Turkish Journal of Electrical Engineering and Computer Sciences

Volume 31 | Number 4 Article 5

7-1-2023

Improving Unet segmentation performance using an ensemble model in images containing railway lines

MEHMET SEVİ

İLHAN AYDIN

Recommended Citation

SEVİ, MEHMET and AYDIN, İLHAN (2023) "Improving Unet segmentation performance using an ensemble model in images containing railway lines," *Turkish Journal of Electrical Engineering and Computer Sciences*: Vol. 31: No. 4, Article 5. https://doi.org/10.55730/1300-0632.4014
Available at: https://journals.tubitak.gov.tr/elektrik/vol31/iss4/5

Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2023) 31: 739 – 750
© TÜBİTAK
doi:10.55730/1300-0632.4014

Improving Unet segmentation performance using an ensemble model in images containing railway lines

Mehmet SEVİ[1*], **İlhan AYDIN**[2]
[1]IT Department, Muş Alparslan University, Muş, Turkey
[2]Department of Computer Engineering, Faculty of Engineering, Fırat University, Elazığ, Turkey

Received: 02.01.2023 • **Accepted/Published Online:** 31.05.2023 • **Final Version:** 28.07.2023

Abstract: This study aims to make sense of the autonomous system and the railway environment for railway vehicles. For this purpose, by determining the railway line, information about the general condition of the line can be obtained along the way. In addition, objects such as pedestrian crossings, people, cars, and traffic signs on the line will be extracted. The rails and the rail environment in the images will be segmented with a semantic segmentation network. In order to ensure the safety of rail transport, computer vision, and deep learning-based methods are increasingly used to inspect railway tracks and surrounding objects. In particular, the extraction of objects around the railway line has become an important task. The dataset contains images of the railway line and its surroundings, which were obtained in changing environmental conditions, at different times of the day, and under poor lighting conditions. In this study, a new method is proposed for the extraction of objects in and around the railway line. The proposed approach first applied Unet-based segmentation methods on the dataset. Then, a method that improves Unet performance based on the ensemble model is proposed. ResNet34, MobileNetV2, and VGG16 backbones were used to improve segmentation performance. The proposed model is based on the ensemble decision-making process, significantly contributing to the semantic segmentation task. Experimental results of the developed model show that it gives 85% accuracy rate and 54% average IoU results.

Key words: Deep learning, semantic segmentation, railway line, Unet, ensemble model

1. Introduction

Due to the widespread use of railways and the fact that fast transportation is indispensable in human life, the timely detection of defects in railways has become a critical issue. This study aims to automate the railway maintenance operations carried out with traditional methods. For this purpose, firstly, the focus is on determining the railway line and its surrounding objects. This study proposes a deep learning-based method to determine the objects in and around the railway line. Segmentation is the unmasking of the image by performing a pixel-based estimation. There are two types of segmentation, semantic and instance segmentation. Semantic segmentation is when pixels in the same object have the same label. Instance segmentation, on the other hand, is labeling different instances of the same objects differently [1]. A method for object-based analysis of the railway line and its surroundings based on semantic segmentation was proposed. Although the proposed method basically uses the Unet approach, it contributes to metrics such as accuracy rate and average IoU. In this study, the effect of an ensemble learning model consisting of different backbones used in the U-Net architecture on semantic segmentation performance was examined. The rails and the rail environment in the images will

*Correspondence: m.sevi@alparslan.edu.tr

be segmented by a network that performs semantic segmentation. Rail transport is one of the most important modes of transport today. With the development of high-speed trains, railway transportation started to be preferred more by people in transportation. As of 2020, 148 thousand people preferred railway transportation in Turkey [2].

