North Asian International Research Journal of Multidisciplinary

IRJIF IMPACT FACTOR: 5.001

ISSN: 2454-2326 Vol.10, Issue-11 November-2024

Index Copernicus Value: 58.12 Indian Citation Index

NAIRJC UGC-Aligned | HEI-Approved Peer Reviewed Refereed Journal (Under New UGC Guidelines)

DOI: doiglobal.org/10.2024/NAIRJCMD.051

IMAGE-BASED POTHOLE DETECTION AND SIZE ESTIMATION ON INDIAN ROADS USING DIGITAL IMAGE PROCESSING TECHNIQUES

DR. JYOTI. Y. KULKARNI

Associate Professor Govt. First Grade College, Kalagi.

ABSTRACT

Potholes are a recurring pavement defect that affects ride quality, vehicle operating cost and road safety. Conventional field inspection requires personnel to identify defects and manually measure their dimensions, making large-scale surveys time-consuming and potentially hazardous. This paper presents an image-based framework for pothole detection and size estimation on Indian roads using digital image processing techniques. The proposed methodology combines road-image acquisition, grayscale transformation, denoising, contrast enhancement, edge extraction, adaptive segmentation, morphological processing, contour analysis and geometric calibration. Pothole length, width and image-plane area are obtained from segmented contours and converted from pixels to physical units through a calibrated reference or perspective-aware scale. The paper emphasizes that monocular pixel measurements cannot be treated as real-world dimensions without calibration because camera height, pitch and perspective alter apparent size. Published benchmark studies are used only for comparative discussion: a 2020 UAV image-processing study reported 80% precision and 74.4% recall; a 2024 YOLO-based dimension-estimation study reported 93.0% precision, 91.6% recall, 87.0% F1-score and 96.3% mAP; and a 2025 segmentation-plus-inverse-perspective-mapping study reported an average dimension-estimation error of 11.30%. The proposed classical image-processing workflow is positioned as a low-cost and interpretable baseline suitable for controlled road imagery, with deep learning suggested as an extension for complex shadows, patches and heterogeneous pavement textures. The study concludes that reliable pothole measurement requires not only segmentation accuracy but also careful geometric calibration and field validation.

***KEYWORDS**: pothole detection; Indian roads; digital image processing; edge detection; contour analysis; image segmentation; dimension estimation; road maintenance*

1. INTRODUCTION

Road-surface condition is an important component of transportation safety and infrastructure management. Potholes develop when pavement materials deteriorate under traffic loading, water ingress, weathering, inadequate drainage, construction defects or repeated maintenance failure. Their irregular geometry creates difficulties for both manual measurement and automated analysis.

In India, the diversity of road surfaces, traffic environments, lighting conditions and maintenance practices makes automated pothole analysis particularly challenging. Images may contain shadows, repaired patches, lane markings, oil stains, standing water, gravel and vehicles. Several of these objects can resemble potholes in a two-dimensional image.

Digital image processing offers a comparatively inexpensive approach because ordinary cameras can capture road scenes and software can extract candidate defect regions. Classical methods are attractive where computational resources are limited and where interpretability is important. However, detection and physical measurement are distinct tasks. A segmented pothole can be measured in pixels immediately, but conversion into centimetres or metres requires a geometric scale.

This paper develops a rigorous framework for detecting potholes and estimating their dimensions from Indian road images. It combines classical image-processing stages with calibration, contour geometry and published benchmark evidence.

2. PROBLEM STATEMENT

Manual pothole surveys are labour intensive and expose personnel to traffic. Simple image thresholding, however, can produce false detections because potholes often have non-uniform colour and illumination. Furthermore, the apparent pixel length of the same pothole changes with camera distance and viewing angle.

The research problem is therefore: How can digital image-processing techniques detect pothole boundaries in road images and estimate pothole size while accounting for noise, shadows, perspective and pixel-to-real-world conversion?

3. OBJECTIVES

- To design an image-processing pipeline for automatic pothole-region extraction from Indian road images.
- To evaluate grayscale, filtering, edge, thresholding and morphological operations for road-surface segmentation.
- To estimate pothole length, width and projected area using contour geometry.
- To define a valid pixel-to-real-world calibration method rather than assuming a constant pixel scale.
- To compare the proposed methodology with published pothole-detection and dimension-estimation research.

