MAUN Fen Bil. Dergi., 13, 2, 351-356 Araştırma Makalesi/ Research Article
MAUN J. of Sci., 13, 2, 351-356 DOI : 10.18586/msufbd.1763332

Efficient and Real-Time Railway Track Fault Classification Using CNN Integrated with Convolutional Block Attention Module

Canan TAŞTİMUR
[1] Bilgisayar Mühendisliği Bölümü, Mühendislik-Mimarlık Fakültesi, Erzincan Binali Yıldırım Üniversitesi, Erzincan, Türkiye

: ctastimur@erzincan.edu.tr 0000-0002-3714-6826

Geliş (Received): 12.08.2025 Düzeltme (Revision): 25.08.2025 Kabul (Accepted): 01.10.2025

ABSTRACT

The occurrence of various types of faults on railway rail surfaces can lead to accidents such as train derailments. In this study, a new approach that produces very fast and robust results is developed by combining the convolutional neural network and the convolutional block attention module for the classification of rail faults. The dataset used in this study contains four different fault types, and experimental studies were conducted on the public rail dataset. The impact of the convolutional block attention module on the performance of the proposed approach and its contribution to the model's generalization ability are examined, and the performance of the proposed approach increases by approximately 5% compared to the performance of the proposed approach without this module. It has been demonstrated that the proposed approach can be used effectively in railway track fault diagnosis by producing fast and effective results.

Keywords: Attention mechanism, Convolutional block attention module, Rail defect, Rail track inspection

Konvolüsyonel Blok Dikkat Modülü ile Entegre CNN Kullanılarak Verimli ve Gerçek Zamanlı Demiryolu Ray Arıza Sınıflandırması

ÖZ

Demiryolu ray yüzeylerinde çeşitli tipte arızaların oluşması tren raydan çıkması gibi kazalara yol açabilir. Bu çalışmada, ray arızalarının sınıflandırılması için evrişimsel sinir ağı ve evrişimsel blok dikkat modülünün birleştirilmesiyle çok hızlı ve sağlam sonuçlar üreten yeni bir yaklaşım geliştirilmiştir. Bu çalışmada kullanılan veri seti dört farklı arıza tipini içermekte olup, kamuya ait demiryolu veri seti üzerinde deneysel çalışmalar yürütülmüştür. Evrişimsel blok dikkat modülünün önerilen yaklaşımın performansına etkisi ve modelin genelleme yeteneğine katkısı incelenmiş ve önerilen yaklaşımın performansı, bu modül olmadan önerilen yaklaşımın performansına kıyasla yaklaşık %5 oranında artmaktadır. Önerilen yaklaşımın hızlı ve etkili sonuçlar üreterek demiryolu hat arıza teşhisinde etkin bir şekilde kullanılabileceği gösterilmiştir.

Anahtar Kelimeler: Evrişimli blok dikkat modülü, Dikkat mekanizması, Ray arızası, Ray yolu denetimi

INTRODUCTION

Rail transport is a significant transportation alternative today, thanks to its safety, low cost, and environmental friendliness. With the increasing need for freight and passenger transportation, the importance of rail infrastructure has become even more evident. Ensuring the safety and continuity of rail systems depends on regular and thorough maintenance of rail lines [1,2]. Failure to detect rail faults early can lead to accidents that result in loss of life and property. Traditional manual fault diagnosis techniques are time-consuming, costly, and prone to personnel errors, as well as personnel requirements. Manual inspections can miss small, rapidly occurring faults [3]. Therefore, interest in artificial intelligence-assisted fault diagnosis methods to eliminate accident risk factors has increased significantly [4-6]. Deep learning-based approaches are quite popular in image classification and fault diagnosis. Convolutional Neural Network (CNN)-based architectures have demonstrated high success in classifying railway track faults [7, 8]. Thanks to these architectures, early warning systems can respond not only to existing faults but also to potential faults [9, 10]. CNN-based models learn complex patterns and can analyze large amounts of data. The Convolutional Block Attention Module (CBAM), with its attention mechanism, allows the model to make more robust detections by focusing more attention on important regions in the image [11]. There are many deep

