

Research Article

Haiping Ding* and Qianlong Tang

Optimization of the road bump and pothole detection technology using convolutional neural network

<https://doi.org/10.1515/jisys-2024-0164>

received March 15, 2024; accepted July 9, 2024

Abstract: In highway bridge engineering, it is essential to use modern image processing methods to effectively detect and classify road bumps and potholes, considering the unique characteristics of bridge surfaces. Accurate identification and handling of road surface irregularities are crucial for preserving the longevity and security of transportation infrastructure. The study proposes a Convolutional Neural Network for Highway Anomaly Detection (BCNN-HAD) that uses the “BridgeGuard-Vision” (BGV) method, a computer vision technology for highway bridges to increase the accuracy and efficiency of automated image processing for more accurate detection of road irregularities in highway surveillance. Through model training using datasets that imitate various environmental conditions frequently seen on highway bridges, the proposed approach obtains dependable characteristics from photos taken close to bridges, improving flexibility and accuracy. The training method considers variations in lighting, weather conditions, and bridge materials, ensuring the model performs well in various real-world situations. In addition, this work explores the combination of sensor fusion techniques, combining data from many sources such as bridge structural health monitoring systems, cameras, accelerometers, and Global Positioning System. This comprehensive method, represented by BGV-YOLOv5, aims to offer a complete understanding of the bridge surroundings, therefore helping to detect road irregularities and further developing the field of bridge health monitoring. Expected results involve developing a personalized and effective system for detecting road bumps and potholes to tackle specific difficulties in highway bridge situations. In addition, the project seeks to provide a structure for smart transportation systems in the field of bridge engineering. This project seeks to address the particular requirements of highway bridge engineers, improving road safety and infrastructure maintenance methods in highway bridge engineering with the overall objective of establishing a safer and longer-lasting transportation network.

Keywords: highway bridge engineering, BridgeGuard-Vision, BCNN-HAD, advanced image processing, sensor fusion, smart transportation systems

1 Introduction

Highway bridge engineering is a field that combines civil engineering with transportation infrastructure. It involves the planning, building, and upkeep of bridges that enable vehicles to travel safely and efficiently over highways and other transportation systems. Highway bridge engineering is concerned with bridges' design,

* **Corresponding author: Haiping Ding**, School of Road and Bridge Engineering, Jiangxi Vocational and Technical College of Communications, Nanchang 330013, Jiangxi, China, e-mail: blueprussian@163.com

Qianlong Tang: Experimental Training Center, Jiangxi Vocational and Technical College of Communications, Nanchang 330013, Jiangxi, China, e-mail: 174801039@csu.edu.cn

construction, and upkeep, which is essential for effective mobility. Identifying bumps and potholes is vital because uneven road surfaces can affect the strength of the road and create safety hazards. These problems speed up deterioration, undermine the vehicle's comfort, and put road safety at risk. Maintaining a bridge's best condition is very important, and automated image processing and computer vision are crucial in effectively recognizing and fixing these abnormalities. Using modern technology, highway bridge engineers improve structures' durability, reduce dangers related to vehicle travel, and maximize traffic efficiency, guaranteeing critical infrastructure's continued usefulness and safety [1]. New approaches have recently surfaced to overcome the limitations of traditional vibration-based methodologies for bridge structural status monitoring. These methods involve using new sensors or novel signal properties. Contemporary real-time Structural Health Monitoring systems can produce a significant amount of data that needs to be analyzed and assessed to identify early irregularities and issue timely notifications and alerts regarding the state of civil infrastructure [2]. Displacement measurements can offer helpful information on the state of structures and how they behave when subjected to operational and environmental loads. Computer vision systems were confirmed as a method of measuring displacement; the study conducted here aims to provide the foundation of a damage detection system that operates in real time [3].

The project aims to find anomalies in the expressway's computerized surveillance system. To do this, edge computing and deep learning will be used to improve the real-time effectiveness of the management and monitoring of an expressway. The video data recorded by the camera system in the smart surveillance network of the roadway is sent to the edge analyzing server for analysis and then passed on to a convolutional neural network (CNN) [4]. Multiple object detection techniques using deep learning have been developed, utilizing CNN to obtain characteristics. This study proposes using YOLO, or You Only Look Once, for a system that detects potholes. Several variations of the YOLO algorithm are trained using a customized dataset that includes water-logged and dry potholes of different shapes and sizes [5]. To showcase the practicality and improved performance of the new picture recognition models, the machine learning object identification software models are evaluated using mAP (mean average precision) and processing speed, commonly used to quantify efficacy [6]. YOLO series detectors are widely regarded as the optimal balance of precision and efficiency, making them highly suitable for real-time tasks. The YOLOX technique is the most recent in the series and is developed by using and enhancing the YOLOv3 algorithm [7]. YOLO, a method for detecting objects in real-time, uses CNNs to identify and categorize photograph potholes. This deep-learning study intends to address the issue of potholes encountered during the journeys of autonomous or self-driving cars. This method instantly detects potholes using highlighted visual indicators [8]. The algorithm's utilization of CNN allows for the simultaneous prediction of object categories and bounding boxes, improving responsiveness and accuracy. Some of the most often used CV algorithms involve CNNs, commonly employed for video and image recognition, and object recognition algorithms, capable of identifying and detecting objects in images and videos [9]. Other computer vision techniques include feature identification algorithms that recognize particular visual characteristics in an image and classification algorithms that divide an image into different areas or items.

In the picture identifying anomalies method shown in Figure 1, primary phases establish a complete object-detection network. First, the input data is prepared through initial preprocessing. Then, convolutional layers to extract important hierarchical characteristics for identifying anomalies. Feature mapping includes (i) pooling layers to help reduce the input size, making computations more efficient, and (ii) interconnected layers that analyze the retrieved features, forming a basis for classification. Post-processing methods improve output predictions, guaranteeing precise anomaly identification. This enhances the network's capability to identify and classify irregularities, making it a crucial framework for reliable visual anomaly detection in many scenarios.

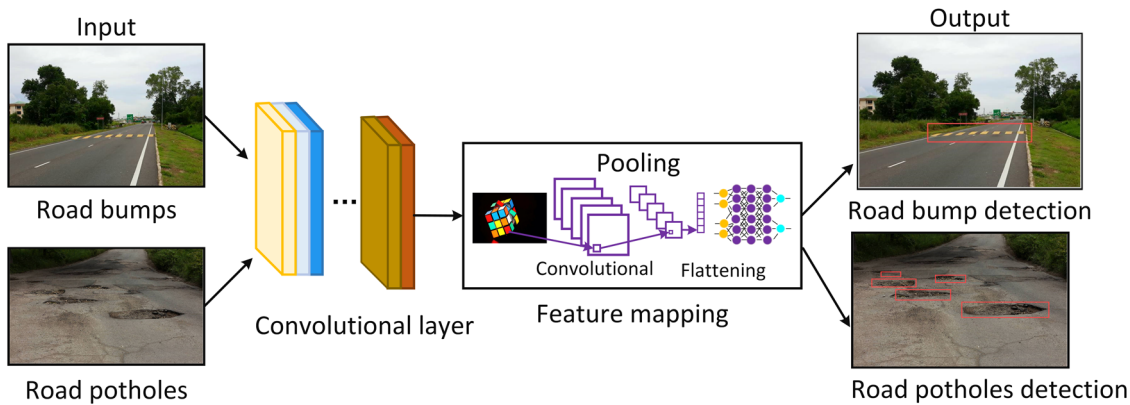


Figure 1: Image anomaly detection process.

