IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 April 2024, accepted 22 April 2024, date of publication 6 May 2024, date of current version 3 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3397002

RESEARCH ARTICLE

Multi-Modal Fusion for Multi-Task Fuzzy Detection of Rail Anomalies

YANG LIYUAN[1,2], GHAZALI OSMAN[2], SAFAWI ABDUL RAHMAN[2], AND MUHAMMAD FIRDAUS MUSTAPHA[2]

[1]Kunming University, Kunming, Yunnan 650214, China
[2]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Shah Alam, Selangor 40450, Malaysia

Corresponding author: Ghazali Osman (ghaza936@uitm.edu.my)

This work was supported in part by Kunming University of China, Yunnan Provincial Department of Education Scientific Research Fund Project, under Grant 2023J0843 (RMB 20,000); and in part by Kunming Science and Technology Bureau Project under Grant 2023-1-NS-031 (RMB 0).

ABSTRACT Due to prolonged exposure to heavy train loads, various anomalies can emerge on the surface of railway tracks, posing a direct threat to safe train operation. The accurate and timely detection of these anomalies is important to ensure safe transportation and advancing intelligent maintenance. However, in the domain of anomaly detection, several challenges have arisen owing to the variability in illumination conditions, imaging blur inherent to the capture devices, and the introduction of noise from environmental factors such as dust particles. These interferences have significantly undermined the accuracy of rail anomaly detection based on target detection techniques. In response to these challenges, this study introduced a novel approach for rail anomaly detection in the presence of image artifacts by utilizing a multi-modal multi-task framework. The objective of this study is to enhance the performance of rail anomaly detection under interference-prone conditions. This study integrated color moment features, HU invariant moment features, and Haralick features to construct a fuzzy detection model for rail anomalies using a multi-task learning (MTL) strategy. The model prioritized the primary task of classifying rail anomalies, with interference level classification and fuzzy logic interference judgment serving as auxiliary tasks within the network. Finally, based on the results of fuzzy logic and interference level detection, a fuzzy judgment was made distinguishing between ''with interference'' and ''without interference.'' Experimental findings consistently demonstrated that the integration of multi-modal features and multi-task learning methodologies significantly improves the accuracy of rail anomaly recognition in the presence of interference, thus establishing an effective approach for rail anomaly identification in challenging scenarios.

INDEX TERMS Rail anomalies, classification, multi-feature fusion, multi-task learning, fuzzy logic, neural network.

I. INTRODUCTION

Railway transportation is a critical infrastructure in modern cities and serves as the backbone of public transportation. However, the continuous impact of heavy-loaded trains on rails, as well as stress, fatigue, and defects on the rails, has led to rail anomalies such as scratches, blocks, cracks, scabs, and even fractures on the surface of rails, which directly threaten the safety of high-speed trains. Machine vision-based object detection technology can more accurately and efficiently detect anomalies in rails, which is beneficial for implementing different maintenance and repair measures against various types of anomalies. The accurate image recognition of rail anomalies is of paramount importance. However, railway tracks are exposed to outdoor environments over extended periods and their surfaces are susceptible to the accumulation of small particles such as dust and soil, among other disturbances. Variations in brightness owing to sunlight can also lead to alterations in the appearance, color, and even the emergence of issues such as shadows, highlights, and reflections in the captured images. However, factors such as equipment vibrations, train movements, or inaccurate camera

The associate editor coordinating the review of this manuscript and approving it for publication was Jesus Felez.

focusing during the image acquisition process can introduce image blurring and other distortions. These interferences collectively impact the identification of rail anomalies, thereby diminishing recognition accuracy. Figure 1 shows the examples of disturbances in rail imaging caused by dust, sunlight, and blurriness.

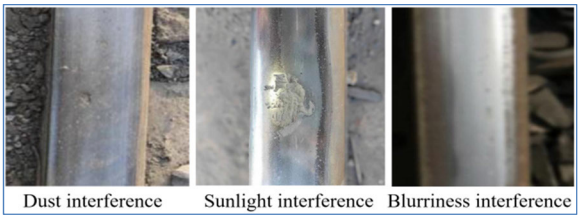


FIGURE 1. Interfered anomalous railway tracks.

In this study, based on a newly developed small-sample dataset, we propose a multi-feature multi-task fuzzy detection method to enhance the performance of railway track anomaly recognition under interference conditions. The main contributions of this paper are as follows:

a. Fusion of color moments, HU moments, and Haralick features of railway track anomalies to improve the recognition accuracy, addressing the issue of low accuracy in single-feature recognition.

b. Introduction of a multi-task learning approach on top of the multi-feature fusion method, utilizing interference information to aid neural network learning. The primary task is railway track anomaly recognition, with a secondary task for simulating interference level recognition and a third task for fuzzy logic classification of ''With Interfered'' and ''Without Interfered.'' These auxiliary tasks collaborate to enhance the performance of the primary task.''

II. RESEARCH REVIEW

For the detection of anomalies on rail surfaces, Zhang et al. [1] introduce an enhanced approach employing YOLOX and image enhancement techniques for the detection of rail surface defects. The method yields a notable 2.42% improvement in the mean average precision (mAP) of the YOLOX network. Yang et al. [2] combined the grayscale features of different parts of the image and proposed a rail surface segmentation method based on sliding window grayscale maximum value, achieving an improved detection accuracy of 2.78%. Wang et al. [3] proposed a wavelet subband minimum mean square (LMS) adaptive filter, which effectively eliminates strong noise in railways and detects crack signals. Wang et al. [4] proposed a deep learning-based region fusion network for detecting rail wear on the running band. Firstly, they utilized a modified MiDaS model to estimate depth maps to guide rail wear detection. Then, an improved mask-based convolutional neural network was employed to segment and extract the running band of the rail from the images. Finally, a dual-channel attention fusion network was constructed, achieving a recall rate of 84.21% for wear detection. Karakose et al. [5], [6] employed cameras installed at the bottom and top of trains to capture images of rail tracks and surfaces. They utilized Canny edge detection and Hough transform methods to identify detection targets, ultimately applying a decision tree classification algorithm for recognizing track types and classifying surface anomalies. Tastimur et al. [7] initially employed Hough transform, morphological operations, and edge detection to detect railway tracks. Subsequently, through image enhancement, Laplacian low-pass filtering, and morphological feature extraction, they identified anomalous regions with a detection accuracy of 94.73%. Santur et al. [8] utilized Principal Component Analysis (PCA), Kernel Principal Component Analysis (KPCA), Singular Value Decomposition (SVD), and Histogram Matching (HM) to extract image features of steel rails. They then employed the Random Forest method to distinguish between ''normal'' and ''abnormal'' rails. The research findings indicated that combining Principal Component Analysis with the HM method enhanced the overall accuracy of the system to 85%. Origlia [9] and Lin [10] improved detection performance using time series methods.

