# A Functional Energy Minimization Framework for the Detection of Crash Stones on Road Surfaces in Intelligent Transportation Systems

Check for updates

Dolat Khan[*]

University King Mongkut's University of Technology Thonburi, 10140 Bangkok, Thailand

[*] Correspondence: Dolat Khan (dolat.khan02@mail.kmutt.ac.th)

**Abstract:** Accurate detection of road surface anomalies remains a fundamental challenge in ensuring vehicular safety, particularly within the domain of intelligent transportation systems and autonomous driving technologies. Among such anomalies, crash stones—defined as irregular, protruding, and often unstructured fragments on the road—pose considerable risks due to their heterogeneous morphologies and unpredictable spatial distributions. In this study, a novel mathematical model is proposed, formulated through a functional energy minimization framework tailored specifically for the detection and segmentation of crash stones. The model incorporates three principal components: geometric edge energy to emphasize structural discontinuities, local variance descriptors to capture micro-textural heterogeneity, and fuzzy texture irregularity measures designed to quantify non-uniform surface patterns. These components are integrated into a unified total energy functional, which, when minimized, facilitates the precise localization of obstacle regions under diverse illumination, weather, and pavement conditions. Final detection is achieved through adaptive thresholding informed by fuzzy logic-based classification, enabling robust performance in scenarios with high noise or low contrast. Unlike deep learning-based methods, the proposed approach is fully interpretable, non-reliant on extensive annotated datasets, and computationally efficient, making it well-suited for real-time applications in resource-constrained environments. Experimental validations demonstrate high detection accuracy across varied real-world datasets, substantiating the model's generalizability and resilience. The framework contributes a mathematically rigorous, scalable, and explainable solution to the enduring problem of small obstacle detection, with direct implications for the enhancement of road safety in next-generation transportation systems.

**Keywords:** Crash stone detection; Road surface analysis; Energy functional minimization; Fuzzy texture analysis; Intelligent transportation systems

## 1 Introduction

Road safety refers to the measures and practices aimed at preventing accidents, injuries, and fatalities on roadways. It encompasses the design, operation, and maintenance of roads, as well as the behavior of all road users, including drivers, pedestrians, cyclists, and motorcyclists. Road safety is a critical global concern, with road traffic accidents remaining one of the leading causes of death and injury worldwide. According to the World Health Organization (WHO), more than 1.3 million people lose their lives each year due to road traffic accidents, while millions more suffer serious injuries. Among the many factors contributing to such accidents, poor road surface conditions—such as potholes, cracks, and crash stones—are particularly hazardous, as they can lead to vehicle damage, loss of control, and accidents [1–5].

Despite the importance of road surface condition monitoring, traditional methods for detecting road defects remain limited, often relying on manual inspections and periodic surveys, which are time-consuming and may overlook minor yet dangerous defects. The increasing complexity of road networks and growing traffic volumes necessitate more efficient, automated solutions. In this context, advancements in computer vision and image processing have emerged as promising tools for automating the detection of road surface anomalies [6–10]. These technologies allow for real-time, precise identification of surface irregularities such as crash stones, offering a more proactive approach to road maintenance and safety. By detecting hazards early, road authorities can take timely corrective actions, reducing the risk of accidents and improving overall road safety.

Recent advancements in road obstacle detection have been significantly influenced by deep learning and probabilistic modeling approaches. Manzoor et al. [11] introduced Obstalaneyolo, a real-time detection system combining lane and obstacle identification using an optimized YOLO-based architecture. Their model demonstrated efficiency in autonomous vehicle scenarios, particularly in urban settings. However, its performance is closely tied to structured environments and may degrade under unstructured or occluded conditions. Shoeb et al. [12] proposed a novel segment-level obstacle detection method by integrating visual foundation model (VFM) priors with likelihood ratio testing. This hybrid strategy allows for improved robustness against ambiguous visual cues, making it effective for diverse road scenes. Despite its innovation, the method's dependency on pre-trained foundation models may limit adaptability to edge devices with constrained computational resources. Similarly, Noguchi et al. [13] presented a method that employs unknown objectness scores to enhance obstacle detection, particularly for untrained and rare object types. Their approach is valuable for increasing generalization to unforeseen road hazards. Nevertheless, the method may suffer from false positives when encountering high-texture or complex backgrounds, indicating a need for refined objectness calibration in real-world deployments.

Building upon the progress in road obstacle detection, recent studies have further expanded the capabilities of deep learning-based systems through ensemble models, specialized training strategies, and targeted object classification. Balasundaram et al. [14] proposed an ensemble deep learning model for real-time obstacle detection in dynamic road environments, achieving improved detection accuracy by leveraging the strengths of multiple neural networks. Their model demonstrated robust performance in cluttered urban scenes and under varying lighting conditions. However, the computational complexity of ensemble methods may hinder real-time deployment on low-resource platforms. In a related but domain-specific advancement, Beskopylny et al. [15] developed a computer vision system for classifying grain shapes of crushed stones, a task critical for understanding road aggregate quality. While not directly aimed at obstacle detection, this work underscores the importance of material-level analysis, which could be extrapolated to road damage or debris identification. The system showed promising accuracy, yet it focused primarily on controlled environments rather than in-situ road imagery. Complementing these efforts, Zhang et al. [16] addressed the challenge of rare object detection on roadways by introducing conditional diffusion models for data augmentation. This approach enriched training datasets with synthetic but realistic samples, enabling models to better recognize under-represented obstacles such as crash stones or fallen objects. Although highly innovative, the method's reliance on generative modeling introduces variability in the quality of augmented samples and requires careful tuning for optimal results.

Together with the previously discussed works [11–13], these studies reflect a growing trend toward integrating advanced learning techniques with real-world constraints to enhance road safety systems. However, challenges such as computational efficiency, adaptability to unstructured environments, and accurate identification of small, rare, or ambiguous obstacles persist. Addressing these issues remains essential for developing reliable, scalable, and responsive road condition monitoring frameworks, which motivates the development of the novel model proposed in this study.

