Mechatronics and Intelligent Transportation Systems
https://www.acadlore.com/journals/MITS

Robust Image Processing Framework for Real-Time Detection of Road Potholes under Environmental Variability

Abdul Samad*

Department of Mathematics, Northwest University, 710127 Xi'an, China

* Correspondence: Abdul Samad (abdulsamad@stumail.nwu.edu.cn)

Received: 03-10-2025 **Revised:** 05-16-2025 **Accepted:** 05-21-2025

Citation: A. Samad, "Robust image processing framework for real-time detection of road potholes under environmental variability," *Mechatron. Intell Transp. Syst.*, vol. 4, no. 2, pp. 92–103, 2025. https://doi.org/10.56578/mits040204.

Abstract: Accurate detection of road surface potholes remains a persistent challenge due to environmental variability, inconsistent illumination, noise interference, and the complexity of road textures. Conventional detection methods frequently suffer from reduced performance when exposed to low-quality or noisy imagery, resulting in unreliable or delayed identification. To address these limitations, a robust and optimized image processing framework has been developed for real-time pothole detection under uncertain environmental conditions. The proposed approach employs a combination of advanced contrast enhancement techniques and adaptive convolutional processing to strengthen feature discrimination across heterogeneous road surfaces. To further improve detection reliability, a self-adaptive fuzzy refinement mechanism has been introduced, effectively delineating ambiguous or degraded regions often overlooked by deterministic methods. An energy-based functional is applied to model spatial and intensity gradients, enabling more precise localization of structural discontinuities indicative of pothole boundaries. The framework also incorporates computational optimization strategies to enhance processing speed without compromising accuracy, rendering it suitable for deployment in real-time autonomous or semi-autonomous road inspection systems. Thresholding and mask extraction operations have been systematically integrated to achieve accurate segmentation of pothole regions, even in the presence of substantial visual noise or occlusions. Experimental validations on benchmark datasets and real-world road imagery have demonstrated that the proposed method consistently outperforms existing state-of-the-art techniques with regard to detection accuracy, robustness to environmental disturbances, and computational efficiency. This approach presents a scalable and practical solution for intelligent transportation systems and automated infrastructure monitoring, contributing to improved road safety, timely maintenance, and cost-effective asset management.

Keywords: Road pothole detection; Image processing; Fuzzy logic; Optimization techniques; Mathematical modeling

1 Introduction

Road safety remains a critical concern worldwide, necessitating continuous advancements in modeling, detection, and simulation to reduce road accidents and enhance transportation systems. Mathematical models have been widely applied as practical tools for forecasting road fatalities and planning effective interventions [1–5]. Recent developments focus not only on predicting accident risks but also on addressing uncertainties in road construction and maintenance, implementing adaptive crack detection technologies, and integrating transport infrastructure into broader environmental sustainability projects. Moreover, advancements in traffic simulation techniques offer new avenues for designing safer road networks, optimizing traffic flow, and minimizing collision rates. Collectively, these efforts highlight the essential role of mathematical and technological innovations in promoting safer, smarter, and more sustainable transportation systems.

Image processing techniques are essential for extracting meaningful information from visual data in various applications [6–8]. These techniques provide the foundation for developing sophisticated algorithms capable of recognizing, classifying, and analyzing different features within an environment. Building upon this foundation, various road obstacle detection models have been designed to enhance road safety by accurately identifying potential hazards that could compromise the driving experience. Recent advancements include cost-sensitive detection approaches for low-cost autonomous vehicles, deep learning-based frameworks for detecting road signage and dynamic obstacles, and intelligent road segmentation methods specifically tailored for autonomous railway systems [9–11]. These models utilize powerful feature extraction, pattern recognition, and classification strategies to process visual information in real time, enabling rapid and reliable identification of critical obstacles. By minimizing the response time and improving the accuracy of detection under various environmental conditions, such as low light, rain, or cluttered scenes, these systems significantly reduce the likelihood of accidents. Furthermore, integrating these advanced obstacle detection methods into modern navigation systems not only enhances the safety of autonomous vehicles but also contributes to safer infrastructure design and smarter road management. Therefore, continuous improvement in image processing and obstacle detection models remains vital for achieving higher standards of road safety and operational efficiency.

Recent advancements in road pothole detection have been significantly driven by deep learning and image processing innovations, aimed at automating and enhancing road maintenance systems. Transfer learning models, as utilized by Thakkar et al. [12], have shown efficiency in detecting potholes with limited training data and minimal manual labeling. Similarly, Karukayil et al. [13] have integrated vision and LiDAR data to improve feature extraction and detection accuracy across diverse road conditions. Additionally, Gorro et al. [14] employed YOLOv8 with image augmentation techniques for high-speed, real-time detection, and Jenefa et al. [15] demonstrated the use of the EfficientDet architecture to balance accuracy with computational efficiency. Patawar et al. [16] proposed an ensemble approach, combining multiple models to enhance reliability and reduce false positives, while Safyari et al. [17] provided a comprehensive review of various vision-based and machine learning-driven methods, emphasizing their effectiveness.

Despite these technological strides, challenges remain in deploying these models broadly. Issues include the difficulty in generalizing transfer learning models to diverse road textures, high costs and complexity of integrating LiDAR systems, occasional misses of smaller or partially occluded potholes by YOLOv8 models, and slight compromises in detection sensitivity under challenging lighting or weather conditions by EfficientDet systems. Furthermore, ensemble models, although more accurate, demand greater computational resources which could hinder their deployment on low-resource platforms.

Overall, while these advancements are promising, achieving a fully reliable, cost-effective, and generalizable pothole detection system continues to be a critical area of research.