- To identify limitations of monocular 2-D analysis and recommend extensions using perspective transformation, stereo vision or deep learning.

4. RESEARCH QUESTIONS

- Which preprocessing operations best suppress road-image noise while preserving pothole boundaries?
- Can contour-based segmentation distinguish potholes from shadows and repaired road patches?
- How should pothole length and width be defined for irregular shapes?
- What calibration is required to convert image pixels into physical dimensions?
- How does a classical image-processing pipeline compare conceptually with modern deep-learning approaches?

5. LITERATURE REVIEW

Pothole detection has evolved from handcrafted image-processing methods toward machine learning, deep learning, stereo vision and three-dimensional reconstruction. Classical approaches commonly use filtering, edge detection, thresholding, morphology and contour extraction. Their principal advantages are low computational cost and transparent processing steps, while their weakness is sensitivity to environmental variation.

Pehere et al. (2020) reported a UAV-based image-processing approach using median filtering and edge detection, with 80% precision and 74.4% recall. This provides a useful classical benchmark because the approach is close to the digital image-processing focus of the present study.

More recent systems combine learning-based detection with geometric measurement. A 2024 study in Machine Learning with Applications used YOLOv5 and in-vehicle technologies for pothole detection and dimension estimation, reporting 93.0% precision, 91.6% recall, 87.0% F1-score and 96.3% mAP.

Khamkaew et al. (2025) combined YOLOv8n-seg with inverse perspective mapping and reported precision of 0.745, mAP@0.5 of 0.708 and an average dimension-estimation error of 11.30%. This result is especially relevant because it demonstrates that accurate detection does not automatically imply exact physical measurement; geometry remains a separate source of error.

A broad review by Ma et al. (2022) documents the transition from 2-D image processing and 3-D modeling toward deep neural approaches. Nevertheless, classical methods remain useful as baselines, for constrained imaging systems and for explaining the geometric stages required for size estimation.

6. RESEARCH GAP

A common weakness in pothole-image studies is the treatment of pixel size as if it were physical size. Without camera calibration or a reference object on the road plane, an image cannot uniquely determine real-world pothole length. Perspective makes distant objects appear smaller and oblique viewing distorts shape.

A second gap is insufficient distinction between detection accuracy and measurement accuracy. A detector can correctly locate a pothole but still estimate its dimensions poorly. The proposed framework therefore evaluates these stages separately.

A third gap concerns Indian road variability. Methods validated on homogeneous pavement may perform differently on patched asphalt, dusty surfaces, water-filled depressions and mixed urban roads. The proposed data protocol therefore includes multiple surface and illumination conditions.

7. PROPOSED METHODOLOGY

The proposed pipeline contains image acquisition, calibration, region-of-interest extraction, grayscale conversion, denoising, contrast enhancement, segmentation, edge detection, morphology, contour extraction, candidate filtering and geometric measurement. Each stage produces an interpretable intermediate result.

For controlled experiments, the camera should be mounted at a known height and pitch. A planar calibration target or road-plane reference of known dimensions is captured under the same geometry. For handheld images, a visible reference marker near the pothole is preferable.

8. IMAGE ACQUISITION FOR INDIAN ROAD CONDITIONS

Images should be collected from urban, semi-urban and rural roads with asphalt surfaces and a range of defect severities. Each physical pothole should receive a unique identifier. Multiple views may be captured, but train/test-style comparisons must not treat repeated photographs of the same pothole as independent field cases.

Metadata should record camera model, resolution, approximate camera height, viewing angle, time of day, surface condition, weather, pothole identifier and manually measured ground-truth dimensions.

Ground truth should be measured safely using a measuring tape, calibrated scale or survey method when traffic conditions permit. Length can be defined as the maximum Feret diameter; width as the maximum perpendicular dimension or minor axis of an oriented bounding rectangle.

9. REGION OF INTEREST

The road region should be isolated before defect detection. In a vehicle-mounted camera, the lower central portion of the image usually contains the relevant pavement. Removing sky, buildings and vehicles reduces false edges and computational load.

A trapezoidal road region can be transformed into a bird's-eye representation if camera parameters are known. This is particularly useful for dimension estimation because measurements on the road plane become more consistent.