The primary contribution of the work is,

- To introduce a BCNN-HAD model that extracts dependable features from photos taken near the bridge, advanced image processing techniques are used to identify and categorize road bumps and potholes, considering the distinct characteristics of bridge surfaces.
- To develop BGV-YOLOv5, a sophisticated computer vision technique designed explicitly for highway bridges to enhance automated image processing for improved identification of road anomalies.
- To improve flexibility and accuracy, the system is trained using datasets that simulate the many environmental conditions commonly found on highway bridges.
- To explore the combination of sensor fusion techniques, data from many sources such as bridge structural health monitoring systems, cameras, accelerometers, and Global Positioning System (GPS) must be combined.
- To provide a structure for intelligent transportation systems in bridge engineering, creating a safer and more durable transportation network.

The study's primary goal in highway bridge engineering is to improve road safety and infrastructure maintenance. The main objective is to develop the BCNN-HAD technique, which uses BridgeGuard-Vision (BGV) and sensor fusion, to effectively identify road abnormalities such as bumps and potholes. The goal is to create a complete system for intelligent transportation, ensuring a safer and more long-lasting transportation network. The research enhances intelligent transportation systems in bridge engineering by introducing the BGV-YOLOv5 technique. This method integrates sensors, sophisticated picture analysis, and CNNs to thoroughly observe the state of bridges. The research contributes to the construction of safer and more lasting transportation networks in the field of bridge engineering by offering a customized system for detecting road anomalies and assuring flexibility in various climatic circumstances.

The study will follow with the sections including Section 2 as a literature review, proposed automated image processing systems procedures are described in Section 3, Section 4 covers the experimental analysis and outcomes, and the study concludes the work in Section 5 covering with a conclusion and plans for future research.

2 Literature review

Peralta-López et al. [10] proposed a deep neural network to automatically analyze color images obtained by a ZED camera to detect road potholes. An efficient design was created to accelerate the training and usage process. Bharat et al. [11] proposed significant technical advancements in road pothole detection. A CNN structure was created manually, starting from scratch, focusing on tackling the difficulties related to pothole categorization and detection. Wu et al. [12] developed a theoretical structure for an automated system that

detects potholes using cell phone vibration data. Along with performing necessary actions (including resampling, redirection, filtering, etc.) upon the collected data, the program uses a sliding window to divide the continuous data stream into separate segments.

Lyu et al. [13] suggested bidirectional short-term memory (Bi-LSTM) to identify bump features. The randomly positioned smartphone during vehicle movement recorded the 3D speed increase, 3D positioning, and GPS data collection. Alhussan et al. [14] introduced a new approach called adaptable mutation and dipper-throated optimizing for selecting features and optimizing the random forest, or RF, classifier. Once again, a thorough statistical analysis is performed on the gathered data to assess the significance and reliability of the proposed method. The findings highlighted the efficiency and advantages of the proposed method.

Ozoglu and Gökgöz [15] described creating a system for identifying potholes. The technology utilizes sensors for vibration and the GPS built into cellphones, eliminating the need for extra devices in automobiles and the associated expenses. In detecting road anomalies based on vibrations, a new method that uses CNNs is a significant advancement in this area. Bosurgi et al. [16] proposed 3D photographs of the pavement surfaces obtained using advanced equipment. Based on computer vision protocols, the method can accurately detect potholes, quantify their size, and assess their depth. Borgalli [17] proposed the systems used to identify various sorts of potholes by employing techniques that rely on data collected from an ultrasonic sensor, gyroscope, and Pi camera. This data provides information about the severity and shape of the pothole under any circumstances. Users can upload images of potholes, and once approved by the admin, they will be processed by a deep-learning framework.

Shaghouri et al. [18] proposed, utilized, and evaluated various deep-learning structures to identify the existence of potholes. To begin, a cell phone attached to the automobile windshield captures multiple photographs of potholes.

The study suggests tackling issues about uneven road surfaces on highway bridges using the BCNN-HAD approach. This entails utilizing BGV technology and sensor fusion techniques to enhance the precision of anomaly detection. The approach involves teaching the model using different datasets to mimic other environmental conditions commonly found on highway bridges, guaranteeing flexibility and accuracy. The inclusive method efficiently identifies and classifies road irregularities and holes, improving road safety and infrastructure upkeep in highway bridge engineering.

3 Proposed methodology

The primary objective of this work is to enhance the domain of computerized methods for image processing. The project intends to create an advanced system that can precisely detect and categorize road surface abnormalities using state-of-the-art computer vision techniques. The emphasis on highway bridges acknowledges these constructions' distinct difficulties, including differing lighting conditions, weather, and various bridge materials.

The research improves the flexibility and accuracy of the BCNN-HAD model in different environmental situations by training it with datasets that simulate common variations on highway bridges. This encompasses elements including illumination, atmospheric conditions, and the composition of the bridge. The suggested approach guarantees that the BCNN-HAD model can successfully function in multiple environmental conditions, enhancing its resilience and dependability in real-life scenarios by subjecting the model to various settings.

3.1 BGV-based convolutional neural network-highway anomaly detection (BCNN-HAD)

The proposed BCNN-HAD method takes as input images captured nearby highway bridges. The system uses these images to recognize and extract features associated with road anomalies like bumps and potholes. The BCNN-HAD method, based on BGV-YOLOv5, enhances automated image processing and improves the

identification of road abnormalities. The system's CNN is designed to extract significant characteristics from these images, helping to detect and categorize road surface irregularities accurately. Although input handling is more involved than other architectures, the goal is to preserve the rapid processing rate of the CNN network by reducing the overall number of network layers. Unlike many studies, the suggested approach effectively identifies speed bumps and potholes on road surfaces in highway bridges, as shown in the flowchart of Figure 2.