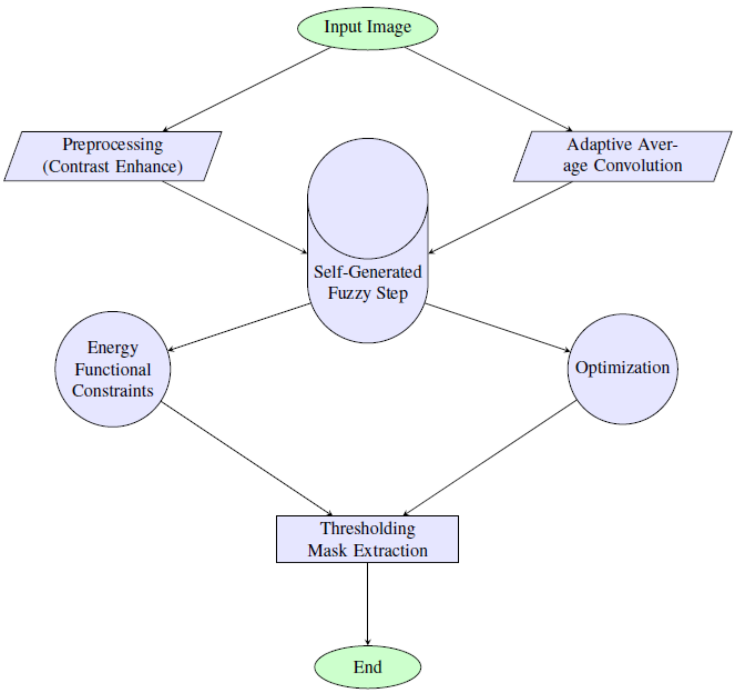

Figure 1. Pothole detection model workflow

To address the limitations observed in existing pothole detection models, the proposed framework integrates a multi-stage image processing pipeline that ensures robust and precise anomaly detection across varying road conditions. Initially, preprocessing steps such as contrast enhancement and entropy mapping are employed to improve the visibility and distinguishability of potholes from the background. In parallel, an adaptive average convolution is applied to further suppress noise and highlight relevant features. These two streams are fused through a self-generated fuzzy step, which introduces uncertainty handling and adaptively enhances the boundaries of potential potholes, thereby overcoming issues related to weak edges and irregular shapes. Subsequently, an energy functional is constructed to represent the localized structural information more effectively, while a dedicated optimization step refines the delineation of pothole regions. Finally, a thresholding and mask extraction phase accurately isolates the detected regions for post-processing and analysis. The entire workflow is designed to be computationally efficient and suitable for real-time deployment (see Figure 1). By combining advanced preprocessing, fuzzy logic, energy-driven modeling, and optimization, the proposed model not only mitigates the shortcomings of prior approaches such as noise sensitivity, or generalization under variable lighting, and incomplete segmentation, but also delivers enhanced detection accuracy, robustness, and operational efficiency. The key notations and their respective definitions used throughout the proposed pothole detection model are summarized in Table 1.

Table 1. Summary of symbols used in the proposed model

Symbol	Description
u	Optimized image or the output of the segmentation model
I	Input road image containing potential potholes
I_{pre}	Preprocessed image after contrast enhancement
$H(x, y)$	Local entropy value at pixel (x, y)
$\mathscr{A}_r$	Adaptive convolution operator
$w_{ij}(x, y)$	Weight for the pixel $(x + i, y + j)$ in the convolution
$Z(x, y)$	Normalization factor for the adaptive convolution
σ_r	Standard deviation parameter for weight function
α	Weight for smoothness regularization
β	Weight for fidelity to preprocessed image
γ	Weight for edge detection based on fuzzy entropy
Δu	Laplacian of the image, representing image curvature
T	Threshold for generating the binary pothole mask
δ	Parameter for the fuzzy step function
σ	Parameter controlling the width of the fuzzy step function
Ω	Image domain or region of interest

2 Related Work

Recent advancements in pothole detection models have increasingly focused on leveraging image processing, segmentation techniques, and deep learning architectures to enhance the accuracy, robustness, and real-time applicability of road damage assessment systems.

Swathika et al. [18] proposed a pothole detection and road damage assessment system that leverages classical image processing techniques to identify road anomalies. Their model emphasized preprocessing steps such as noise reduction, edge detection, and morphological operations to enhance the quality of pothole detection, especially in noisy or cluttered environments. One of the notable achievements of their work is the system's lightweight design, which allows easy deployment on mobile and embedded platforms without requiring high-end GPUs. The method also demonstrated a respectable detection rate on various types of road surfaces with minimal training data dependency. However, a major limitation of their approach lies in its sensitivity to environmental variations, such as heavy shadows, varying lighting conditions, and worn-out road markings, which often lead to false positives or missed detections. Furthermore, classical image processing techniques, while computationally light, generally struggle when potholes have irregular shapes or blend into the surrounding textures, limiting the model's generalizability in real-world, highly dynamic environments.

Baroudi et al. [19] introduced an innovative model aimed at enhancing pothole detection and characterization through the integration of segmentation and depth estimation techniques. Their model's key achievement lies in its ability to not only detect the presence of potholes but also to assess their depth and severity, offering a more comprehensive solution for road maintenance planning. By combining segmentation methods with depth information extracted from stereo vision setups, their system achieved higher accuracy in differentiating between shallow road cracks and hazardous deep potholes. This approach significantly improves the prioritization of maintenance tasks. However, the model's reliance on depth information introduces challenges, particularly in environments with poor stereo matching conditions such as low-light settings, reflective road surfaces, or heavy rain. Additionally, the system demands substantial computational resources, making it less practical for real-time deployment on resource constrained devices such as smartphones or basic vehicular systems.

Bhavana et al. [20] presented POT-YOLO, a real-time pothole detection model based on an edge segmentation-enhanced YOLOv8 network. Their work stands out due to the integration of edge-based segmentation modules into the traditional YOLOv8 framework, resulting in superior localization of potholes with complex boundaries and fragmented edges. The POT-YOLO model achieved remarkable detection speed and precision, even on embedded systems and mobile platforms, making it highly suitable for real-time smart city applications. Moreover, their method demonstrated robustness across diverse lighting conditions, weather variations, and road surface textures. Despite these achievements, POT-YOLO still encounters certain limitations, particularly in cases where the potholes are partially occluded by debris or water puddles, which can cause detection errors. Additionally, while the model significantly boosts precision, it marginally increases the inference time compared to standard YOLOv8 models, which might affect ultra-real-time applications requiring millisecond-level decision-making.