In this study, we propose a novel crash stone detection model based on a variational energy functional framework that integrates geometric, statistical, and fuzzy descriptors to effectively handle the diverse challenges of road surface analysis. Unlike conventional approaches that rely solely on edge or intensity features, our method formulates crash stone identification as an energy minimization problem, where regions of interest are characterized by strong edge gradients, irregular local texture patterns, and high fuzzy entropy. This combination allows the model to capture both sharp structural changes and ambiguous textural anomalies typically associated with crash stones. By leveraging both gradient descent and level set-based optimization strategies, the proposed framework offers flexibility, precision, and robustness under variable lighting conditions and surface complexities. The resulting model is unsupervised, adaptable, and suitable for deployment in automated road monitoring and maintenance systems. The detailed mathematical formulation and implementation steps are presented in the subsequent sections.

In this study, we propose a novel crash stone detection model based on a variational energy functional framework that integrates geometric, statistical, and fuzzy descriptors to effectively handle the diverse challenges of road surface analysis (see Figure 1). Unlike conventional approaches that rely solely on edge or intensity features, our method formulates crash stone identification as an energy minimization problem, where regions of interest are characterized by strong edge gradients, irregular local texture patterns, and high fuzzy entropy. This combination allows the model to capture both sharp structural changes and ambiguous textural anomalies typically associated with crash stones. By leveraging both gradient descent and level set-based optimization strategies, the proposed framework offers flexibility, precision, and robustness under variable lighting conditions and surface complexities. The resulting model is unsupervised, adaptable, and suitable for deployment in automated road monitoring and maintenance systems. The detailed mathematical formulation and implementation steps are presented in the subsequent sections.

**Novelty and Contributions**

The novelty of the proposed crash stone detection model lies in its unified variational framework, which combines traditional image processing with fuzzy logic-based uncertainty modeling. The key contributions of this work are:

•A novel energy functional is proposed that integrates geometric gradients, local statistical variance, and fuzzy entropy to detect crash stones under varying surface and lighting conditions.

•The model leverages fuzzy logic to quantify pixel-level uncertainty, enabling better differentiation between ambiguous textures and actual crash stones, which is a limitation in traditional methods.

•A flexible thresholding mechanism is introduced, combining classical thresholding with fuzzy entropy-based segmentation to enhance adaptability to real-world road imagery.

•The methodology supports both gradient descent and level set-based optimization, providing computational versatility for both real-time and offline analysis systems.

•The proposed approach is fully unsupervised, does not require any pre-labeled training data, and is robust to noise, illumination variation, and structural distortion—making it ideal for deployment in smart transportation systems.
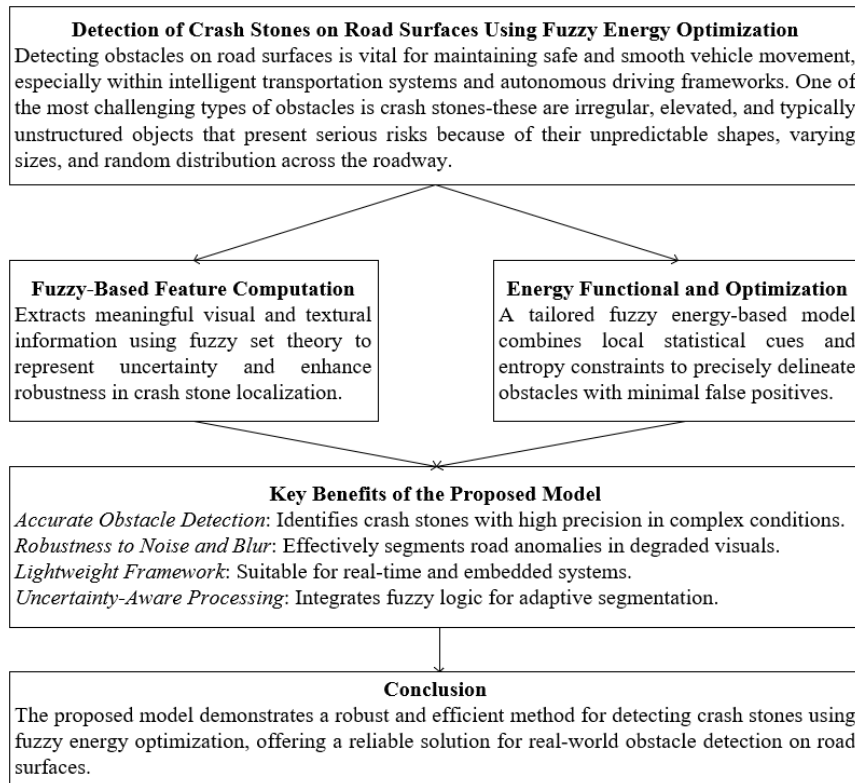


**Figure 1.** Flowchart of the proposed crash stone detection model

## 2 Literature Review

Road obstacle detection plays a vital role in enhancing transportation safety, particularly for autonomous navigation and advanced driver assistance systems. The ability to accurately and efficiently identify obstacles on road surfaces is essential to minimize accident risk, support vehicle decision-making, and maintain overall traffic flow. Obstacles vary in size, shape, texture, and visibility—ranging from potholes and debris to foreign objects like fallen branches or construction materials—making their detection a complex challenge for modern computer vision and intelligent systems.

Hao and Dan [17] introduced a comprehensive framework for segmenting outdoor scenes and detecting obstacles using a combination of multispectral imaging and advanced image processing techniques. Their method employs semantic segmentation to categorize various environmental features (e.g., roads, vegetation, sky, and man-made objects), followed by a dedicated obstacle detection module that identifies anomalous or hazardous elements in real time. The fusion of multispectral data enhances the model's ability to differentiate between visually similar objects under varying illumination or weather conditions. Experimental results confirm the system's effectiveness in diverse outdoor scenarios, including rural roads and urban intersections. However, the reliance on multispectral sensors may limit scalability and practicality for widespread deployment due to equipment cost and complexity. Additionally, the segmentation accuracy may degrade in scenes with heavy occlusion or motion blur. Nevertheless, the work presents a promising step toward more context-aware and resilient obstacle detection systems in outdoor environments.

