Fastener and rail surface defects detection with deep learning techniques

Merve Yilmazer [a,1,*], Mehmet Karakose [b,2]

[a] Computer Engineering Department, Munzur University, Tunceli, Türkiye
[b] Computer Engineering Department, Firat University, Elazig, Türkiye
[1] merveyilmazer@munzur.edu.tr; [2] mkarakose@firat.edu.tr
* corresponding author

ARTICLE INFO

Article history
Received August 6, 2023
Revised March 13, 2024
Accepted March 21, 2024
Available online May 31, 2024

Keywords
Defect detection
Deep learning
Autonomous drone
Faster R-CNN
ResNet101v2

ABSTRACT

The railways, which are frequently used by countries for both passenger and freight transportation, should be checked periodically. Controls made with classical methods are slow and there are often overlooked faults. This work suggests a novel deep learning-based technique for identifying fastener and railway track surface defects. Within the scope of the proposed method, firstly, the railroad track was observed using an autonomous drone. Shaky images in the recorded video were removed with a video stabilization algorithm. Frames were created and labeled from the video, and rail and fastener regions were detected using the Faster R-CNN deep neural network. Fault detection was performed through ResNet101v2-based classification using different datasets for identifying surface detects in rails and different datasets for the detection of fasteners. The proposed method was experimentally shown to have a 98% accuracy rate for detecting rail surface flaws and a 95% accuracy rate for detecting fastener flaws. A user interface was developed to display the identified faulty images on computers, tablets, and mobile phones via a mobile application. The system, which was previously proposed in a different study, was retrained by going through the video stabilization step, thus improving the fault detection rate, and the method was also included in the user interface module. This study contributes to the processing of ever-increasing video data with deep learning-based methods. It is also anticipated that it will benefit researchers working in the field of railway non-contact fault detection.

1. Introduction

Railway transportation, which contributes significantly to economic and industrial development, is widely used by countries. Although it is considered the most reliable transportation system, inadequate infrastructure maintenance causes railway defects. The railway consists of many parts including fasteners, ballast, sleepers, and rails, failure to fix the faults in a timely manner can cause accidents that can lead to great loss of life and property. There are various types of defects such as rail surface defects [1], fastener defects [2], switch crossing and level crossing point defects [3]. These faults should be detected and repaired before they reach a level that could cause accidents. Railway tracks are the components of the railway where abrasions and breaks occur due to friction and environmental factors on which the train wheel moves. The fasteners, which are the components of the railway line, are used to connect the rails to each other and also to connect the rails and their sleepers to each other. If they are broken or missing, malfunctions occur that threaten rail safety.

Most of the non-contact methods proposed for the detection of faults use rail visual data. Non-contact fault detection is performed using image processing or deep learning algorithms on images obtained with various techniques and various cameras [4]–[6]. Some of the methods proposed in the literature for contactless railway fault detection using image data are given below.

A YOLOv5 deep neural network-based approach was presented by Luo *et al.* [7] to identify rail surface flaws. Before model training, they enriched the data sets with data augmentation techniques such as random cropping and rotation. They reduced regression errors and accelerated algorithmic convergence by using Soft-SIoUNMS loss. As a result, they showed that their proposed YOLOv5-based model detected rail surface defects with 96.9% average precision. Using UAV photos, Singh *et al.* [8] trained the YOLOv4 algorithm to recognize rail sleepers. They demonstrated that the 99.08% average precision (mAP) of their suggested method for detecting railroad sleepers.

Cheng *et al.* [4] aimed to improve model training success by improving the quality of images used for detecting railway defects. They proposed HybridGAN by combining DeblurGANv2 and ESRGAN to improve image quality. They trained the YOLOv4 network with and without HybridGAN applied forms of the dataset. They experimentally demonstrated that there is an improvement in mAP values in the dataset with HybridGAN. Qiu *et al.* [9] proposed YOLOv8-FAM based synthetic data generation and Automated detection system to detect railway faulty fasteners. Thus, they showed that the cost of collecting faulty fastener data was reduced by 40% and that they were able to detect faults better than other methods. Liu *et al.* [10] proposed a DCNN-based method to detect railway faulty fasteners. They used SSD and RCNN models in the two-stage method. They determined the fastener regions on the images. They were able to detect faults with an average precision of 95.38% in the specified areas. In addition, new approaches have been proposed for the detection of fastener-related defects in CNN and ResNet50 [11], U-net, ResNer50, Fully Convolutional Network (FCN) [12], YOLOv5 [13] studies.

Railroad rails are the component on which the train moves. It is exposed to heat, light and friction. For this reason, it is a component where defects often occur. Chen *et al.* [14] used both camera images and ultrasound B-scan images to identify rails in their proposed work. In the method, they first designed a segmentation algorithm including filtering and rail surfacing using edge detection. Then, they used BoTNet 50 network to extract features from the five-class data set. As a result, they were able to identify rail defects with 96.97% accuracy with their proposed Camera and Ultrasound Data Fusion (CUFuse) model.