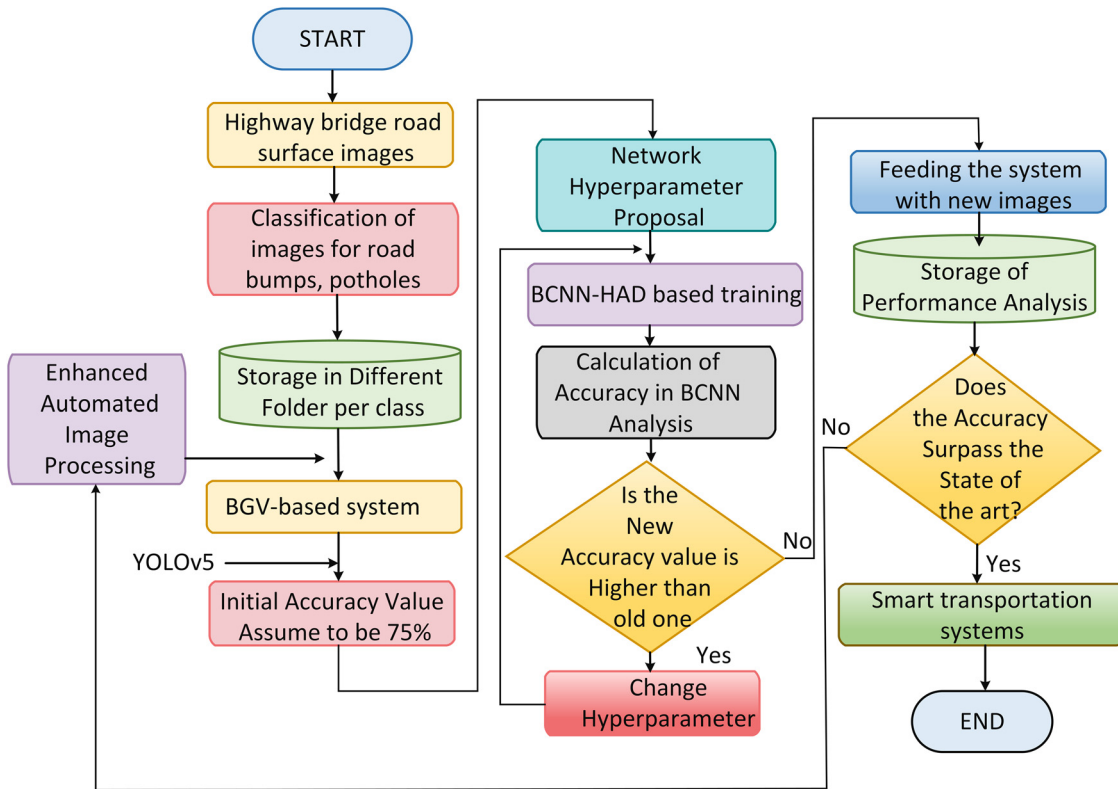


Figure 2: Overall flow of proposed BCNN-HAD work.

3.1.1 Material and selection

BGV-YOLOv5 (BridgeGuard Vision–You Only Look Once) could be a suitable foundation for implementing BCNN-HAD in highway bridge engineering. BGV-YOLOv5 is well-known for its capability to rapidly identify items in live situations, making it a convenient option for promptly identifying irregularities on the road, such as bumps and potholes. The architecture is created to be efficient, allowing for fast image processing, which is essential for applications that require quick reactions in real-time anomaly detection. BGV-YOLOv5 guarantees precise object detection, which is necessary for accurately classifying and recognizing road abnormalities.

The BGV-YOLOv5 technique, which processes the entire image at once, helps to speed up processing and achieve real-time performance, making it suitable for highway bridge engineering. In addition, the framework is supported by a solid and engaged community, which provides plenty of help, detailed documentation, and a wide range of tools for developers. Although BGV-YOLOv5 is a notable choice, the final decision relies on the project's unique needs, the dataset's peculiarities, and the development team's experience with the selected framework.

The BCNN-HAD technique utilizes BGV, which combines sophisticated image processing and sensor fusion. It is proficient at identifying road irregularities close to highway bridges by training on varied datasets, guaranteeing flexibility in different environmental situations. The approach attempts to improve road safety

and infrastructure upkeep, contributing to a more efficient transportation network. The suggested BCNN-HAD approach helps enhance traffic safety in highway bridge engineering by utilizing modern image processing techniques. It precisely identifies road irregularities like bumps and potholes close to bridges, making it easier to schedule maintenance promptly. This improves the general state of the road, reduces safety risks, and guarantees a safer and more secure transportation system.

3.2 BGV-YOLOv5-based system

3.2.1 Data collection

BGV-YOLOV5 starts by gathering input data, which includes high-quality photos taken close to highway bridges. These pictures concentrate on the road's outside and function as the unprocessed information for examination. The bridge's structure may be monitored for vibrations and oscillations using vibration sensors. When vibration patterns are out of the ordinary, it could be a sign of structural problems like fractures or weaknesses.

Under load, strain gauges determine how much strain is applied to different bridge parts. This is useful for determining whether the material is being stressed beyond its intended capacity.

The dynamic reaction of the bridge to traffic, wind, and other environmental conditions may be better understood using accelerometers, which measure acceleration and movement.

The precise location of bridge components may be followed by GPS, which can detect even the smallest movements or tilts. Finding possible structural changes or problems with the foundation requires this. It is feasible to accurately monitor deformation over time by putting GPS sensors at different sites. Continuous GPS data may track any displacement or movement in real time, allowing for prompt reactions to any problems. A bridge health monitoring system that integrates GPS and sensor data provides a comprehensive, up-to-the-minute picture of the bridge's structural condition and the effects of the surrounding environment. This comprehensive method improves predictive maintenance, guarantees safety, and increases the bridge's lifetime by facilitating prompt interventions.

3.2.2 Pre-processing

The input photos go through preprocessing stages to improve quality and standardize characteristics. This might involve changing the size, standardizing, and modifying lighting conditions.

3.2.3 Features extraction with CNN

The essential aspect of BGV-YOLOv5 focuses on utilizing CNN. CNN captures significant attributes from the processed images, highlighting traits important for identifying road irregularities like bumps and potholes. Figure 3 shows the proposed workflow of the system.

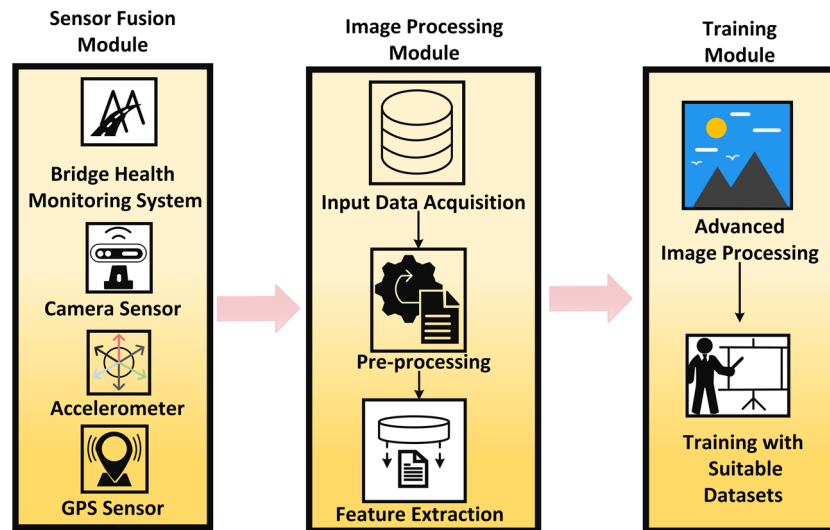


Figure 3: BridgeGuardVision system.