The surface anomalies of steel rails are difficult to detect owing to the influence of light changes, camera shaking, and stains. To address this issue, Luo et al. [11] employed Gabor filtering to denoise the images and transform the color model to HSV, followed by an improved Faster R-CNN for defect recognition. The experimental results indicated that the recognition accuracies for cracks, blocks, and scratches were 91.87%, 92.75%, and 91.52%, respectively. Meanwhile, He et al. [12] proposed a rail surface defect detection algorithm based on background differences that includes four steps: (1) rail area extraction, (2) background modeling and difference, (3) threshold segmentation, and (4) image filtering. This method partially addresses the adverse effects of image lighting changes, uneven reflection, and lack of features during the rail surface defect segmentation process. The experimental results showed that this method could identify block defects with recall and accuracy rates of 96% and 80.1%, respectively. Min et al. [13] addressed the influence of vibration interference by combining the Hough transform and least squares method to extract the rail surface area, followed by the super-entropy theory and fuzzy theory for defect segmentation. They also established a sample feature database by extracting Harr-like and low-level features and designed a defect classifier based on the Classification and Regression Tree (C4.5) and AdaBoost algorithms. They found that the average recognition rate for rail surface defects was 97.02%.

These methodologies have demonstrated promising results in the detection of anomalies in rails. However, rails are susceptible to environmental interferences such as variations in sunlight, noise, and blurriness. Overcoming these interferences presents distinct challenges. For instance, addressing the challenge of mitigating sunlight interference involves enhancing the model's robustness to changes in illumination. Overcoming noise interference entails the accurate identification of small targets, such as dust particles. Addressing

blurry interference involves improving image restoration performance in specific environmental conditions. This study comprehensively considers the presence of these three interferences, leveraging interference-level information to assist network training. This approach aims to address disturbances caused by sunlight, dust, and blurriness in anomaly detection, thereby enhancing the performance of rail anomaly detection.

III. METHODOLOGY

This study presents a multi-task railway track anomaly fuzzy detection method based on feature fusion. Figure 2 illustrates the workflow of the research methodology, which consist of the following key steps:

a. Dataset Acquisition and Preprocessing: We established and preprocessed a novel dataset of railway track anomalies. Computer-simulated interference including illumination, noise, and blurriness was introduced into the image data.
b. Feature Extraction: We extracted the color moment, HU moment, and Haralick features from the image data.
c. Multi-feature and multi-task blur detect: The multi-feature fusion and multi-task learning fuzzy detection models were incrementally constructed and subsequently trained.
d. Comparative result: Comparison of recognition results of different models.

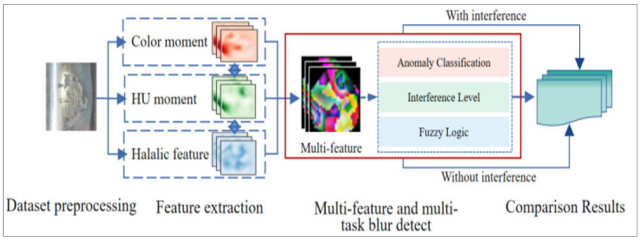


FIGURE 2. The flowchart of methodology.

A. IMAGE DATASET

In the field of railway engineering, there are various datasets for the classification and detection of railway defects available, including datasets for railway scene classification and detection [14] and datasets for ultrasonic detection [15]. However, there is lack of a realistic dataset for surface anomalies on railway tracks. This is because the distribution of abnormal steel tracks is random, and the collection of data on such abnormalities is difficult to be conducted due to the operational constraints of high-speed railways. To address this issue, researchers collected and labeled 100 × 50 resolution images of six categories of anomalies [16] and built an RSDDs dataset [17] containing grayscale images of 195 steel track surface anomalies in Type-I and Type-II segmented masks. Additionally, the NEU data set [18] focusing mainly on the surface anomalies of hot-rolled strip steel was also used. However, these datasets were collected using high-speed linear scanning cameras with low-resolution and coarse-grained annotations. Thus, they cannot adequately train robust deep-learning algorithms for real-world applications.

The rail anomaly dataset used in this study was collected by the author's team from various railway lines in Yunnan Province, China, including the Neijiang-Liupanshui and Kunming-Hekou Railways. Images were captured using camera equipment positioned at a vertical distance of approximately 20 cm from the damaged surface (rail track surface). This dataset represents a novel and specific focus on rail surface anomalies. Four types of anomalies, namely scratch, block, crack, and scab formation, were selected as the research objects. A total of 427 images depicting rail anomalies were utilized for experimentation, including 200 images of normal rails, 56 images of scratches, 60 images of blocks, 56 images of cracks, and 55 images of scab formations. This study separately introduced varying levels of luminance, noise, and blur interference in the images. In total, 12,810 images with added interference information were generated. Among these, there were 4,270 images with luminance interference, 4,270 with noise interference, and 4,270 with blur interference. Table 1 presents a statistical summary of these images. The following preprocessing steps were applied to the image data.

a. The images were cropped to a size of 448 × 448 pixels, with the steel rail anomaly region as the center (Region of Interest or ROI).
b. Simulations were conducted for sunlight variation with 10 levels of luminance interference, dust interference with 10 levels of Gaussian noise interference, and camera shaking with 10 levels of blur interference. Table 2 provides the details of the settings for the 10 interference levels.
c. The RGB color model of the ROI images was converted to YCbCr.

Table 3 shows examples of images from the dataset with and without interference (level 10).

TABLE 1. Statistical summary of image data.

With Interference Image															
Type	Normal			Abnormal											
				Scratch			Block			Crack			Scab		
Quantity	200			56			60			60			55		
Without Interference Image															
Type	Normal			Abnormal											
				Scratch			Block			Crack			Scab		
Interference	L	N	F	L	N	F	L	N	F	L	N	F	L	N	F
Quantity	2000	2000	2000	560	560	560	600	600	600	560	560	560	550	550	550
Total	6000			6810											

L=Luminance interference ; N=Noise interference; F=Fuzzy interference

B. FEATURE EXTRACTION

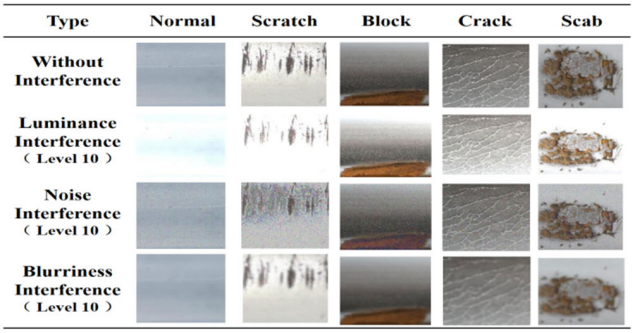
1) HARALICK FEATURE

Haralick, which was initially proposed as a method to quantify the relationships between adjacent pixels in images, has

TABLE 2. Coefficients for different levels of interference.

Coefficient	Undisturbed	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8	Level 9	Level 10
Luminance	1	0.5	0.6	0.7	0.8	0.9	1.1	1.2	1.3	1.4	1.5
Noise	0	3	6	9	12	15	18	21	24	27	30
Blurriness	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5

TABLE 3. Examples of images.



found extensive applications in image processing [19] and computer vision [20], among other multidimensional data classification domains. Recently, it has gained significant traction in the field of medicine [21], [22]. Haralick can extract rich texture information from images and capture texture characteristics that are significant in human perception. Moreover, Haralick features often exhibit robustness against variations in the image illumination and brightness.