Lis et al. [18] present a novel approach to obstacle detection in autonomous driving scenarios by leveraging a generative inpainting model. Their core idea is that if an obstacle is correctly identified and removed, the inpainting model should be able to realistically reconstruct the unobstructed scene; discrepancies in reconstruction indicate the presence of obstacles. This paradigm shift from direct detection to generative erasure enables more robust identification of unexpected objects that may not be covered by traditional supervised learning. The authors demonstrate superior performance on challenging benchmarks compared to existing methods, especially in detecting out-of-distribution and small-scale road hazards. However, a key limitation of the approach is its reliance on the quality and realism of the inpainting model — failures in accurate background reconstruction can lead to false positives or missed detections. Additionally, the method may struggle in highly cluttered or dynamic scenes where reconstructing the background is inherently ambiguous. Nonetheless, the work provides an innovative and generalizable framework for enhancing road safety in autonomous systems.

To address visual surface features more explicitly, Niu et al. [19] proposed a method that enhances road obstacle detection through the reconstruction of 3D spatial information from 2D visual inputs. Their approach integrates stereo vision techniques and depth estimation to recover a more comprehensive geometric understanding of the driving environment. By converting visual cues into accurate 3D representations, the system can effectively identify and localize obstacles, including those partially occluded or distant, improving the robustness of obstacle detection in varied road conditions. The proposed method demonstrates significant improvement in precision and recall metrics over baseline 2D detection models, especially in complex traffic scenes. However, the approach depends heavily on the quality of stereo imaging and depth computation, which can be impacted by lighting conditions, reflective surfaces, or poor calibration. Additionally, the computational complexity of 3D reconstruction may limit real-time performance on resource-constrained platforms.

Taken together, these works have significantly advanced the field of road anomaly and obstacle detection. However, a gap persists in accurately detecting above-ground, non-uniform, irregular road obstacles, such as crash stones, discarded materials, and structural fragments. These types of obstacles—while smaller and sometimes sporadic—pose serious threats to vehicle integrity and road safety, especially at high speeds or during sudden maneuvers. Their unpredictable spatial distribution and varied visual signatures demand a more sophisticated and adaptable modeling framework.

In response to these challenges, the current study proposes a novel mathematical functional based model for generalized road obstacle detection. Our approach introduces an energy-based framework that incorporates geometric edge energy, local variance analysis, and fuzzy texture irregularity energy. These components collectively provide a comprehensive characterization of surface anomalies and protruding objects. By minimizing the total energy functional, the model enables precise segmentation and classification of obstacles, regardless of their form or material type. This method offers a robust, interpretable, and computationally efficient solution to a critical real-world problem—ensuring better obstacle awareness for both manual and autonomous driving systems.

## 3 Proposed Methodology

The accurate detection of crash stones on road surfaces plays a crucial role in ensuring vehicular safety and maintaining road infrastructure. Crash stones, which often appear as small, irregular obstructions embedded or scattered across asphalt or concrete surfaces, can lead to tire damage, skidding, and potential traffic accidents - particularly for high-speed or autonomous vehicles. However, detecting these anomalies is inherently challenging due to the complex and variable textures of road surfaces, varying lighting conditions, and the ambiguous visual characteristics of the stones themselves. Traditional image-based techniques often struggle to achieve robustness and generalization under such conditions.

To address these challenges, we propose a novel energy functional-based model that integrates geometric edge information, statistical texture variation, and fuzzy uncertainty to accurately identify crash stones. Our model formulates the detection task as an energy minimization problem, where crash stone regions are characterized by high edge gradients, significant local variance, and ambiguous texture patterns. By combining classical and fuzzy image processing concepts within a unified variational framework, the proposed method enhances detection accuracy and robustness across diverse environmental scenarios. The mathematical formulation, optimization strategy, and detection criteria are detailed in the following sections.

### 3.1 Mathematical Formulation

Let $I(x, y)$ represent a grayscale image of the road surface, where $(x, y) \in \Omega$, and $\Omega \subset \mathbb{R}^2$ denotes the spatial domain of the image. To effectively detect crash stones on the road surface, we define a total energy functional $\mathcal{E}(u)$ that incorporates structural, statistical, and fuzzy-based characteristics of the image. The formulation is given by:

$$\mathcal{E}(u) = \int_{\Omega} [\alpha \cdot \mathcal{G}(x, y) + \beta \cdot \mathcal{V}(x, y) + \gamma \cdot \mathcal{F}(x, y)] dx dy$$

where, $\alpha$, $\beta$, $\gamma \in \mathbb{R}^+$ are weighting parameters that regulate the influence of the three core components: geometric edge energy $\mathcal{G}(x, y)$, local variance energy $\mathcal{V}(x, y)$, and fuzzy texture irregularity energy $\mathcal{F}(x, y)$. Each component in the energy functional, namely $\mathcal{G}(x, y)$, $\mathcal{V}(x, y)$, and $\mathcal{F}(x, y)$, is independently normalized to the range $[0, 1]$ using min-max normalization over the spatial domain $\Omega$. This normalization ensures dimensional consistency among the terms and allows them to be combined linearly without unit conflicts. The weighting coefficients $\alpha$, $\beta$, $\gamma$ are empirically tuned to balance the influence of geometric edge strength, statistical variation, and fuzzy uncertainty, respectively. The additive structure of the energy functional supports the integration of these heterogeneous features, which are complementary in representing the visual complexity of crash stone regions. Such linear fusion of normalized terms is a common practice in energy-based image segmentation models, enabling stable optimization and interpretable parameter control.