3.2.4 Advanced image processing

BGV-YOLOv5 uses sophisticated image processing methods after extracting features. These methods are designed to identify and categorize abnormalities on road surfaces, taking into account the unique characteristics of bridge surfaces.

3.2.5 Training using varied datasets

To improve flexibility and precision, the system is trained using datasets that replicate a range of environmental variables commonly found on highway bridges. This includes differences in illumination, weather conditions, and bridge materials.

3.2.6 Sensor fusion module

BGV-YOLOv5 investigates methods of sensor fusion, which involve merging data from many sources. This might include data from bridge structural health monitoring systems, cameras, accelerometers, and GPS devices. Incorporating many data sources helps achieve a more complete picture of the bridge environment.

3.2.7 Comprehension understanding and abnormality detection

The final step of the workflow is getting a thorough awareness of the bridge's surroundings using BGV-YOLOv5. This knowledge helps efficiently identify road abnormalities, contributing to the broader area of bridge condition monitoring.

The study highlights the importance of YOLO series detectors, specifically YOLOX, in identifying road irregularities by balancing accuracy and efficiency. YOLOX, being the most recent in the series, improves on YOLOv3, using CNNs for real-time detection. This enhances the ability to quickly and accurately deal with road problems. BGV-YOLOv5 has an essential function in the BCNN-HAD system as it is the fundamental technology

for detecting objects in real time. It is suitable because it can quickly and accurately detect road irregularities close to highway bridges, including bumps and potholes. The design guarantees effective image processing, making it suitable for applications that require fast reactions in highway bridge engineering. Furthermore, the assistance from an engaged community, comprehensive documentation, and various tools make BGV-YOLOv5 a dependable option for the suggested system.

3.3 Conventional neural network architecture

Initially, the camera captures a color image with dimensions of $a \times b \times c$ (where $a = 672$ columns and $b = 376$ rows, along with $c = 3$ indicating the RGB color index). Every pixel in an image contains an 8-bit resolution, and the color illustration adheres to the RGB paradigm.

$$I(u, v, w) \in 0 \leq W \leq 255 \vee \{u \mid 1 \leq u \leq a\}, \{v \mid 1 \leq v \leq b\}, \{w \mid 1 \leq w \leq c\}. \quad (1)$$

In equation (1) mentioned before, I denote a matrix with three dimensions. The initial layer, the “ImageInputLayer,” uses an information n normalization process. This requires subtracting the average image (mean) of the simulated set, in which the highest value of n is 714, from every image that is input, as specified in equations (2) and (3). The entered images are colored, with an elevation of 376 pixels, an expanse of 672 pixels, and three color channels saved in the RGB color space. These photos are sorted and stored within a tensor with a dimension of (376, 672, 3).

$$\text{mean}M = \frac{M_1 + M_2 + \dots + M_n}{n}, \quad (2)$$

$$\hat{M}_n = M - \text{means}. \quad (3)$$

The following layer, known as the “Convolution2Dlayer,” has convolutional processing, including 37 filters with a specific width and height of 3×3 ($f_1, f_2, \dots, f_x = 37$). The layer has a stride of 1 and does not include pooling. However, it incorporates padding of the “same” kind to capture image features and avoid sub-sampling surrounding information. The layer’s data are stored in a tensor with dimensions (376, 672, 3, 37).

$$\text{CONV}_x(u, v, 1) = \sum_{du=-g}^g \sum_{dv=-h}^h f_x(du, dv, 1) \hat{M}_n(u - dv, v - dw, 1), \quad (4)$$

$$\text{CONV}_x(u, v, 2) = \sum_{du=-g}^g \sum_{dv=-h}^h f_x(du, dv, 2) \hat{M}_n(u - dv, v - dw, 2), \quad (5)$$

$$\text{CONV}_x(u, v, 3) = \sum_{du=-g}^g \sum_{dv=-h}^h f_x(du, dv, 3) \hat{M}_n(u - dv, v - dw, 3). \quad (6)$$

From equations (4)–(6), f_x is the filtering window with dimensions $(2g + 1) \times (2h + 1)$ and dx and dy are the indicators indicating move across the window on both the u and v axes. Figure 4 shows the all-functional layers of the CNN architecture and their functions.

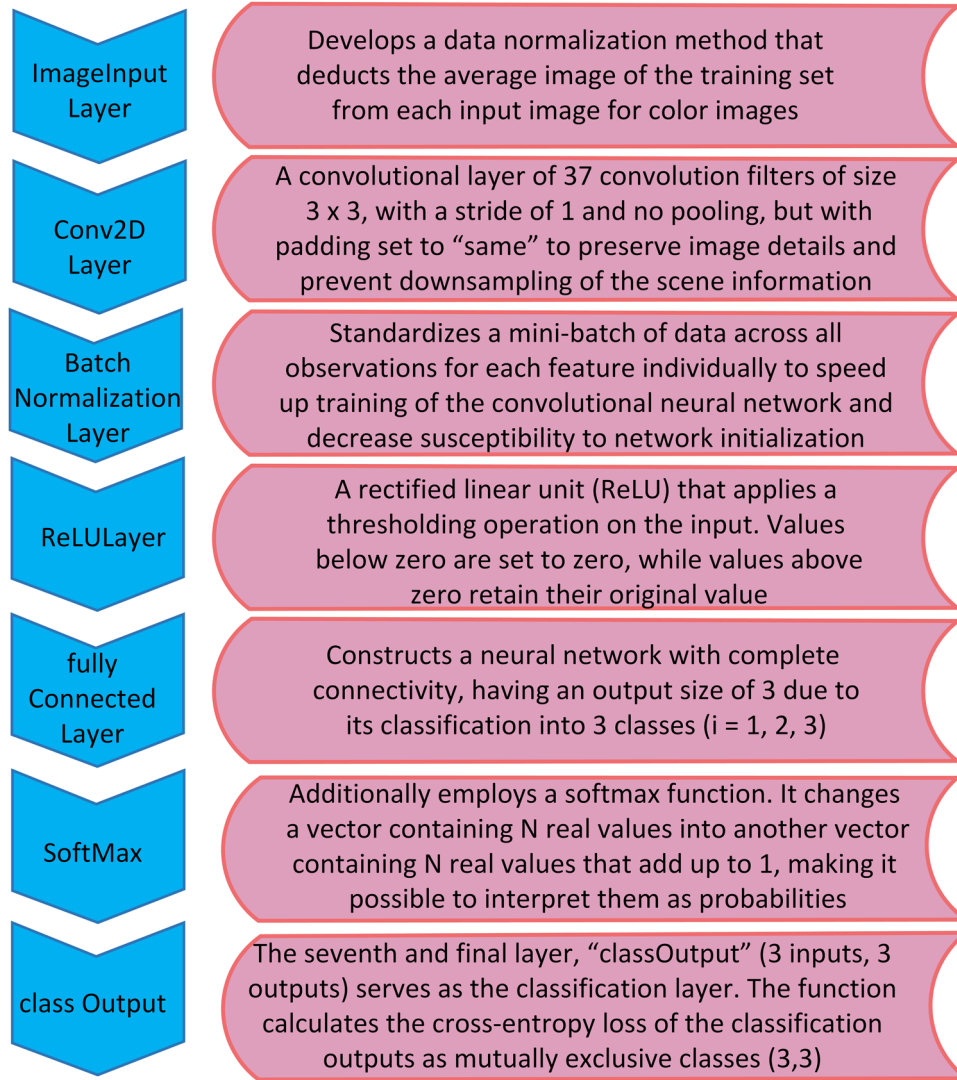


Figure 4: Various layers of CNN architecture and their functions.