learning-based studies in the literature on the classification of rail faults to ensure railway rail safety.
Zhang et al. [1] used VGG19 and ResNet50 on a dataset of 5000 rail images to classify rail faults. They achieved 96% and 97% accuracy. In this study, there was an overfitting problem due to the imbalance between the classes. In another study [2], MobileNet, a transfer learning approach, was used to classify 10,000 rail images. 94% accuracy was achieved. They also used both visual data and IoT sensor data in their study. Faghih-Roohi et al. [3] presented a time-series signal analysis approach with CNN and LSTM for rail fault detection. Although the signals from a high-speed train were presented in the dataset, they continued the approach solely based on images. A study was conducted with ResNet architecture for the analysis of cracked rail faults [4].
Gao et al. [12] used only CNN architecture to classify faults in rail images. They used a dataset containing 12,000 rail images. They achieved 93.2% accuracy. Lee et al. [13] used AlexNet for a rail fault dataset with 4 classes. They frozen all layers and set only the last layer to the number of classes and achieved 88% accuracy. In [14], they made binary classification as normal and cracked on a dataset of 8,000 rail crack faults using ResNet50 architecture. They achieved 95.8% accuracy, and processing with a single fault type without considering other fault types can be considered as a shortcoming. Park et al. [15] classified a dataset containing 10,000 rail images with CNN architecture. They achieved 91.7% accuracy and tested it in real-time with a Raspberry Pi. Rao et al. [16] They classified four types of faults using the VGG16 architecture. 3000 ray images were found in the training dataset and achieved an accurate rate of 89%.
Li et al. [17] added CBAM to the Inception-ResNet architecture to classify different types of rail faults with 94.5% accuracy. The dataset used consisted of 32,000 images with four classes. Wang et al. [18] performed both rail fault classification and segmentation. EfficientNet and CBAM modules were integrated, achieving a classification accuracy of 93.7%. CBAM increased the ability to detect vulnerable rail faults. In another study [19], rail faults were classified in a similar manner by adding CBMA to the VGG16 architecture. They achieved 91.4% accuracy on an 800-image dataset. Park et al. [20] integrated CBAM with CNN to classify early-stage rail crack faults. 97.1% accuracy was achieved, and it was stated that CBAM increased the ability to detect non-classic areas in rail faults.

Table 1. Literature comparison for rail defect classification

Ref.	Method	Accuracy (%)	Dataset size
[1]	VGG19, ResNet50	96.00 / 97.00	5,000
[2]	MobileNet + IoT	94.00	10,000
[12]	CNN	93.20	12,000
[14]	ResNet50	95.80	8,000
[15]	CNN	91.70	10,000
[16]	VGG16	89.00	3,000
[17]	Inception-ResNet + CBAM	94.50	32,000
[18]	EfficientNet + CBAM	93.70	1,838
[19]	VGG16 + CBAM	91.40	800
[20]	CNN + CBAM	97.10	3,000

THE PROPOSED APPROACH

Timely and accurate detection of rail faults is crucial for ensuring railway transportation safety. Rail is one of the most crucial components that require attention to prevent disruptions to rail transportation. In this study classifying rail surface faults, the CBAM module was added to the pure CNN architecture. This improved the model's attention mechanism and improved feature extraction. It also produces highly successful results in rail fault cases where fine details can be detected. Detailed information about the proposed method's architecture is provided in Fig. 1.
The proposed model employs three consecutive convolution layers with a filter size of 3x3. These layers are 32, 64, and 128, respectively, and dimensionality reduction is achieved by applying max pooling after each of these convolution layers. The CBAM module is implemented after the third convolution layer, ensuring attention is focused on the channel and spatial regions. The channel attention mechanism uses global and max pooling, and the attention score is generated using dense layers. The spatial attention mechanism is achieved using convolution layers. This architecture is continued with global average pooling and a dense layer with 128 neurons in the final layer to improve classification performance.