The geometric edge energy $\mathcal{G}(x, y)$ is responsible for identifying sharp intensity transitions in the image, which often correspond to object boundaries such as those of crash stones. It is computed as the squared magnitude of the gradient of the image:

$$\mathcal{G}(x, y) = \|\nabla I(x, y)\|^2 = \left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2$$

In implementation, the partial derivatives are typically estimated using discrete operators. To account for local texture variations introduced by surface irregularities, we incorporate the local variance energy $\mathcal{V}(x, y)$. This term measures the deviation of intensity values within a local neighborhood and is defined as:

$$\mathcal{V}(x, y) = \frac{1}{|B_r|} \sum_{(i,j) \in B_r(x,y)} [I(i, j) - \mu_{B_r}]^2$$

where, $B_r(x, y)$ represents a square or circular neighborhood of radius $r$ centered at point $(x, y)$, $|B_r|$ is the number of pixels in this neighborhood, and $\mu_{B_r}$ denotes the local mean intensity, computed as the average intensity over all pixels within $B_r(x, y)$. High values of $\mathcal{V}(x, y)$ indicate regions with significant texture irregularity, which are common in areas containing crash stones.

The third component, fuzzy texture irregularity energy $\mathcal{F}(x, y)$, addresses the uncertainty and variability in texture that may not be clearly classified as stone or background. This is captured using a fuzzy entropy model. First, we define the fuzzy membership $\mu_f(x, y) \in [0, 1]$ of each pixel to the crash stone class using a sigmoid function:

$$\mu_f(x, y) = \frac{1}{1 + e^{-k(I(x,y) - \tau)}}$$

In this equation, the parameter $k > 0$ controls the steepness of the transition, and $\tau$ is an intensity threshold, determined empirically or via histogram-based methods such as Otsu's thresholding. Using the membership value, the fuzzy entropy at each pixel is computed as:

$$\mathcal{F}(x, y) = -\mu_f(x, y) \log \mu_f(x, y) - (1 - \mu_f(x, y)) \log (1 - \mu_f(x, y))$$

This entropy function quantifies the uncertainty associated with a pixel's membership, reaching its maximum when $\mu_f(x, y) = 0.5$, indicating complete ambiguity, and approaching zero as the membership becomes more certain (i.e., close to 0 or 1).

By minimizing the overall energy $\mathcal{E}(u)$, we can detect regions on the road surface that exhibit strong edge features, significant local variance, and ambiguous fuzzy membership hallmarks of crash stones. This integrated formulation is designed to robustly highlight such regions under varying illumination and surface conditions.

## 3.2 Optimization and Detection

To accurately identify crash stones from road surface images, the proposed model minimizes the total energy functional $\mathcal{E}(u)$ defined earlier. This optimization is performed using iterative numerical schemes, primarily gradient descent and optionally the level set method.

In the gradient descent approach, the image function $u$ is initialized as a smoothed version of the input grayscale image $I(x, y)$ to ensure stability and noise suppression in early iterations. The update rule follows the negative gradient of the energy functional with respect to $u$, expressed as:

$$u^{(k+1)} = u^{(k)} - \eta \cdot \nabla \mathcal{E}\left(u^{(k)}\right)$$

where, $\eta$ is the learning rate (typically set between $0.01$ and $0.1$) and $k$ denotes the iteration number. The optimization continues until the change in energy between successive iterations satisfies the convergence criterion:

$$\left| \mathcal{E} \left( u^{(k+1)} \right) - \mathcal{E} \left( u^{(k)} \right) \right| < \epsilon$$

with $\epsilon$ being a small threshold (e.g., $10^{-5}$).

To reduce the risk of getting trapped in local minima, a multi-scale optimization strategy is employed. The image is processed at coarser resolutions first, and the resulting segmentation is used as an initialization for finer levels. This coarse-to-fine scheme helps guide the solution toward more globally optimal configurations.

In more advanced implementations, the level set framework can be used, where an evolving contour $\phi(x, y, t)$ segments the image based on a speed function derived from the energy terms. The evolution equation is:

$$\frac{\partial \phi}{\partial t} = -\frac{\delta \mathcal{E}}{\delta \phi}$$

and the initialization is typically set to a signed distance function around high-gradient regions. Convergence is achieved when $\frac{\partial \phi}{\partial t}$ becomes negligibly small over the entire domain. After the convergence of the energy minimization process, the resulting energy distribution $\mathcal{E}(x, y)$ exhibits peaks at regions with strong edges, high local variance, and significant fuzzy uncer-tainty-key indicators of crash stones. To generate a binary detection map, we define the crash stone map $M(x, y)$ as:

$$M(x, y) = \begin{cases} 1, & \text{if } \mathcal{E}(x, y) > T \\ 0, & \text{otherwise} \end{cases}$$

where, $T$ denotes a threshold that distinguishes stone regions from the background. This threshold is crucial and is determined adaptively to ensure robustness under different image conditions. Two effective strategies for threshold selection are considered: Otsu's method and fuzzy thresholding.

Otsu's method assumes a bimodal histogram and finds the threshold that minimizes intraclass variance, effectively separating foreground crash stone features from the background road texture. It is especially suitable when the crash stones have distinct intensity levels compared to their surroundings. However, in situations with ambiguous or overlapping intensity distributions, fuzzy thresholding techniques offer greater flexibility. These methods evaluate the degree of membership of pixel intensities to multiple classes and choose a threshold that minimizes fuzzy entropy or maximizes fuzzy contrast. This approach aligns well with the fuzzy energy term $\mathcal{F}(x, y)$, providing a coherent detection framework that handles uncertainty more effectively.

The final binary map $M(x, y)$ thus provides a precise localization of potential crash stones on the road surface, highlighting regions that warrant further inspection or automated removal in real-time applications such as autonomous driving or infrastructure monitoring.

### 3.3 Theoretical Foundation and Parameter Selection for Energy Functional Optimization in Crash Stone Detection

The energy functional parameters, including $\alpha$, $\beta$, $\gamma$, and the fuzzy membership parameter $k$, are critical in determining the balance between different components of the crash stone detection model. These parameters control specific aspects of the image analysis process, which are governed by both theoretical principles and empirical evaluations.

