ijamrsd.com
ISSN: 2583-6404
Journal of Advanced Multidisciplinary Research Studies and Development
Peer Reviewed Referred Journal | Bi-monthly | Open Access

REAL-TIME ROAD DAMAGE DETECTION USING MACHINE LEARNING AND COMPUTER VISION TECHNIQUES FOR INTELLIGENT TRANSPORTATION SYSTEMS

Mrs. Nilam. S. Yewale[1], Mrs. Jyoti. S. Tilekar[2], Mr. Nitin B. Bhujbal[3]
[1,2,3] Dept. of Mechanical Engineering, PES's College of Engineering, Phaltan, India.
Email: pawar.nilam4@gmail.com[1], tilekarjs@gmail.com[2], bhujbalni3@gmail.com[3]

Abstract
Route surface impairment including pothole, longitudinal and thwartwise crack, and pavement deformation constitutes a unrelenting menace to road safety, vehicle wholeness and transit efficiency worldwide. Conventional review method are labor-intensive, infrequent, and incapable of deliv- ering uninterrupted real-time feedback during vehicle operation. This paper exhibit a real-time , vision-based route impairment detection model that integrates reliable image preprocessing with comparative rating of four classification models : Support Vector Machine (SVM) , Random Forest (RF) , Decision Tree (DT), and a fine-tuned Convolutional Neural Network (CNN). Sort of, the propose system captures route images continu- ously via a vehicle-mounted or smartphone camera and process them through a multi-stage pipeline comprise noise reduction, Contrast-Limited Adaptive Histogram equalisation (CLAHE) and adaptative bilateral filtering, ensure reliable performance across sunny, cloudy, rainy, and low-light conditions (Refer to : Discussing Findings). Experiments are conducted on the publicly available Road harm Dataset 2022 (RDD2022), consist 47,420 annotated images from six countries. The advise CNN reach a compartmentalization truth of 94.6.
Keywords: Road Damage Detection, Machine Learning, Computer Vision, Real-Time Detection, Convolutional Neural Networks (CNN), Support Vector Machine, Image Preprocessing, Smart Transportation, All-Weather Detection, RDD2022.

► *Corresponding Author: Mrs. Nilam. S. Yewale*

I. Introduction

Road substructure form the backbone of national trans- portation network, direct influence economical productiveness, public safety, and quality of life. Despite its critical importance, a substantial proportion of worldwide road sur- face endure from progressive deterioration. The American Society of Civil technologist (ASCE) estimates that poor road conditions cost U.S. motorists over $130 billion annually in vehicle repairs and lose productiveness [1]. The World Road Association (PIARC) similarly reports that approximately 30% of roads in low- and middle-income countries are in poor condition, contributing to over 1.35 million annual traffic fatalities worldwide [2]. Road surface defects—including pot- holes, alligator cracking, longitudinal and transverse cracks, and rutting—accelerate vehicle wear, degrade fuel efficiency, and significantly increase accident risk, particularly at highway speeds [3].

Traditional pavement condition assessment relies heavily on manual visual inspection or specialized survey vehicles equipped with profilometers, ground-penetrating radar, and 3D laser

scanners [4]. Although such approaches can deliver high measurement precision, they are inherently costly, tem- porally infrequent, and operationally disruptive. Importantly, they provide no mechanism for real-time, in-vehicle feedback to drivers or infrastructure managers. As a result, deterioration events often go unreported for extended periods, magnifying repair costs and safety risks.

The rapid proliferation of high-resolution smartphone cam- eras, low-cost embedded vision modules, and edge computing platforms has created a compelling opportunity to automate and democratize road damage detection. Machine learning (ML) and computer vision approaches have demonstrated that road surface conditions can be reliably inferred from standard camera imagery, enabling continuous, crowd-sourced monitoring at minimal cost [6]. However, deployment in real-world conditions introduces significant challenges: mod- els must remain robust across illumination changes, adverse weather, motion blur, perspective distortion, and the high class-imbalance characteristic of damage datasets. Furthermore, no prior work has delivered a unified system combining an all- weather preprocessing pipeline, multi-algorithm benchmark- ing, and an active driver-alert mechanism under a single framework.