Convolutional Block Attention Module

Classical CNN model demonstrates superior performance in object classification and fault diagnosis. However, pure CNN models still cannot learn contextual information in images. Furthermore, because CNN assign equal importance to each feature in an image, irrelevant features can sometimes be trained with the same model, resulting in increased algorithmic complexity. To address these shortcomings, CBAM has increased the representational power of CNNs by applying both spatial and channel-level attention mechanisms [21, 22]. One of the key advantages of CBAM is its lightweight design. It introduces minimal computational overhead while significantly improving the performance of backbone CNN models. This makes it suitable for real-time fault detection systems in industrial applications. CBAM focuses on where and what important information is located in feature maps [23]. This allows the model to prioritize important information and achieve high accuracy when detecting anomalies in complex images. The channel attention mechanism examines the importance of channels in a feature map. This attention is fed by avg-pooling and max-pooling in a Multilayer Perceptron (MLP) network. The spatial attention mechanism, on the other hand, evaluates which regions are more meaningful. This double attention is applied sequentially, thus revealing which information and areas within the feature map are important. Moreover, CBAM's attention maps can be visualized to interpret the model's decision-making process, offering transparency and explainability — a desirable feature in safety-critical domains like railway fault diagnostics. The mathematical expression of the channel attention mechanism is in (1) and the spatial attention mechanism is in (2).

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (1)$$

$$M_s(F) = \sigma(f^{7x7}([AvgPool(F)\,; MaxPool(F)])) \quad (2)$$

Here, σ is the sigmoid activation function. F and f represents the input feature map and kernel filter, respectively.

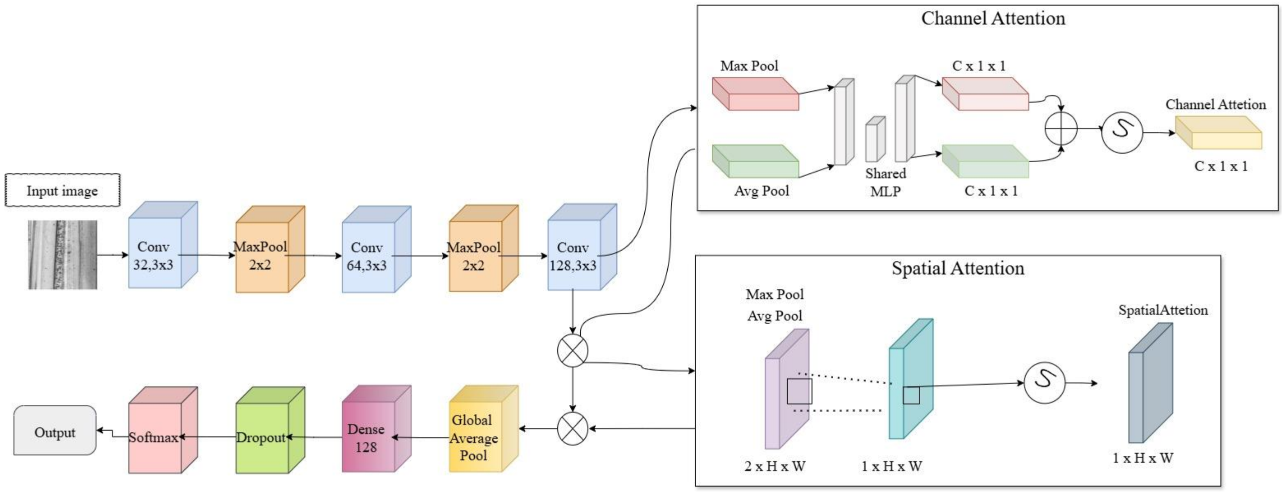

Figure 1. The overview of the proposed method.