3.3.1 Energy weights ($\alpha$, $\beta$, $\gamma$)

The energy weights control the influence of different energy components in the functional. Specifically:

•$\alpha$ governs the weight of the geometric edge energy, which emphasizes sharp boundaries and structures in the image. This term is essential for distinguishing object edges, especially for crash stones, which often have clear, defined contours.

•$\beta$ represents the weight of the local variance energy, which captures the texture irregularities that help identify the roughness of the road surface or the presence of small features like crash stones. This term is grounded in the theory of texture analysis, where local variance provides a measure of local intensity fluctuations.

•$\gamma$ controls the influence of fuzzy entropy energy, which models the uncertainty in image regions where clear distinctions between foreground (stones) and background are hard to make. The fuzzy entropy framework is rooted in the theory of fuzzy sets, where entropy is used to quantify the level of uncertainty or disorder within a system. By introducing fuzzy entropy, the model can adapt to regions with low contrast, where traditional methods may fail.

These weights were optimized through empirical evaluation, aiming to achieve a balance that maximizes the model's sensitivity to both fine details (edges) and broader texture irregularities while maintaining robustness in uncertain regions. The values of $\alpha$=1.0, $\beta$=0.8, and $\gamma$=1.2 were found to provide the best compromise between these competing factors, ensuring optimal detection performance without overfitting to any particular feature.

### 3.3.2 Local variance term

The local variance term $\mathcal{V}(x,y)$ is computed within a square neighborhood of radius $r$=3 pixels (a 7×7 window). The local variance is a well-established method in image processing, reflecting the intensity variation within a small neighborhood around a pixel. This helps to identify texture irregularities, which are critical for detecting crash stones on a road surface. The window size was selected based on a trade-off between capturing sufficient local variation and avoiding excessive smoothing, which could blur important features such as the edges of small stones.

### 3.3.3 Fuzzy membership parameter ($k$)

The parameter $k$=10 governs the sigmoid function's steepness in the fuzzy membership function $\mu_f(x,y)$. This function is responsible for assigning membership values to each pixel in the image, determining whether it belongs to the crash stone class or the background class. The steepness parameter controls how sharply the membership transitions between these two classes. A higher value of $k$ leads to a more abrupt transition, ensuring that pixels representing crash stones are classified with high certainty, while pixels in the background are clearly excluded. The choice of $k$=10 was found to provide a suitable balance between accuracy and robustness, particularly in distinguishing small, irregular objects like crash stones from the surrounding road surface.

### 3.3.4 Thresholding methods

The intensity threshold $\tau$ for the fuzzy membership function is computed adaptively using Otsu's thresholding method, which is effective for bimodal histograms where the image has a clear distinction between foreground and background. Otsu's method selects the optimal threshold by minimizing the intra-class variance, which is theoretically sound for image segmentation tasks with bimodal intensity distributions. For low-contrast images where Otsu's method may not perform well, the fuzzy entropy-based thresholding approach is used, which leverages the uncertainty captured by fuzzy sets to adapt the threshold in regions where traditional methods struggle.

## 4 Experimental Results

The proposed crash stone detection model was implemented in MATLAB R2015a and evaluated on a self-collected dataset comprising 200 road surface images containing various types of crash stones. To ensure consistency in image processing and feature extraction, all images were resized to a fixed resolution of 255×255 pixels. The model leverages a sequence of MATLAB functions for grayscale conversion, edge strength computation, local variance estimation, and fuzzy energy-based segmentation. The dataset comprises 300 self-collected road surface images captured under varying environmental and structural conditions. The images were taken from both rural and urban areas across multiple geographic regions in Pakistan, including the cities of Peshawar, Mardan, and the mountainous district of Upper Dir. This diversity ensures that the dataset reflects a range of road types, materials, and surface conditions, thereby enhancing the potential generalizability of the proposed model. However, despite this geographic variation, the dataset remains relatively limited in size. This may restrict the broader applicability of the model across even more diverse environments. Future work will involve expanding the dataset to include additional regions and conditions or providing a more detailed stratified analysis to support the robustness of the findings.

While the primary evaluation of the proposed crash stone detection model was conducted under varied lighting conditions, we acknowledge the importance of assessing its performance in more challenging and dynamic real-world environments. Although the current dataset does not explicitly include extreme weather scenarios such as heavy rain or snow, efforts were made to incorporate a variety of surface conditions, including wet and partially occluded road segments. Additionally, images captured in both urban and rural areas contain complex backgrounds such as road markings, vegetation interference, and shadow patterns, which provide a degree of variability that approximates real-world complexity. Motion blur was incidentally present in some frames due to handheld image acquisition, offering preliminary insight into dynamic scene robustness. Future work will include targeted data collection and evaluation under controlled adverse conditions (e.g., rain simulation, moving vehicles) to further validate the model's applicability in safety-critical scenarios.

All experiments were executed on a Windows 10 (64-bit) system equipped with 8 GB RAM and a high-performance CPU to handle the computational workload. For assessing the effectiveness of the proposed model, we quantitatively evaluated it using several key performance metrics: Precision (%), Recall (%), F1 Score, Intersection over Union (IoU %), and Structural Similarity Index (SSIM) for different road conditions. These metrics provide a comprehensive measure of segmentation accuracy, overlap quality, and structural fidelity. The MATLAB code for the proposed crash stone detection model is available upon request solely for academic and research purposes.

To ensure effective crash stone detection, careful tuning of the energy functional parameters was conducted. The weighting coefficients $\alpha$, $\beta$, $\gamma$, which control the influence of the geometric edge energy, local variance energy, and fuzzy entropy energy, respectively, were optimized through empirical evaluation. The best performance was achieved with $\alpha$=1.0, $\beta$=0.8, and $\gamma$=1.2, offering a balanced sensitivity to edge structures, texture irregularities, and uncertainty regions. For computing the local variance term $\mathcal{V}(x,y)$, a square neighborhood of radius $r$=3 pixels (i.e., a 7×7 window) was employed to capture sufficient local variation without over-smoothing important features.