Gibert et al. determined the materials belonging to 10 classes used on the railway line with a semantic segmentation-based method. They achieved an accuracy rate of 93% with the Deep Convolutional Neural Network (DCNN) method they developed [3]. Periodic monitoring of the perimeter of the railway line is essential to ensure railway safety. Traditional image processing methods only study the railway line of a simple background. He et al. designed the Deep Semantic Segmentation Convolutional Neural Networks (DSSCNN) method to make sense of the railway scene [4]. The method they designed was more successful than fully convolutional networks (FCN). In another study, a pixel-based semantic segmentation model was proposed to detect railway facilities around the railway line [5]. As a result of the study, a mean IoU value of 0.49 was obtained. Semantic segmentation is not only used to make sense of the railway line and its surroundings. It has been used to detect anomalies on the railway [6]. Wang et al. proposed a new surface defect detection network based on Mask Region-Based Convolutional Neural Networks (Mask R-CNN) to detect rail defects [7]. Experimental evaluation showed that the proposed model achieves an average accuracy (MAP) of 98.70% on the dataset and can locate the defect more accurately. Li et al. proposed a semantic segmentation-based algorithm to detect rail fasteners. In the study, they proposed a semantic segmentation-based model for rail fastener detection that detects and classifies rail fastener states by combining PSPNet and vector geometry measurements. As a result of the study, fasteners in different situations were detected at a rate of 93% [8]. Mobile mapping methods for railway safety have also been developed in the literature. Grandio et al. mapped the railway line by a method based on semantic segmentation. The methodology of the study is to segment both linear and punctual elements from the railway infrastructure and test it in four scenarios. I- 90 km long railway; II- 2 km long point clouds of poor quality; III) - 400 m long high-quality point clouds; IV- The 1.4 km long railway was recorded with the aerial mapping system. The longest was used for training and testing, with an average accuracy of more than 90%. Other scenarios were used for testing only [9]. In other studies in the literature, the RailSem19 dataset was used mostly as in this study. It is used in many areas such as segmenting the railway line, detecting foreign objects around it, and generating defective data [10]. Alexandrescu et al. proposed a railway semantic partitioning method using two deep Unet architectures, Unet and ResUnet++ [11]. Alexandrescu et al. used the RailSem19 dataset in their proposed method. In the dataset, they used only rail and background classes in their studies. They obtained mean IoU of 0.54 with Unet model for the rail class. Zhang et al. proposed DFA-Unet, a segmentation algorithm based on an improved Unet network architecture. The model used the same encoder-decoder structure as Unet. RailSem19 dataset was used as the dataset. The results showed that the model showed a 2.48% improvement in the mean IoU metric compared to Unet [12]. The proposed work for the analysis of railway tracks and surrounding objects falls within the field of computer vision. Images taken from real environments will be used in deep learning-based image segmentation algorithms after various preprocessing. The railway track and the rail environment in the images will be visually separated with deep networks capable of semantic segmentation. Objects such as people or automobiles can intrude on railway lines. These unauthorized entries both threaten the lives of living things and cause train accidents. Foreign objects are quite common, especially on railway lines close to settlements. These objects make it difficult to detect the railway line and its components, and may also pose a threat to railway transportation security. Therefore, it is extremely important to detect these intrusions in a timely manner.

The proposed ensemble approach in Unet segmentation is critical and is precisely the motivation of this study. This study proposes a segmentation system based on a newly developed hybrid loss function and ensemble decision-making process that is tailored for accurate Unet segmentation. In the proposed segmentation-based method, we propose a solution for visual analysis of the railway line and surrounding objects using models with different Unet backbones. The main contributions of the proposed solution are given below:

- A hybrid loss function is proposed in the study.
- The proposed model is based on the ensemble decision-making process. It has improved the performance of models that make decisions individually.
- The performance of the proposed segmentation model is comparable to other pretrained state-of-the-art networks.
- Foreign objects around the railway line were detected result of the study segmentation for transportation safety.

In the study, the railway and its surrounding objects were determined by semantic segmentation-based methods. It is aimed to improve Unet segmentation performance by using the Unet-based ensemble model.

In the second part of the study, the dataset and the proposed method, the experimental results in the third part, and the general results in the last part will be mentioned.

2. The proposed method

A Unet deep learning-based model is proposed to obtain homogeneous objects around the railway line. With deep learning methods, which are one of the subbranches of artificial intelligence, it is quite possible to detect defects autonomously from images. Deep learning techniques help to automatically learn different problem-specific features compared to traditional techniques [13, 14]. In other words, deep learning discovers the parameters that need to be defined in machine learning and can give more accurate results than machine learning. Convolutional neural networks are based on a deep learning strategy. Normally, large-featured learning networks such as convolutional neural networks need to have large datasets for training. Another important point is that general image training in classical convolutional neural networks is done over class labels. However, some problems require pixel-based approaches. In areas that require a sensitive approach, such as health or safety areas, class information of each pixel is needed. This is where image segmentation comes into play. Semantic segmentation methods were used in this study. Semantic segmentation determines which class each pixel in an image belongs to [15].

Firstly, the Unet model was used for the segmentation process. Then, the ensemble learning model was applied to the dataset by using various backbones with Unet architecture. Ronneberger et al. first announced the Unet Convolutional Neural Network approach in 2015 [16]. Unet is an automatic segmentation method and is all about defining the boundary. Unet architecture gives more successful results than classical models. It can also give good results even with a small number of training images. Unet gets its name from its U-like architecture, as seen in Figure 1.