In this study, Haralick features were primarily based on a Color Mapping Co-occurrence Matrix (CMCM). CMCM is a statistical feature used to describe texture characteristics in color images by examining the correlations and distributions between colors [23]. It describes the frequency of occurrence of different color values in given directions (0^o, 45^o, 90^o, 135^o) and distances (stride).

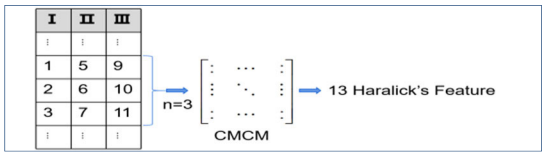


FIGURE 3. Haralick for CMCM.

As illustrated in Figure 3, each pixel of a color image I is represented by three color channels, such as R(Red), G (Green), and B (Blue), i.e., $I(x, y) = [R(x, y), G(x, y), B(x, y)]$, where (x, y) represents the pixel's coordinates, with $x \in 1,2,\ldots, N\}$ and $y \in 1,2,\ldots, M\}$. The image is divided into $n \times n$ pixel patches. Let P denote the number of patches in the image, such that $P = (N\text{-}n + 1)$ x $(M\text{-}n + 1)$. For each patch, denoted as k ($k = 1,2,\ldots, P$), a color mapping co-occurrence matrix ($CMCM\_k$) of size $B \times B$ is computed. Here, B represents the number of bins based on histograms

TABLE 4. Haralick for CMCM.

Parameters	Physical Meaning
Energy (Energ)	Measures the uniformity and regularity of image texture
Variance (Sosvh)	Characterizes irregularities and variations in image texture
Contrast (Contr)	Assesses the degree of contrast between different grayscale levels in the image
Correlation (Corrm)	Quantifies the correlation between GLCM elements, reflecting the linear characteristics of image texture
Homogeneity (Homom)	Signifies the degree of consistency among texture elements in the image
Sum Average (Savgh)	Reflects the average distribution of texture elements with the same grayscale level in the image
Sum Variance (Svarh)	Measures the distribution variance of texture elements with the same grayscale level
Sum Entropy (Entro)	Measures the overall entropy of GLCM elements, representing the overall uncertainty of image texture
Difference Variance (Dvarh)	Reflects the variance of differences between different grayscale levels in the image
Difference Entropy (Denth)	Reflects the variance of differences between different grayscale levels in the image, indicating the uncertainty of differences
Information Measure of Correlation 1 (Inf1h)	Related to correlation but incorporates the concept of entropy in its calculation, providing a more comprehensive description of the correlation and uncertainty of image texture
Information Measure of Correlation 2 (Inf2h)	Related to correlation but employs different mathematical formulas to provide a more precise description of the correlation and uncertainty of image texture

for each color channel. The calculation method for $CMCM\_k$ is as follows:

a. Quantize the colour values from each channel $\omega(\omega=I,II,III)$ of patch $k$into B bins using a uniform quantization scheme.
b. Construct a joint histogram of color pairs (i, j) within patch k, which represents the frequency of occurrences of color pairs between different color channels within the patch. Here, i and j represent the quantized color values for channel 1 and channel 2, channel 1 and channel 3, and channel 2 and channel 3, respectively.
c. Normalize the joint histogram by dividing each element by the total number of color pairs in the patch to obtain a probability matrix.
d. Add the probability matrices of all patches to obtain the final CMC of size $B \times B$.

Based on the CMCM, texture features were computed using Haralick's 13 parameters (Tab. 4).

2) COLOR MOMENT FEATURE

Color moments can effectively represent color distribution in an image [24]. The mathematical basis of this method is that any color distribution can be represented by its moments. Moreover, the color distribution information is mainly concentrated in low-order moments, and features are described using the first-order moment (mean), second-order moment (variance), and third-order moment (skewness) of the color (Table 5). The mean (E) represents the average of the color distribution, specifically the average value of all pixels in the image. It gauges the overall brightness, or saturation, of the image, with larger values indicating a brighter image. Variance (σ) characterizes the dispersion of the color distribution, i.e., the variability of pixel values in the image. It measures the range of the color distribution, with larger values indicating a broader color distribution range. Skewness (S) quantifies the degree of asymmetry in the color distribution of pixel values in the image. It assesses the extent of skewness in

TABLE 5. Formulas for feature calculation.

Color moments	Haralic Feature
$E_i = \sum_N^{j=1} \frac{1}{N} P_{ij}$	$Energ = \sum_i \sum_j p(i,j)^2$
$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_N^{j=1}(p_{ij} - E_i)^2\right)}$	$Sosvh = \sum_i \sum_j (i-\mu)^2 p(i,j)$
$S_i = \sqrt[3]{\left(\frac{1}{N}\sum_N^{j=1}(p_{ij} - E_i)^3\right)}$	$Contr = n^2 \left\{\sum_{i=1}^{Ng}\sum_{j=1}^{Ng} p(i,j)\right\} \lvert i-j\rvert = n$
HU moments	$Corrm = \frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
$M_1 = y_{20} + y_{02}$	$Homom = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
$M_2 = (y_{20} - y_{02})^2 + 4y_{11}^2$	$Savgh = \sum_{i=2}^{2Ng} i p_{x+y}(i,j)$
$M_3 = (y_{30} - 3y_{12})2 + (3y_{21} - y_{03})^2$	$S\,\mathrm{var}\,h = \sum_{i=2}^{2Ng} (i - f_S)^2 p_{x+y}(i)$
$M_4 = (y_{30} + y_{12})^2 + (y_{21} + y_{03})^2$	$Senth = -\sum_{i=2}^{2Ng} p_{x+y}(i) \log\{p_{x+y}(i)\} = f_S$
$M_5 = (y_{30} - 3y_{12})(y_{30} + y_{12})\left((y_{30} + y_{12})^2 - 3(y_{21} + y_{03})^2\right)$	$Entro : -\sum_i \sum_j p(i,j) \log(p(i,j))$
$\ldots + (3y_{21} - y_{03})(y_{21} + y_{03})\left(3(y_{30} + y_{12})^2 - (y_{21} + y_{03})^2\right)$	$D\ \mathrm{var}\ h = \sum_{i=0}^{Ng-1} i^2 p_{x-y}(i)$
$M_6 = (y_{20} - y_{02})\left((y_{30} + y_{12})^2 - (y_{21} + y_{03})^2\right)$	$Denth = -\sum_{i=0}^{Ng-1} p_{x-y}(i) \log\{p_{x-y}(i)\}$
$\ldots + 4y_{11}(y_{30} + y_{12})(y_{21} + y_{03})$	$Inf1h = \frac{HXY - HXY1}{\max\{HX, HY\}}$
$M_7 = (3y_{21} - y_{03})(y_{30} + y_{12})\left((y_{30} + y_{12})_2 - 3(y_{21} + y_{03})^2\right)$	$Inf2h = (1 - \exp[-2[HXY2 - HXY]])^{\frac{1}{2}}$
$\ldots - (y_{30} - 3y_{12})(y_{21} + y_{03})\left(3(y_{30} + y_{12})^2 - (y_{21} + y_{03})^2\right)$	

the color distribution, with larger values indicating a greater concentration of colors in one region rather than an even distribution across the entire image [25].