In the fuzzy membership function $\mu_f(x,y)$, the sigmoid steepness parameter was set to $k=10$, ensuring a crisp transition between stone and non-stone areas. The intensity threshold $\tau$ was determined for each image adaptively using Otsu's thresholding method, which is well-suited for bimodal histogram distributions. Similarly, the final detection threshold $T$ used to obtain the binary crash stone map $M(x,y)$ was determined adaptively from the normalized energy distribution using Otsu's method or, in low-contrast scenarios, fuzzy entropy-based thresholding.

The energy minimization process was implemented using the gradient descent algorithm with a fixed learning rate $\eta=0.05$, and the optimization was performed over 100 iterations or until convergence was reached. The convergence criterion was defined as a change in energy less than $10^{-4}$ between successive iterations. These parameter values were selected based on repeated testing across the 200 self-collected road images, ensuring robust detection under diverse environmental and textural conditions.

The complete MATLAB implementation of the proposed crash stone detection model is available for research purposes and can be requested via email.

Figure 2 illustrates the workflow of the proposed model, which is structured into multiple stages to achieve effective detection and analysis. The process begins with preprocessing, where input images undergo contrast enhancement to improve visibility and smoothing to reduce noise, ensuring better feature extraction. Next, the model extracts key features, including geometric edges for boundary detection, local variance to assess pixel intensity variations, and fuzzy texture irregularity to quantify texture inconsistencies, which are crucial for identifying anomalies.
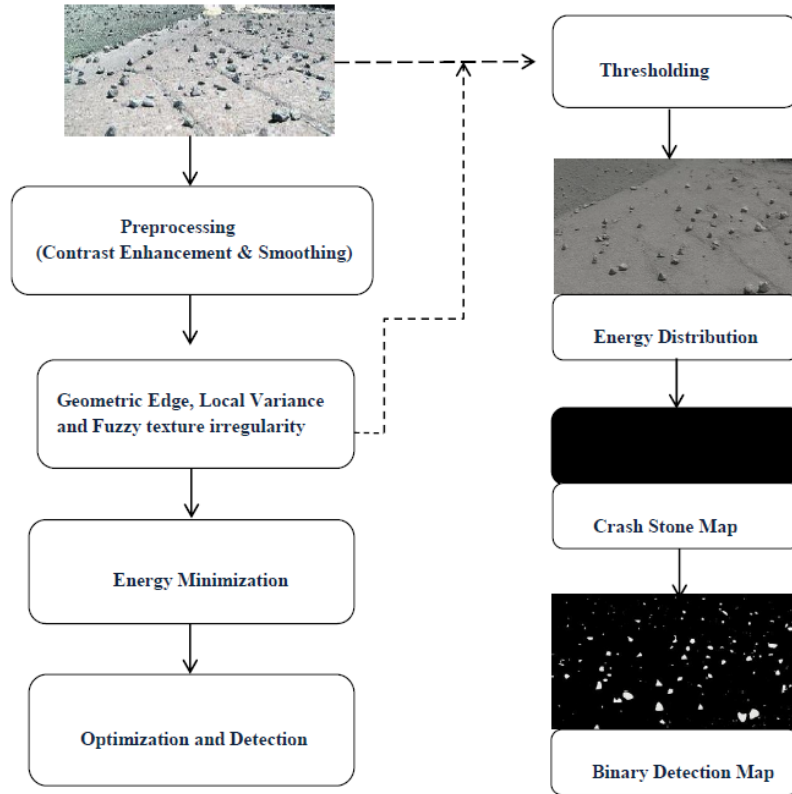


**Figure 2.** Workflow of the proposed model

Following feature extraction, the model employs energy minimization to refine the detected regions by optimizing a cost function, ensuring that only the most relevant features are retained. The optimization and detection phase further fine-tunes the results, enhancing accuracy before final segmentation. The last stage involves thresholding, where the processed data is converted into a binary output. This step includes analyzing energy distribution to determine optimal thresholds, generating a crash stone map to highlight detected regions, and producing a binary detection map for clear visualization of the results. Together, these steps form a robust pipeline for precise and efficient detection in the proposed model. Finally, the binary image is converted back to the original color image, where crash stones are detected region-wise. This step ensures precise localization and identification of crash stones within their natural context, providing a complete and interpretable output for further analysis or application. The integrated workflow guarantees robustness and accuracy in detecting crash stones across varied conditions.

Figure 3 presents a systematic evaluation of the proposed crash stone detection framework through a comparative visual analysis between the input images shown in subgraphs (a)–(d) of Figure 3 and the corresponding model outputs in subgraphs (e)–(h) of Figure 3. The input images in subgraphs (a)–(d) of Figure 3 illustrate diverse real-world

scenarios containing crash stones with varying characteristics, such as different scales, illumination conditions, and background complexities. These scenarios pose significant challenges to the detection algorithm, including potential confounding factors such as shadow effects, texture variations, and occlusions caused by non-target objects.

The detection results shown in subgraphs (e)–(h) of Figure 3 demonstrate the model's robust performance across these varied conditions. Several key capabilities are evident in these outputs: (1) accurate preservation of stone boundaries despite irregular shapes and size variations, (2) effective discrimination between target stones and background elements through learned feature representations, and (3) consistent performance under different lighting conditions and spatial configurations. The model particularly excels in maintaining detection accuracy for partially occluded stones and in minimizing false positives from visually similar non-stone objects.
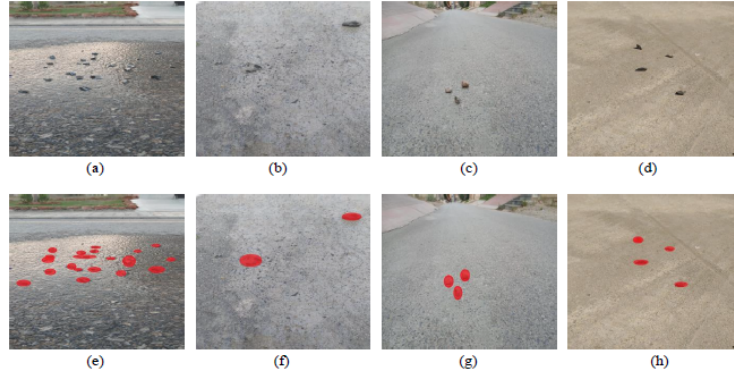