10. GRAYSCALE CONVERSION

RGB road images are converted to grayscale for edge- and intensity-based processing. A standard luminance transformation can be written as $I = 0.299R + 0.587G + 0.114B$. Grayscale reduces dimensionality while preserving major intensity transitions.

Colour information may still be retained in a parallel branch because some patches and stains can be separated using colour-space features. The proposed baseline, however, focuses on grayscale digital processing.

11. NOISE REDUCTION

Road images may contain sensor noise, gravel texture and small surface irregularities. Gaussian filtering reduces high-frequency noise but can blur boundaries. Median filtering is effective against impulse-like noise and preserves edges relatively well.

Filter size should be selected empirically. Excessive smoothing can merge small potholes with the surrounding pavement, while insufficient smoothing generates fragmented contours.

12. CONTRAST ENHANCEMENT

Uneven illumination is a major challenge. Histogram equalization can improve global contrast, while Contrast Limited Adaptive Histogram Equalization (CLAHE) enhances local contrast without excessively amplifying noise.

CLAHE is particularly useful when the road contains both shaded and sunlit areas. Nevertheless, strong shadows may still require adaptive rather than global thresholding.

13. EDGE DETECTION

The Canny edge detector is proposed as the primary boundary extraction method because it combines smoothing, gradient estimation, non-maximum suppression and hysteresis thresholding. Sobel gradients can be used as a simpler comparison.

Pothole edges are often incomplete because parts of the boundary have weak contrast. Therefore, edge detection alone is insufficient; morphological closing and region analysis are required to create coherent candidates.

14. SEGMENTATION AND THRESHOLDING

Adaptive thresholding calculates local thresholds and is preferable to a single global value when illumination varies across the image. Otsu's method can serve as a global baseline where the intensity histogram is sufficiently separable.

Dark-region segmentation should be treated cautiously because shadows and repaired asphalt may also be dark. Candidate regions must therefore be filtered using shape, texture, location and boundary evidence.

15. MORPHOLOGICAL PROCESSING

Morphological opening removes small isolated components, while closing fills narrow gaps and connects nearby edge fragments. Dilation can strengthen weak boundaries, and erosion can remove small protrusions.

A sequence such as closing followed by opening can produce smoother pothole masks. Kernel dimensions should be proportional to image resolution rather than fixed arbitrarily across all datasets.

16. CONTOUR EXTRACTION

Connected components or contours are extracted from the binary mask. Each contour is described by area, perimeter, bounding rectangle, convex hull, circularity and aspect ratio.

Very small contours can be discarded as road texture. Candidate filters should be validated against labeled examples because aggressive shape constraints may remove genuine irregular potholes.

17. POTHOLE LENGTH AND WIDTH IN PIXELS

For an irregular contour C, the simplest axis-aligned bounding box may overestimate dimensions when the pothole is rotated. An oriented minimum-area rectangle is preferable. Its longer side can be treated as image-plane length and its shorter side as image-plane width.

Alternatively, maximum Feret diameter can define length: $L_{px} = \max \|p_i - p_j\|$ for contour points p_i and p_j. Width can then be measured along the principal perpendicular direction.

Projected image area A_{px} is the number of pixels enclosed by the contour. This is not physical area until geometric calibration is applied.

18. PIXEL-TO-REAL-WORLD CALIBRATION

A constant conversion s centimetres per pixel is valid only under controlled planar geometry where the pothole and reference occupy approximately the same road plane and perspective effects have been corrected. If a known reference length L_{real} corresponds to L_{ref} pixels after rectification, $s = L_{real}/L_{ref}$.

Then estimated pothole length is $L_{real\_est} = s \times L_{px}$ and width is $W_{real\_est} = s \times W_{px}$. Area becomes $A_{real\_est} = s^2 \times A_{px}$.

For oblique vehicle images, homography or inverse perspective mapping should transform the road plane to a top-down coordinate system before applying metric scale. This is one of the most important requirements of the proposed method.

19. PERSPECTIVE TRANSFORMATION

A planar homography maps image coordinates to road-plane coordinates using a 3×3 projective transformation. Four or more non-collinear calibration correspondences can estimate the homography. Once rectified, pothole contours can be measured in a common road-plane coordinate system.