EXPERIMENTAL RESULTS

In this study, the public dataset presented in [24] is used for the classification of railway rail faults. This dataset consists of 1838 faulty rail images with four classes: broken rail, gap, and loose joint. These images were collected from the field and are high-resolution. The dataset was divided into two parts, 70% training and 30% testing, for the purpose of training and evaluation of the model. Sample images from the dataset are shown in Figure 2. The model's performance has evaluated using accuracy, precision, recall, and f1score metrics. Furthermore, AUC values specific to each class are calculated using ROC curves and presented visually. Parameters such as the number of epochs, batch size, learning rate and optimizer used in training the model are given in detail in Table 2.

Table 2. Hyperparameter value

Optimizer	Learning rate	Epochs	Batch size
Adam	0.001	50	64

By examining the visual features and classification performance results in the images, it is concluded that the model can clearly distinguish each class using UMAP and t-SNE maps. The confusion matrix is used to evaluate the confusion between classes, and the overall accuracy level has visualized. Furthermore, the performance of the model with and without CBAM has been evaluated separately.

Healthy Joint Squats Another squat

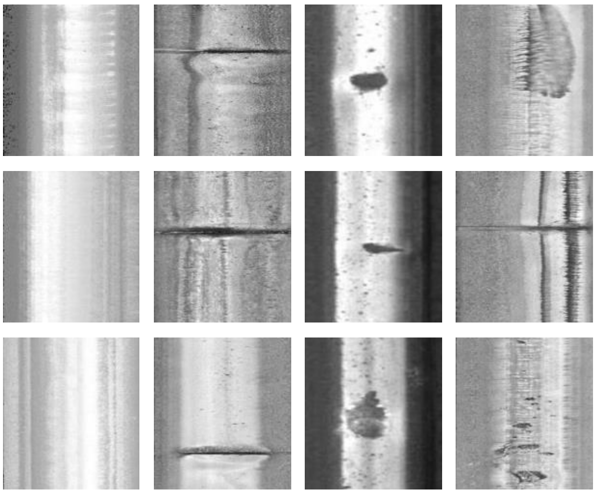

Figure 2. Sample input images selected from the publicly.

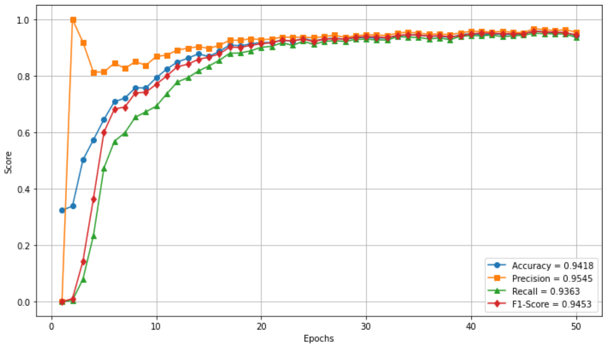

Figure 3. The overall performance of the proposed method without CBAM.

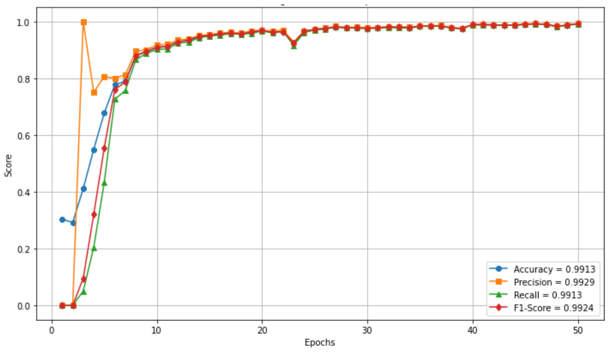

Figure 4. The overall performance of the proposed method with CBAM.