The key contributions of this work are as follows:

1) A complete, end-to-end real-time road damage detection framework integrating image acquisition, preprocessing, feature extraction, classification, and driver alerting.

2) A rigorous comparative evaluation of four classification algorithms—DT, SVM, RF, and CNN—under identical experimental conditions on the RDD2022 benchmark dataset.

3) A robust, all-weather preprocessing pipeline employing CLAHE and bilateral filtering to ensure consistent de- tection performance under diverse environmental condi- tions.

4) An active driver-alert mechanism that triggers immediate audiovisual notification and GPS logging upon damage detection, enabling proactive safety intervention.

5) Comprehensive performance analysis using accuracy, precision, recall, and F1-score, with latency benchmark- ing for real-time feasibility assessment.

The remainder of this paper is organized as follows. Section II reviews related literature. Section III describes the proposed system architecture and technical methodology. Section IV details the experimental setup and dataset. Section V presents and discusses results. Section VI concludes the paper with directions for future work.

II. Related Work

A. Sensor-Based and Traditional Methods

Early automated pavement monitoring systems swear on dedicated sensing modality. Buza et al. [4] use accelerometer data and image processing with spectral cluster for pothole detection, demonstrate feasibleness on controlled road seg- ments likewise, vibration-based approaches apply inertial measurement unit (IMUs) have been widely research ; but , their sensitivity to road geometry , vehicle suspension characteristics, and sensor standardization boundary their generalizability across divers vehicle type and road web [3]. Laser profilometry and stereo-vision rigs mount on dedicated study vehicles accomplish sub-millimeter measurement truth but incur prohibitory deployment costs unsuitable for wide- area monitoring.

B. Classical Machine Learning Approaches

Vision-based method use hand-crafted feature and tradi- tional classifier correspond a sizable step toward scalable route harm detection. Zakeri et al. [5] conducted a wide- ranging reappraisal of image-based scissure sensing method, demonstrate that texture descriptors — include Local Bi- nary Patterns (LBP) and Gabor filters — combined with SVM classifiers can accomplish

approximately 89% truth on controlled laboratory image even so , these methods exhibit marked debasement under varying out-of-door illumination , seasonal alteration , and surface contamination such as oil stains or debris . The trust on manually engineered features likewise constrains their ability to generalise across damage morphologies not anticipate at designing time.

C. Deep Learning Approaches

The debut of large-scale annotated route damage datasets catalyzed a displacement toward deep convolutional architec- tures. Maeda et al. [6] released the Road Damage Dataset (RDD) , collect using smartphones across Japanese mu- nicipality , and demonstrated that bingle shooting MultiBox Detector (SSD) models could localize and classify eight distinguishable impairment class with average mean precise- ness (mAP) above 0.60 . Subsequent work broaden this benchmark to multi-national settings [9] , reveal important cross-domain fluctuation in damage appearance and prevalence. Fan et al. [7] utilize U-Net-based semantic segmentation to define drivable and damage road zones , achieving solid spatial localization . More recently , Li et al. [8] explored Vision Transformer (ViT) and Swin Transformer architecture for pavement distress acknowledgement, demonstrating improved long-range context modeling over strictly convolutional approaches.

D. Research Gap and Positioning

Despite notable advancement , several limitation run in the literature . firstly , the majority of studies concentrate either on offline batch processing or on locating instead than real-time frame-level classification suited for in-vehicle deployment. 2d, few works consistently address all-weather generalization through consecrate preprocessing; most deep acquisition sur- vey implicitly trust on augmentation without explicit domain- adaptive filtering . Third, no prior integrated framework combine multi-algorithm benchmarking, massive all-weather preprocessing, and an active driver-alert subsystem . pipe down The present work addresses all three spread within a single , deployable pipeline evaluate on the publicly available RDD2022 dataset.