Inverse perspective mapping does not recover true depth. It improves planar length, width and area estimation. Pothole depth requires stereo, structured light, LiDAR, photogrammetry or another source of three-dimensional information.

20. SEVERITY CLASSIFICATION

For maintenance prioritization, detected potholes can be grouped into size-based categories using locally adopted engineering thresholds. The paper intentionally does not invent universal severity cut-offs because agencies may define severity differently.

A defensible system should store measured length, width and area and allow the responsible road authority to configure classification thresholds.

21. ALGORITHM

- Input calibrated RGB road image.
- Select or transform the road region of interest.
- Convert ROI to grayscale.
- Apply median or Gaussian denoising.
- Enhance local contrast using CLAHE if required.
- Generate candidate mask using adaptive thresholding and/or Canny edges.
- Apply morphological closing/opening.
- Extract connected contours.
- Reject implausible candidates using minimum area and road-position rules.
- Compute oriented bounding geometry and pixel area.
- Transform contour into calibrated road-plane coordinates.
- Calculate physical length, width and projected area.
- Output pothole location, dimensions, confidence/quality flag and annotated image.

22. EVALUATION METRICS

Detection performance should be reported using precision, recall and F1-score. Precision = TP/(TP+FP), Recall = TP/(TP+FN), and F1 = 2PR/(P+R). Intersection over Union can measure overlap between predicted and manually annotated pothole masks.

Dimension estimation requires separate metrics. For measured length, absolute percentage error is |L_est - L_true|/L_true × 100. Mean Absolute Percentage Error can summarize multiple potholes, although it should be used carefully for very small ground-truth values.

Width and area errors should be reported independently. Agreement plots or correlation alone are insufficient because high correlation can coexist with systematic measurement bias.

23. PUBLISHED COMPARATIVE RESULTS

Study	Method	Precision	Recall	F1 / mAP	Dimension result
Pehere et al. (2020)	UAV + image processing	80%	74.4%	Not reported	Not reported
2024 MLWA study	YOLOv5 + vehicle camera	93.0%	91.6%	F1 87.0%; mAP 96.3%	On-site statistical validation
Khamkaew et al. (2025)	YOLOv8n-seg + IPM	74.5%	—	mAP@0.5 70.8%	11.30% average dimension error
Park & Nguyen (2025)	Two-camera vision	—	—	—	Measurement error within 5% reported

Table 1 presents published benchmark values for context. These figures belong to the cited studies and are not claimed as experimental results of the proposed Indian-road image-processing pipeline.

24. INTERPRETATION OF PUBLISHED RESULTS

The published results show a clear distinction between detection and measurement. The 2024 YOLOv5 study reported strong detection metrics, while the 2025 IPM study explicitly quantified dimension error at 11.30%. A two-camera 2025 measurement study reported errors within 5%, illustrating the benefit of richer geometric information.

The 2020 UAV study's 80% precision and 74.4% recall demonstrate that classical image processing can provide useful detection but may struggle with environmental variability. The proposed method therefore uses classical processing as an interpretable baseline while incorporating calibration and perspective correction more explicitly.

25. PROPOSED RESULTS TABLE FOR INDIAN ROAD DATASET

Method	Precision	Recall	F1	IoU	Length error	Width error
Canny + Morphology	To be measured	To be measured	To be measured	To be measured	To be measured	To be measured
Adaptive Threshold + Contours	To be measured	To be measured	To be measured	To be measured	To be measured	To be measured
Combined Proposed Pipeline	To be measured	To be measured	To be measured	To be measured	To be measured	To be measured

These cells must be populated only after the Indian-road dataset is processed and compared with manually measured ground truth. Published numbers from other datasets should not be inserted into this table.

26. SOURCES OF MEASUREMENT ERROR

Measurement error may arise from inaccurate contour boundaries, calibration error, road non-planarity, camera vibration, lens distortion and uncertainty in manual ground truth. Water-filled potholes can hide the true boundary, while shadows can enlarge the detected region.

The system should therefore attach a quality flag to measurements when calibration is uncertain or the pothole lies outside the reliable road-plane region.