3) HU MOMENTS FEATURE

HU moments [26] are a type of internal transform method used for shape analysis. Moment features are another type of internal transform method that is used for shape analysis. The first moment is related to the shape; the second moment shows the extent to which a curve extends around the average value of a line; and the third moment measures the symmetry of the mean. HU uses the second and third normalized central moments to construct seven (7) invariant moments (Table 5). Among these, *M1*represents the weighted average concerning image intensity, typically indicative of the overall grayscale of the image. *M2* quantifies the shape contrast of the image, specifically the distribution differences of pixels in the horizontal and vertical directions. *M3* measures the shape symmetry of the image, reflecting the distribution differences of pixels along diagonal directions. *M4* assesses the stretching and skewness of the image's shape, primarily employed to quantify the disparities in pixel distribution along the main and minor diagonals. *M5* describes the skew contrast of the image, primarily capturing differences in pixel distribution during clockwise and counterclockwise rotations. *M6* characterizes the angular contrast of the image, reflecting disparities in pixel distribution at different angles. *M7* evaluates the shape invariance of the image, indicating differences in pixel distribution under scaling transformations.

The HU moment [27] feature does not change with position, size, and orientation, which are less affected by factors such as noise. It's also invariant to the image translation, scaling, and rotation. Therefore, this study used seven (7) invariant moments as parameters for shape features. The Canny edge detection algorithm was employed to extract the edges of anomalous regions in rail images, followed by the utilization of the OTSU algorithm [27] to determine the threshold for binary segmentation. Subsequently, the anomalous regions were segmented, and seven (7) HU moments [26] were calculated.

Table 5 presents the mathematical expressions for calculating the parameters of the Color Moments, HU Moments, and Haralick Features.

C. MULTIPLE FEATURES AND MULTIPLE TASKS

1) MULTI-FEATURE FUSION

Feature fusion can enhance detection performance [28]. In this study, a feature concatenation approach was employed to combine the extracted features, including the Color Moment (9 dimensions), HU Moment (7 dimensions), and Haralick features (13 dimensions). Initially, the Mahalanobis Distance method was used to eliminate outlier values, followed by dimensionality reduction using Principal Component Analysis (PCA) to reduce the feature dimensions from 29 to 15. Subsequently, the StandardScaler function was applied to standardize the feature values, meaning that the data were scaled such that each feature had a mean of 0 and a variance of 1. This standardization helps mitigating the impact of variations in the feature thresholds, ultimately enhancing the accuracy and stability of the detection model. The processed data were then passed to the neural network model, which recognized four types of rail anomalies: scratches, blocks, cracks, and scabs.

The network architecture comprises two hidden layers (Fully Connected Layers) and 1 Dropout Layer. The Dropout Layer is positioned after the first hidden layer and serves the purpose of regularizing the target classification. Briefly, it randomly drops the output of the hidden layer to enhance the generalization capability. In this study, a dropout rate of 0.2 was configured. This implies that during each forward pass, there is a 0.2 probability of setting certain neuron outputs to zero, thereby reducing the risk of overfitting. Figure 4 illustrates the network structure for multi-feature fused rail anomaly recognition.

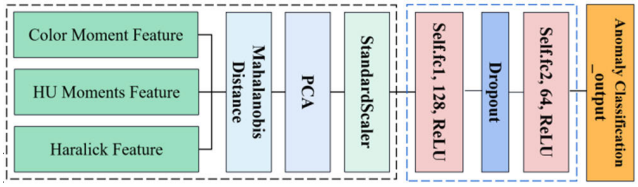


FIGURE 4. Network architecture for multi-feature fusion.

2) MULTI-TASK LEARNING

Multi-task learning comprises the recognition of rail-surface anomalies and the identification of interference level information. The network architecture consists of one input layer,

two hidden layers (Fully Connected Layer), one Dropout layer, and two output layers. The input layer of the network receives feature-fused data, which includes color moment features, HU moment features, and Haralick features, as indicated by the first dashed box in Figure 4. Both the anomaly recognition task and the interference level recognition task share the same feature extractor, namely, two hidden layers and one dropout layer, as depicted by the dashed box in Figure 5. Initially ion this study, the feature extractor extracts features, and the processing results were outputted through additional fully connected layers (target_output_layer and interference_output_layer). These fully connected layers branch into two segments; one dedicated to rail anomaly recognition and the other dedicated to interference-level recognition tasks. Subsequently, cross-entropy loss functions were separately defined for anomaly recognition and interference level recognition, and a softmax activation function was employed to transform the network's raw outputs into class probability distributions. This ensures that the sum of the probabilities for all output classes equals one. The network architecture for multi-task learning is shown in Figure 5.

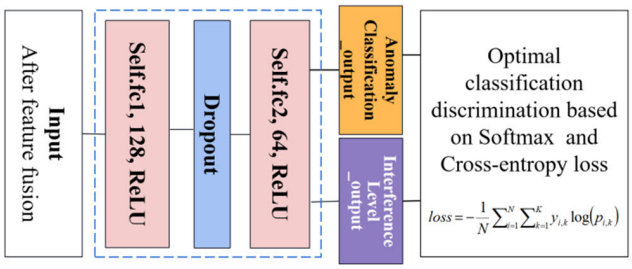


FIGURE 5. Network architecture for multi-task learning.

The cross-entropy loss function is a form of loss function employed to assess the disparity between the probability distribution of the predicted values and probability distribution of the actual labels. This loss function is characterized by its simplicity in gradient computation, rapid convergence in gradient descent, and insensitivity to outliers [29]. The mathematical formula is as follows:

$$loss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{i,k}\log\left(p_{i,k}\right) \tag{1}$$

The symbol N denotes the number of samples in the data set, where $y_{i,k}$ represents the true label of the k^{th} category for the i^{th} sample, and $p_{i,k}$ is the predicted probability of the k^{th} category for the i^{th} sample.

The Softmax activation function is a frequently employed activation function in tasks involving multi-class classification. It takes a vector containing real-valued numbers, typically referred to as logits, as input, and transforms it into an output that represents probabilities for different categories, with each category having a corresponding probability [30]. The mathematical formula is as follows:

$$S(z)\_k = e^{\wedge}(z\_k)/\sum(e^{\wedge}(z\_i)) \tag{2}$$

In this context, $S(z)\_k$ represents the output probability for the k-th category, z_k corresponds to the k th element of the input vector z, e denotes the base of the natural logarithm (approximately equal to 2.71828), and $\sum$ represents the summation symbol.

3) FUZZY LOGIC

The network architecture of the fuzzy detection model comprises one input layer, two hidden layers, two dropout layers, two neural network layers dedicated to fuzzy-logic tasks, and three output layers. The input consists of feature fusion data with three branches for three distinct tasks: rail anomaly recognition (Target_output), interference level recognition (Interference_output), and fuzzy logic tasks. Similar to the multi-task learning network structure, the fuzzy logic tasks in this study shared the same fused feature data as the input with the anomaly recognition and interference level recognition tasks and shared the feature extractor. In addition, a dropout layer was included to prevent overfitting. Two neural network layers (fuzzy_fc1 and fuzzy_fc2) dedicated to fuzz- logic tasks were added to this foundation to capture data fuzziness. Finally, a Sigmoid activation function was employed to map the network's output to a probability distribution range, where 0 represents low fuzziness or uncertainty, and 1 represents high fuzziness or uncertainty. The Mean Squared Error (MSE) loss function was used to measure the difference between the model's output and target values. Figure 6 illustrates the network structure of the fuzzy logic model.