Li *et al.* [15] introduced a new approach for fast and efficient detection of rail fasteners based on the YOLOv5. They also used the K-Means++ clustering algorithm to enhance the positioning competence of the method. Test results showed that their proposed method gave accurate results with an average sensitivity (mAP) rate of 97.4%. In smart fault detection systems, which are developed as an alternative to traditional railway inspection systems, fault detection can be made by using different types of data other than image data. Shafique *et al.* [16] used acoustic signals for error detection in their proposed study. They collected data using acoustic signals from the Pakistani railway line. They divided these data into 3 categories and classified them with various classification techniques such as CNN, SVM, random forest, logistic regression and decision tree. They showed that the best results were acquired by random forest (RF), decision tree (DT) with 97% accuracy. Aydin *et al.* [17] suggested a hybrid method based on YOLOv4-Tiny and CNN. In their proposed method, they located the railway fasteners with YOLOv4-Tiny. Then, they classified the fasteners with CNN. They experimentally demonstrated that their recommended method can detect fastener defects with 98.57% accuracy. Zheng *et al.* [18] introduced a multi-object identification approach based on DCNN to detect detection fastener and rail surface problems. Researchers using YOLOv5 to localize fasteners and rails employed Mask R-CNN to find rail surface flaws and a ResNet based method to detect rail fastener faults. They classified fasteners using three classes as normal, loose and broken.

In a proposed method based on machine vision, Zhou [19] used rail images obtained with a CCD camera to detect surface defects in rail. In order to reduce the effect of ambient brightness on the images, they converted the images from RGB format to HSI format. Undesirable noise on the image is removed by Gaussian filtering. Wu *et al.* [20] proposed RGBNet architecture, which can define thin edge features in order to find wear on the rail surface. They have shown that the model they recommend can accurately detect over 90% precision and recall values. Zheng *et al.* [21] Aiming to detect errors that occur on the sleepers using the YOLOv3 algorithm and K-Means clustering algorithms, the researchers trained the YOLOv3 deep neural network in 15000 iterations with 64 batch-size and 0.001 learning rate. They showed that they could detect faults with an average precision (mAP) of 86.29% and an accuracy of 91.72%. Hu *et al.* [22] proposed a new method named YoLoX-Nano railway defect detection. They added CSPDarknet's Path Fusion Feature Pyramid Network and three output feature maps to the method for the feature extraction stage. They determined the mAP value of the proposed system as 98.07%. Su *et al.* [23] first used the K-means algorithm to determine the dimensions of the fasteners in their recommended method to determine the fastener defects. Then, they showed in their studies that they could detect errors with a 96.1% accuracy rate by training the YOLOv5.

There are many studies, apart from the methods mentioned above, on the determination of the defects that occur in the railway, which consists of a combination of many components as a result of processing the image data. Objects (obstacles) on the railway were detected with 85.2% accuracy with the 2D Singular Spectrum Analysis (SSA) parsing tool and Faster R-CNN-based method [24]. A fastener detection system based on an SVM classifier has been developed [25]. A U-Net-based railway sleeper defect detection system with an f1-score of 86.5% has been proposed [26]. ResNet50-based rail surface defect detection was performed from railway images obtained with 3D Laser Cameras [27]. A fastener detection method with a precision of 97% was proposed with a two-stage Mask R-CNN called FishTwoMask R-CNN [28]. They proposed a YOLOv4-based railway rail surface error detection method [29].

In this study, rail fastener defects and rail surface defects were determined. The data used in the study was created with an autonomous drone. Thus, a data set was created in a short time with less human effort. By viewing the railway as a bird's eye view, more reliable data were obtained for fault detection. The quality of the images has been increased with video stabilization. In addition, an interface has been developed that will enable the detected faults to be sent to the railway fault control unit with a mobile application. In the other parts of the study, the details of the proposed method, method outputs, and performance evaluation are given. In addition, a general evaluation of the method was made by comparing the other methods in the literature with the proposed approach in this study. Advantageous aspects of the proposed method are given; Using autonomous drones for data acquisition, Improvement of data with video stabilization, Fault detection with the combination of object detection, Classification and segmentation algorithms in deep learning, Developing a user interface to work on the mobile application in order to facilitate the visibility of detected faults.

2. Method

Security, monitoring, and fault detection are now carried out using Video Surveillance Systems (VSS). Monitoring video data by expert personnel to detect abnormal situations is a method that requires a long time and has a high error rate. For this reason, artificial intelligence-based video and image processing methods such as deep learning have recently been recommended to extract meaningful information from video and image data and detect abnormal/faulty situations. In the study, a deep learning-based method is proposed to detect defects occurring on the railway line. In the proposed method, firstly, the video obtained from the railway line monitored by the autonomous drone was transferred to the computer, stabilized in order to eliminate the vibrational, and then frames were created. A faster R-CNN object detection algorithm was trained in order to determine the rail and fastener parts on the images. A 3-class data set was created in order to detect defects from the determined fastener parts. For the determination of rail surface defects, 3 classes of the data set used in Ref [30] were used. In addition, the method proposed in Ref [31] was included in the user interface module, which is the last stage of the retrained model, by going through the video stabilization stage within the scope of this study. Python OpenCV library is used for video stabilization. Stabilized video was obtained by calculating the optical flow in the video given as input for stabilization. As seen in the block diagram in Fig. 1, the data sets consisting of 3 classes for the rail surface defect and 3 classes for the fastener were classified using the ResNet101V2 deep transfer model and fault detection was made. The steps followed in the recommended method are presented in the form of an artificial intelligence-based video processing library to be used in future studies. Algorithm 1 (Fig. 2) was applied to obtain the artificial intelligence-based Python library.