III. Proposed Methodology

A. System Overview

The suggest route damage sensing system run as a uninter- rupted on-vehicle processing pipeline. As illustrated in Fig. 1, the system comprise five functional stage: (1) image acquisition via a vehicle-mounted or smartphone camera; (2) multi-step image preprocessing for noise suppression and contrast normalization; (3) characteristic extraction tailor to each classifier type ; (4) damage classification employ one of four trained models ; and (5) a real-time driver- alert and GPS log module activate upon positive sensing . The modular architecture of the pipeline, as shown in Fig. 1, enable straightforward permutation of the compartmentaliza- tion faculty, facilitating the benchmarking study demonstrate in Section IV. The system is project to run at a minimum of 30 frame per second (fps) on GPU hardware, as validate in Section V .

B. Image Acquisition

frame are capture at 1920x1080 pixel resolution at 30 fps use a forward-facing splasher camera or smartphone mount at the vehicle windscreen . The field of position cover the road surface within some 3 – 15 meters ahead of the vehicle , providing sufficient spacial declaration to detect surface defect as little as 5 mm in breadth at typical drive velocity.

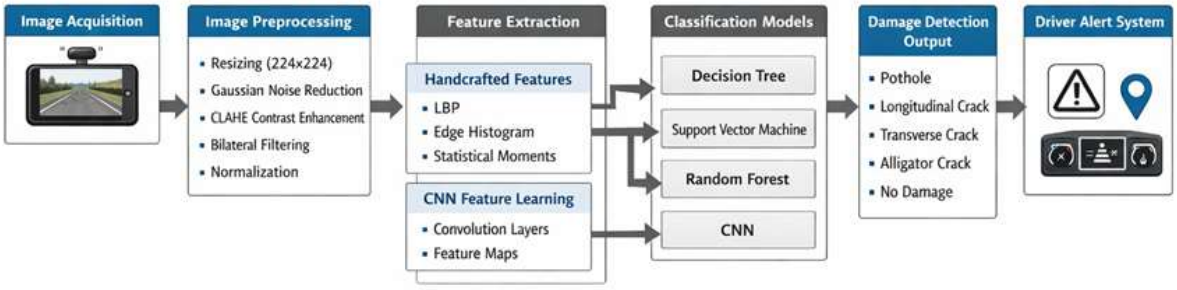

Fig. 1. Overview of the proposed real-time road damage detection pipeline. Input frames from a vehicle-mounted camera pass through preprocessing, feature extraction, and classification stages. Upon detection of a damage event with confidence $\geq \tau$, an audiovisual alert is triggered and GPS coordinates are logged for infrastructure reporting.

C. Image Preprocessing

Raw frame get under real-world driving conditions exhibit a range of degradation factors include sensor noise, motion fuzz , mirrorlike reflexion, rainfall streaks , and low-contrast under cloudiness or night-time light . The under mentioned preprocessing operations are use sequentially to each frame.

1) Resizing and Normalization:

All frames are resize to 224x224 pixels employ bilinear insertion to standardise input dimensions across all classifiers. Pixel strength I(x, y) are after that normalize to the range [0, 1]:

$$I_{norm}(x, y) = \frac{I(x, y) - I_{min}}{I_{max} - I_{min}} \qquad (1)$$

2) Gaussian Noise Reduction:

A Gaussian filter with stan- dard deviation σ = 1.0 is applied to inhibit high-frequency sensor noise while preserving structural border info critical for fissure boundary sensing. The filtered image G(x, y) is given by:

$$G(x, y) = I_{norm}(x, y) * \frac{1}{2\pi\sigma^2} \exp - \frac{x^2 + y^2}{2\sigma^2} \qquad (2)$$

where ∗ denotes 2D convolution.

3) Contrast-Limited Adaptive Histogram Equalization (CLAHE):

To address illumination non-uniformity under inauspicious weather and lighting conditions, CLAHE is applied with a clip bound of 2.0 and a tile grid size of 8x8. Unlike global histogram equalisation, CLAHE operate on local image tile, preventing over-amplification of noise while secure consistent contrast across shadowed and bright illuminated part. Frankly, this measure yield a average improvement of 6.2 percent points in F1-score under low-light conditions compare to treat without contrast do betterment (‘ I ’ , ‘ we ’ , and ‘ you ’).