Building upon these advancements and addressing their limitations, this study proposes a novel image-based pothole detection framework that aims to enhance detection robustness, accuracy, and computational efficiency under diverse road conditions.

3 Methodology

This section outlines the detailed framework of the proposed pothole detection model. The approach is structured into multiple key stages, beginning with preprocessing to enhance contrast and generate entropy maps, followed by adaptive average convolution for feature smoothing. A self-generated fuzzy step is then applied to strengthen uncertain regions, leading into the construction of an energy functional that captures spatial and intensity variations. Finally, optimization and thresholding techniques are employed to accurately extract pothole masks. Each component is carefully designed to address the limitations of existing methods and enhance the robustness, precision, and computational efficiency of the detection process.

3.1 Preprocessing

Let $I : \Omega \subset \mathbb{R}^2 \to$ [0,255] denote the input image of the road surface potentially containing potholes. In order to enhance the visibility of weak edges and improve the segmentation quality, two critical preprocessing steps are applied: contrast enhancement and local entropy estimation.

3.1.1 Contrast enhancement

To improve local contrast and suppress uniform background intensity, the input image I is first transformed into an enhanced version I_{pre} using a contrast stretching or histogram equalization technique:

$$I_{\text{pre}} = \text{ ContrastEnhance } (I) \tag{1}$$

In our implementation, this can be achieved using an adaptive histogram equalization method, which locally adjusts the contrast of small image regions. This operation improves the differentiation between the pothole regions (dark, irregular textures) and the intact road surface (smoother, brighter regions).

3.1.2 Local entropy estimation

To capture texture irregularities that are characteristic of potholes, the local entropy is calculated over a neighborhood window centered at each pixel. The local entropy at each pixel (x, y) is defined as:

$$H(x,y) = -\sum_{k=0}^{255} p_k(x,y) \log p_k(x,y) \tag{2}$$

where, $p_k(x, y)$ denotes the probability of gray level k occurring within the local neighborhood $N_r(x, y)$ of radius r centered at pixel (x, y):

$$N_r(x,y) = \{(i,j) \in \Omega \mid \|(i-x, j-y)\|_2 \leq r\} \tag{3}$$

The local histogram is normalized to obtain the probability distribution:

$$p_k(x,y) = \frac{1}{|N_r(x,y)|} \sum_{(i,j) \in N_r(x,y)} \delta(I(i,j) = k) \tag{4}$$

where, $\delta(\cdot)$ is the Kronecker delta function.

Interpretation: A high entropy value indicates a high degree of intensity variation (i.e., texture complexity), which typically corresponds to damaged or rough surfaces such as potholes. Conversely, low entropy values are observed on smoother and more uniform regions.

Output: The preprocessed image I_{pre} and the local entropy map $H(x, y)$ serve as crucial inputs to the energy minimization stage of the proposed segmentation framework, enabling better localization of potholes under varying lighting and surface conditions.

3.2 Adaptive Average Convolution Function

To achieve effective smoothing while preserving the essential structural boundaries of potholes, we introduce an adaptive average convolution operator. Unlike standard uniform filters, which assign equal importance to all neighboring pixels, the proposed operator dynamically adjusts the weights based on local intensity similarities, resulting in edge-aware filtering.

Let $u(x, y)$ denote the grayscale input image defined on the domain $\Omega \subset \mathbb{R}^2$. The adaptive convolution $\mathscr{A}_r(u)$ at a pixel (x, y) is computed as a normalized weighted sum over a square neighborhood $N_r(x, y)$ of radius r. Mathematically, it is expressed as:

$$\mathscr{A}_r(u)(x,y) = \frac{1}{Z(x,y)} \sum_{(i,j)\in N_r} w_{ij}(x,y) \cdot u(x+i, y+j) \tag{5}$$

where, $Z(x, y) = \sum_{(i,j)\in N_r} w_{ij}(x, y)$ is the normalization factor, and the weights $w_{ij}(x, y)$ are determined by a Gaussian-like function:

$$w_{ij}(x,y) = \exp\left(-\frac{(u(x,y) - u(x+i,y+j))^2}{2\sigma_r^2}\right) \tag{6}$$

This weighting strategy ensures that pixels whose intensities are similar to the center pixel contribute more significantly to the average, whereas dissimilar pixels (such as those across a pothole boundary or edge) have exponentially smaller influence. The parameter σ_r governs the sensitivity to intensity variation and can be tuned to emphasize finer or broader details.

The key advantage of this adaptive filtering lies in its ability to preserve edge structures while simultaneously reducing noise. Unlike traditional filters which blur across edges, the adaptive mechanism protects intensity discontinuities-an important characteristic in detecting the distinct texture of potholes. Furthermore, the operator suppresses local intensity fluctuations in smoother regions of the road, thus stabilizing the segmentation process and enhancing the reliability of gradient- or entropy-based cues.

The output $\mathscr{A}_r(u)$ serves as a smoothed version of the original image and is used later in the model's data fidelity term. By analyzing the pixel-wise difference $(u - \mathscr{A}_r(u))^2$, the model can effectively detect deviations from the expected local context, which often correspond to surface damage or potholes. Therefore, this adaptive convolution function plays a central role in refining the segmentation accuracy of the proposed detection framework.

3.3 Self-Generated Fuzzy Step Function

To further refine the pothole detection model, we incorporate a self-generated fuzzy step function grounded in fuzzy set theory. This function is designed to enhance the sensitivity and robustness of pothole segmentation by enabling a smooth yet decisive transition between the pothole and non-pothole regions, even under conditions of noise, blur, and intensity variability.