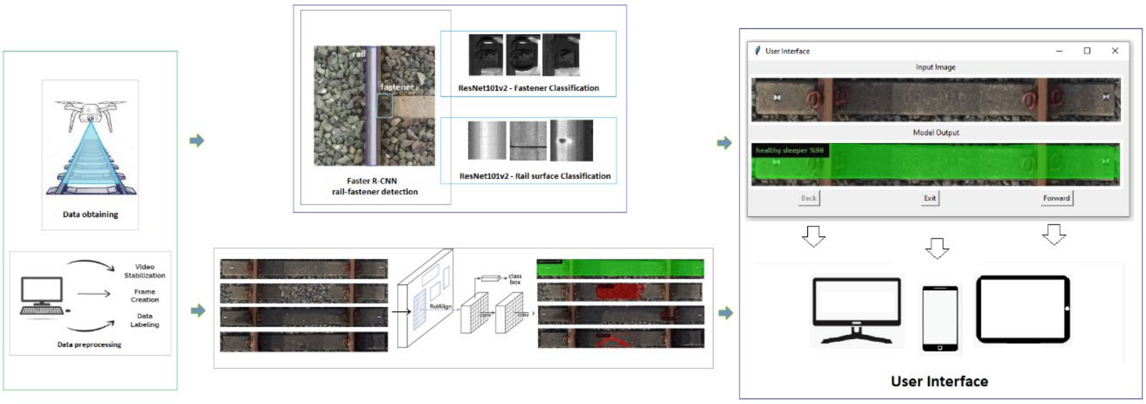


Fig. 1. Block schematic diagram of the method

First of all, a video data set was created by viewing the railway line from a bird's eye view with an autonomous drone. A video stabilization algorithm was applied to eliminate drone-induced shakes in the video containing the railway line. Frames were produced from the video with vibrations. Then, the frames allocated for training and testing were labeled, and the Faster R-CNN deep neural network was trained. Thus, the rail and fastener regions were determined.

```
Algorithm 1: Artificial intelligence-based Python video processing library
import cv2
import tensorflow as tf
model=tf.keras.models.load_newmodel('newmodel.h5')
def stabilize_video(video):
    stabilize_video=video
    return stabilize_video
def extract_frame(stabilize_video)
    frame=cv2.video_capture(stabilize_video)
    fps=frame.get(cv2.CAP_PROP_FPS)
    frame_term=int(1/fps)
      while True:
           ret, video=frame.read()
           if not ret:
                break
      frame.release()
      cv2.destroyAllWindows()
  def  video_processing(frame)
      fault_detection=model.predict(frame)
      return fault_detection
```

Fig. 2. Algorithm 1: Artificial intelligence-based Python video processing library

A 3-class data set named Deformed, Healthy, and Missing was created from the determined fastener regions. Since there were not enough deformed images in the rail regions determined for rail surface defects, the 3-class data set named Healthy, Joint, Squat given in [30] was used. Rail and fastener data sets were trained with ResNet101V2 deep neural network, the classification process was carried out, and fault detection was carried out. In a different previously proposed study, the method developed for multiple fault detection by segmentation using the Mask R-CNN architecture [31] was retrained through the video stabilization phase and incorporated into the user interface module. In order to perform artificial intelligence-based video processing and error detection, the proposed method was converted into a Python library to be used in future studies. In order to ensure ease of use, a user interface has been developed that can send faulty parts of the railway to computers, tablets, and phones via a mobile application.

Within the scope of the article, Faster R-CNN, ResNet101v2, and Mask R-CNN, which are all deep learning-based, were used. Faster R-CNN was used to determine the rails and fasteners on the railway line. ResNet101v2 was used for the classification of rails and fasteners. Mask R-CNN was used to identify multiple railway components by segmentation.

A faster R-CNN object detection algorithm provides higher detection accuracy and faster test results compared to other R-CNN groups. As seen in the diagram in Fig. 3, a separate bounding box is drawn for each class determined in the model consisting of RoI and fully connected layers [32].

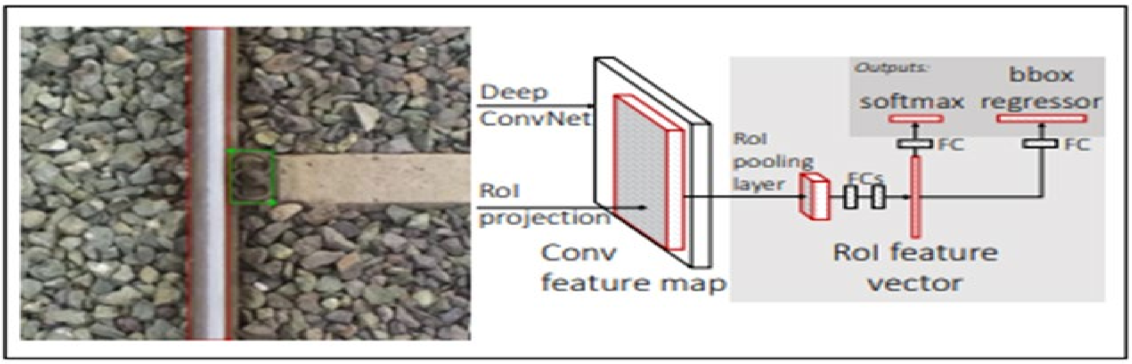


Fig. 3. Faster R-CNN architecture [32]

Mask R-CNN is a design that consists of two stages: object recognition and object identification using masking, with the goal of detecting objects via sample segmentation. Lcls classification loss, Lbox bounding box loss, and Lmask mask loss are examples. The total loss function is the sum of the three loss functions. Thus, the total loss function is expressed as L= Lcls + Lbox + Lmask [33].

ResNet101v2 is a deep transfer learning algorithm that allows classification with low complexity despite its high depth. It is also easy to optimize the model, which increases the accuracy of the methods with the large number of layers. For this reason, it is the deep transfer learning model that is frequently used for image classification [34].

