEE Transactions on Intelligent Transportation Systems*, 2024, **25**, pp. 18954-18963, doi: 10.1109/TITS.2024.3416508.
[14] S. Thiruppathiraj, U. Kumar, S. Buchke, Automatic pothole classification and segmentation using android smartphone sensors and camera images with machine learning techniques, *IEEE Region 10 Conference (TENCON)*. November 16-19, 2020. Osaka, Japan. IEEE, 2020, pp. 1386-1391, doi: 10.1109/tencon50793.2020.9293883.
[15] V. Pereira, S. Tamura, S. Hayamizu, H. Fukai, Semantic Segmentation of Paved Road and Pothole Image Using U-Net Architecture, *International Conference of Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*. September 20-21, 2019. Yogyakarta, Indonesia. IEEE, 2019, pp. 1-4, doi: 10.1109/icaicta.2019.8904105.
[16] J. Guan, X. Yang, L. Ding, X. Cheng, V. C. S. Lee, C. Jin, Automated pixel-level pavement distress detection based on stereo vision and deep learning, *Automation in Construction*, 2021, **129**, 103788, doi: 10.1016/j.autcon.2021.103788.
[17] A. Gagliardi, V. Staderini, S. Saponara, An embedded system for acoustic data processing and AI-based real-time classification for road surface analysis, *IEEE Access*, 2022, **10**, 63073-63084.
[18] Aparna, Y. Bhatia, R. Rai, V. Gupta, N. Aggarwal, A. Akula, Convolutional neural networks based potholes detection using thermal imaging, *Journal of King Saud University - Computer and Information Sciences*, 2022, **34**, 578-588, doi: 10.1016/j.jksuci.2019.02.004.
[19] Y. Li, C. Liu, Q. Gao, D. Wu, F. Li, Y. Du, ConTrack distress dataset: a continuous observation for pavement deterioration spatio-temporal analysis, *IEEE Transactions on Intelligent Transportation Systems*, 2022, **23**, 25004-25017, doi: 10.1109/TITS.2022.3201968.
[20] N. Çınar, M. Kaya, An automated pothole detection *via* transfer learning, *International Conference on Decision Aid Sciences and Applications (DASA)*. March 23-25, 2022, Chiangrai, Thailand. IEEE, 2022, 1355-1358, doi: 10.1109/DASA54658.2022.9765021.
[21] C. F. Ahmad, A. T. Al-Sayegh, A. Cheema, W. Qayyum, R. Ehtisham, S. Saghir, A. Ahmad, Classification of different size of potholes based on surface area using convolutional neural network, *Discover Applied Sciences*, 2024, **6**, 492, doi: 10.1007/s42452-024-06207-3.
[22] D. Li, Z. Duan, X. Hu, D. Zhang, Y. Zhang, Automated classification and detection of multiple pavement distress images based on deep learning, *Journal of Traffic and Transportation Engineering (English Edition)*, 2023, **10**, 276-290, doi: 10.1016/j.jtte.2021.04.008.
[23] Y. Gupta, F. Chauhan, K. Singla, Analysis of different deep learning algorithms for road surface damage detection, *International Conference on Disruptive Technologies (ICDT)*. May 11-12, 2023, Greater Noida, India. IEEE, 2023, 13-17, doi: 10.1109/ICDT57929.2023.10150453.
[24] O. A. Egaji, G. Evans, M. G. Griffiths, G. Islas, Real-time machine learning-based approach for pothole detection, *Expert Systems with Applications*, 2021, **184**, 115562, doi: 10.1016/j.eswa.2021.115562.
[25] M. Aiman Dani Asmadi, S. Zainuddin, H. Mohd Nasir, I. Syafiza Md Isa, N. Emileen Abd Rashid, I. Pasya, Doppler radar-based pothole sensing using spectral features in k-nearest neighbors, *Bulletin of Electrical Engineering and Informatics*, 2025, **14**, 276-286, doi: 10.11591/eei.v14i1.8398.
[26] X. Ma, D. Yue, S. Li, D. Cai, Y. Zhang, Road potholes detection from MLS point clouds, *Measurement Science and Technology*, 2023, **34**, 095017, doi: 10.1088/1361-6501/acdb8d.
[27] S. Patra, A. I. Middya, S. Roy, PotSpot: Participatory sensing based monitoring system for pothole detection using deep learning, *Multimedia Tools and Applications*, 2021, **80**, 25171-25195, doi: 10.1007/s11042-021-10874-4.

[28] M. Noroozi, A. Vinjimoor, P. Favaro, H. Pirsiavash, Boosting self-supervised learning *via* knowledge transfer, *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018, 9359-9367, doi: 10.1109/CVPR.2018.00975.
[29] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, "ALBERT: A Lite BERT for Self-supervised Learning of Language Representations," in 8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020, OpenReview.net, Available: https://openreview.net/forum?id=H1eA7AEtvS
[30] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A Simple Framework for Contrastive Learning of Visual Representations," in Proceedings of the 37th International Conference on Machine Learning, H. D. III and A. Singh, Eds., in Proceedings of Machine Learning Research, vol. 119. PMLR, Jan. 2020, pp. 1597–1607. Available: https://proceedings.mlr.press/v119/chen20j.html
[31] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, Dilip Krishnan, "Supervised Contrastive Learning," in Advances in Neural Information Processing Systems, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, Eds., Curran Associates, Inc., 2020, pp. 18661–18673
[32] J. S. Bridle, Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. *Neurocomputing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1990, 227-236, doi: 10.1007/978-3-642-76153-9_28.
[33] P. H. Le-Khac, G. Healy, A. F. Smeaton, Contrastive representation learning: a framework and review, *IEEE Access*, 2020, **8**, 193907-193934.
[34] B. Ginsburg, I. Gitman, Y. You, Large batch training of convolutional networks with layer-wise adaptive rate scaling, 2018.
[35] R. Sathya, B. Saleena, CNN-MAO: convolutional neural network-based modified aquilla optimization algorithm for pothole identification from thermal images, *Signal, Image and Video Processing*, 2022, **16**, 2239-2247, doi: 10.1007/s11760-022-02189-0.
[36] P. P, G. B.K, Robust pothole detection in adverse weather conditions using thermal imaging and image processing, *23rd IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*. May 28-31, 2024, Aurora, CO, USA. IEEE, 2024, 1-6, doi: 10.1109/ITherm55375.2024.10709366.