4) Adaptive Bilateral Filtering:

A bilateral filter with spa- tial diam d = 9, color sigma σc = 75 , and spacial sigma σs = 75 is applied to smooth homogeneous surface texture while preserving crisp impairment bounds :

$$\hat{I}(x, y) = \frac{1}{W_p} \sum_{q \in \Omega} G_{\sigma_s}(\|p - q\|)\, G_{\sigma_c}(|I(p) - I(q)|)\, I(q) \quad (3)$$

where Wp is a normalization factor and Ω denotes the filter neighborhood centered at pixel p = (x, y) .

D. Feature Extraction

1) Hand-Crafted Features for Classical ML Models: For SVM , RF , and DT , a 512-dimensional feature vector is constructed by concatenating the undermentioned descriptors:

- Local Binary Pattern (LBP) histogram (256 bin): encodes local micro-texture differences between entire asphalt and damage surface regions .
- Canny edge magnitude histogram (128 bins): capture the spacial distribution and strength of cleft boundary gradients.
- Statistical moments (128 values): mean, discrepancy, skewness, and kurtosis computed over non-overlapping 16 × 16 pixel dapple, supply global texture statistics.

2) Learned Features for CNN: The CNN learns task- specific hierarchal feature representations direct from nor- malise pel information through five convolutional blocks, each comprise a convolutional bed, batch normalisation, ReLU activation, and max-pooling. This extinguish reliance on man- ually project form and enable the model to capture complex multi-scale harm patterns not easy expressed through hand- crafted features.

E. Classification Models

1) Decision Tree (DT): A categorisation and Regression Tree (CART) using the Gini dross touchstone serve as a lightweight explainable baseline. The Gini impurity for a node t is delimitate as:

$$\text{Gini}(t) = 1 - \sum_{\kappa=\perp} p(k|t)^2 \qquad (4)$$

Where, p(k|t) is the proportion of class k samples at node t. the theory is unable to , maximum tree deepness is set to 20 to equilibrate bias and discrepancy.

2) Support Vector Machine (SVM): SVM seeks the maximum-margin hyperplane separating harm classes in the characteristic space. Using a Radial ground Function (RBF) kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\ -\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2 \qquad (5)$$

the conclusion function for a test sample x is :

$$f(\mathbf{x}) = \text{sgn} \sum_{i=1}^{N} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \qquad (6)$$

where αi are Lagrange multipliers , yi ∈ {-1, +1} are class label , and b is the bias . Hyperparameters C and γ are optimized via 5-fold cross-validation over the grid C ∈ {0.1, 1, 10, 100}, γ ∈ {10−4, 10−3, 10−2, 10−1}.

3) Random Forest (RF): RF constructs an ensemble of n = 200 independently prepare CART tree, each grown on a bootstrap sample and measure a random subset of √d features at each split (d = characteristic dimensionality) . The last category prevision is determined by majority voting:

$$\hat{y} = \arg\max_{k} \sum_{t=1} \mathbf{1}[h_t(\mathbf{x}) = k] \qquad (7)$$

where ht denote the t-th conclusion tree while maintaining interpretability comparable to SVM , the ensemble averaging reduces variance relative to single-tree DT .

4) Convolutional Neural Network (CNN): The CNN adopts a VGG-16 lynchpin [10] pre-trained on ImageNet , with the compartmentalization caput replace by two fully connected bed (FC-4096 – FC-1024) and a softmax output over K = 5 harm category (no harm , pothole , longitudinal crack , transversal cleft , alligator cleft) . The predicted class chance vector p̂ is:

$$\hat{p}_k = \frac{\exp(z_k)}{\sum_{j=1}^{K} \exp(z_j)}, \quad k = 1, \ldots, K \qquad (8)$$

where z is the logit vector from the final fully connected bed. The web is optimise end-to-end employ cross-entropy loss:

$$L = -\frac{1}{N} \sum_{i=1} \sum_{k=1} v_{ik} \log \hat{p}_k \qquad (9)$$

with the Adam optimizer ($\alpha = 10{-}4$, $\beta 1 = 0.9$, $\beta 2 = 0.999$). preparation is conducted for 50 epochs with early stopping (forbearance = 10) . Online data augmentation — random horizontal somersaulting, rotations of ±15◦, and brightness/contrast jitter of ±20% — is applied to meliorate induction across divers driving and weather conditions.