Object detection is used to identify semantic structures from digital images. In this study, an object detection algorithm was used to determine the regions with rails and fasteners from a railway line followed by an autonomous drone. Although there are many developed object detection algorithms, Faster R-CNN has been preferred due to its scalability, multi-object detection, object detection using a regional recommendation network, and high-speed and high-sensitivity object detection. The classification method was used for rail and fastener faults in different forms, each of which consists of three classes. ResNet101v2 was used for classification because of its high generalization and classification performance.

In order to quantify the performance of the transfer learning model after multi-class classification, the accuracy, recall, precision, and f1-score values given in Formula (1), (2), (3), and (4) were calculated. Calculation results are given in the Results and Discussion section. Explanations of the metrics given in the equations are given in Table 1.

$$accuracy = \frac{\sum_{i=1}^{N} TP(C_i)}{\sum_{i=1}^{N}\sum_{j=1}^{N} C_{i,j}} \quad (1)$$

$$recall(C_i) = \frac{TP(C_i)}{TP(C_i)+FN(C_i)} \quad (2)$$

$$precision(C_i) = \frac{TP(C_i)}{TP(C_i)+FP(C_i)} \quad (3)$$

$$f1 - score(C_i) = 2 * \frac{precision(C_i)*recall(C_i)}{recision(C_i)+recall(C_i)} \quad (4)$$

Accuracy is determined by dividing the total number of correct predictions (true positives, TP) by the total number of predictions made, providing a measure of the model's overall correctness. Recall, also known as the true positive rate, quantifies the model's ability to correctly identify positive examples for a specific class C_i by comparing the number of true positives to the total of true positives and false negatives (FN). Precision measures the ratio of true positive predictions to all positive predictions (true positives and false positives, FP), demonstrating the model's dependability in predicting positive cases. The F1-score is a metric that combines precision and recall by computing their harmonic mean. It provides a balanced estimate of a model's performance for a certain class C_i.

Table 1. Meaning of Equation Abbreviations

Abbreviation	Full Name	Meaning
TP	True Positive	Saying right to right
TN	True Negative	Saying wrong to wrong
FP	False Positive	Saying wrong to right
FN	False Negative	Saying right to wrong

2.1. Dataset

500 frames were obtained from the video obtained with the drone. 400 of these images were used to train the Faster R-CNN, and 100 were used to test it. The 603 fastener data determined by Faster R-CNN is divided into 3 groups: Deformed, healthy, and Missing. Since there were insufficient error types in the images obtained for rail surface defects, a data set consisting of a total of 1176 images with 3 classes named Healthy, Joint, and Squats, created in Reference [30], was used.

3. Results and Discussion

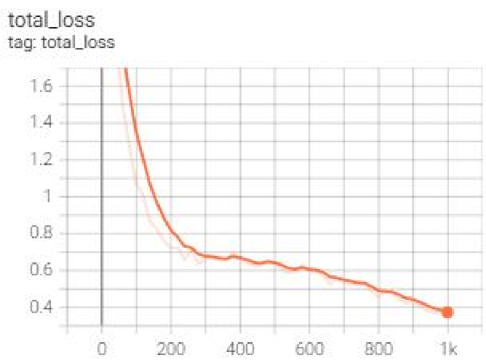
Railroad images recorded by drone tracking were transferred to the computer and tagged for detection of rail and fastener regions using the Labelme data labeling tool. The Faster R-CNN deep neural network was trained with the training data by separating the labeled data for training and testing. The training parameters and the number of images used in the data set are given in Table 2.

Table 2. Faster R-CNN Training Details

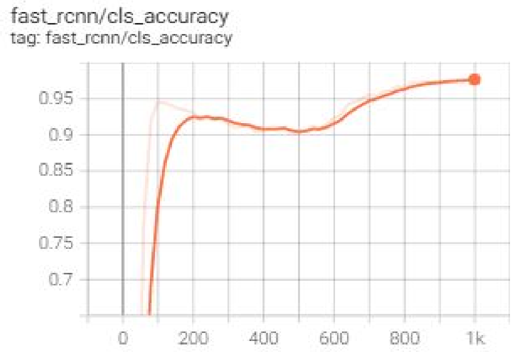
Name	**Value**
Learning Rate	0.001
Training dataset	400
test dataset	100
Image size	900x900
Number of iterations	1000
Number of classes	2

Examples of the training total loss function graph, training accuracy graph, and model test output image are given in Fig. 4. Accordingly; the training was completed in 1000 iterations with a loss function of 0.367 and a training accuracy of 0.976. The model trained using the test dataset was tested, and it was seen that fasteners and rail sections could be detected successfully.