Figures 3 and 4 show the effect of the CBAM module on the proposed method. The accuracy, recall, precision, and f1 score showed a significant increase with CBAM.

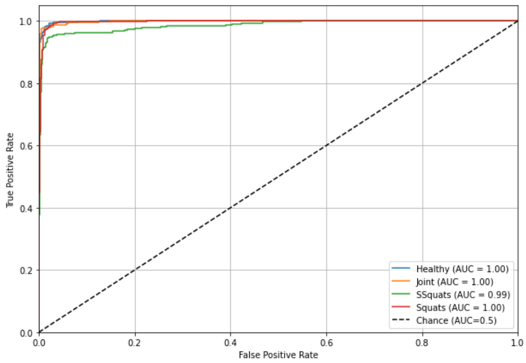

Figure 5. The ROC curve of the proposed method without CBAM.

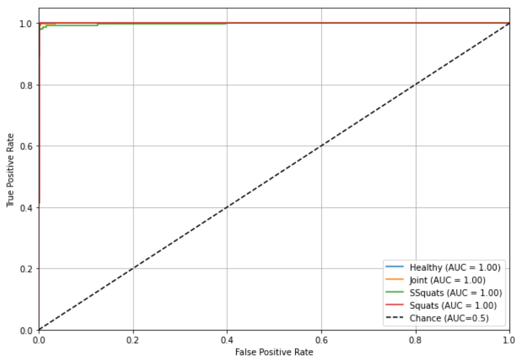

Figure 6. The ROC curve of the proposed method with CBAM.

When comparing the ROC curves in Fig. 5 and 6, there is an increase in the curve area with CBAM.

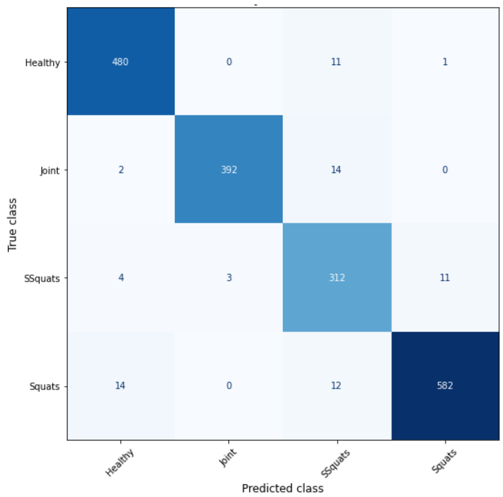

Figure 7. The confusion matrix of the proposed method without CBAM.

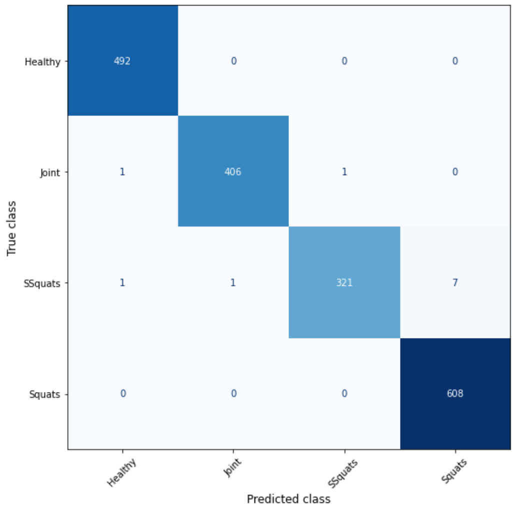

Figure 8. The confusion matrix of the proposed method with CBAM.

Figures 7 and 8 show the effect of the CBAM module on the confusion matrix. While there were 72 misclassifications in Fig. 7 without CBAM, this number decreased to 11 in Fig. 8 with CBAM.