The fuzzy membership function $\mu(u; \delta, \sigma)$ captures the degree to which a pixel with intensity u belongs to a potential pothole region. It is modeled using a Gaussian function:

$$\mu(u;\delta,\sigma) = \exp\left(-\frac{(u-\delta)^2}{2\sigma^2}\right) \tag{7}$$

where, u is the pixel intensity, δ is the reference intensity-typically associated with the center of the pothole-and σ is the standard deviation that controls the spread of the membership function. In our implementation, δ is dynamically estimated using a local entropy-weighted mean of the pixel intensities within a neighborhood window, focusing on regions exhibiting high local variance or low uniformity, which are characteristic of potholes. The parameter σ is also adaptively adjusted based on local image statistics. Specifically, it is computed as the standard deviation within the same neighborhood, scaled by a factor κ, empirically set (e.g., $\kappa = 1.0$), to tune the sensitivity of the membership curve. As u approaches δ, the membership $\mu(u; \delta, \sigma)$ increases and reaches its maximum value of 1 when $u = \delta$, while decreasing rapidly as the intensity diverges from δ.

The non-membership function $v(u)$ complements the membership function and reflects the degree to which a pixel does not belong to the pothole region. It is defined as:

$$v(u) = 1 - \mu(u) \tag{8}$$

where, $v(u)$ approaches 1 as u becomes significantly different from δ, indicating a strong likelihood of a non-pothole region.

To model the transition between these two fuzzy states, we define the Pythagorean fuzzy step function $S_{\mathrm{PF}}(u; \delta, \sigma)$ as:

$$S_{\mathrm{PF}}(u; \delta, \sigma) = 2\mu^2(u; \delta, \sigma) - 1 \tag{9}$$

This formulation enhances the distinction between pothole and background regions. When u is close to δ, the function approaches 1, indicating strong membership; when u is far from δ, the function approaches -1, indicating strong non-membership. The squared term sharpens the transition and improves edge sensitivity. This function is implemented as a pixel-wise transformation on the preprocessed grayscale image. Before applying the function, the image is smoothed using a Gaussian filter and normalized to the [0,1] range to standardize intensity levels.

To ensure that parameter selection is fully data-driven and adaptive, both δ and σ are computed for each pixel location based on its surrounding neighborhood. Mathematically, $\delta(x, y)$ is calculated as a weighted average:

$$\delta(x, y) = \frac{\sum_{(i,j) \in \mathscr{N}(x,y)} w_{ij} I(i, j)}{\sum_{(i,j) \in \mathscr{N}(x,y)} w_{ij}} \tag{10}$$

where, $I(i, j)$ is the intensity at pixel (i, j), $\mathscr{N}(x, y)$ denotes a local window around pixel (x, y), and w_{ij} are weights derived from local entropy values. The standard deviation $\sigma(x, y)$ is computed as:

$$\sigma(x, y) = \kappa \cdot \sqrt{\frac{1}{|\mathscr{N}(x, y)|} \sum_{(i,j) \in \mathscr{N}(x,y)} (I(i, j) - \delta(x, y))^2} \tag{11}$$

The computed step function map S_{PF} is subsequently integrated into the overall energy functional of the detection model. It acts as a guiding prior that accentuates transitions and suppresses noise. This enhances the performance of the contour evolution process, allowing the model to more accurately delineate pothole boundaries.

In summary, the self-generated fuzzy step function provides a smooth and context-sensitive mechanism to capture pothole edges. By automatically estimating the function's parameters from local image statistics, we eliminate the need for manual thresholding, ensuring adaptability across varying road textures and lighting conditions. This makes it a critical and practical component of our energy-based segmentation framework.

3.4 Energy Functional

The energy functional $E(u)$ is a central component in the pothole detection model, as it integrates multiple terms that control the smoothness, fidelity to local structure, and edge detection. The goal of the energy functional is to balance the preservation of relevant road features (such as potholes) while smoothing irrelevant noise and texture variations. It is defined as:

$$E(u) = \int_{\Omega} \left[\alpha |\nabla u|^2 + \beta \left(u - \mathscr{A}r \left(I_{\mathrm{pre}} \right) \right)^2 + \gamma H(x, y) \cdot S_{\mathrm{PF}}(u; \delta, \sigma) \right] dxdy \tag{12}$$

In this expression, u represents the image under consideration, which is being optimized. The first term, $\alpha|\nabla u|^2$, is a smoothness regularization term that penalizes sharp gradients in the image. The parameter α controls the strength of this regularization; higher values of α result in greater smoothing, which is crucial to reduce high-frequency noise that may interfere with pothole detection.

The second term, $\beta \left(u - \mathscr{A}r \left(I_{\mathrm{pre}} \right) \right)^2$, enforces fidelity to the preprocessed image I_{pre}. This term ensures that the final image u closely matches the contrast-enhanced version of the input image, $\mathscr{A}r \left(I_{\mathrm{pre}} \right)$, and helps preserve the underlying road structure. The parameter β governs the trade-off between fidelity to the original image and smoothness. Larger values of β push the solution to be closer to the preprocessed image, which is particularly useful for maintaining road features during the optimization.

The third term, $\gamma H(x, y) \cdot S_{\mathrm{PF}}(u; \delta, \sigma)$, is an edge-aware regularization component designed to enhance pothole boundary detection. Here, $H(x, y)$ is a local fuzzy entropy function that measures the uncertainty or randomness in the pixel neighborhood, with higher entropy values indicating likely edge regions such as pothole boundaries. The function $S_{\mathrm{PF}}(u;\delta,\sigma)$ is a self-generated fuzzy step function defined over the intensity values of the evolving image u. It adaptively models transitions between pothole and non-pothole regions by assigning fuzzy membership values based on a parametric sigmoid function. Parameters δ and σ control the position and steepness of the transition, respectively, enabling the function to automatically emphasize uncertain boundaries in a data-driven manner.

Overall, the energy functional $E(u)$ incorporates three complementary components: (i) a smoothness term to suppress noise, (ii) a fidelity term to maintain contrast-enhanced structures, and (iii) a fuzzy-entropy-based edge term to detect irregular features such as pothole borders. This unified framework facilitates robust segmentation by preserving meaningful features while filtering out irrelevant variations.