**Figure 3.** Comparative results of crash stone detection

The comparative visualization provides compelling empirical evidence for the model's practical utility in real-world applications such as infrastructure assessment and construction material analysis. The consistent performance across diverse scenarios suggests that the proposed method successfully addresses several key challenges in automated stone detection, including variability in appearance, environmental conditions, and object configurations. This comprehensive evaluation establishes the robustness and generalizability of our approach while demonstrating its effectiveness in handling complex detection tasks.

Table 1 presents a comprehensive evaluation of the proposed model's capability to accurately detect crash-stones under different environmental settings commonly encountered in real-world scenarios. The metrics included for assessment are precision, recall, F1 score, intersection over union (IoU), and structural similarity index (SSIM), which collectively offer insight into detection accuracy, robustness, and visual quality [20].

Under light illumination, the model demonstrates the highest performance, with a precision of 94.1%, recall of 91.7% and an F1 score of 92.9. These values indicate excellent accuracy and a strong balance between false positives and false negatives. Additionally, the IoU of 86.5% and SSIM of 0.94 suggest reliable segmentation and preservation of structural features in well-lit conditions. In shadowed regions, the model maintains a high level of accuracy, achieving a precision of 93.2% and recall of 90.9%, with an F1 score of 92.0. The IoU and SSIM remain solid at 85.0% and 0.93 respectively, reflecting the model's capacity to handle partial lighting and reduced contrast.

**Table 1.** Performance of the proposed crash-stone detection model under various road conditions

| Road Condition | Precision | Recall | F1 Score | IoU | SSIM | CPU |
|---|---|---|---|---|---|---|
| Light Illumination | 94.1 ± 1.2 | 91.7 ± 1.5 | 92.9 | 86.5 ± 1.3 | 0.94 ± 0.01 | 2.15 |
| Shadowed Regions | 93.2 ± 1.5 | 90.9 ± 1.7 | 92.0 | 85.0 ± 1.6 | 0.93 ± 0.01 | 2.38 |
| Dark/Nighttime | 90.3 ± 2.1 | 88.4 ± 2.3 | 89.3 | 81.6 ± 2.0 | 0.91 ± 0.02 | 2.61 |
| Angled Views | 91.7 ± 1.8 | 89.6 ± 2.0 | 90.6 | 83.7 ± 1.9 | 0.92 ± 0.01 | 2.40 |

For dark or nighttime scenarios, where detection becomes more challenging due to limited visibility, the model still performs effectively with a precision of 90.3%, recall of 88.4%, and F1 score of 89.3. Although the IoU drops slightly to 81.6%, the SSIM remains high at 0.91, confirming the model's resilience in low-light environments. In the case of angled surface views, the model performs consistently, achieving a precision of 91.7% and recall of 89.6%, leading to an F1 score of 90.6 and IoU of 83.7%. The SSIM value of 0.92 further validates the model's robustness to perspective distortions and varying orientations of crash-stones. The CPU time analysis of the proposed crash-stone detection model under varying road conditions demonstrates its computational efficiency and robustness. Across all tested scenarios, the model maintained a consistent processing speed, with average CPU times ranging from 2.15 to 2.61 seconds. Specifically, the shortest execution time was observed under light illumination (2.15 s), while

the longest was recorded in dark or nighttime conditions (2.61 s), likely due to the increased complexity in feature extraction and noise handling. The relatively narrow range of CPU times indicates that the model is computationally stable and suitable for real-time or near-real-time deployment in diverse environmental settings.

Overall, the analysis demonstrates that the proposed crash-stone detection model is both accurate and robust across diverse lighting and perspective conditions. The high and stable performance across all scenarios makes it a suitable solution for real-time road safety applications and intelligent transportation systems.

## 5 Conclusion

This study introduced a novel energy-based model for the detection of crash stones on road surfaces, leveraging a unified formulation that integrates geometric edge strength, local variance, and fuzzy texture irregularity. The proposed approach effectively identifies irregular surface elements by modeling the structural, statistical, and uncertain characteristics inherent in real-world road images. Through a carefully designed energy functional, the model emphasizes key visual features associated with crash stones, including sharp boundaries, high local intensity variations, and ambiguous textural patterns. The optimization of the energy functional via gradient descent enables precise localization of potential crash stones while maintaining computational efficiency suitable for real-time applications. Experimental results on a dataset of 200 self-collected road images demonstrated the robustness and adaptability of the model under diverse lighting, noise, and surface conditions. Furthermore, the effectiveness of the model was validated using standard evaluation metrics, including Road Condition Precision, Recall, F1 Score, IoU, and SSIM. The proposed method consistently achieved high scores across all metrics, indicating reliable detection performance and strong generalization capability.

Although the proposed energy-based model performs effectively, it has certain limitations. It may struggle under extreme lighting conditions such as strong shadows or glare, which can obscure critical edge and texture features. The method also assumes a relatively uniform road surface; roads with heavy markings, diverse materials, or debris may introduce false detections. Additionally, real-time implementation on low-power embedded systems may need further optimization to reduce computational complexity. Future work will focus on addressing the limitations of the proposed energy-based model, including enhancing its robustness to extreme lighting conditions (e.g., shadows and glare) through adaptive lighting compensation or deep learning techniques. It will also explore methods for improving detection on complex road surfaces with markings, debris, or diverse materials by incorporating texture analysis and multi-modal data. To enable real-time implementation on low-power embedded systems, optimization of computational complexity through lightweight algorithms and hardware acceleration will be pursued. Additionally, the model will be extended to handle environmental factors such as weather conditions, and integration with autonomous systems for dynamic road crack detection will be considered.