Unet consists of encoder and decoder parts. In the encoder part, the content in the image is captured. The maximum pooling layers are also in this section. The decoding part, on the other hand, is the symmetric expansion path used to achieve precise localization using delegated convolutions, also called a decoder. This work presents a semantic segmentation model based on ensemble decision-making for railway images combining

3 Unet architectures with pretrained backbone networks. One of the proposed segmentation models has an Unet architecture with 34 layers of ResNet34 pretrained on [17] as the backbone. Other backbones are MobileNetV2 and VGG16. These backbones are pretrained on ImageNet. All the mentioned backbones weights are pretrained on the ImageNet dataset [18]. The advantage of this is shortening the learning procedure done by the last few layers of the network, speeding up the convergence, and getting high performance from the dataset.

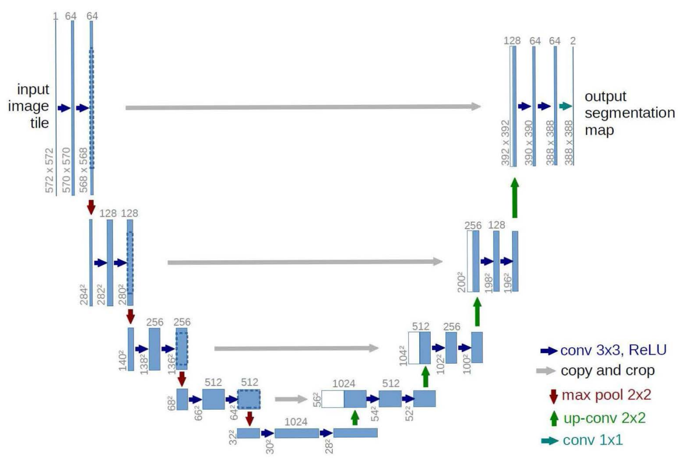

Figure 1. Unet architecture [16].

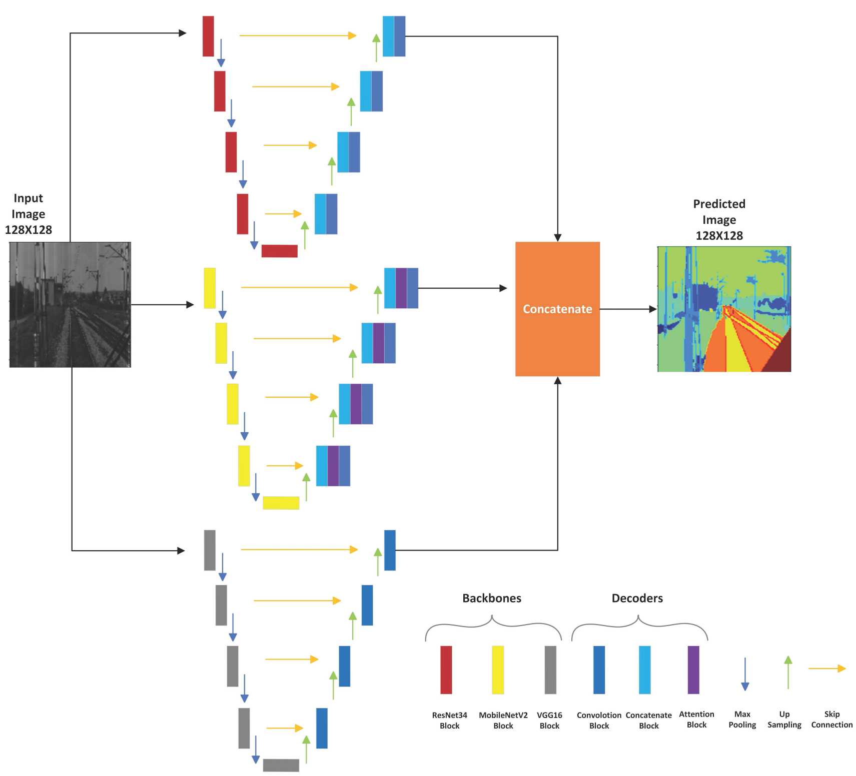

Figure 2. Architecture of the proposed model.