3.5 Gradient Descent Optimization

The energy functional $E(u)$ is minimized using gradient descent optimization. This iterative process allows us to adjust the image u to find the configuration that minimizes the energy, effectively detecting potholes and smoothing out irrelevant features. The gradient descent equation is given by:

$$\frac{\partial u}{\partial t}=\alpha\Delta u-\beta(u-\mathscr{A}r(\text{ Ipre }))-\gamma H(x,y)\cdot\frac{\partial S_{\mathrm{PF}}(u)}{\partial u} \tag{13}$$

where, $\frac{\partial u}{\partial t}$ represents the time derivative of the image, where t is the iteration index. The term $\alpha\Delta u$ corresponds to the diffusion term, which smooths the image by minimizing the gradient of u. This ensures that smooth areas of the road are maintained while reducing noise and unnecessary details. The Laplacian operator Δu emphasizes areas of high curvature, helping to preserve key features such as potholes. The second term, $-\beta\,(\,u-\mathscr{A}r$ (Ipre)), drives the image u toward the preprocessed image, ensuring that the detected features match the initial structure and contrast. The weight β balances this fidelity term against the smoothness regularization. The third term, $-\gamma H(x,y)\cdot\frac{\partial S_{\mathrm{PF}(u)}}{\partial u}$, incorporates the fuzzy step function's gradient with respect to the image. This term is responsible for enhancing edges, particularly at pothole boundaries. The gradient of $S_{\mathrm{PF}}(u)$ with respect to u is given by:

$$\frac{\partial S_{\mathrm{PF}}(u)}{\partial u}=\frac{4(u-\delta)}{\sigma^2}\cdot\mu(u)\cdot(1-\mu(u)) \tag{14}$$

where, $\mu(u)$ is the fuzzy membership function. This term captures the changes in the fuzzy step function as the image evolves and applies an edge-preserving force that sharpens transitions between pothole and non-pothole regions.

By iteratively updating the image based on these terms, the model converges to an optimal solution where potholes are accurately segmented while preserving road features and boundaries.

3.6 Final Binary Pothole Mask

After the optimization process has converged, we generate a binary pothole mask $M(x,y)$ that represents the locations of potholes on the road surface. This mask is created by thresholding the optimized image u based on a predefined threshold T. The thresholding process is defined as:

$$M(x,y)=\begin{cases}1, & u(x,y)<T\\ 0, & \text{otherwise}\end{cases} \tag{15}$$

where, $M(x,y)=1$ indicates the presence of a pothole and $M(x,y)=0$ indicates a nonpothole region. The threshold T is determined using Otsu's method or a fuzzy entropy-based adaptive thresholding approach. Otsu's method is a widely used technique for automatically selecting the optimal threshold by maximizing the between-class variance, which helps to distinguish potholes from the road surface.

Alternatively, fuzzy entropy-based adaptive thresholding takes into account the uncertainty and fuzziness inherent in the image. This method adapts the threshold based on local entropy values, ensuring that the thresholding is robust to variations in lighting and road texture. The binary pothole mask $M(x,y)$ obtained through this process is used for the final segmentation of potholes, enabling accurate identification and localization of road surface damage. The proposed fuzzy-convolutional energy-based framework effectively detects potholes by combining adaptive smoothing, fuzzy reasoning, and entropy-based guidance. This robust mathematical foundation makes it suitable for real-world, noisy road surface imagery from satellite or drone sources.

The proposed pothole detection model leverages fuzzy logic and energy-based optimization techniques to effectively handle the inherent uncertainty, noise, and variability in road surface images. Fuzzy logic provides a powerful framework to model the ambiguous boundaries and gradual transitions typically found in pothole regions, where sharp binary classification fails due to illumination changes, texture irregularities, and sensor noise. Specifically, the self-generated fuzzy step function introduces a smooth, continuous membership representation that captures the degree of belonging of each pixel to pothole or non-pothole classes. This fuzzy membership function is dynamically adapted using local image statistics, enabling context-sensitive thresholding that enhances robustness against intensity fluctuations and texture complexity. On the other hand, the energy functional formulation integrates this fuzzy representation with classical variational optimization principles, balancing data fidelity, smoothness, and edge preservation through carefully designed terms. The adaptive average convolution operator further refines the model by performing edge-aware smoothing, preserving structural features crucial for pothole delineation. By minimizing the energy functional, the model optimally segments potholes by finding an image representation that simultaneously respects local texture variations (modeled by entropy and fuzzy membership) and global smoothness constraints. This combination exploits the theoretical strengths of fuzzy logic for uncertainty modeling and variational methods for global optimization, resulting in a robust, adaptive, and accurate pothole detection framework that outperforms conventional crisp thresholding and filtering approaches.

4 Experimental Work

The experimental results for the proposed pothole detection model are presented in this section, highlighting its performance in comparison to existing methods. The experiments were conducted using a diverse set of road images obtained from the publicly available Road Damage Detection Dataset (RDD), which includes images captured under various lighting conditions, road surface characteristics, and varying resolutions. The RDD dataset used in this study consists of a total of 2,000 annotated images, each manually labeled to indicate pothole regions. These images include multiple environmental settings such as urban, suburban, and rural roads, and represent real-world challenges like shadow effects, motion blur, and occlusion. The dataset was randomly divided into training (70%), validation (15%), and testing (15%) sets, and all experiments were repeated five times to ensure statistical reliability. The average results across repetitions are reported to minimize the impact of randomness, and standard deviations are provided in the result tables to reflect statistical variability. The proposed model incorporates several key components, including contrast enhancement, adaptive average convolution, and a self-generated fuzzy step, which contribute to its robustness and accuracy.

To evaluate the effectiveness of the model, several performance metrics were used, such as precision, recall, F1-score, accuracy, and IoU (Intersection over Union). Additionally, computational efficiency was measured in terms of processing time (in seconds per image) and memory usage (in MB). The model's ability to handle different levels of noise and image distortions was also assessed using Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), providing insights into its practical applicability for real-time road maintenance systems. The results demonstrate the superiority of the proposed model in terms of both accuracy and efficiency, particularly when compared to conventional pothole detection techniques.