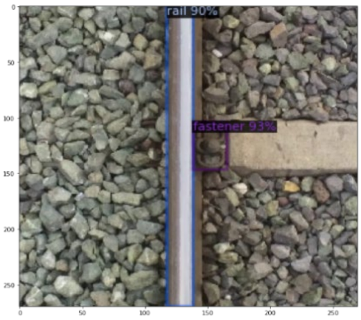
Vol. 10, No. 2, May 2024, pp. 253-264



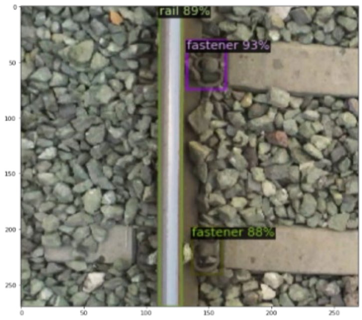
(a) Faster R-CNN Total Loss



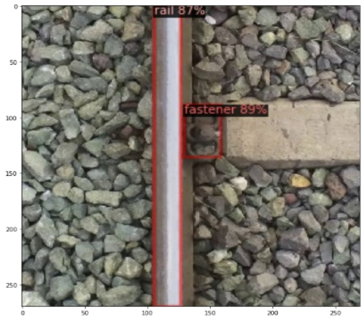
(b) Faster R-CNN Accuracy



(c) Model results image-1



(d) Model results image-2

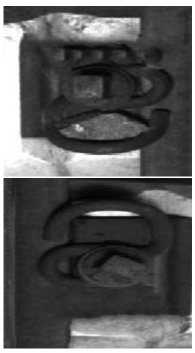


(e) Model results image-3

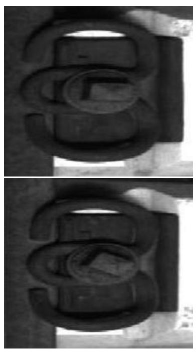
Fig. 4. Object detection (rail and fastener) output images

After determining the rail and fastener using the Faster R-CNN object detection algorithm, the ResNet101v2 deep transfer network was trained separately for both component types using the data sets for the fasteners and rails, examples of which are given in Fig. 5.

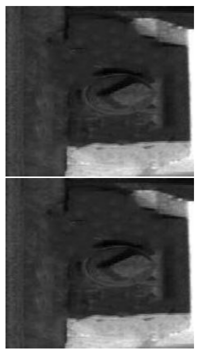
Fastener



(a) Deformed

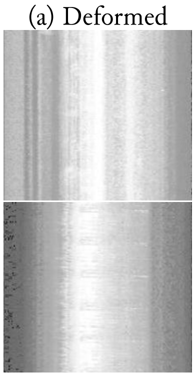


(b) Healthy

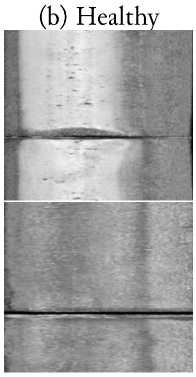


(c) Missing

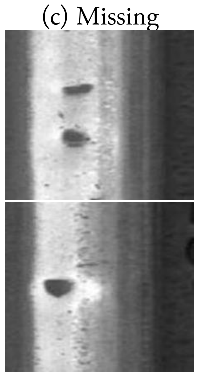
Rail surface



(d) Healthy



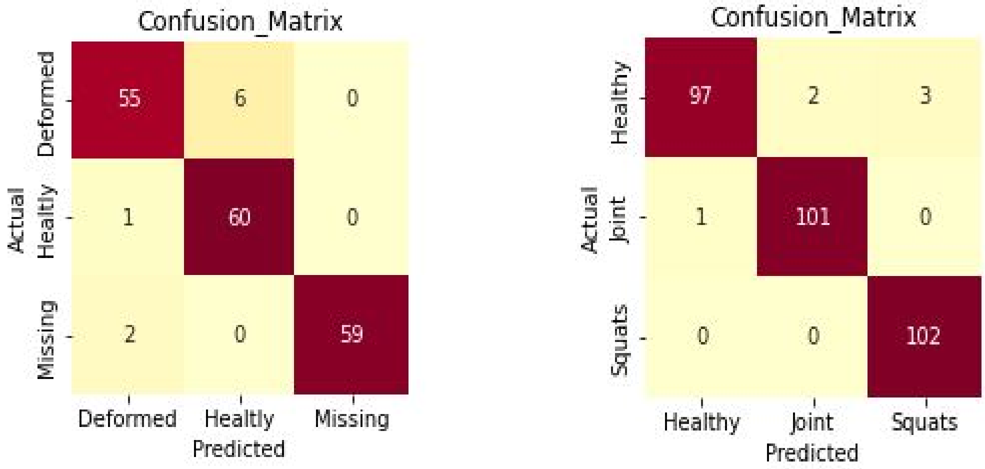
(e) Joint



(f) Squats

Fig. 5. Example images of connector and rail dataset

603 visual railway data of 227x227 size was used for fastener classification. For the rail surface classification, 1176 visual data of 224x224 dimensions were used. For both components, 15 iterations of the ResNet101v2 deep transfer network were trained. The effectiveness of the separate training for rail and fastener was measured by plotting the multi-class Confusion matrix and calculating the success criteria given in Equations (1), (2), (3), and (4), which are used for multi-class classification. Accordingly, Confusion Matrixes drawn for the fastener in Fig. 6 (a) for the rail surface defect in Fig. 6 (b) are given.



(a) Fastener confusion matrix (b) Rail surface confusion matrix

Fig. 6. Confusion matrix (Fastener and Rail)

ResNet101v2 deep neural network was trained using rail and fastener datasets, and then Equation (1) was calculated, and overall accuracy rates were determined. Accordingly, the accuracy rates were measured as 95% and 98% for rail and fastener, respectively.

The fastener dataset, consisting of the Deformed, Healthy, and Missing classes, was tested with the images allocated for testing. Model evaluation metrics were calculated. For the Deformed class, 95% Precision, 90% Recall, and 92% F1-score values were obtained. For the Healthy class, 91% Precision, 98% Recall, and 94% F1-score values were obtained. 100% Precision, 97% Recall, and 98% F1-score values were obtained for the Missing class. The recorded values are given in Fig. 7.