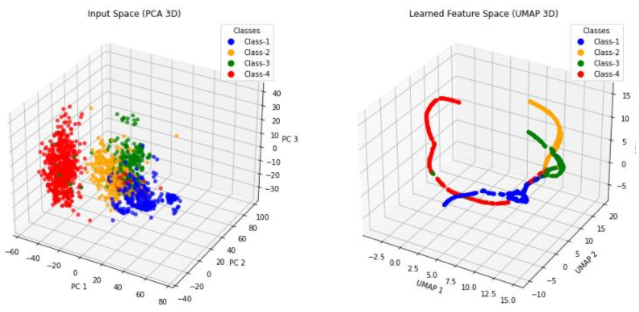

Figure 9. The UMAP visualization of the proposed method without CBAM.

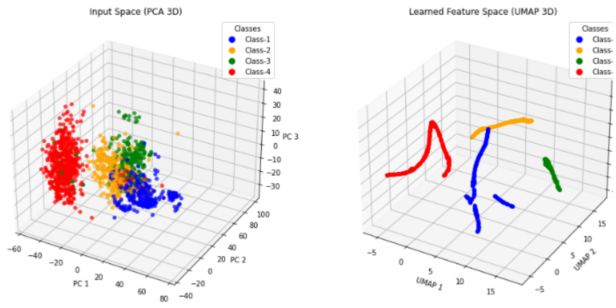

Figure 10. The UMAP visualization of the proposed method with CBAM.

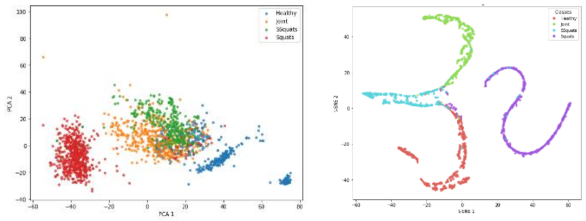

Figure 11. The t-SNE visualization of the proposed method without CBAM.

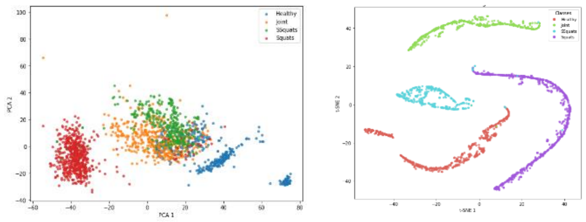

Figure 12. The t-SNE visualization of the proposed method with CBAM.

Figures 9 and 10 represent how classes are separated in the feature space with UMAP. Figures 11 and 12 illustrate this situation with t-SNE. Figures 9 and 11 show that the model without CBAM does not separate the samples sufficiently, while Fig. 10 and 12 show that the samples are sufficiently separated from each other with CBAM.

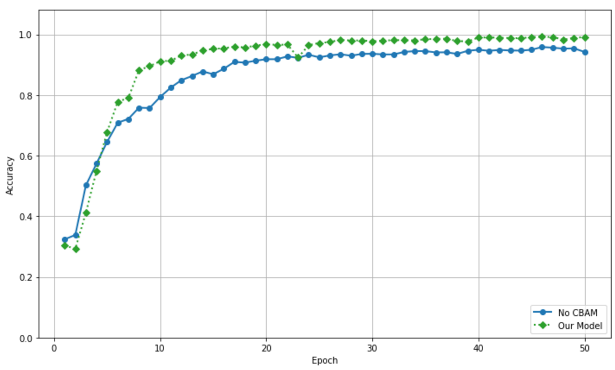

Figure 13. Comparison of classification accuracy between the proposed method with and without the CBAM module.

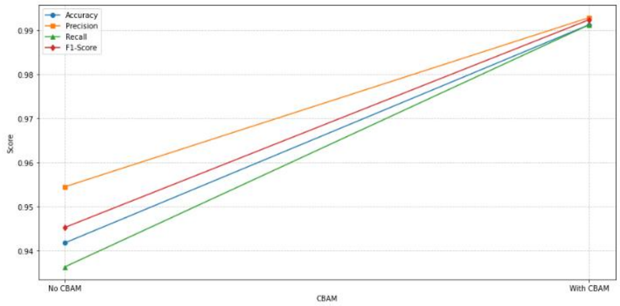

Figure 14. Comparison of overall performance between the proposed method with and without the CBAM module.