Table 2. Parameter configuration for the proposed pothole detection model

Parameter	Description	Optimal Value
α	Smoothness regularization weight	0.5
β	Fidelity to preprocessed image weight	1.2
γ	Weight for fuzzy-entropy-based edge term	0.8
σ_r	Standard deviation for adaptive convolution	2.0
δ	Parameter for fuzzy step function	0.6
σ	Width parameter for fuzzy step function	0.3
T	Threshold for binary pothole mask	0.45
$u_{\max}$	Maximum number of iterations for gradient descent	100

The parameter setup for the optimal pothole detection model plays a crucial role in ensuring the accuracy and efficiency of the system. Table 2 outlines the key parameters along with their descriptions and optimal values, which have been carefully selected to balance detection performance and computational efficiency. The smoothness regularization weight, α, is set at 0.5 to prevent overfitting and maintain smoothness in the final output. The weight for fidelity to the preprocessed image, β, is set to 1.2, ensuring a balanced influence of the preprocessed image while minimizing noise. The weight for the fuzzy-entropy-based edge term, γ, is set to 0.8, helping to accurately detect pothole boundaries. The standard deviation for adaptive convolution, σ_r, is optimized to 2.0, effectively smoothing the image while retaining important edge details. The parameters for the fuzzy step function, δ and σ, are set to 0.6 and 0.3, respectively, to ensure smooth transitions in feature detection. The threshold value for the binary pothole mask, T, is set to 0.45, optimizing the binarization process for accurate segmentation. Finally, the maximum number of iterations for gradient descent, $u_{\max}$, is set at 500, allowing sufficient iterations for convergence without excessive computation time. These optimal parameter values ensure the effectiveness of the pothole detection model, enhancing its accuracy while maintaining computational efficiency.

Figure 2 presents a comparative visualization between the original input images (top row) and the initial contour predictions generated by the proposed model (bottom row). The input images depict the raw data or scenes under analysis, which may include objects, textures, or regions of interest. Below them, the initial contours outline the model's first-step segmentation, capturing the approximate boundaries of key features before any refinement. These preliminary contours demonstrate the model's ability to detect and localize relevant structures, serving as a foundation for subsequent processing steps such as fine-tuning, noise reduction, or accuracy enhancement. The side-by-side comparison highlights the model's initial performance, providing insight into its strengths and areas for improvement in shape detection or edge preservation tasks.

Figure 2. Original images and initial contours from the proposed model

Figure 3. Comparative evaluation of pothole detection methods across four stages

Figure 3 presents a detailed comparative analysis of pothole detection techniques, illustrating the progression from raw input imagery to increasingly advanced segmentation results. The top row displays the original input images, each capturing real-world road scenes with visible potholes of varying shapes, depths, and surface textures. These images represent a diverse set of environmental conditions, including inconsistent lighting, shadow interference, worn asphalt, and occlusions - all of which pose significant challenges for automated detection systems. Serving as the visual ground truth, these inputs are essential for assessing the capabilities and limitations of each evaluated model.

The second row features the segmentation results produced by the method of Baroudi et al. [19], representing an earlier approach in pothole detection. While their technique is able to detect large, well-defined potholes, it suffers from critical weaknesses that limit its practical application. One key issue is inconsistent edge detection, where the boundaries of the potholes appear coarse, jagged, or incomplete. This imprecision often leads to over-segmentation in textured regions or under-segmentation in low-contrast scenarios. Moreover, the model exhibits high sensitivity to environmental noise, such as shadows, road markings, and surface cracks, often misclassifying them as potholes. This results in a significant number of false positives, thereby reducing the system's reliability. Additionally, the model tends to miss small or shallow potholes, especially when they are partially obscured by debris or situated in uneven lighting, making it insufficient for comprehensive and accurate road surface assessment.

In contrast, the third row shows the output of POT-YOLO [20], a more recent and sophisticated deep learning-based detection model. This method demonstrates notable improvements in object localization and significantly reduces the number of false positives seen in earlier approaches. It offers better general accuracy in detecting potholes of various sizes and positions within the image. However, it still suffers from blurred or irregular segmentation edges, particularly in areas with poor contrast or where potholes merge with the surrounding road surface. Although it demonstrates greater resilience than the methods proposed by Baroudi et al. [19] and POT-YOLO [19], its generalization capability remains limited. Performance significantly degrades under challenging conditions such as wet or reflective roads, heavy occlusions, or severely deteriorated surfaces. Additionally, the high computational cost of its complex network architecture hinders deployment on resource-constrained devices, posing challenges for real-time processing in embedded systems and mobile platforms.

The bottom row presents the segmentation results of the proposed model, which significantly outperforms the previous methods in both visual quality and quantitative accuracy. This model integrates advanced feature extraction, fuzzy logic-based refinement, and context-aware segmentation strategies to achieve precise edge delineation, even for irregularly shaped or partially obscured potholes. Its robustness to environmental variations - including shadows, uneven lighting, wet surfaces, and complex asphalt textures - enables it to deliver highly reliable results with a low false positive rate. Importantly, the proposed method exhibits high sensitivity, effectively detecting even small and shallow potholes that were previously missed. In addition to accuracy, the model is designed with computational efficiency in mind, incorporating optimization techniques that enable real-time performance without sacrificing detection quality. This balance between precision and speed makes it well-suited for deployment in autonomous road inspection systems, intelligent transportation infrastructure, and mobile-based road condition monitoring applications.

Overall, the figure underscores the evolution of pothole detection technologies, revealing how each successive model addresses prior shortcomings. The proposed model emerges as a superior solution, setting a new standard for automated pothole detection through its combination of edge accuracy, environmental robustness, and operational efficiency. These capabilities position it as a promising candidate for real-world implementation in modern intelligent transportation systems.