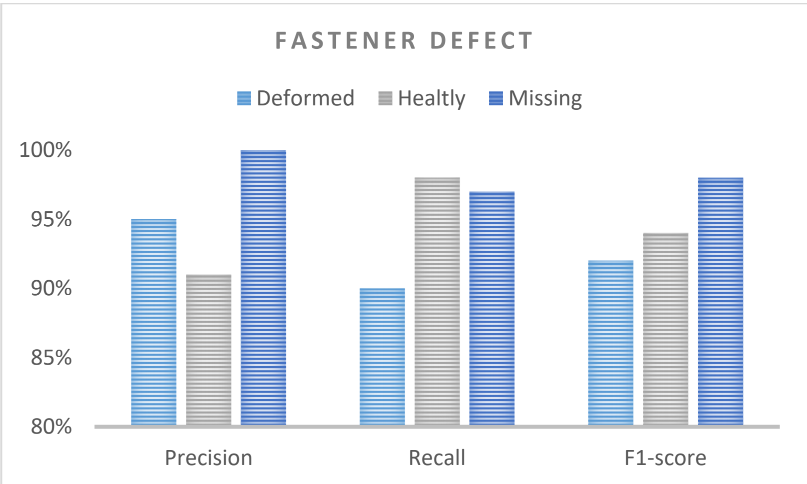


Fig. 7. Fastener evaluation metrics graph

The rail surface dataset, consisting of the Healthy, Joint, and Squats classes, was tested with the images reserved for testing. Model evaluation formulas were calculated. For the Healthy class, 99% Precision, 95% Recall, and 97% F1-score values were obtained. 98% Precision, 99% Recall, and 99%

F1-score values were obtained for the Joint class. For the Squats class, 97% Precision, 100% Recall, and 99% F1-score were achieved. The recorded values are given in Fig. 7.

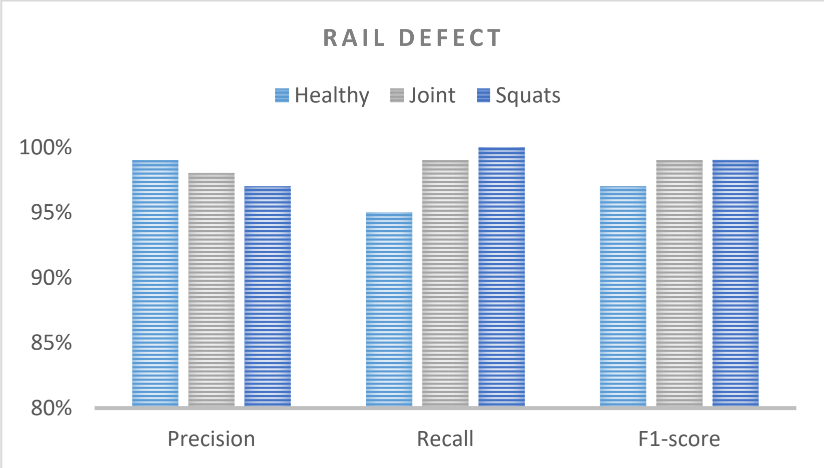


Fig. 8. Rail surface evaluation metrics graph

In order to detect multiple faults with Mask R-CNN, the data in the proposed study was retrained by passing the video stabilization block. As a result of the training, it was observed that the model accuracy increased from 95% to 97.4%. Segmentation results were transferred to the user interface. The input image and model output image can be displayed on the interface. The comparison between the proposed method and similar research in the literature is given in Table 3.

Table 3. Comparison of Literature and Proposed Method

References	Fault type	Used method	Accuracy
[11]	Fastener defects	CNN and ResNet50	70.0 %
[35]	Rail defects	Deep convolutional neural network (DCNN)	93.35 %
[36]	Fastener	YOLOv4-Hybrid model	94.4 %
[37]	Rail surface	YOLOX	96.1 %
[1]	Rail defects	FCN-8 deep learning network	81.0 %
Proposed Method	Rail surface defect – Fastener defect	Faster R-CNN and Transfer Learning (ResNet101v2)	rail surface: 98.0 % fastener : 95.0 %

Railway transportation systems are frequently used by countries. Fault detection is of great importance to ensure safe and continuous transportation. In time, non-contact flaw detection methods have replaced the difficult manual maintenance on kilometers-long railways. Various flaw detection studies have been recommending in the literature for fault detection in railways consisting of many components. In this study, visual railway data is obtained by autonomous drone, saving time and human labor. The data was improved with video stabilization. For fault detection, object detection, classification, and segmentation methods were used in deep learning. In addition, a user interface was developed where the detected faults can be sent to the railway inspection unit via a mobile application. Better detection results were obtained compared to other proposed studies, with 98.0% accuracy in determining rail surface defects and 95.0% accuracy in determining fastener defects. Within the scope of the study, the proposed method contributed to the field of video processing and fault detection in industrial areas. In addition, an artificial intelligence-based video processing Python library was created to facilitate the use of the method in future studies. Thus, in future studies, video processing applications will be developed in various fields (Industrial, security, etc.).

4. Conclusion

Fault detection studies should be carried out periodically for safe and continuous transportation in railway transportation systems, which are frequently used in urban and intercity transportation today. As an alternative to the proposed non-contact technique for flaw detection, in the study, a new approach, including object detection and classification and defect detection stages, is proposed. In addition, the fault detection method with segmentation has been improved and included in the model. Rail and fastener sections were determined by training the railway picture acquired with the autonomous drone with the Faster R-CNN. A three-class data set was created for fasteners, and a three-class open-access dataset was used for rail surface defects. Fault detection in both railway components was made using the ResNet101v2 deep neural network, which allows classification in deep transfer learning. The user interface where input and model output images can be observed has been developed. Fasteners were classified as Deformed, Healthy, and Missing with 95% accuracy. Rail surface defects were classified as Healthy, Joint, and Squats with 98% accuracy. Improved detection accuracies were obtained compared to the methods proposed in the relevant field. Within the scope of the article, video segmentation, object detection, and classification techniques were used. A Python library was created as an artificial intelligence accelerator in order to contribute to the application of deep learning in industrial areas and to enable its use in such studies. In future studies, the artificial intelligence-based video processing library proposed within the scope of the study will be used for video processing in different industrial areas.