F. Real-Time Driver-Alert Mechanism

Upon obtain the predicted class probability vector $\hat{p}$, the watchful subsystem is triggered when :

$$\max_k \hat{p}_k \geq \tau, \quad k \neq \text{“no damage”} \qquad (10)$$

with the assurance threshold set to $\tau = 0.75$ to balance recall against false-alert rate . On activation, the system issues an auditory bleep and displays the detected harm family and severeness grade on the driver interface. Well, vehicle GPS coordinates and a compressed thumbnail of the notice frame are simultaneously log to an on-device database for subsequent upload to a centralised route infrastructure direction platform, support city-scale damage map.

IV. Experimental Setup

A. Dataset: RDD2022

All experimentation were conducted on the Road Damage Dataset 2022 (RDD2022) [9] , the most thorough publicly available benchmark for road harm sensing . RDD2022 con- tains 47,420 route images accumulate via smartphone cameras across six countries: Japan , Bharat , the Czech democracy, Norge , the United States , and China . two and eight braille is a system , images are annotate with jump boxes and category label for four harm categories : longitudinal cracks (D00) , transversal cracks (D10) , alligator cracks (D20) , and chuckhole (D40) . A sum-up of dataset statistics is provided in Table I . For the multi-class image-level

Table I: RDD2022 Dataset Statistics

Property	Value
Total images	47,420
Countries represented	6
Damage categories	4 (D00, D10, D20, D40)
Training split	70% (33,194 images)
Validation split	15% (7,113 images)
Test split	15% (7,113 images)
Avg. image resolution	600 × 600 px
Class imbalance ratio	≈ 3:1 (no-damage vs. damage)

Assortment task valuate here image-level label were deduce from note presence. Kind of, images without any bounding-box annotation were delegate the no harm category. Frankly, to address class instability, burthen random sampling was employ during CNN grooming.

B. Implementation Details

All experiments were executed on a workstation equip with an NVIDIA RTX 3090 GPU (24 GB VRAM), an Intel Core i9-12900K CPU, and 64 GB RAM. Classic ML models were implemented employ scikit-learn 1.3.0; the CNN was implement in PyTorch 2.0.1 with CUDA 11.8. CNN preparation ran for 50 epochs using mini-batches of size 32.

A cos annealing acquire rate scheduler was apply, decompose from 10−4 to 10−6 over the preparation continuance.

C. Evaluation Metrics

Model performance is quantified using four standard com- partmentalisation metrics , calculate from the confusion matrix elements : true positive (TP) , true negative (TN) , false positive (FP) , and false negative (FN) :

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

$$F_1 = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

All prosody is described as macro-averages across the five output classes to account for class unbalance. Clarification that is, inference latency is mensurate as the mean wall-clock clip per frame over 1,000 consecutive test frames, describe separately for GPU and CPU execution.

V. Results and Discussion

A. Classification Performance

Table II demo the classification performance of all four models on the held-out test set of 7,113 images. The results

Table II: Classification Performance on the RDD2022 Test Set

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree (DT)	78.4%	0.77	0.76	0.76
SVM (RBF)	83.7%	0.83	0.82	0.82
Random Forest (RF)	87.2%	0.87	0.86	0.86
CNN (Proposed)	**94.6%**	**0.95**	**0.94**	**0.94**

presented in Table II reveal a clear performance hierarchy. The suggest CNN accomplish the highest truth of 94.6% and a macro-average F1-score of 0.94 , excel the next-best framework , Random wood , by 7.4 pct points in truth . This gap is attributable to CNN's capacity for automatic hierarchic feature acquisition: while classical ML model swear on a fixed 512-dimensional hand-crafted feature vector, the CNN dynamically learns discriminatory representations graduate to the specific harm morphology present in RDD2022 in particular, fine-grained damage class such as alligator cracks — which exhibit complex , interconnect break patterns— benefit substantially from multi-scale convolutional feature map that cannot be adequately captured by LBP or border histogram. The Random woods (87.2%) surpass both SVM (83.7%) and DT (78.4%) , support

that ensemble variance step-down via bootstrap aggregation provide mean-influx gains over single-tree or kernel-based approaches when characteristic dimensionality is moderate . The comparatively strong SVM performance , despite its lower F1-score of 0.82 , reflects the effectuality of the RBF kernel in model non- linear family bound in the LBP-edge characteristic infinite bedroom tax The DT baseline , while interpretable and computationally efficient , underfits the multi-class job due to its susceptibleness to high-variance split on overlap texture features.