The third layer, "batchNormalizationLayer," normalizes a group of data inside a small batch, treating each feature individually. This helps speed up the development of the CNN and makes it less sensitive to how the network is set up. Multiple normalization layers should be incorporated between convolutional regions and anomalies for best results. Here, we employ a ReLU layer after a convolutional layer stored in a (376, 672, 3, 37) matrix.

$$\hat{C}_x(u, v, w) = \frac{\text{CONV}_x(u, v, w) - \mu_x}{\sigma_x}, \quad (7)$$

$$\hat{CN}_x(u, v, w) = \alpha_x \hat{C}_x(u, v, w) + \beta_x. \quad (8)$$

In equations (7) and (8), where μ_x represents the average, α_x represents the amplification factor and β_x represents the offset for each feature.

The fourth layer, known as the "ReluLayer," consists of a linear unit that has been rectified (ReLU) and applies the thresholding work to the input. Numbers below zero are set to zero, while numbers above zero stay the same. When evaluating ReLU to the tanh and the sigmoid functions, ReLU does not have saturation zones. These saturation areas indicate that the outcomes of the neurons tend to remain relatively stable, resulting in a limited range of deviations from the initial conditioning gradient and little changes to network parameters.

This results in a need for more development in the instruction process. Equation (9) demonstrates that the ReLU function is significant in contemporary neural networks, which are stored in a tensor of dimensions (376, 672, 3, 37).

$$\text{CapReLU}(u, v, w) = \max(0, \hat{\text{CN}}_x(u, v, w)). \quad (9)$$

The fifth layer, referred to as “fullyConnectedLayer,” defined by equation (10), creates a neural network that is fully connected with a result size of 3. This is because it is separated into three categories ($i = 1, 2, 3$). The output value generated by the ReLU layer is converted into a series of integers using the `vec()` method. This implies that a matrix is transformed into a vector by merging its components. This procedure is often referred to as leveling. This level resizes the inputs by a set of weights K_i and then supplies a bias to the vector h_i . The outcome is subsequently sent through a feedback loop of transit work. This indicates that our completely linked neural network comprises 28,046,592 components and yields three outcomes (28, 046, 592).

$$\text{FCL}_i = \text{vec}(\text{ReLU}(u, v, w) \times K_i + h_i). \quad (10)$$

The sixth level (three inputs and three results), “Soft Max,” likewise employs a SoftMax function, which is commonly utilized as the last level of classifiers utilizing artificial neural networks, as shown in equation (11). It changes a vector of L exact numbers into an additional vector of L real values that add up to 1, allowing them to be understood as probabilities. For this situation, the value of L is 3.

$$\text{Softmax}(\text{FCL})_i = \frac{e^{\text{FCL}_i}}{\sum_{j=1}^L e^{\text{FCL}_j}}. \quad (11)$$

The last layer, called “classOutput” with three inputs and three outcomes, functions as a categorization layer. The formula (12) computes cross-entropy loss for segmenting the outputs into separate categories (3,3).

$$\text{loss} = -\frac{1}{X} \sum_{n=1}^X \sum_{i=1}^Y q_i t_{ni} \ln(p_{ni}), \quad (12)$$

where X represents the sample count, Y represents the class numbers, q_i represents the weight of class I , and represents the signal that the n th study is part of the i th class and one represents the output for sample n for class I , which is the value from the softmax function.

The suggested BCNN-HAD technique using BGV-YOLOv5, including the “BGV” approach, is its capability to offer specialized computer vision technology designed explicitly for highway bridges. BGV improves image processing by extracting reliable features from images taken near bridges, raising road anomaly identification accuracy. This integration guarantees that the model can easily adjust and accurately perform under different environmental conditions typically seen in highway bridge situations. Based on BGV-YOLOv5, the suggested solution utilizes a complete process for identifying highway anomalies. The procedure starts by gathering and preparing photos taken near highway bridges. Then, a CNN extracts features. Sophisticated picture processing and training with different datasets guarantee flexibility under other environmental circumstances. The system includes a sensor fusion module for a comprehensive understanding. The use of BGV-YOLOv5 in anomaly detection helps improve bridge status monitoring. The suggested technique standardizes image input by deducting the mean image from the dataset during preprocessing. This guarantees consistent characteristics and enhanced quality. The architecture’s Rectified Linear Unit (ReLU) layer is essential, as it introduces non-linearity by converting negative values to zero. This aids in learning features, avoiding saturation areas, and improving the efficiency of the CNN by enabling a more comprehensive range of training gradients.

3.4 Network hyperparameter proposal

In CNNs, hyperparameters are important in the design and include aspects like the input picture dimensions, the number of convolutional methods used, the size of convolutional filter systems, and the count of outputs of the neural network. These conditions are set up before the training process. Conversely, network variables in

the suggested system have values dynamically chosen by the optimization technique, as indicated by (BGV-YOLOv5). Efficient training of CNNs depends greatly on computer technology, especially when using GPUs for iterative procedures. Increasing the network's size alone might not result in improved performance or feature learning, even though optimizing the network architecture, adjusting hyperparameters, and starting with the right database are crucial. When categories are easily identifiable, unnecessarily complicated designs may not be needed. The learning of filters for various categories can be enhanced using a larger input size during training despite the potential for higher computing costs. Given that prediction is the network's primary objective, it is important to note that utilizing high-quality hardware is a prerequisite for achieving optimal performance.

The BCNN-HAD approach uses a neural network structure with important layers, including Convolution2D, batchNormalization, and fully connected. The Convolution2D layer uses 37 filters with a 3×3 window, without pooling but with padding. The batchNormalization layer standardizes data within tiny batches, improving CNN development. The fully connected layer constructs a neural network that is fully linked and produces three outcomes using a rectified linear unit (ReLU) activation function. Combining these layers helps enhance the anomaly detection capabilities of the BCNN-HAD approach. Network parameters are essential for CNN performance in the suggested approach. These settings, such as the input image's dimensions, filters' sizes, and the number of outcomes, are established before training. Appropriately selected hyperparameters are crucial for optimal feature learning and computational efficiency during training. The thoughtful use of hyperparameters contributes to the overall success and effectiveness of the suggested CNN-based anomaly detection method.