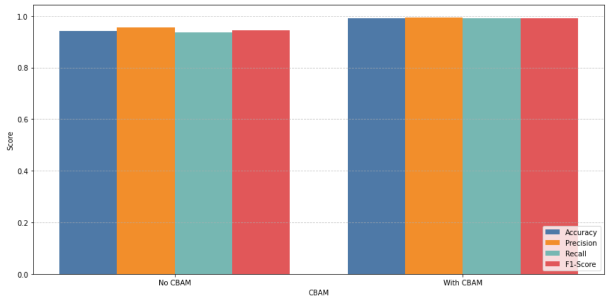

Figure 15. Bar chart comparison of the overall performance of the proposed method with and without the CBAM module.

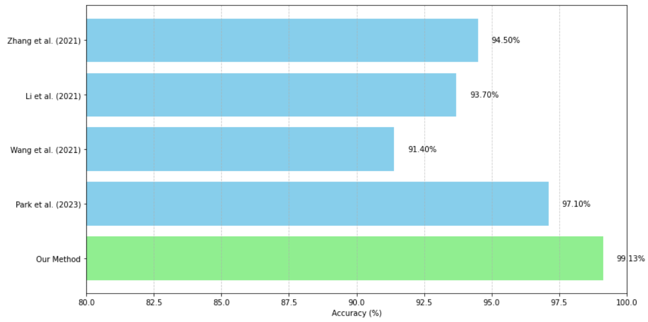

Figure 16. Accuracy comparison of the proposed method with existing state-of-the-art approaches.

Figure 13 clearly demonstrates that the CBAM module significantly increases classification accuracy (from 94.18% to 99.13%). Fig. 14 shows the comparative analysis of accuracy, F1, precision, and recall metrics, demonstrating the positive impact of the CBAM module on all these metrics. This comparison is also illustrated with a bar chart in Fig. 15. Fig. 16 shows that the proposed method, when supported by CBAM, outperforms other current methods in the literature. This demonstrates the model's competitiveness and that the CBAM integration provides a significant contribution.

CONCLUSIONS

In this study, a deep learning-based approach classifies rail faults to ensure the reliability of railway fault detection. The proposed model supports the CNN architecture with the CBMA module. The CBAM module makes significant fault regions in the images more meaningful, allowing the network to focus more attention on these regions. Experimental studies have shown that the CBAM module significantly improves performance. When the images presented for the separation of rail faults in the dataset are examined, it is determined that the CBAM module provides much clearer separation between the examples. It was demonstrated that the proposed model with CBAM increased overall success by 5% compared to the 99.13% accuracy of the version without CBAM (94.18%) and reduced the number of misclassifications from 72 to 11 examples. This study has made a significant contribution to railway rail fault diagnosis and stands out as a valuable solution for industrial applications. It appears to offer an effective and rapid approach to classifying railway track faults.

REFERENCES

[1] Y. Zhang, X. Wang, and Q. Liu, “Rail track fault detection using deep convolutional neural networks,” IEEE Access, vol. 8, pp. 130461–130470, 2020, doi: 10.1109/ACCESS.2020.3009424.

[2] L. Chen, H. Zhang, and P. Li, “Visual inspection in railway systems: A deep learning approach,” Sensors, vol. 19, no. 2, pp. 1–18, 2019, doi: 10.3390/s19020359.

[3] M. Faghih-Roohi, K. Cheng, R. Newman, et al., “Rail defect detection using ultrasonic technology and signal processing: A review,” Insight - Non-Destructive Testing and Condition Monitoring, vol. 56, no. 1, pp. 20–28, 2014.