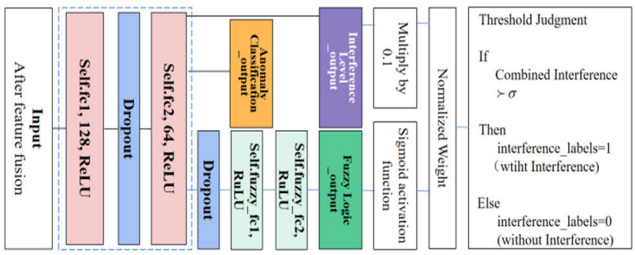


FIGURE 6. Network architecture for fuzzy logic.

The fuzzy-logic task in this study involved producing a scalar value between 0 and 1 based on the fuzziness or uncertainty of the input data. This scalar value was then compared to a predefined threshold to determine the presence of interference. In this study, the class probabilities for the anomaly recognition task were initially computed to identify the maximum class probability value and its corresponding predicted class. Subsequently, a scalar value ranging from 0 to 1, representing the output of the fuzzy logic task, was generated through forward propagation. Finally, the presence of interference was determined by comparing the maximum class probability value to a predefined threshold. If the maximum class probability value is below the threshold, the interference label is set to 1 (indicating interference); otherwise, it is set to 0 (indicating no interference). By introducing the fuzzy logic task, the model can simultaneously learn to handle

different types of tasks, engage in joint training with the target classification and interference classification tasks, and enhance its understanding of the input data, thereby improving the classification performance.

The fuzzy logic task constitutes a comprehensive assessment of uncertainty, primarily focusing on brightness, noise, and blur interference. Therefore, this study combined the results of the fuzzy logic task with interference level recognition through weighted fusion to form a distinct classification task of ''with interference'' and ''without interference.'' The threshold of the fuzzy logic task acted as a parameter for balancing the fuzzy logic task and interference level prediction results.

The mathematical representation of the Mean Squared Error (MSE) loss function is as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{3}$$

N represents the number of samples, y_i is the actual target value for the i th sample, $\hat{y}_i$ is the i model's prediction, $\sum_{i=1}^{N}$ denotes the summation over all samples, and $1/N$ signifies the average.

The mathematical representation of the Sigmoid activation function is:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

$\sigma(x)$ represents the output of the sigmoid activation function, and x represents the input value.

IV. RESULTS AND DISCUSSION

The performance of the model was assessed through K-fold cross-validation using a 5-fold cross-validation approach. The average results across these five(5) folds were used as outcome data. To address the imbalance in sample quantities, the weighted-random-sampler oversampling method was employed to augment the samples for the less represented types of rail anomalies. As this study focuses on enhancing the classification performance of rail anomalies under interference conditions through a multi-feature and multi-task approach, the experimental outcomes pivot around the utilization of accuracy in anomaly classification as the primary evaluation metric.

A. RESULTS OF MULTI-FEATURE FUSION

Firstly, the multi-feature fusion approach under interference-free was compared with the results obtained from single-feature recognition. Subsequently, luminance, noise, and blur interference were randomly introduced, followed by another round of experimental comparisons. Table 6-7 presents the accuracy of rail anomalies classification under both interference-free and interference-laden conditions using single-feature and multi-feature methods.

From Table 6-7, it is evident that the multi-feature fusion method outperforms the single-feature recognition results in rail anomaly detection. Under the condition without added interference, the multi-feature fusion method achieved recognition accuracy of 95.20%, which further improved to 98.82% after converting the RGB color model to YCbCr. In contrast, the other three single-feature methods, except for the color feature, exhibited a relatively modest recognition accuracy, with the color feature recognition accuracy exceeding 90%. Following the random introduction of interference, the recognition accuracy of all methods notably decreased; however, the multi-feature fusion method continued to surpass the single-feature methods in terms of accuracy.

TABLE 6. Results under interference-free conditions.

	Single-Feature Method			Multi-Feature Method
	Color Feature	HU Feature	Haralick Feature	Feature Fusion
RGB Color Model	92.05%	64.71%	32.00%	95.20%
YCbCr Color Model	95.29%	68.60%	67.06%	98.82%

TABLE 7. Results under interference conditions.

	Single-Feature Method			Multi-Feature Method
	Color Feature	HU Feature	Haralick Feature	Feature Fusion
Random Luminance Interference	62. 91%	58. 45%	50. 59%	67. 53%
Random Noise Interference	51. 18%	64. 08%	47. 06%	67. 06%
Random Fuzzy Interference	50. 01%	56. 59%	47. 06%	65. 18%

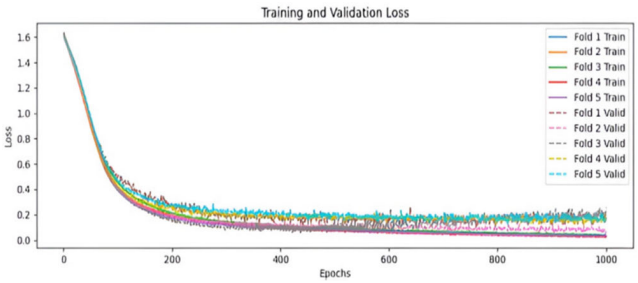


FIGURE 7. 5-Fold cross-validation loss.

Figure 7 illustrates the model loss for the multi-feature fusion method. Both training and validation loss curves exhibited a smooth descent without significant oscillations and sharp increases, indicating favorable convergence performance of the model.

B. RESULT OF MULTI-TASK LEARNING

To validate the effectiveness of the multi-task learning approach, this study conducted a comparative analysis between single-feature multi-task learning methods and multi-feature multi-task learning methods and their corresponding single-task learning counterparts. Initially, 427 original images were manually subjected to luminance, noise, and blur interference levels ranging from 1 to 10.

TABLE 8. Results of multi-feature multi-task.