4 Experiment and results

4.1 Dataset

A total of 174 images of yellow roadway cones and 35-speed bumps that were not seen were gathered and stored in folder 0. Folder 1 had 231 pictures of potholes, while folder 2 had 304 pictures of smooth streets. The pictures were taken with a resolution of 376×672 pixels. The sun's position was a significant consideration in the selection of the photos. For instance, it was sometimes hard to get excellent pictures in the bright afternoon light at about six o'clock. As a result, several photos and images that didn't have any bumps or imperfections were not included to ensure a fair representation in all categories. The collection includes three categories of images: ones that show speed bumps, potholes, and ones that don't have either. In addition, the photos were captured from the outside of a vehicle. The BCNN-HAD uses a simplified seven-layer structure, less resource-intensive on the computer's hardware than Inception V2, which has 12 layers. Results using the test dataset of speed bump images and the trained CNN model [19].

4.2 Performance metrics

4.2.1 Accuracy

In highway bridge technology, precision is crucial for the proposed BCNN-HAD, based on BGV-YOLOv5 CNN, for detecting highway anomalies. Accuracy assesses the general accuracy of the model's predictions, which guarantees dependable detection of road surface irregularities, such as bumps and potholes. A high level of accuracy reflects the model's capacity to correctly categorize anomalies and non-anomalies that contribute to the longevity and safety of the transportation system.

4.2.2 Precision

Precision is essential in the BCNN-HAD system as it evaluates the correctness of optimistic predictions, especially in identifying road anomalies. Precision refers to the algorithm's accuracy in accurately detecting and classifying events like road irregularities, such as bumps and potholes. A high precision score demonstrates the system's efficiency in reducing incorrect identifications, ensuring that detected irregularities are indeed on the road.

Considering the terms, TP is the True Positives (accurately detected abnormalities); TN is the True Negatives (accurately detected non-anomalies); FP is the False Positives (mistakenly detected abnormalities); FN is the False Negatives (erroneously classified non-anomalies), and the equations for the performance metrics given by,

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}, \quad (13)$$

$$Precision = \frac{TP}{TP + FP}. \quad (14)$$

4.2.3 Recall

Recall, sometimes referred to as sensitivity, is crucial for BCNN-HAD to assess its capability to detect all occurrences of road anomalies. An elevated recall score suggests that the algorithm accurately identifies a significant portion of actual road surface abnormalities. This measurement is essential for reducing incorrect results and ensuring the system recognizes occurrences of bumps and potholes, thereby improving safety and infrastructure upkeep.

4.2.4 F1-score

The *F1-Score*, a mathematical mean of recall and accuracy, offers a fair evaluation of BCNN-HAD's performance. This measurement considers both incorrect positive and negative results, thoroughly assessing the model's precision in detecting road irregularities. A high *F1-Score* implies a balanced trade-off between precision and recall, displaying the model's effectiveness in precise anomaly identification while minimizing errors. To summarise, accuracy ensures overall correctness, precision highlights the accuracy of optimistic predictions, recall analyses the ability to identify anomalies thoroughly, and the *F1-Score* evaluates BCNN-HAD's performance. Together, these measurements help achieve the system's objective of establishing a more secure and long-lasting transportation network in highway bridge engineering.

$$Recall = \frac{TP}{TP + FN}, \quad (15)$$

$$F1 = \frac{2 \cdot \text{precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (16)$$

Equations (13)–(16) offer a numerical representation of the primary assessment criteria. Figures 5 and 6 represent the analysis of various performance metrics of the proposed BCNN-HAD. Regarding highway bridge

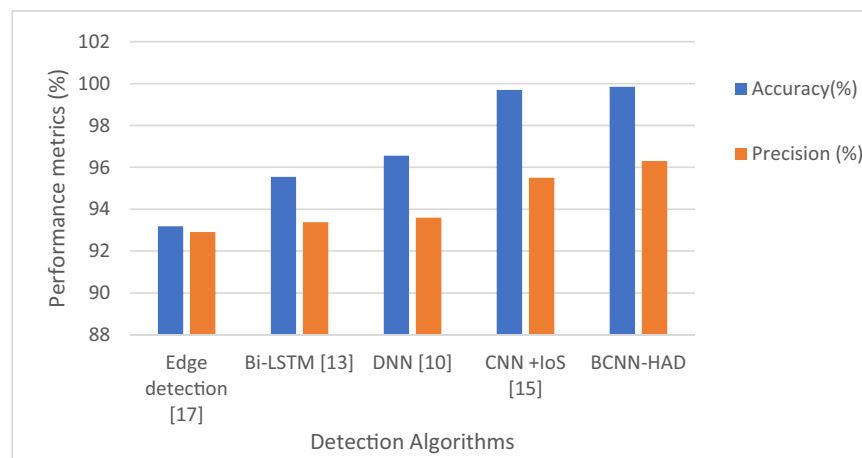


Figure 5: Evaluation of accuracy and precision based on the BCNN-HAD method.

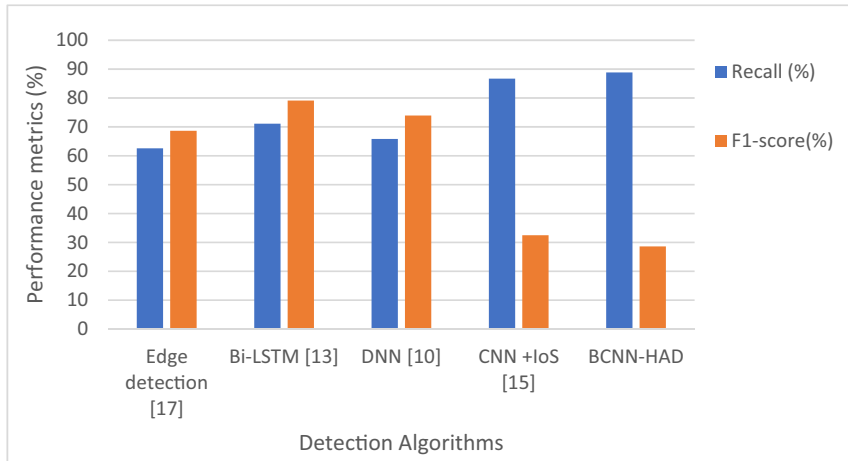


Figure 6: Evaluation of accuracy and precision based on the BCNN-HAD method.

engineering and BCNN-HAD, these metrics assist in evaluating the model's capacity to effectively detect and handle irregularities on the pavement, which helps to maintain the durability and safety of traffic infrastructure.

4.3 Training and validation time

“Training and validation time” is a measurement that quantifies the duration required to train and validate the CNN model using the suggested BGV-YOLOv5 algorithm. This measure is essential for comprehending the computational effectiveness and pace of the training procedure, which is vital for the actual execution and utilization of the model.