As seen in Figure 2, 3 pretrained backbones using the Unet architecture make a joint decision for the predicted image. It is used as the encoder part of the Unet architecture in 3 backbones. The decoder path is defined automatically. These backbones were used as encoders, the first half of Unet. Then, the decoder layers are trained with the dataset. Next, the decoding layers are trained with the augmented dataset. It helps save the training procedure and enhances Unet's ability to learn from small data. First of all, all three backbones are trained independently of each other. As seen in Figure 2, we propose an architecture that can handle ResNet34, MobileNetV2, and VGG16 as a decoder. In the encoder part of Unet, each backbone used its own blocks. In the encoder path, the traditional stack of convolutional and 2x2 max pooling layers was used. Max pooling simply takes the maximum input value in the given stride. In the decoder part, each backbone has its own unique layer structure. The decoder consists of upsampling and concatenation followed by regular convolution operations. In Figure 2, a 2 x 2 convolution layer, concatenation, or attention block was used, which halves the number of feature channels after upsampling in the expansion path on the right. Finally, the proposed model concatenates merge output from each Unet model.

The following loss functions were used to understand how well Unet models the given data. Two types of loss functions are combined in the study, Dice loss function and Focal loss function. In the proposed model, a hybrid loss model is obtained by combining the most widely used loss functions in the literature. The loss functions used in the proposed hybrid model are explained below. Dice loss is written using the definition of precision (1) and recall (2). The calculation equation of the Dice loss is shown in (3). The Dice loss function is the most widely used segmentation evaluation metric [19], it directly optimizes. In the studies in the literature, it has been observed that the accuracy rate of the Unet model is maximum when the beta is taken as 1 [20].

$$Precision = \frac{TruePositive}{TruePositive + FalseNegative} \tag{1}$$

$$Recall = \frac{TruePositive}{TruePositive + FalsePositive} \tag{2}$$

$$Dice\,(Precision, Recall) = 1 - \left(1+\beta^2\right)\frac{Precision * Recall}{\beta^2 * Precision * Recall} \tag{3}$$

(4) establishes a criterion measuring the categorical Focal loss between the ground truth (gt) and the predicted (pr) [21]. The α value is taken as 0.25 and the γ value as 2.

$$Focal\,(gt, pr) = -gt * \alpha * (1-pr)^{\gamma} * log(pr) \tag{4}$$

We employ a hybrid loss consisting of contributions from both dice loss and focal loss. Then, these two loss functions are processed as in (5), and the final loss value is sent to the model training.

$$TotalLoss = DiceLoss + FocalLoss \tag{5}$$

Another performance evaluation metric that is frequently used in semantic segmentation is the average IoU metric. The IoU metric is calculated by dividing the predicted image and reference image intersection by the predicted image and reference image combination. It is seen in (6). IoU value is calculated separately for each class. By taking the average of these values, the average IoU value is obtained [22]. The accuracy rate measures the number of correct pixel predictions (7).

$$IoU = \frac{Target \cap Prediction}{Target \cup Prediction} \tag{6}$$

$$AccuracyRate = \frac{TruePositive + TrueNegative}{TruePositive + TrueNegative + FalsePositive + FalseNegative} \tag{7}$$

RailSem19 dataset was used in the study. RailSem19 is a dataset specifically designed for semantic segmentation in the railway domain. It contains high-resolution images captured from a mobile mapping system mounted on a railway vehicle. The dataset includes various railway scenes captured from different perspectives, including tracks, catenaries, platforms, and surrounding infrastructure. There are 19 classes in the dataset. Additionally, we added the background class as the 20th class. Some important classes are rail, tramway, human, traffic sign, car, rail vehicles, metal rails, drivable rails, road, pedestrian crossing, vegetation, and terrain classes. The annotations enable the training and evaluation of deep learning models for semantic segmentation tasks related to railway infrastructure. RailSem19 is a valuable resource for developing and evaluating algorithms and models related to railway infrastructure analysis, such as object detection, scene understanding, and condition monitoring. It can facilitate research and development in railway maintenance, safety, and automation applications. The dataset consists of 8500 images containing the railway and its surrounding objects and references images of these images [23]. The images that make up the dataset consist of images that include complex rail scenarios and difficult weather conditions. It supports the proposed model to work in difficult and different conditions. Figure 3 shows examples from the dataset. 80% of the dataset was used for training, 10% for validation and 10% for testing.

Figure 3. Examples from the dataset.

3. Experimental results

The proposed models were performed in an experimental environment with 11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40GHz 2.42 GHz and NVIDIA GeForce MX450 graphics card. Unet segmentation process training parameters: epoch number is set to 120, and the batch size is set to 16. Since no progress was observed in the performance metrics after the 50th epoch, the number of epochs was fixed at 50. The performance and loss graphs of the training and validation stages of the Unet model are shown in Figure 4.