		Single-Task Method			Multi-Task Method		
		Luminance Interference	Noise Interference	Fuzzy Interference	Luminance Interference	Noise Interference	Blurriness Interference
Single-Feature	Color Feature	82.08%	85.36%	91.91%	85.24%	86.06%	90.82%
	HU Feature	47.83%	56.62%	58.90%	54.53%	56.98%	59.47%
	Haralick Feature	50.59%	46.84%	47.19%	50.41%	46.85%	47.06%
Multi-Feature	Feature Fusion	80.33%	86.17%	92.27%	86.26%	91.10%	96.07%

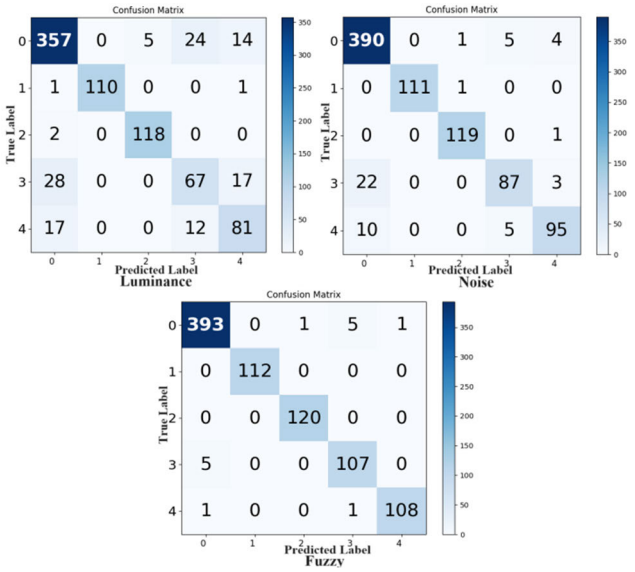
Subsequently, 4270 images, each with different interferences, underwent feature extraction and feature fusion. These processed images were then input into both the single-task and multi-task detection models for comparative experiments. Table 8 presents a comparison of the accuracy of anomaly classification between single-task and multi-task methods under both single-feature and multi-feature conditions.

From Table 8, it is evident that under conditions of luminance, noise, and blur interference, the accuracy of multi-task anomaly recognition was consistently the highest, reaching 86.26%, 91.10%, and 96.07%, respectively. This is notably higher, by 5.93%, 4.93%, and 3.8%, compared to single-task anomaly recognition. The accuracy of fused feature anomaly recognition was also higher than that of single-feature anomaly recognition. Figure 8 displays the confusion matrix of the multi-feature multi-task detection for recognizing the four types of anomalies under luminance, noise, and blur interference. In the confusion matrix, numbers 0-4 represent normal, scratches, blocks, cracks, and scabs.

Furthermore, it is noteworthy that color features, regardless of single or multi-task scenarios, exhibited significant accuracy in anomaly classification under all types of interference, with single-feature accuracy surpassing the 80% threshold. It is also evident that, despite the use of a simple sharing mechanism and keeping other conditions constant, the multi-task classification method, which introduces interference levels as a secondary label, demonstrated improvements in recognition accuracy compared to single-task classification learning methods. This implies that the multi-task learning approach utilizing interference information is effective in the task of rail anomaly recognition.

C. FUZZY JUDGMENT OF INTERFERENCE

The classification of blurred interference yielded two categories: ''With Interference'' and ''Without Interference.'' Initially, a scalar value between 0 and 1 was obtained through a fuzzy logic task. Subsequently, the interference prediction results were used to adjust the fuzzy logic output, reinforcing the influence of the interference prediction results on the interference situation. Through multiple experiments, a fuzzy logic target expectation value of 0.3 was selected. The weights for the fuzzy logic and interference results were allocated in a ratio of 0.85:0.15, with a threshold set at 0.3.



0-Normal; 1-Scratch; 2-Block; 3-Chap; 4-Scab

FIGURE 8. Confusion matrix for multi-feature multi-task.

This means that results weighted above 0.3 are categorized as ''With Interference,'' while those below 0.3 are categorized as ''Without Interference.'' The average interference results across different interference levels are presented in Figure 9.

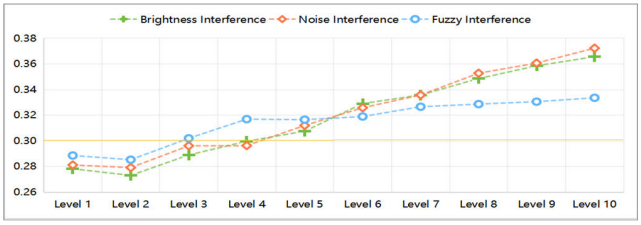


FIGURE 9. Results for different levels of interference.

Figure 9 illustrates how the interference intensity increases with increasing interference levels. With a threshold of 0.3, both brightness and noise interference were classified as ''interfered'' beyond Level 5. In contrast, fuzzy interference began showing signs of ''interference'' from level 3 onwards, and its intensity exhibited a relatively gradual incline with increasing interference levels. This indicates that the detection model maintains a more stable recognition performance when facing varying intensities of the fuzzy interference.

D. PERFORMANCE COMPARISON

1) SAME FEATURE EXTRACTION + DIFFERENT CLASSIFICATION MODELS

To further substantiate the effectiveness of the multi-task ensemble methodology proposed in this study, experiments were conducted whereby identical interference information was introduced into the image data. Subsequently, identical color, HU moments, and Haralick features were extracted, followed by feature fusion and normalization. The resulting multi-feature fusion feature vectors were then separately

TABLE 9. Results of different classification models.

Classification Models	Accuracy Without Interference	Accuracy With Interference		
		Luminance	Noise	Blurriness
Multi-feature Multi-task Fuzzy Method	98. 82%	86. 26%	91. 10%	96. 07%
SVM	97. 65%	70. 93%	72. 37%	74. 91%
Random Forest	94. 59%	82. 67%	88. 47%	92. 85%
SOTA	97. 07%	73. 30%	77. 17%	79. 51%

TABLE 10. Results of different detection methods.

Method	Accuracy without interference	Accuracy with interference		
		Luminance	Noise	Blurriness
Multi-feature multi-task Fuzzy Method	98. 82%	86. 26%	91. 10%	96. 07%
Laplacian low-pass filtering+SVM	99. 02%	97. 63%	83. 51%	91. 24%
Histogram Matching +Random Forest	99. 69%	88. 29%	90. 39%	96. 38%
SOTA	94. 55%	96. 63%	94. 45%	97. 40%

inputted into the network architecture of the proposed method, as well as into detection models such as SVM, Random Forest, and SOTA (State-Of-The-Art) neural networks based on the PyTorch framework, for assessment using accuracy in anomaly classification as the benchmark to evaluate model performance. The classification results are listed in Table 9.

Upon comparing the four detection models, it is evident that the proposed multi-feature multi-task fuzzy detection mode consistently exhibited the highest anomaly classification accuracy, regardless of the presence or absence of interference. In scenarios with luminance, noise, and blur interference, the classification accuracy achieved by this model surpassed those of the next-best detection models by 3.59%, 2.63%, and 3.22%, respectively. It is worth noting that SVM and SOTA outperformed Random Forest in scenarios without interference, but Random Forest performed better than SVM and SOTA when faced with luminance, noise, or blur interference. This suggests that the multi-feature multi-task fuzzy detection method proposed in this study demonstrated superior robustness, particularly in the presence of interference.

2) DIFFERENT FEATURE EXTRACTION + DIFFERENT CLASSIFICATION MODELS

To validate the performance of the multi-feature multi-task method proposed in this study, different combinations of feature extraction and classification models were utilized for comparison against the proposed approach. For instance, the Laplacian low-pass filtering-based SVM approach, the random forest method based on Histogram Matching, and the SOTA method. Different methods were applied to classify anomalies in rails on the same dataset, evaluating their classification performance under various interference types. The detection outcomes are presented in Table 10, obtained through K-fold cross-validation (K=5), thereby ensuring a robust assessment of the method's efficacy in detecting rail anomalies amidst diverse interferences.