Acknowledgment

The Authors would like to express our sincere gratitude to TUBITAK (The Scientific and Technological Research Council of Turkey) for their financial support under Grant No: 5220154, which made this research possible.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. This work was supported by the TUBITAK (The Scientific and Technological Research Council of Turkey) under Grant No: 5220154.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

References

[1] P. Bojarczak and P. Lesiak, “UAVs in rail damage image diagnostics supported by deep-learning networks,” *Open Eng.*, vol. 11, no. 1, pp. 339–348, Jan. 2021, doi: 10.1515/eng-2021-0033.

[2] M. Yilmazer and M. Karakose, “Mask R-CNN Architecture Based Railway Fastener Fault Detection Approach,” in *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, Mar. 2022, pp. 1363–1366, doi: 10.1109/DASA54658.2022.9765024.

[3] M. Yilmazer and M. Karakose, “Railway Condition Monitoring and Fault Detection Based on YOLOv4,” in *2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, Nov. 2022, pp. 439–443, doi: 10.1109/3ICT56508.2022.9990640.

[4] M.-Y. Cheng, R. R. Khasani, and K. Setiono, “Image quality enhancement using HybridGAN for automated railway track defect recognition,” *Autom. Constr.*, vol. 146, p. 104669, Feb. 2023, doi: 10.1016/j.autcon.2022.104669.

[5] M. G. Minguell and R. Pandit, “TrackSafe: A comparative study of data-driven techniques for automated railway track fault detection using image datasets,” *Eng. Appl. Artif. Intell.*, vol. 125, p. 106622, Oct. 2023, doi: 10.1016/j.engappai.2023.106622.

[6] N. Mukhtiar, K. Khuwaja, D. P. Tocut, and I. Tarca, “An Experimental Study on the Railway Track Surface Fault Detection with Automation,” in *2023 17th International Conference on Engineering of Modern Electric Systems (EMES)*, Jun. 2023, pp. 1–4, doi: 10.1109/EMES58375.2023.10171771.

[7] H. Luo, L. Cai, and C. Li, "Rail Surface Defect Detection Based on An Improved YOLOv5s," *Appl. Sci.*, vol. 13, no. 12, p. 7330, Jun. 2023, doi: 10.3390/app13127330.

[8] A. K. Singh, A. K. Dwivedi, N. Nahar, and D. Singh, "Railway Track Sleeper Detection in Low Altitude UAV Imagery Using Deep Convolutional Neural Network," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, Jul. 2021, pp. 355–358, doi: 10.1109/IGARSS47720.2021.9554542.

[9] S. Qiu *et al.*, "Automated detection of railway defective fasteners based on YOLOv8-FAM and synthetic data using style transfer," *Autom. Constr.*, vol. 162, p. 105363, Jun. 2024, doi: 10.1016/j.autcon.2024.105363.

[10] J. Liu, H. Liu, C. Chakraborty, K. Yu, X. Shao, and Z. Ma, "Cascade Learning Embedded Vision Inspection of Rail Fastener by Using a Fault Detection IoT Vehicle," *IEEE Internet Things J.*, vol. 10, no. 4, pp. 3006–3017, Feb. 2023, doi: 10.1109/JIOT.2021.3126875.

[11] P. Chandran, J. Asber, F. Thiery, J. Odelius, and M. Rantatalo, "An Investigation of Railway Fastener Detection Using Image Processing and Augmented Deep Learning," *Sustainability*, vol. 13, no. 21, p. 12051, Oct. 2021, doi: 10.3390/su132112051.

[12] M. Chen, W. Zhai, S. Zhu, L. Xu, and Y. Sun, "Vibration-based damage detection of rail fastener using fully convolutional networks," *Veh. Syst. Dyn.*, vol. 60, no. 7, pp. 2191–2210, Jul. 2022, doi: 10.1080/00423114.2021.1896010.

[13] L. Wang, Q. Zang, K. Zhang, and L. Wu, "A rail fastener defect detection algorithm based on improved YOLOv5," Feb. 2024, doi: 10.1177/09544097241234380.

[14] Z. Chen, Q. Wang, Q. He, T. Yu, M. Zhang, and P. Wang, "CUFuse: Camera and Ultrasound Data Fusion for Rail Defect Detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 21971–21983, Nov. 2022, doi: 10.1109/TITS.2022.3189677.

[15] X. Li, Q. Wang, X. Yang, K. Wang, and H. Zhang, "Track Fastener Defect Detection Model Based on Improved YOLOv5s," *Sensors*, vol. 23, no. 14, p. 6457, Jul. 2023, doi: 10.3390/s23146457.

[16] R. Shafique *et al.*, "A Novel Approach to Railway Track Faults Detection Using Acoustic Analysis," *Sensors*, vol. 21, no. 18, p. 6221, Sep. 2021, doi: 10.3390/s21186221.

[17] İ. Aydin, M. Sevi, E. Akin, E. Güçlü, M. Karaköse, and H. Aldarwich, "A Deep Learning-Based Hybrid Approach to Detect Fastener Defects in Real-Time," *Teh. Vjesn. - Tech. Gaz.*, vol. 30, no. 5, pp. 1461–1468, Oct. 2023, doi: 10.17559/TV-20221020152721.