[4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[5] C. Li, J. Zhao, Y. Liu, X. Li, and X. Han, "Rail surface defect detection based on deep learning," IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1–14, 2021, doi: 10.1109/TIM.2021.3066373.

[6] W. Li, M. Chen, J. Liu, and Y. Xu, “Automated rail defect classification using attention-based deep networks,” Transp. Res. Part C Emerg. Technol., vol. 127, pp. 103115, 2021, doi: 10.1016/j.trc.2021.103115.

[7] S. A. Yadav, R. Tripathi, and M. Shukla, “Railway Track Fault Detection and Classification Using Hybrid Deep LearningTechniques,” IEEE Access, vol. 9, pp. 145892–145905, 2021.

[8] M. Z. Afzal, H. T. H. Nguyen, and Y. Kim, “Vision-based railway track monitoring using deep neural networks,” IEEE Sensors Journal, vol. 20, no. 23, pp. 14162–14170, 2020.

[9] F. J. Perez-Peña et al., “Real-time rail inspection system using computer vision,” Transportation Research Part C: Emerging Technologies, vol. 103, pp. 146–158, 2019.

[10] H. Tang, Z. Li, B. Wang, et al., “Deep learning-based surface defect detection for high-speed railway track inspection,” Neurocomputing, vol. 417, pp. 307–319, 2020.

[11] S. Woo, J. Park, J. Lee, and I. Kweon, “CBAM: Convolutional block attention module,” in Proc. Eur. Conf. Comput. Vis. (ECCV), 2018, pp. 3–19.

[12] Gao, B., Zhang, C., & Li, Y. (2020). "DeepRail: Railway defect classification using convolutional neural networks." Journal of Rail Transport Planning & Management, 16, 100221.

[13] Patel, V., Kumar, A., & Sharma, R. (2021). "Residual networks for track surface fault detection in Indian railways." Transportation Research Part C, 127, 102908.

[14] Lee, S., Choi, J., & Kim, D. (2019). "Rail defect classification using transfer learning on AlexNet." Sensors, 19(21), 4751.

[15] Park, D., Yoon, S., & Lee, B. (2022). "Real-time lightweight CNN for railway surface defect classification." IEEE Sensors Journal, 22(5), 4102–4110.

[16] Rao, H., & Singh, M. (2018). "Multi-class defect detection on railway track using VGGNet and augmentation." Procedia Computer Science, 132, 785–792.

[17] Zhang, X., Liu, Y., Wang, H. (2021). “Deep learning-based railway surface defect classification with attention mechanisms,” IEEE Access, vol. 9, pp. 125441–125452.

[18] Li, J., Chen, W., & Wu, Y. (2022). “Multiscale attention network for rail surface defect detection,” Sensors, 22(4), 1547.

[19] Wang, M., Lu, Q., & Liu, H. (2021). “EfficientNet and CBAM for railway track damage classification,” Transportation Research Part C, 132, 103391.

[20] Park, D., Choi, H., Kim, Y. (2023). “Rail track anomaly detection via GAF-based image analysis and CBAM-enhanced CNNs,” Engineering Applications of Artificial Intelligence, 123, 105558.

[21] Song, K., & Yan, Y. (2021). “A multi-attention deep learning model for rail surface defect classification,” Measurement, 178, 109376.

[22] Zhou, X., Wang, Q., & Zhang, Y. (2022). “Lightweight attention networks for surface defect classification in rail inspection,” Neural Computing and Applications, 34(7), 5617–5632.

[23] Bakır, F., Aksoy, F., & Demir, C. (2023). “Deep Attention Mechanisms for Track Surface Defect Detection,” Expert Systems with Applications, 208, 118202.

[24] Aydin, I., Akin, E., & Karakose, M. (2021). Defect classification based on deep features for railway tracks in sustainable transportation. Applied Soft Computing, 107706.