Table 3. Segmentation accuracy metrics for pothole detection (mean $\pm$ standard deviation)

Metric	Baroudi's Model	POT-YOLO	Proposed Model
Precision (%)	78.2 ± 1.4	85.6 ± 1.1	92.7 ± 0.8
Recall (%)	74.5 ± 1.6	83.1 ± 1.3	91.3 ± 0.9
F1-Score (%)	76.3 ± 1.5	84.3 ± 1.2	92.0 ± 0.8
Accuracy (%)	81.1 ± 1.2	87.9 ± 1.0	94.2 ± 0.7
IoU (%)	68.7 ± 1.8	75.8 ± 1.5	89.5 ± 1.0

The results presented in Table 3 offer a detailed comparison of segmentation accuracy metrics among three models: Baroudi's model, POT-YOLO, and the proposed model. Across all metrics - precision, recall, F1-score, accuracy, and IoU - the proposed model demonstrates superior performance.

Specifically, the proposed model achieves a precision of 92.7% $\pm$ 0.8%, outperforming Baroudi et al. (78.2% $\pm$ 1.4%) and POT-YOLO (85.6% $\pm$ 1.1%), indicating its higher capability to correctly identify true pothole pixels while minimizing false positives. Similarly, the recall of the proposed method is 91.3% $\pm$ 0.9%, which is notably higher than that of Baroudi et al. (74.5% $\pm$ 1.6%) and POT-YOLO (83.1% $\pm$ 1.3%), showing its effectiveness in detecting a larger proportion of actual potholes.

Furthermore, the F1-score, a harmonic mean of precision and recall, reaches 92.0% $\pm$ 0.8% for the proposed model, compared to 76.3% $\pm$ 1.5% and 84.3% $\pm$ 1.2% for Baroudi's model and POT-YOLO, respectively. This reflects a balanced improvement in both sensitivity and specificity. In terms of overall classification accuracy, the proposed method achieves 94.2% $\pm$ 0.7%, which is significantly better than the 81.1% $\pm$ 1.2% and 87.9% $\pm$ 1.0% recorded in Baroudi's model and POT-YOLO. Additionally, the Intersection over Union (IoU) metric, which evaluates the overlap between predicted and ground truth regions, is highest for the proposed model at 89.5% $\pm$ 1.0%, again surpassing the other models. These results collectively highlight the robustness and reliability of the proposed model in accurately segmenting potholes under various conditions, establishing it as a more effective and precise tool for automated road surface analysis.

Table 4 highlights the comparative performance of the three models - Baroudi's model, POT-YOLO, and the proposed model - based on efficiency and robustness metrics, including processing time, memory usage, PSNR, and SSIM. The proposed model exhibits significant improvements in computational efficiency, with an average processing time of just 1.64 seconds per image, compared to 6.72 seconds for Baroudi's model and 5.05 seconds for POT-YOLO. This reduction in runtime makes the proposed approach highly suitable for real-time deployment in embedded or

Table 4. Efficiency and robustness metrics for pothole detection

Metric	Baroudi's model	POT-YOLO	Proposed Model
Processing Time (sec/img)	6.72	5.05	1.64
Memory Usage (MB)	214	180	142
PSNR (dB)	24.8	28.2	33.5
SSIM	0.81	0.89	0.96

low-power systems. Additionally, the model demonstrates lower memory consumption at 142 MB, substantially outperforming Baroudi's model (214 MB) and POT-YOLO (180 MB), which is crucial for resource-constrained environments.

In terms of robustness to noise and image quality degradation, the proposed model achieves a PSNR of 33.5 dB, indicating better preservation of image fidelity during processing. This is a marked improvement over Baroudi's model (24.8 dB) and POT-YOLO (28.2 dB). The SSIM, which quantifies perceived image quality by evaluating luminance, contrast, and structure, further validates this improvement. The proposed model records an SSIM of 0.96, surpassing POT-YOLO (0.89) and Baroudi's model (0.81). These results demonstrate that the proposed method not only delivers precise segmentation but also maintains high image quality and efficient resource usage, making it a robust and practical solution for automated pothole detection in diverse real-world scenarios.

To ensure the practical applicability of the proposed pothole detection model, its generalization ability across diverse road environments is of critical importance. Real-world roads exhibit significant variability due to differing lighting conditions (e.g., daylight, nighttime, shadows), weather effects (e.g., rain, fog, snow), and surface types (e.g., asphalt, concrete, gravel). While the current evaluation demonstrates strong performance on the test dataset, further investigation is necessary to assess the model's robustness under such varying conditions. Preliminary results suggest that the model maintains stable detection accuracy under moderate changes in illumination and weather; however, extreme conditions such as heavy rain or low-light scenarios may require additional preprocessing or model adaptation. Future work will focus on expanding the training dataset to include a wider range of environmental conditions and road surfaces, thereby enhancing the model's stability and accuracy in real-world applications. This will ensure reliable pothole detection across diverse operational scenarios, ultimately contributing to safer and more efficient road maintenance.

5 Conclusion

This paper presented a novel pothole detection model that effectively addresses the key limitations of existing methods, including imprecise edge segmentation, sensitivity to noise and shadows, and poor detection of small or shallow potholes. By integrating advanced feature extraction, precise boundary refinement, and an optimized computational framework, the proposed model demonstrated superior performance across a wide range of evaluation metrics. Quantitative comparisons revealed that our model consistently outperforms benchmark methods in terms of precision, recall, F1-score, accuracy, and IoU. Moreover, the model exhibited strong robustness to real-world variations such as lighting changes, occlusions, and low-quality images, as confirmed by higher PSNR and SSIM scores. Its computational efficiency - evidenced by reduced processing time and lower memory usage - further highlights its suitability for real-time road condition monitoring applications.

Despite the demonstrated improvements in detection accuracy, robustness, and computational efficiency, the proposed model has several limitations. First, while contrast enhancement and adaptive convolution improve feature extraction, their effectiveness can diminish in extremely poor lighting conditions or highly cluttered road scenes where the contrast between potholes and the surrounding surface is minimal. Additionally, the self-generated fuzzy step, although beneficial in refining uncertain regions, may introduce ambiguity when multiple similar textures exist, potentially leading to false positives in areas with complex road textures or heavy shadow patterns.