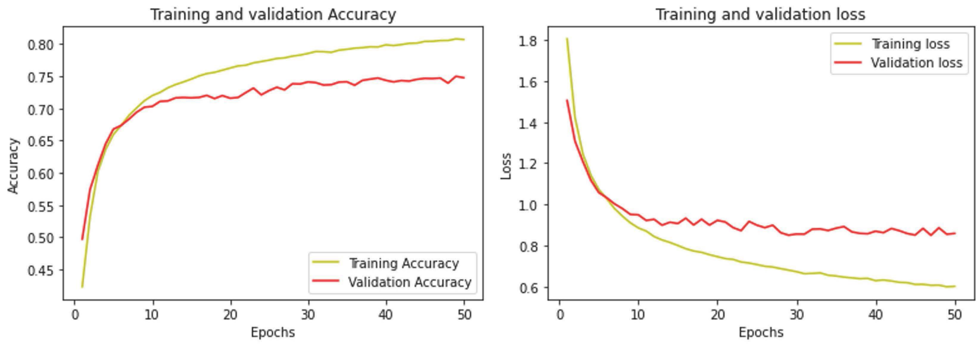

Figure 4. Validation and training graphs of the Unet model.

As seen in Figure 4, the accuracy rate of 0.7478 and average IoU of 0.3625 was obtained as a result of the validation process. Estimation was performed on the dataset using the trained Unet model. The result of the process is shown in Figure 5. Figure 5 shows the testing image, the reference image of the image, and the predicted image, respectively.

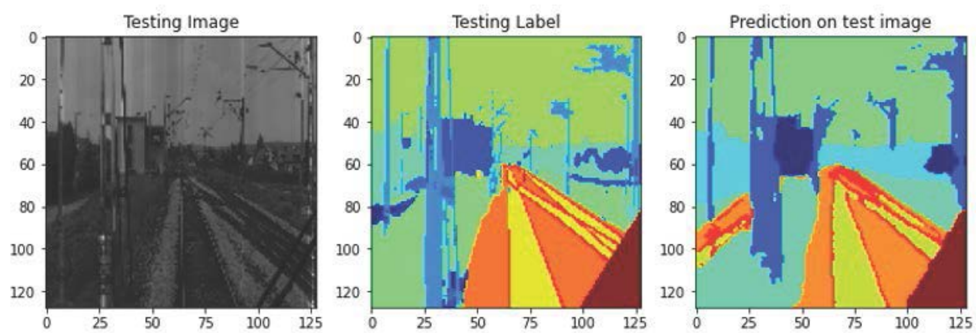

Figure 5. Predictions of the Unet model.

A model based on the ensemble decision-making process is proposed as a result of training the backbones using the Unet architecture to increase the average IoU and validation values obtained. ResNet34, MobileNetV2, and VGG16 were used as backbones. The MobileNetV2 backbone obtained the values in Figure 6 as a result of the training.

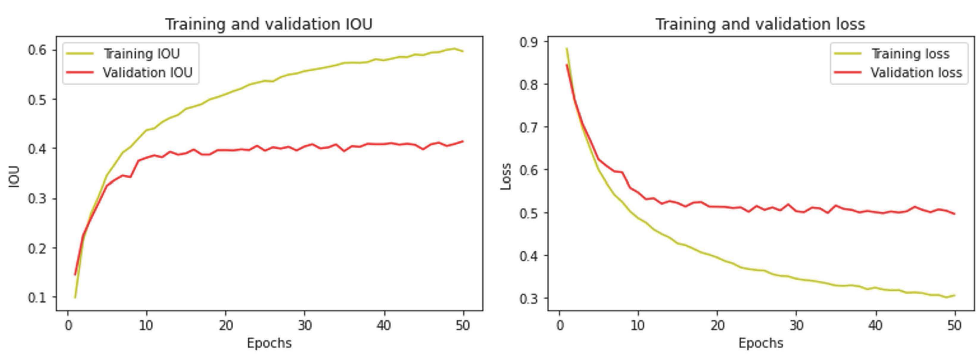

Figure 6. Training and validation results of the MobileNetV2 backbone used.

As a result of the validation, the MobileNetV2 backbone achieved an average IoU value of 0.4131. The MobileNetV2 backbone obtained an average IoU value of 0.5957 as a result of the training process. The ResNet34 backbone obtained the training and validation values in Figure 7 as a result of the training.