The experimental findings reveal that these composite methods yield superior anomaly classification outcomes. Particularly, the SOTA method exhibits outstanding performance across each interference type, albeit registering notably lower accuracy in interference-free scenarios. This discrepancy primarily stems from the limited anomaly data volume in the absence of interference, totaling 427 instances compared to 12,810 instances in the presence of interference. The diminished accuracy of the SOTA method on smaller-scale datasets is attributed to the limited information available. Additionally, compared to the other three methods, the SOTA method exhibits prolonged detection times. Conversely, the other two composite methods demonstrate commendable performance under luminance and blurriness interferences. The proposed approach consistently delivers stable output results under both interference and interference-free conditions, achieving accuracy rates of 86.26%, 91.10%, and 96.07% across the three interference scenarios, respectively.

While conventional shallow learning models such as Support Vector Machines, Random Forests, and Laplacian low-pass filtering have demonstrated commendable performance to a certain extent, deep learning models, exemplified by knowledge-distillation-based recognition models, hold promise for surpassing them. Future investigations could delve into the applicability of deep learning models in railway anomaly detection tasks, thereby augmenting both detection accuracy and generalization capabilities.

E. RESULTS OUTPUT

To test how effective the multi-feature multi-task approach is in identifying different types of anomalies (such as normal, scratch, block, chap, and scab) under ''multi-interference'' conditions, 200 images were subjected to random additions of luminance, noise, and blurriness interferences. After that, different models were used to test the images, and the results of the anomaly classification outcomes of these different models are presented in Table 11.

Upon contrasting the anomaly classification outcomes of the four different methods, it becomes evident that varying detection methods yield different detection efficiencies for distinct anomaly types. The approach employed in this study demonstrates favorable detection results for the normal, scratch, and block anomaly types, whereas results for the chap and scab anomaly types are relatively lower. Upon analysis, it speculated that the three extracted features may not adequately capture other significant features specific to

TABLE 11. Results of multi-interferences.

Method	Normal	Scratch	Block	Chap	Scab	Totle
Multi-feature multi-task Fuzzy Method	92. 99%	100%	98. 08%	83. 39%	87. 38%	92. 74%
Laplacian low-pass filtering+ SVM	89. 02%	88. 87%	88. 71%	89. 69%	87. 73%	90. 08%
Histogram Matching+Random Forest	88. 89%	93. 98%	92. 13%	93. 52%	93. 06%	92. 31%
SOTA	93. 44%	96. 01%	94. 83%	95. 53%	92. 94%	94. 68%

scab and chap anomaly types, thus resulting in lower accuracy. In subsequent research endeavors, further exploration of the chap and scab anomaly types is warranted, focusing on identifying feature extraction methodologies better suited for these anomalies. Additionally, the integration of additional image processing techniques or deep learning methodologies may facilitate the extraction of richer and more discriminative features. Moreover, the exploration of alternative ensemble learning techniques holds promise for enhancing the model's generalization capabilities. Nonetheless, the Multi-feature multi-task Fuzzy Method employed in this study continues to exhibit satisfactory accuracy in detecting railway track anomalies, demonstrating notable robustness and generalization capabilities across various data subsets (K-fold cross-validation).

Figure 10 depicts the image input under sunlight interference (as shown in Fig. 1), along with the outputs of anomaly classification, interference level, and fuzzy logic tasks of the multi-feature multi-task method.

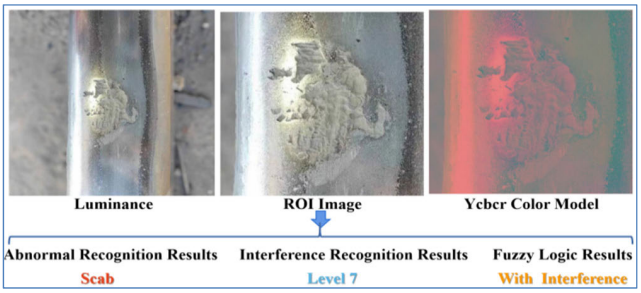


FIGURE 10. Image input and detection result output.

V. CONCLUSION

This study introduced a multi-feature multi-task fuzzy detection method, which integrates color, HU moments, and Haralick features. The method fuses these features to address the challenges posed by sunlight, dust, and blurriness interference in rail anomaly detection. By concurrently accomplishing the primary task of classifying rail anomalies and leveraging interference level information to assist network training, the proposed approach mitigates the impact of environmental interferences. Experimental results indicated that the multi-feature fusion method outperformed single-feature methods in both single-task and multi-task learning scenarios. Moreover, incorporating interference level information for network training enhanced the performance of rail anomaly detection by 5.93%, 4.93%, and 3.8% under luminance, noise, and fuzzy interferences, respectively. Compared to alternative detection methods, the multi-feature fusion multi-task blur detection method excelled in anomaly detection tasks while providing valuable perturbation information. This approach has introduced a novel perspective to the realm of rail anomaly detection research under interference conditions.

REFERENCES

[1] C. Zhang, D. Xu, L. Zhang, and W. Deng, "Rail surface defect detection based on image enhancement and improved YOLOX," *Electron.*, vol. 12, p. 2672, 2023.
[2] H. Yang, J. Liu, G. Mei, D. Yang, X. Deng, and C. Duan, "Research on real-time detection method of rail corrugation based on improved ShuffleNet V2," *Eng. Appl. Artif. Intell.*, vol. 126, Nov. 2023, Art. no. 106825.
[3] K. Wang, H. Li, Y. Qian, C. Xu, and Y. Yang, "An improved rail crack detection method based on the wavelet subband LMS adaptive filtering and wavelet analysis," in *15th Int. Conf. Graph. Image Process. (ICGIP 2023)*, vol. 2024, pp. 240–248.
[4] Q. Wang, T. Gao, Q. He, Y. Liu, J. Wu, and P. Wang, "Severe rail wear detection with rail running band images," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 38, no. 9, pp. 1162–1180, Jun. 2023.
[5] M. Karakose, O. Yamanand, K. Murat, and E. Akin, "A new approach for condition monitoring and detection of rail components and rail track in railway," *Int. J. Comput. Intell. Syst.*, vol. 11, pp. 830–845, Aug. 2018.
[6] M. Karakose, O. Yaman, M. Baygin, K. Murat, and E. Akin, "A new computer vision based method for rail track detection and fault diagnosis in railways," *Int. J. Mech. Eng. Robot. Res.*, vol. 61, pp. 22–27, Jun. 2017.
[7] C. Tastimur, M. Karaköse, E. Akin, and I. Aydin, "Rail defect detection with real time image processing technique," in *Proc. IEEE 14th Int. Conf. Ind. Informat. (INDIN)*, Jul. 2016, pp. 411–415.
[8] Y. Santur, M. Karaköse, and E. Akin, "Random forest based diagnosis approach for rail fault inspection in railways," in *Proc. Nat. Conf. Electr., Electron. Biomed. Eng. (ELECO)*, Dec. 2016, pp. 745–750.
[9] A. Origlia, S. Di Martino, and E. Battista, "Rail anomalies detection: A comparative analysis of three self-supervised models on real data," *Comput. Ind.*, vol. 148, Jun. 2023, Art. no. 103909.
[10] H. Lin and C. Wang, "DIGWO-N-BEATS: An evolutionary time series prediction method for situation prediction," *Inf. Sci.*, vol. 664, Apr. 2024, Art. no. 120316.
[11] H. Luo and G. G. Xu, "Rail surface defect detection based on image enhancement and deep learning," *J. Railway Sci. Eng.*, vol. 18, pp. 623–629, Sep. 2021.
[12] Z. D. He, Y. N. Wang, J. Liu, and F. Yin, "Image segmentation of high-speed rail surface defects based on background difference," *Chin. J. Sci. Instrum.*, vol. 37, pp. 640–649, Oct. 2016.
[13] Y. Z. Min, T. D. Cheng, and H. F. Ma, "Track defect identification method based on multi-feature fusion and AdaBoost algorithm," *J. Railway Sci. Eng.*, vol. 14, pp. 2554–2562, Mar. 2017.
[14] O. Zendel, M. Murschitz, M. Zeilinger, D. Steininger, S. Abbasi, and C. Beleznai, "RailSem19: A dataset for semantic rail scene understanding," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 1221–1229.
[15] S. E. M. Iem-Rm, "B-Scan ultrasonic image analysis for internal rail defect detection," in *Proc. Int. Conf. World Congr. Railway Res. (WCRR)*, vol. 2003, 2003, pp. 1–6.
[16] S. Faghih-Roohi, S. Hajizadeh, A. Núñez, R. Babuska, and B. De Schutter, "Deep convolutional neural networks for detection of rail surface defects," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 2584–2589.
[17] J. Gan, Q. Li, J. Wang, and H. Yu, "A hierarchical extractor-based visual rail surface inspection system," *IEEE Sensors J.*, vol. 17, no. 23, pp. 7935–7944, Dec. 2017.
[18] K. Song and Y. Yan, "A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects," *Appl. Surf. Sci.*, vol. 285, pp. 858–864, Nov. 2013.