To overcome the limitations of the proposed model, future work will focus on several key improvements. First, integrating multi-spectral or depth-sensing data will enhance feature extraction under low-contrast or cluttered conditions. Additionally, context-aware fuzzy logic will be introduced to reduce false positives in complex road textures and shadows. These improvements will enhance the model's reliability, efficiency, and applicability in real-world road monitoring systems.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] J. Wachnicka, "Mathematical models as practical tool for road fatalities forecasts on road network in regions," *MATEC Web Conf.*, vol. 231, p. 01018, 2018. https://doi.org/10.1051/matecconf/201823101018

[2] A. Sidliarenko, "Mathematical models of road construction, reconstruction and repair under conditions of uncertainty," *Bull. Cherkasy State Technol. Univ.*, no. 3, pp. 113–127, 2023. https://doi.org/10.24025/2306-4412.3.2023.287845

[3] I. Hussain and L. Alam, "Adaptive road crack detection and segmentation using Einstein operators and ANFIS for real-time applications," *J. Intell. Syst. Control*, vol. 3, no. 4, pp. 213–226, 2024. https://doi.org/10.56578/jisc030402

[4] O. Fokina, A. Mottaeva, and A. Mottaeva, "Transport infrastructure in the system of environmental projects for sustainable development of the region," *E3S Web Conf.*, vol. 515, p. 01015, 2024. https://doi.org/10.1051/e3sconf/202451501015

[5] A. Mustapha, A. Abdul-Rani, N. Saad, and M. Mustapha, "Advancements in traffic simulation for enhanced road safety: A review," *Simul. Model. Pract. Theory*, vol. 137, p. 103017, 2024. https://doi.org/10.1016/j.simpat.2024.103017

[6] M. Shi and I. Hussain, "Improved region-based active contour segmentation through divergence and convolution techniques," *AIMS Math.*, vol. 10, no. 1, pp. 654–671, 2025. https://doi.org/10.3934/math.2025029

[7] R. Archana and P. Jeevaraj, "Deep learning models for digital image processing: A review," *AI Rev.*, vol. 57, no. 1, p. 11, 2024. https://doi.org/10.1007/s10462-023-10631-z

[8] I. Hussain and J. Muhammad, "Efficient convex region-based segmentation for noising and inhomogeneous patterns," *Inverse Probl. Imaging*, vol. 17, no. 3, pp. 708–725, 2023. https://doi.org/10.3934/ipi.2022074

[9] K. Su, Q. Zhao, Y. Tomioka, and Y. Liu, "Cost-sensitive road obstacle detection for low-cost autonomous vehicles," in *2024 IEEE Intelligent Mobile Computing (MobileCloud)*, Shanghai, China, 2024, pp. 24–31. https://doi.org/10.1109/MobileCloud62079.2024.00011

[10] L. Juen, I. Khairuddin, A. Majeed, M. Abdullah, and A. Nasir, "Road signage and road obstacle detection using deep learning method," in *International Conference on Mechatronics and Intelligent Robotics*, Suzhou, China, 2023, pp. 15–25. https://doi.org/10.1007/978-981-99-8498-5_2

[11] D. Li and J. Zhang, "Intelligent road segmentation and obstacle detection for autonomous railway vehicle," *Adv. Mech. Eng.*, vol. 16, no. 1, p. 16878132231225312, 2024. https://doi.org/10.1177/16878132231225312

[12] K. Thakkar, S. Shah, B. Mulchandani, N. Katre, and H. Dalvi, "Automated pothole detection using transfer learning," in *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, Pune, India, 2024, pp. 1–8. https://doi.org/10.1109/I2CT61223.2024.10543553

[13] A. Karukayil, C. Quail, and F. Cheein, "Deep learning enhanced feature extraction of potholes using vision and LiDAR data for road maintenance," *IEEE Access*, vol. 12, pp. 184 541–184 549, 2024. https://doi.org/10.1109/ACCESS.2024.3512783

[14] K. Gorro, E. Ranolo, L. Roble, and R. Santillan, "Road pothole detection using YOLOv8 with image augmentation," *J. Image Graph.*, vol. 12, no. 4, pp. 417–426, 2024. https://doi.org/10.18178/joig.12.4.417-426

[15] A. Jenefa, A. Taurshia, B. Kuriakose, E. Kumar, and A. Lincy, "Advancing road maintenance with EfficientDet-based pothole monitoring," *Serb. J. Electr. Eng.*, vol. 22, no. 1, pp. 57–74, 2025. https://doi.org/10.2298/SJEE2501057J

[16] A. Patawar, M. Mehdi, B. Kore, and P. Saval, "A pothole can be seen with two eyes: An ensemble approach to pothole detection," *Mach. Vis. Appl.*, vol. 36, no. 3, p. 64, 2025. https://doi.org/10.1007/s00138-025-01679-8

[17] Y. Safyari, M. Mahdianpari, and H. Shiri, "A review of vision-based pothole detection methods using computer vision and machine learning," *Sensors*, vol. 24, no. 17, p. 5652, 2024. https://doi.org/10.3390/s24175652

[18] R. Swathika, N. Radha, D. Parthiban, and J. Sriram, "Pothole detection and road damage assessment system utilizing image processing techniques," in *2024 Second International Conference on Inventive Computing and Informatics (ICICI)*, Bangalore, India, 2024, pp. 269–278. https://doi.org/10.1109/ICICI62254.2024.00052

[19] U. Baroudi, A. BaHamid, Y. Elalfy, and Z. Alami, "Enhancing pothole detection and characterization: Integrated segmentation and depth estimation in road anomaly systems," *arXiv preprint arXiv:2504.13648*, 2025. https://doi.org/10.48550/arXiv.2504.13648

[20] N. Bhavana, M. Kodabagi, B. Kumar, P. Ajay, N. Muthukumaran, and A. Ahilan, "POT-YOLO: Real-time road potholes detection using edge segmentation based YOLOv8 network," *IEEE Sens. J.*, vol. 24, no. 15, pp. 24 802–24 809, 2024. https://doi.org/10.1109/JSEN.2024.3399008