Training Time is the time needed to train a CNN model using the training dataset. It is an important measurement, and shorter training durations are often favored for quicker model growth. Nevertheless, finding a middle ground between minimizing training time and guaranteeing that the model attains both high accuracy and robustness is essential. Validation Time is the duration required to assess the model's effectiveness, which was trained upon a distinct validation dataset. Just like the time spent on training, it is preferable to have shorter validation durations since they enable faster iterations in improving the model. Verification is crucial for assessing the model's ability to generalize unfamiliar data.

Training and Validation Combined Time refers to the whole duration of both the training and validation processes. This measurement offers a thorough perspective on the time spent designing and assessing the model. It helps calculate the total computational resources used for both the training and validation stages, providing information about the efficiency of the model creation process. Equation (17) shows the total training time taken by the epoch (T_i) as,

$$T_i = a \times i + b. \quad (17)$$

Here, T_i is the total time at epoch i , a is the slope of the line, indicating the rate at which time rises per epoch, i is the epoch number, and b is the y-intercept, reflecting the initial time at epoch 1. Figure 7 shows the training and validation evaluation curve compared with other object detection methods.

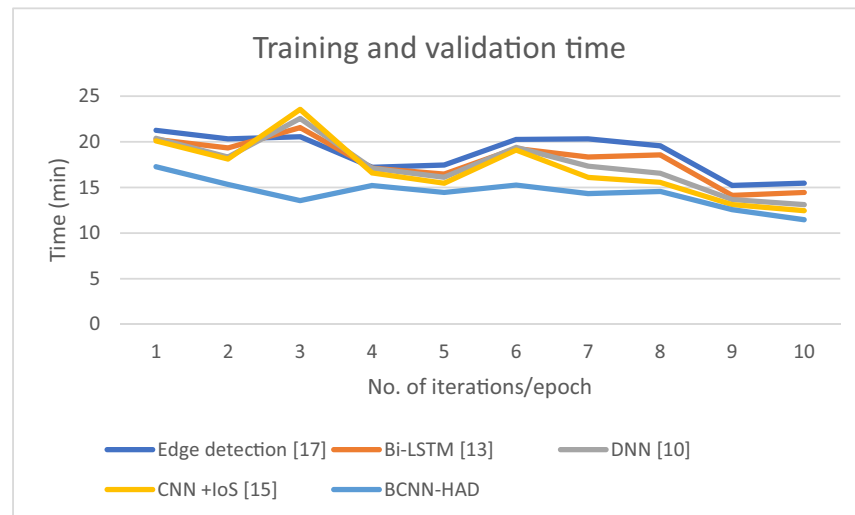


Figure 7: Evaluation of training and validation time for BCNN-HAD.

Extended training and validation durations may suggest a model requiring significant computational resources or more advanced hardware. Reduced durations indicate effectiveness in the model structure, optimization methods, or parallel processing skills. To summarise, “training and validation time” is a practical measure that affects the practicality and scalability of implementing the suggested BGV-YOLOv5 algorithm in real-life scenarios. Reducing this time while ensuring high-quality outcomes is essential for successful deployment.

The proposed BCNN-HAD method and BGV-YOLOv5 are essential in transportation infrastructure and road safety. By effectively detecting road anomalies, they enhance infrastructure maintenance and road safety in highway bridge engineering. This comprehensive approach aligns to establish a safer and longer-lasting transportation network. The study assesses the efficiency of the BCNN-HAD technique utilizing parameters such as mean average precision (mAP) and processing speed. mAP evaluates the precision of object detection by taking into account precision-recall curves. The processing speed measures how fast the system detects abnormalities in real-time situations. These measurements offer a thorough assessment of the accuracy and effectiveness of the suggested approach in identifying irregularities on highway bridges.

4.4 Effectiveness of sensor fusion techniques











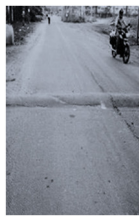






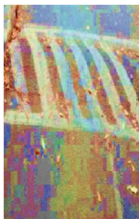


Sensor fusion combines information from several sensors to improve the overall comprehension and precision of the surroundings. When identifying road bumps and potholes in bridge engineering, combining information from several sensors is important to improve the system’s capabilities. Within highway bridge engineering, sensor fusion techniques in the suggested system are crucial for improving the identification of road anomalies. Combining data from several sources, such as bridge structural health monitoring systems, cameras, accelerometers, and GPS, is essential in understanding the bridge environment. Evaluation criteria include the correlation and synchronization of sensor data, managing redundancy, real-time processing capabilities, improving accuracy through fusion, adaptability to environmental changes, influence on maintenance decision-making, and scalability to handle larger data volumes. Effective sensor fusion guarantees that the system adjusts to various situations, offers practical insights for maintenance choices, and stays efficient as the monitoring infrastructure grows. A thorough assessment of these factors enables engineers to improve the sensor fusion algorithms, guaranteeing the dependability and efficiency of road anomaly detection in practical highway bridge scenarios.

Sensor integration is essential in BGV-YOLOv5 for detecting highway anomalies by merging data from systems that monitor structural health, cameras, accelerometers, and GPS. This inclusive method improves the comprehension of the area around the bridge, which helps detect road irregularities more effectively. Combining many sources ensures a comprehensive viewpoint for designing intelligent transportation systems in highway bridge engineering.

4.5 Evaluation of different suggestions of current technology

The suggested BGV-YOLOv5-based system for identifying road irregularities is compared to various object identification models designed to detect irregularities on roadways. This evaluation entails analyzing the advantages and disadvantages of the BGV-YOLOv5 system in terms of accuracy, completeness, and overall effectiveness. Table 1 shows the comparative analysis of the proposed BCAA-HAD system with other proposals.

Table 1: Results using the test dataset of speed bump images and the trained CNN model

S. no.	Input image	Preprocessing	Segmented image	Output image	Output description
1					Image of a speed bump taken during the day. With great precision, our model can identify the speed bump
2					With the help of brightness and contrast adjustments, the collected and enhanced speed bump picture is accurately forecasted
3					Orientation is changed by augmenting the captured picture of the speed bump. Using our algorithm, about 70% of the pixels may be accurately classified as speed bumps
4					Showing an image of an unlabeled speed bump. Even though there are no unlabeled speed bumps in our training picture dataset, our model still manages to segment around 25% of the pixels accurately
5					Accurate segmentation is not achieved when the speed bump picture is shot from a vertical perspective. Twenty percent or so of the pixels are incorrectly labeled as speed bumps even though they belong to a different class

Our advanced neural network utilizes a modest seven layers, which is less demanding on the processing hardware than Inception V2, which has 12 levels. Our approach attains an accuracy of 99.13%, as indicated in Table 2.

Table 2: Study of various state-of-the-art proposals

Various proposals	Datasets	Anomaly detection
Edge detection [17]	8,905 Images	Potholes
Bi-LSTM [13]	6,433 Images	Road bump
DNN [10]	899 Images	Potholes
CNN + IoS [15]	712 Images	Potholes
Proposed BCNN-HAD	655 Images	Both potholes + bump

Unlike other works with the same approach, this study chooses not to use strides or combined techniques to gather all the accessible information. In addition, BCNN-HAD includes a “BatchNormalization” layer by layer to speed up the training process and reduce the impact of network initialization, a feature not present in other systems. Furthermore, the proposed method uses a “SoftMax” layer that enables the interpretation of the fully linked network’s output as probabilities.