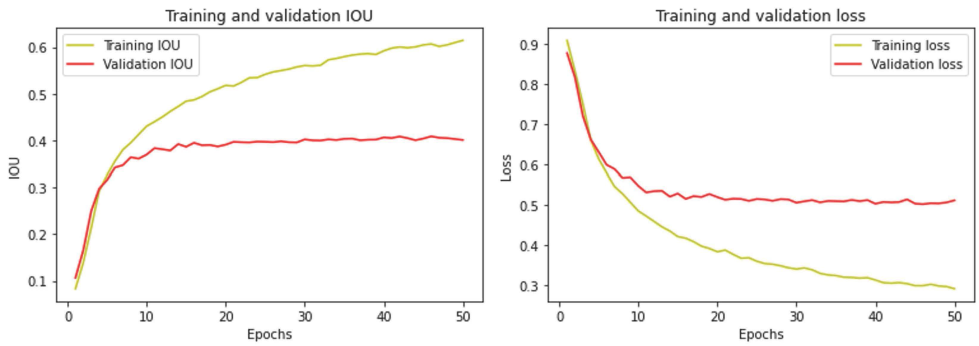

Figure 7. Training and validation results of the ResNet34 backbone used.

As a result of the validation, the ResNet34 backbone achieved an average IoU value of 0.4009. The ResNet34 backbone obtained an average IoU value of 0.6145 as a result of the training process. Finally, the Unet model is trained on the VGG16 backbone dataset. The training and validation values in Figure 8 were obtained.

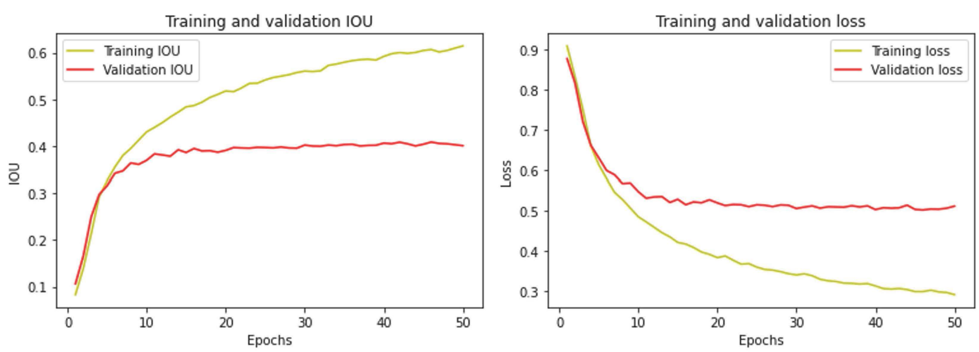

Figure 8. Training and validation results of the VGG16 backbone used.

As a result of the validation, the VGG16 backbone obtained an average IoU value of 0.6117. The VGG16 backbone obtained an average IoU value of 0.4056 as a result of the training process. The training IoU and loss graphics are shown in Figures 6, 7 and 8. As in Figures 6, 7, and 8, high IoU with low loss means a low error in the dataset. When the loss starts to decrease, IoU starts to increase in loss and IoU graphic of Unet backbones. This means the model is learning and working fine. As seen in Table 1, the proposed hybrid total loss function contributed positively to the verification performance of Unet backbones. The testing process was tested with 3 Unet backbones to evaluate the performance of the trained dataset. The test results in Table 2 were obtained by using the weights obtained as a result of the training.

As can be seen in Table 2, when MobileNetV2, ResNet34, and VGG16 are used as the backbone of the semantic segmentation network, better results are obtained than the classical Unet architecture. The models chosen as the backbone exceeded the average IoU values they obtained alone when they entered the ensemble decision-making process. It can be seen in Table 2.

Table 1. The effect of hybrid loss function on training performance.

Unet Backbone	Focal Loss / Average IoU	Dice Loss / average IoU	Total Loss / average IoU
ResNet34	0.5483	0.5988	0.6145
MobileNetV2	0.5231	0.5875	0.5957
VGG16	0.5713	0.6049	0.6117

Table 2. Average IoU values of ensemble model and Unet models with different backbones.

Model	Average IoU (Test)
Unet	0.3625
Unet (MobileNetV2)	0.5289
Unet (ResNet34)	0.4685
Unet (VGG16)	0.4106
Ensemble Model (Equal Proportion)	0.5308

The IoU value obtained as an ensemble decision in Table 2 was obtained by influencing the decision-making process of the 3 models with the same weight. In order to maximize the performance of the ensemble decision, it is aimed to influence the ensemble decision by multiplying the weights obtained as a result of the training with a coefficient. The trained backbones and their coefficients on test performance are shown in Table 3.

Table 3. Ensemble model performances with different coefficients.