[18] D. Zheng *et al.*, "A Defect Detection Method for Rail Surface and Fasteners Based on Deep Convolutional Neural Network," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–15, Jul. 2021, doi: 10.1155/2021/2565500.

[19] Q. Zhou, "A Detection System for Rail Defects Based on Machine Vision," *J. Phys. Conf. Ser.*, vol. 1748, no. 2, p. 022012, Jan. 2021, doi: 10.1088/1742-6596/1748/2/022012.

[20] Y. Wu, Y. Qin, Y. Qian, F. Guo, Z. Wang, and L. Jia, "Hybrid deep learning architecture for rail surface segmentation and surface defect detection," *Comput. Civ. Infrastruct. Eng.*, vol. 37, no. 2, pp. 227–244, Feb. 2022, doi: 10.1111/mice.12710.

[21] Y. Zheng, S. Wu, D. Liu, R. Wei, S. Li, and Z. Tu, "Sleeper Defect Detection Based on Improved YOLO V3 Algorithm," in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, Nov. 2020, pp. 955–960, doi: 10.1109/ICIEA48937.2020.9248299.

[22] J. Hu *et al.*, "High Speed Railway Fastener Defect Detection by Using Improved YoLoX-Nano Model," *Sensors*, vol. 22, no. 21, p. 8399, Nov. 2022, doi: 10.3390/s22218399.

[23] Z. Su, K. Han, W. Song, and K. Ning, "Railway fastener defect detection based on improved YOLOv5 algorithm," in *2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Oct. 2022, vol. 2022-Octob, pp. 1923–1927, doi: 10.1109/IAEAC54830.2022.9929911.

[24] R. Kapoor, R. Goel, and A. Sharma, "An intelligent railway surveillance framework based on recognition of object and railway track using deep learning," *Multimed. Tools Appl.*, vol. 81, no. 15, pp. 21083–21109, Jun. 2022, doi: 10.1007/s11042-022-12059-z.

[25] Y. Qiu, H. Liu, J. Liu, and B. Shi, "Center-Triplet Loss for Railway Defective Fastener Detection," *IEEE Sens. J.*, vol. 24, no. 3, pp. 3180–3190, Feb. 2024, doi: 10.1109/JSEN.2023.3339883.

[26] M. A.-M. Khan, S.-H. Kee, and A.-A. Nahid, "Vision-Based Concrete-Crack Detection on Railway Sleepers Using Dense U-Net Model," *Algorithms*, vol. 16, no. 12, p. 568, Dec. 2023, doi: 10.3390/a16120568.

[27] Y. Santur, M. Yilmazer, M. Karakose, and E. Akin, "A New Rail Surface Defects Detection Approach Using 3D Laser Cameras Based on ResNet50," *Trait. du Signal*, vol. 39, no. 4, pp. 1339–1345, Aug. 2022, doi: 10.18280/ts.390427.

[28] A. Saini, D. Singh, and M. Alvarez, "FishTwoMask R-CNN: Two-stage Mask R-CNN approach for detection of fishplates in high-altitude railroad track drone images," *Multimed. Tools Appl.*, vol. 83, no. 4, pp. 10367–10392, Jan. 2024, doi: 10.1007/s11042-023-15924-7.

[29] T. Bai, J. Gao, J. Yang, and D. Yao, "A Study on Railway Surface Defects Detection Based on Machine Vision," *Entropy*, vol. 23, no. 11, p. 1437, Oct. 2021, doi: 10.3390/e23111437.

[30] I. Aydin, E. Akin, and M. Karakose, "Defect classification based on deep features for railway tracks in sustainable transportation," *Appl. Soft Comput.*, vol. 111, p. 107706, Nov. 2021, doi: 10.1016/j.asoc.2021.107706.

[31] M. Yilmazer, M. Karaköse, İ. Aydin, and E. Akin, "Mask R-CNN Derin Sinir Ağı ile Demiryolu Hattı Bileşenlerinde Çoklu Hata Tespiti," *Çukurova Üniversitesi Mühendislik Fakültesi Derg.*, vol. 37, no. 4, pp. 1103–1112, Dec. 2022, doi: 10.21605/cukurovaumfd.1230955.

[32] R. Girshick, "Fast R-CNN," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, vol. 2015 Inter, pp. 1440–1448, doi: 10.1109/ICCV.2015.169.

[33] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, vol. 2017-Octob, pp. 2980–2988, doi: 10.1109/ICCV.2017.322.

[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-December, pp. 770–778, Dec. 2016, doi: 10.1109/CVPR.2016.90.

[35] X. Giben, V. M. Patel, and R. Chellappa, "Material classification and semantic segmentation of railway track images with deep convolutional neural networks," in *2015 IEEE International Conference on Image Processing (ICIP)*, Sep. 2015, vol. 2015-Decem, pp. 621–625, doi: 10.1109/ICIP.2015.7350873.

[36] F. Guo, Y. Qian, and Y. Shi, "Real-time railroad track components inspection based on the improved YOLOv4 framework," *Autom. Constr.*, vol. 125, p. 103596, May 2021, doi: 10.1016/j.autcon.2021.103596.

[37] Y. Min, J. Guo, and K. Yang, "Research On Real-time Detection Algorithm Of Rail-surface Defects Based On Improved YOLOX," *J. Appl. Sci. Eng.*, vol. 26, no. 6, pp. 799–810, 2022. [Online]. Available at: http://jase.tku.edu.tw/articles/jase-202306-26-6-0006.