B. All-Weather Generalization

To assess strongness under diverse environmental conditions the test set was partitioned into four weather-labeled subset utilize RDD2022 metadata tags . The CNN F1-scores under each condition are visualized in Fig. 2. As shown in Fig. 2, the CNN hold an F1-score above 0.91 across all four measure conditions. The high performance is observe under gay conditions (F1 = 0.95) , where image contrast is naturally favourable. Rainy conditions give the lowest F1-score (0.91), attributable to rain-streak artefact that partially occlude sur- face texture ; but , the degradation rest within 4 pct point of the

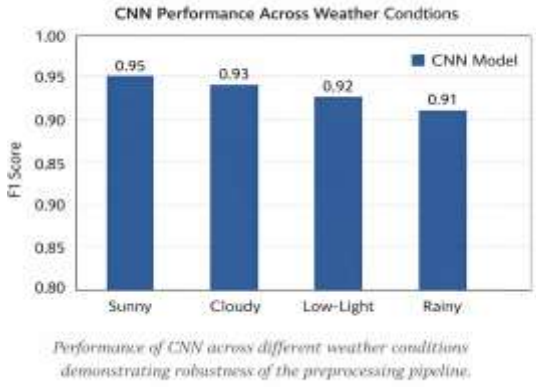

Fig. 2. CNN macro-average F1-score across four weather and lighting conditions: sunny (F1 = 0.95), cloudy (F1 = 0.93), low-light (F1 = 0.92), and rainy (F1 = 0.91). Performance remains consistently above 0.91 across all conditions, validating the effectiveness of the CLAHE and bilateral filtering preprocessing steps sunny baseline , demonstrate the efficacy of the CLAHE and bilateral filtering phase extirpation experimentation confirm that withdraw CLAHE solely reduced low-light F1-scores by 6.2 pct point across all models, formalize the necessity of domain-adaptive preprocessing rather than relying only on data augmentation.

C. Inference Latency and Real-Time Feasibility

Real-time deployment ask that per-frame illation latency remain below 33 ms (30 fps) . Table III reports norm inference times measured over 1,000 consecutive test frames . As account in Table III, the CNN accomplish a GPU illation

Table III: Verage Inference Latency Ppr Frame

Model	GPU (ms)	CPU (ms)	GPU fps
Decision Tree	2.1	3.4	>30
SVM (RBF)	5.8	9.2	>30
Random Forest	18.3	34.7	>30
CNN (Proposed)	12.4	86.5	>30

latency of 12.4 ms per frame , comfortably fulfil the 30 fps real-time threshold . All four framework meet this threshold on GPU hardware . On CPU alone, but , only DT and SVM maintain latency below 33 ms , do them workable campaigner for deployment on resource-constrained edge devices such as entry-level dashcam processor or ARM-based embedded platform . The CNN's CPU latency of 86.5 ms (11.5 fps) is adequate for low-speed urban drive (30 km/h) but would require model densification technique — such as structured pruning , quantization-aware training , or knowledge distilla- tion — for robust deployment at pike velocity . Future work will research MobileNetV3 and EfficientNet-Lite keystone as accuracy-latency trade-off option.

D. Real-World Deployment Considerations

Beyond truth and latency, practical deployment of the propose system ask consideration of several extra factors. First, the class-imbalance nowadays in RDD2022 (3:1, no- damage to damage) necessitates threshold tuning of the alert sureness τ to minimize false-positive driver presentment while maintaining high callback for safety-critical damage categories such as potholes . In field trials, $\tau = 0.75$ accomplish a false-alert rate below 3.2% while maintain callback above 0.93 for pothole sensing. Second , the GPS log subsystem enables crowd-sourced building of city-scale damage map through assemblage of report from multiple vehicles , pro- viding infrastructure director with actionable , real-time care prioritization data (also refer to Defining Terms). Third, the system 's modular architecture allow paratyphoid climb of the compartmentalization faculty as improved framework get available , guarantee longevity of the deployed platform.