Backbone and coefficient	Average IoU (Test)
0.3x MobileNetV2 + 0.2xResNet34 + 0.0xVGG16	0.5430
0.3x MobileNetV2 + 0.2xResNet34 + 0.1xVGG16	0.5348
0.3x MobileNetV2 + 0.2xResNet34 + 0.2xVGG16	0.5331
0.3x MobileNetV2 + 0.2xResNet34 + 0.1xVGG16	0.5354

Table 3 shows the effect of Unet backbones with different coefficients on the ensemble decision. The MobileNetV2 backbone, as seen in Table 2, achieved the best test performance alone on the RailSem19 dataset. The ResNet34 backbone also showed the best performance after MobileNetV2. These performances had a direct impact on the ensemble decision-making process. The best performance of the ensemble decision was obtained with the coefficients of 0.3x MobileNetV2 + 0.2xResNet34 + 0.0xVGG16. The ensemble model with these coefficients obtains a 0.5430 average IoU value. A model obtaining an average IoU value greater than 0.5 is considered to have a good performance [24]. The best coefficients for Unet backbones determined were found in the experiments with nested loops.

As seen in Figure 9, the ensemble decision model produced an image similar to the reference image as a prediction result. Ensemble decision-making has improved the average IoU metric as well as improved accuracy.

As seen in Table 4, the ensemble decision model achieved a better accuracy rate than the traditional Unet model, which makes predictions alone. Table 5 shows the performances of Unet models with different backbones included in the decision-making process with their coefficient ratios. The coefficients that maximize the test rate have been determined. As can be seen in Tables 3 and 5, the proposed ensemble model achieved 0.5430 average IoU and 0.8589 accuracy rate. Average IoU and accuracy are standard metrics for semantic segmentation. The accuracy rate metric is biased in distinguishing classes such as the background that covers the majority of the image. Therefore, the average IoU metric is more important for evaluating semantic segmentation performance.

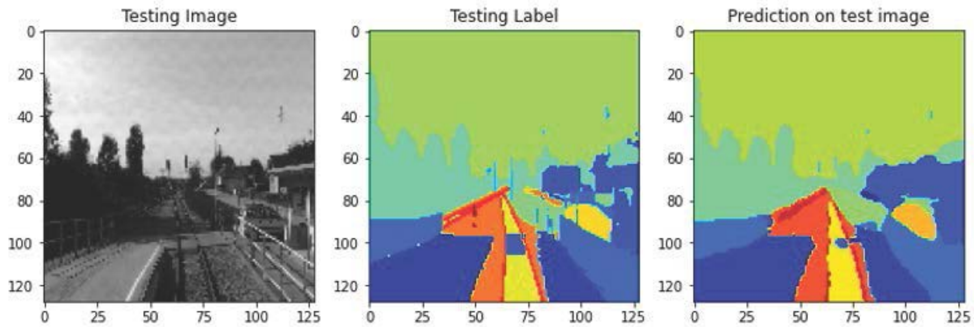

Figure 9. Image obtained by ensemble decision as a result of testing.

Table 4. Accuracy rate of ensemble model (equal proportion) and Unet models with different backbones.

Model	Accuracy rate (test)
Ensemble model (Equal proportion)	0.8575
Unet (MobileNetV2)	0.8423
Unet (ResNet34)	0.8466
Unet (VGG16)	0.8269

Table 5. Accuracy rate values of ensemble model (different coefficients) and Unet models with different backbones.

Backbone and coefficient	Accuracy rate (test)
0.3x MobileNetV2 + 0.3xResNet34 + 0.2xVGG16	0.8589
0.3x MobileNetV2 + 0.3xResNet34 + 0.1xVGG16	0.8586
0.3x MobileNetV2 + 0.2xResNet34 + 0.2xVGG16	0.8572
0.3x MobileNetV2 + 0.1xResNet34 + 0.0xVGG16	0.8459

4. Conclusion

This work presented a methodology to effectively segment the RailSem19 dataset. In the study, 8500 images were examined. In other studies in the literature, RailSem19 has applied deep learning methodologies to segment images of the dataset, but none of them focused on using the segmentation models of Unet backbones MobileNetV2, ResNet34, and VGG16 together. This shows the contribution of this study to the railway sector. Table 6 shows that the performance of the model proposed in the study compares with other studies in the literature.

Table 6. Comparison of the proposed model with studies in the literature.