The research evaluates machine learning models for identifying objects, specifically focused on YOLO series detectors such as YOLOX, for jobs requiring real-time processing. BGV-YOLOv5, using CNNs, demonstrates practicality and efficiency in detecting road irregularities, improving road safety in highway bridge engineering, and contributing to intelligent transportation systems.

5 Conclusion

Overall, the suggested BCNN-HAD, based on BGV-YOLOv5, is an advanced option in highway bridge engineering. The “BGV” program, designed for highway bridges, shows its effectiveness in automated image processing for accurately detecting road irregularities like bumps and potholes. The model may be flexible and accurate in real-life scenarios by using sophisticated image processing methods and combining data from several sensors, especially in different environments. The inclusive method exemplified by BGV-YOLOv5 helps to a more profound comprehension of the bridge environment, aiming to improve the identification of road abnormalities and progress the field of bridge health monitoring. The expected results of this study comprise creating a customized and effective system designed to identify road surface irregularities specific to highway bridge situations. This method tackles highway bridge engineers’ difficulties in enhancing road safety and infrastructure maintenance standards. In addition, by examining sensor fusion techniques and including various data sources such as structural health monitoring systems, cameras, accelerometers, and GPS, the project aims to provide a robust framework for intelligent transportation systems in the field of bridge engineering. The primary purpose is to create a transportation network that is safer and longer-lasting, in line with the broader goals of improving the resilience of infrastructure and assuring the longevity of highway bridges. Further study could improve the model’s ability to adapt to changing climate conditions and expand its usefulness to different bridge designs. Working with transportation authorities and industry stakeholders could help incorporate the created system into the current infrastructure, leading to a broader influence and better methods for maintaining highway bridges.

The experimental analysis uncovers restrictions in the study, such as the requirement for crowdsourced data to improve system effectiveness. Possible enhancements may involve increasing the dataset used for training models and dealing with datasets that have imbalances. These constraints might affect the practical use by impacting the system’s overall effectiveness and adaptability in real-life situations.

Funding information: The authors received no financial support for the research, authorship, and/or publication of this article.

Author contributions: Qianlong Tang: editing data curation and supervision. Haiping Ding: writing – original draft preparation.

Conflict of interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

Data availability statement: This article does not cover data research. No data were used to support this study. <https://www.kaggle.com/datasets/mrongsubho/pothole-vehicle-and-speedbraker-detection-yolov5>.

References

- [1] Jansen A, Geißler K. Multi-feature anomaly detection for structural health monitoring of a road bridge using an autoencoder. In 10th International Conference on Structural Health Monitoring of Intelligent Infrastructure–SHMII. Vol. 10, 2021.
- [2] Moallemi A, Burrello A, Brunelli D, Benini L. Exploring scalable, distributed real-time anomaly detection for bridge health monitoring. *IEEE Internet Things J.* 2022;9(18):17660–74.
- [3] Lydon D, Kromanis R, Lydon M, Early J, Taylor S. Use a roving computer vision system to compare anomaly detection techniques for health monitoring of bridges. *J Civ Struct Health Monit.* 2022;12(6):1299–316.
- [4] Wang J, Wang M, Liu Q, Yin G, Zhang Y. Deep anomaly detection in expressway based on edge computing and deep learning. *J Ambient Intell Hum Comput.* 2022;13(3):1293–305.
- [5] Chitale PA, Kekre KY, Shenai HR, Karani R, Gala JP. Pothole detection and dimension estimation system using deep learning (yolo) and image processing. In 2020, the 35th International Conference on Image and Vision Computing New Zealand (IVCNZ). IEEE; 2020, November. p. 1–6.
- [6] Pham SVH, Nguyen KVT. Productivity assessment of the Yolo V5 model in detecting road surface damages. *Appl Sci.* 2023;13(22):12445.
- [7] KC S. Enhanced pothole detection system using YOLOX algorithm. *Auton Intell Syst.* 2022;2(1):22.
- [8] Archana KV, Komarasamy G. A novel deep learning-based brain tumor detection using the Bagging ensemble with K-nearest neighbor. *J Intell Syst.* 2023;32(1):1696–9. doi: 10.1515/jisys-2022-0206.
- [9] Hassan M, Kanwal A, Jarrah M, Pradhan M, Hussain A, Mago B. Smart city intelligent traffic control for connected road junction congestion awareness with deep extreme learning machine. In 2022 International Conference on Business Analytics for Technology and Security (ICBATS). IEEE; 2022, February. p. 1–4.
- [10] Peralta-López JE, Morales-Viscaya JA, Lázaro-Mata D, Villaseñor-Aguilar MJ, Prado-Olivarez J, Pérez-Pinal FJ, et al. Speed bump and pothole detection using a deep neural network with images captured through a zed camera. *Appl Sci.* 2023;13(14):8349.
- [11] Bharat R, Ikotun A, Ezugwu A, Abualigah L, Shehab M, Zitar RA. A real-time automatic pothole detection system using convolution neural networks. *Appl Comput Eng.* 2023;6(1):879–86.
- [12] Wu C, Wang Z, Hu S, Lepine J, Na X, Ainalis D, et al. An automated machine-learning approach for road pothole detection using smartphone sensor data. *Sensors.* 2020;20(19):5564.
- [13] Lyu H, Xu K, Jiao D, Zhong Q. Bump feature detection of the road surface based on the Bi-LSTM. *Open Geosci.* 2023;15(1):20220478.
- [14] Alhussan AA, Khafaga DS, El-Kenawy ESM, Ibrahim A, Eid MM, Abdelhamid AA. Pothole and plain road classification using adaptive mutation dipper throated optimisation and transfer learning for self-driving cars. *IEEE Access.* 2022;10:84188–211.
- [15] Ozoglu F, Gökgöz T. Detection of road potholes by applying convolutional neural network method based on road vibration data. *Sensors.* 2023;23(22):9023.
- [16] Bosurgi G, Modica M, Pellegrino O, Sollazzo G. An automatic pothole detection algorithm using pavement 3D data. *Int J Pavement Eng.* 2022;1–15.
- [17] Borgalli R. Smart pothole detection and mapping system. *J Ubiquitous Comput Commun Technol.* 2020;2(3):136–44.
- [18] Shaghouri AA, Alkhatib R, Berjaoui S. Real-time pothole detection using deep learning. *arXiv preprint arXiv:2107.06356.* 2021;1–10.
- [19] Agushaka JO, Ezugwu AE. Evaluation of several initialization methods on arithmetic optimization algorithm performance. *J Intell Syst.* 2021;31(1):70–94. doi: 10.1515/jisys-2021-0164.