[19] F. Z. Benabdallah and L. Djerou, "Active contour extension basing on Haralick texture features, multi-gene genetic programming, and block matching to segment thyroid in 3D ultrasound images," *Arabian J. Sci. Eng.*, vol. 48, no. 2, pp. 2429–2440, Feb. 2023.
[20] P. Anil, C. Jagrati, G. Sumit, K. Aditi, S. PARAM, S. Rachna, P. Lchetan, and R. Kumar, "A machine learning model for differentiating lytic lesion of LCH and Tuberculosis using Haralick texture features and k-nearest neighbour (kNN) as a classifier," *J. Nucl. Med.*, vol. 64, p. 592, 2023.
[21] J. Tarquino, S. Arabyarmohammadi, R. E. Tejada, A. Madabhushi, and E. Romero, "Intra-nucleus Mosaic pattern (InMop) and whole-cell Haralick combined-descriptor for identifying and characterizing acute leukemia blasts on single cell peripheral blood images," *Cytometry Part A*, vol. 103, no. 11, pp. 857–867, 2023.
[22] A. Pandey, J. Chaudhary, S. Garg, A. Khurana, P. Sharma, R. Seth, C. Patel, and R. Kumar, "A machine learning model for differentiating lytic lesion of LCH and Tuberculosis using Haralick texture features and k-nearest neighbour (kNN) as a classifier," *J. Nucl. Med.*, vol. 64, p. 592, 2023.
[23] G. Osman and M. S. Hitam, "Skin colour classification using linear discriminant analysis and colour mapping co-occurrence matrix," in *Proc. Int. Conf. Comput. Appl. Technol. (ICCAT)*, Jan. 2013, pp. 1–5.
[24] R. Nagi and S. S. Tripathy, "Plant disease identification using fuzzy feature extraction and PNN," *Signal, Image Video Process.*, vol. 17, no. 6, pp. 2809–2815, Sep. 2023.
[25] K. Tian, "Multi model fusion of binary information mining on impact velocity prediction of abnormal projectile," *J. Ballistics/Dandao Xuebao*, vol. 35, no. 2, pp. 126–132, 2023.
[26] V. L. Narla and G. Suresh, "Multiple feature-based tomato plant leaf disease classification using SVM classifier," in *Proc. 3rd Int. Conf. MIND*, 2021, pp. 443–455.
[27] Y. Jusman, M. K. Anam, S. Puspita, and E. Saleh, "Machine learnings of dental caries images based on Hu moment invariants features," in *Proc. Int. Seminar Appl. Technol. Inf. Commun. (iSemantic)*. IEEE, 2021, pp. 296–299.
[28] X. Li, H. Lin, J. Du, and Y. Yang, "Computer vision-based driver fatigue detection framework with personalization threshold and multi-feature fusion," *Signal, Image Video Process.*, vol. 18, no. 1, pp. 505–514, Feb. 2024.
[29] A. Mao, M. Mohri, and Y. Zhong, "Cross-entropy loss functions: Theoretical analysis and applications," 2023, *arXiv:2304.07288*.
[30] K. Shen, J. Guo, X. Tan, S. Tang, R. Wang, and J. Bian, "A study on ReLU and softmax in transformer," 2023, *arXiv:2302.06461*.

YANG LIYUAN was born in Yunnan, China, in 1985. She received the master's degree in transportation planning and management from Central South University, Changsha, Hunan, China, in 2010. She is currently pursuing the Ph.D. degree in information management with Universiti Teknologi MARA (UiTM), Kelantan Branch, Malaysia.

She is also a Lecturer with Kunming University, Kunming, Yunnan. Her research interests include intelligent transportation, image processing, pattern recognition, and deep learning.

GHAZALI OSMAN received the B.Sc. degree from Universiti Kebangsaan Malaysia, the Postgraduate Diploma degree in computer science from Universiti Teknologi Malaysia, the M.Sc. degree in information management from Universiti Teknologi MARA, and the Ph.D. degree in artificial intelligence from Universiti Malaysia Terengganu. He is currently an Associate Professor with the School of Information Science, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Kelantan Branch, Malaysia. He is also an Associate Member with the Chartered Tax Institute of Malaysia. His research interests include artificial intelligence, image processing, pattern recognition, and information system management. He is a Professional Technologist of the Malaysian Board of Technologists (MBOT).

SAFAWI ABDUL RAHMAN received the master's degree in information technology from Universiti Kebangsaan Malaysia and the Ph.D. degree in information technology from Universiti Teknologi MARA. He is currently an Associate Professor with the School of Information Science, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Puncak Perdana Campus, Selangor, Malaysia. His research interests include artificial intelligence behaviors, knowledge management, and information systems. He is a Professional Technologist of the Malaysian Board of Technologists (MBOT).

MUHAMMAD FIRDAUS MUSTAPHA received the B.Sc. and M.Sc. degrees in computer science from Universiti Teknologi Malaysia (UTM), in 2006 and 2009, respectively, and the Doctor of Philosophy (Ph.D.) degree in computer science from Universiti Teknologi MARA (UiTM), Shah Alam, Malaysia, in 2018. He is currently a Senior Lecturer with UiTM, Kelantan Branch, Malaysia. His current research interests include artificial intelligence, computer vision, machine learning, and deep learning.