27. INDIAN ROAD-SPECIFIC CHALLENGES

Indian road scenes can contain heterogeneous pavement repairs, dust, open drainage edges, speed breakers, lane paint, debris, water and mixed traffic. A purely darkness-based detector is particularly vulnerable in these environments.

Seasonal conditions are also important. Monsoon images may contain water-filled depressions and reflections, while dry-season roads may contain high-contrast cracks and dust. Dataset collection should therefore span multiple conditions rather than relying on a single road segment.

28. PRACTICAL APPLICATIONS

- Municipal and highway road-condition surveys.
- Preliminary defect measurement for maintenance prioritization.
- Vehicle-mounted road inspection systems.
- Mobile applications for trained field inspectors.

• UAV-assisted surveys where legally and operationally appropriate.

• Creation of geotagged pothole inventories for maintenance planning.

29. LIMITATIONS

A single uncalibrated image cannot provide reliable physical pothole dimensions. The proposed method therefore requires calibration, a reference object or a known camera-road geometry. It estimates surface dimensions but not true pothole depth.

Classical segmentation is sensitive to lighting and road texture. Deep-learning segmentation may generalize better when trained on diverse data, but it introduces dataset and computational requirements. Neither approach eliminates the need for field validation.

30. FUTURE SCOPE

Future research should combine classical geometric measurement with deep-learning segmentation. YOLO-seg or Mask R-CNN can provide robust pothole masks, while inverse perspective mapping supplies physical dimensions.

Stereo cameras can estimate depth and volume. Smartphone inertial sensors and GPS can add geolocation and roughness information. A larger Indian-road dataset spanning climatic zones, pavement types and seasons would support stronger external validation.

Integration with municipal GIS could automatically prioritize repairs according to measured size, traffic importance and repeated citizen reports.

31. CONCLUSION

Digital image processing provides a low-cost and interpretable foundation for pothole detection and size estimation. The proposed pipeline combines grayscale conversion, denoising, contrast enhancement, edge and threshold segmentation, morphology, contour analysis and calibrated geometry.

The most important methodological conclusion is that pixel length is not equivalent to physical length. Reliable measurement requires a scale defined on the road plane and, for oblique imagery, perspective correction. Detection metrics and dimensional error must also be reported separately.

Published research demonstrates the feasibility of automated pothole analysis, ranging from classical UAV image processing to YOLO segmentation, inverse perspective mapping and two-camera measurement. For Indian roads, a hybrid approach that combines robust segmentation with explicit geometric calibration is likely to provide the most practical path toward scalable road-condition assessment.

REFERENCES

1. Pehere, S., Sanganwar, P., Pawar, S., & Shinde, A. (2020). Detection of pothole by image processing using UAV. Journal of Science & Technology, 5(3), 101–110. https://doi.org/10.46243/jst.2020.v5.i3.pp101-110
2. Ma, N., Fan, J., Wang, W., Wu, J., Jiang, Y., Xie, L., & Fan, R. (2022). Computer vision for road imaging and pothole detection: A state-of-the-art review of systems and algorithms. Transportation Safety and Environment / preprint literature.
3. Augmenting roadway safety with machine learning and deep learning: Pothole detection and dimension estimation using in-vehicle technologies. (2024). Machine Learning with Applications, 16, 100547. https://doi.org/10.1016/j.mlwa.2024.100547
4. Khamkaew, N., Jitpairod, K., Rinchumphu, D., Nonjui, Y., Me-ngoen, T., Khansila, N., & Thueakchaikham, B. (2025). Pothole detection and dimension estimation via image transformation and scaling with Thai road data integration. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-4/W7-2025, 75–81. https://doi.org/10.5194/isprs-annals-X-4-W7-2025-75-2025
5. Park, S. S., & Nguyen, N. N. (2025). Two-camera vision technique for measuring pothole area and depth. Measurement, 247, 116809. https://doi.org/10.1016/j.measurement.2025.116809
6. Katageri, M., Mandal, M., Gandhi, M., Koregaonkar, N., & Sengupta, S. (2016). Automated management of pothole related disasters using image processing and geotagging. arXiv:1610.08808.
7. Canny, J. (1986). A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6), 679–698.
8. Gonzalez, R. C., & Woods, R. E. (2018). Digital image processing (4th ed.). Pearson.