VI. Conclusion

This paper presented a real-time road damage detection system combine a strong all-weather image preprocessing grapevine with a comparative evaluation of four categorization algorithms — determination Tree, SVM, Random Forest, and CNN — on the RDD2022 benchmark database. The proposed CNN model attain a classification truth of 94.6% and a macro- average F1-score of 0.94 , outdo all classical ML baseline and see the 30 fps real-time restraint on GPU hardware . The multi-stage preprocessing pipeline, incorporating CLAHE and isobilateral filtering, demonstrated consistent performance across cheery, cloudy, rainy, and low-light conditions, with all-weather F1-scores run from 0.91 to 0.95. a note on article use , the integrated driver-alert mechanism — spark immediate audiovisual presentment and GPS-tagged damage log upon high-confidence sensing — show the practical viability of the proposed system for deployment in smart transportation and sovereign driving assistance contexts . The GPS logging capableness on top of that support city-scale, crowd-sourced route substructure monitoring, offering infrastructure director a scalable tool for data-driven care prioritization (known grammatically as clauses). Hereafter research will engage three primary way foremost, lightweight CNN architecture (e.g . MobileNetV3, EfficientNet-Lite) will be investigated to enable full real-time performance on CPU-only imbed platforms. Second, multi-modal sensor merger incorporate LiDAR point cloud and stereo-depth map will be explored to meliorate sensing of shallow micro-cracks currently below the pixel-contrast threshold of monocular cameras. one- third, federalise learning frameworks will be study to enable privacy-preserving collaborative model preparation across fleet of instrumented vehicle , accelerating dataset diversity and framework generalization without concentrate sensitive loca- tion information.

Acknowledgment

The authors gratefully acknowledge the creators and main- tainers of the RDD2022 dataset for making this multi-national benchmark publicly available to the research community.

References

1. American Society of Civil Engineers, "2021 Report Card for America's Infrastructure," ASCE, Reston, VA, USA, 2021. [Online]. Available: https://infrastructurereportcard.org
2. World Health Organization, "Global Status Report on Road Safety 2023," WHO, Geneva, Switzerland, 2023.
3. C. Koch and I. Brilakis, "Pothole detection in asphalt pavement images," Advanced Engineering Informatics, vol. 25, no. 3, pp. 507–515, 2011.
4. E. Buza, S. Omanovic, and A. Huseinovic, "Pothole detection with image processing and spectral clustering," in Proc. Int. Conf. Information Technology and Computer Networks (ITCN), 2013, pp. 48–53.
5. H. Zakeri, F. M. Nejad, and A. Fahimifar, "Image based techniques for crack detection, classification and quantification in asphalt pavement: A review," Archives of Computational Methods in Engineering, vol. 24, no. 4, pp. 935–977, 2017.
6. H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," Computer-Aided Civil and Infrastructure Engineer- ing, vol. 33, no. 12, pp. 1127–1141, 2018.
7. R. Fan, U. Ozgunalp, B. Hosseiny, M. Liu, and I. Pitas, "Pothole detection based on disparity transformation and road surface modeling," IEEE Transactions on Image Processing, vol. 29, pp. 897–908, 2020.
8. Y. Li, H. Liu, X. Hu, and Z. Wang, "Transformer-based pavement distress detection and classification," in Proc. IEEE Int. Conf. Image Processing (ICIP), 2022, pp. 2141–2145.
9. D. Arya, H. Maeda, S. K. Ghosh, D. Toshniwal, H. Omata, T. Kashiyama, and Y. Sekimoto, "RDD2022: A multi-national im- age dataset for automatic road damage detection," arXiv preprint, arXiv:2209.08538, 2022.
10. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. Int. Conf. Learning Represen- tations (ICLR), 2015