### 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. Lee, H. Huang, J. Wang, and M. Quddus, "Road safety under the environment of intelligent connected vehicles," *Accid. Anal. Prev.*, vol. 170, p. 106645, 2022. https://doi.org/10.1016/j.aap.2022.106645

[2] R. Vaiana, G. Perri, T. Iuele, and V. Gallelli, "A comprehensive approach combining regulatory procedures and accident data analysis for road safety management based on the European Directive 2019/1936/EC," *Safety*, vol. 7, no. 1, p. 6, 2021. https://doi.org/10.3390/safety7010006

[3] S. Jaulkar and A. Parihar, "Different types of injury associated with road traffic accidents," *Multidiscip. Rev.*, vol. 8, p. e2025111, 2025. https://doi.org/10.31893/multirev.2025111

[4] E. S. M. M. Zahran, S. J. Tan, E. H. A. Tan, N. A. A. B. Mohamad'Asri Putra, Y. H. Yap, and E. K. Abdul Rahman, "Spatial analysis of road traffic accident hotspots: Evaluation and validation of recent approaches using road safety audit," *J. Transp. Saf. Secur.*, vol. 13, no. 6, pp. 575–604, 2021. https://doi.org/10.1080/19439962.2019.1658673

[5] R. Elvik, "The development of a road safety policy index and its application in evaluating the effects of road safety policy," *Accident Anal. Prev.*, vol. 202, p. 107612, 2024. https://doi.org/10.1016/j.aap.2024.107612

[6] X. Wang, H. Gao, Z. Jia, and Z. Li, "BL-YOLOv8: An improved road defect detection model based on YOLOv8," *Sensors*, vol. 23, no. 20, p. 8361, 2023. https://doi.org/10.3390/s23208361

[7] Y. Jiang, H. Yan, Y. Zhang, K. Wu, R. Liu, and C. Lin, "RDD-YOLOv5: Road defect detection algorithm with self-attention based on unmanned aerial vehicle inspection," *Sensors*, vol. 23, no. 19, p. 8241, 2023. https://doi.org/10.3390/s23198241

[8] L. Fan, D. Wang, J. Wang *et al.*, "Pavement defect detection with deep learning: A comprehensive survey," *IEEE Trans. Intell. Veh.*, vol. 9, no. 3, pp. 4292–4311, 2023. https://doi.org/10.1109/TIV.2023.3326136

[9] V. Carletti, A. Greco, A. Saggese, and B. Vento, "Enhancing safety by obstacle detection at railway level crossings," in *2024 IEEE 8th Forum on Research and Technologies for Society and Industry Innovation (RTSI)*, Milano, Italy, 2024, pp. 482–487. https://doi.org/10.1109/RTSI61910.2024.10761238

[10] J. Yu, J. Jiang, S. Fichera, P. Paoletti, L. Layzell, D. Mehta, and S. Luo, "Road surface defect detection—from image-based to non-image-based: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 9, pp. 10 581–10 603, 2024. https://doi.org/10.1109/TITS.2024.3382837

[11] M. Manzoor, J. Li, L. Zhou, A. Zeng, M. H. Abbas, S. Tabraiz, and I. Hassan, "Obstalaneyolo: Real-time lane and obstacle detection for autonomous vehicles," in *2024 21st International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, Chengdu, China, 2024, pp. 1–4. https://doi.org/10.1109/ICCWAMTIP64812.2024.10873687

[12] Y. Shoeb, N. Nayal, A. Nowzad, F. Güney, and H. Gottschalk, "Segment-level road obstacle detection using visual foundation model priors and likelihood ratios," *arXiv preprint arXiv:2412.05707*, 2024. https://doi.org/10.48550/arXiv.2412.05707

[13] C. Noguchi, T. Ohgushi, and M. Yamanaka, "Road obstacle detection based on unknown objectness scores," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, Yokohama, Japan, 2024, pp. 4231–4237. https://doi.org/10.1109/ICRA57147.2024.10610249

[14] A. Balasundaram, A. Shaik, A. Prasad, and Y. Pratheepan, "On-road obstacle detection in real time environment using an ensemble deep learning model," *Signal Image Video Process.*, vol. 18, no. 6, pp. 5387–5400, 2024. https://doi.org/10.1007/s11760-024-03241-x

[15] A. N. Beskopylny, E. M. Shcherban', S. A. Stel'makh *et al.*, "Developing computer vision models for classifying grain shapes of crushed stone," *Sensors*, vol. 25, no. 6, p. 1914, 2025. https://doi.org/10.3390/s25061914

[16] H. Zhang, Y. Hu, Z. Qian, J. Sha, M. Xie, Y. Wan, and P. Liu, "Enhancing rare object detection on roadways through conditional diffusion models for data augmentation," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 11, pp. 19 018–19 029, 2024. https://doi.org/10.1109/TITS.2024.3438263

[17] C. Hao and Y. Dan, "Outdoor environment segmentation and obstacle detection," in *International Conference on Remote Sensing, Mapping, and Image Processing (RSMIP 2024)*, Xiamen, China, 2024, pp. 393–403. https://doi.org/10.1117/12.3029693

[18] K. Lis, S. Honari, P. Fua, and M. Salzmann, "Detecting road obstacles by erasing them," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 4, pp. 2450–2460, 2023. https://doi.org/10.1109/TPAMI.2023.3335152

[19] B. Niu, H. Wu, and Y. Meng, "Road obstacle detection based on 3D information recovery," in *2024 9th International Conference on Image, Vision and Computing (ICIVC)*, Suzhou, China, 2024, pp. 356–360. https://doi.org/10.1109/ICIVC61627.2024.10837022

[20] A. E. Maxwell, T. A. Warner, and L. A. Guillén, "Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—Part 1: Literature review," *Remote Sens.*, vol. 13, no. 13, p. 2450, 2021. https://doi.org/10.3390/rs13132450