Reference	Model	Average IoU (Test)	Number of classes
Our Model	0.3x MobileNetV2 + 0.2xResNet34 + 0.0xVGG16	0.5430	29
[25]	Unet	0.2960	19
[26]	SegNet3	0.5364	4
[27]	UNet16	0.4258	2

As seen in Table 6, the proposed ensemble decision model was compared with other studies in the literature based on the average IoU metric. The proposed model segmented more classes compared to other studies in the literature. The literature studies mentioned focus on various aspects of railway safety and analysis using semantic segmentation and deep learning techniques. These studies collectively highlight the importance of semantic segmentation in railway safety and analysis, covering areas such as material identification, anomaly detection, defect detection, and mapping of railway infrastructure. Compared to literature studies, our work focuses on effectively segmenting the RailSem19 dataset using the segmentation models of Unet backbones MobileNetV2, ResNet34, and VGG16. This approach is a novel contribution to the railway sector, as none of the previous studies in the literature used this combination. Our proposed ensemble model achieved an average IoU value of 54.30% and segmented more classes compared to other studies. Overall, our work contributes to the railway safety field by emphasizing the importance of detecting faults and intrusions in railways using deep learning-based image segmentation algorithms. First, the article is limited to articles in Turkish and English. This has led to the exclusion of some related articles. Finally, the development process took longer as a strong test environment was not available. In this study, the importance of detecting faults in railways in terms of railway safety has been emphasized. As the use of railway transportation will increase day by day, the problems arising will increase day by day. In future studies, it is planned to transmit possible foreign objects (human, car..) to the railway conductor in the segmented images around the railway.

Acknowledgments

This work was supported by The FUBAP (Fırat University Scientific Research Projects Unit) under grant no: ADEB.2022.02.

References

[1] Hafiz AM, Bhat GM. A survey on instance segmentation: state of the art. International journal of multimedia information retrieval 2020; 9 (3): 171-189.

[2] Yurdakul EM. Türkiye'de Ulaşım Altyapısı ve Ekonomik Kalkınma İlişkisi: ARDL Sınır Testi. Kent Akademisi 2022; 15 (2): 666-680 (in Turkish).

[3] Giben X, Patel VM, Chellappa R. Material classification and semantic segmentation of railway track images with deep convolutional neural networks. In 2015 IEEE International Conference on Image Processing (ICIP); Quebec City, QC, Canada; 2015. pp. 621-625. doi: 10.1109/ICIP.2015.7350873

[4] He Z, Tang P, Jin W, Hu C, Li W. Deep semantic segmentation neural networks of railway scene. In 2018 37th Chinese control conference (CCC); Wuhan, China; 2018. pp. 9095-9100. doi: 10.23919/ChiCC.2018.8483877

[5] Furitsu Y, Deguchi D, Kawanishi Y, Ide I, Murase H et al. Semantic segmentation of railway images considering temporal continuity. In Pattern Recognition: 5th Asian Conference; Auckland, New Zealand; 2019. pp. 639-652.

[6] Jahan K, Umesh JP, Roth M. Anomaly Detection on the Rail Lines Using Semantic Segmentation and Self-supervised Learning. 2021 IEEE Symposium Series on Computational Intelligence (SSCI); Orlando, FL, USA; 2021. pp. 1-7. doi: 10.1109/SSCI50451.2021.9659920

[7] Wang H, Li M, Wan Z. Rail surface defect detection based on improved Mask R-CNN. Computers and Electrical Engineering 2022; 102: 108269.

[8] Li L, Sun R, Zhao S, Chai X, Zheng S, Shen R. Semantic-Segmentation-Based rail fastener state recognition algorithm. Mathematical Problems in Engineering 2021; 1-15.

[9] Grandio J, Riveiro B, Soilán M, Arias P. Point cloud semantic segmentation of complex railway environments using deep learning. Automation in Construction 2022; 141: 104425.

[10] Zouaoui A, Mahtani A, Hadded MA, Ambellouis S, Boonaert J et al. RailSet: A Unique Dataset for Railway Anomaly Detection. 2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS); Genova, Italy; 2022. pp. 1-6. doi: 10.1109/IPAS55744.2022.10052883.

[11] Alexandrescu AR, Manole A. A Dynamic Approach for Railway Semantic Segmentation. Studia Universitatis Babes-Bolyai, Informatica 2022; 67(1).

[12] Zhang Y, Li K, Zhang G, Zhu Z, Wang P. DFA-UNet: Efficient Railroad Image Segmentation. Applied Sciences 2023; 13 (1): 662.

[13] Çetintaş D, Tuncer T. Determining the Type of Document Read Using Eye Movement Properties by Hybrid CNN Method. Traitement du Signal 2022; 39 (4).

[14] Aydın İ, Sevi M, Salur MU, Akın E. Defect classification of railway fasteners using image preprocessing and a lightweight convolutional neural network. Turkish Journal of Electrical Engineering and Computer Sciences 2022; 30 (3): 891-907.

[15] Menteşe E, Hançer E. Histopatoloji görüntülerde derin öğrenme yöntemleri ile çekirdek segmentasyonu. Avrupa Bilim ve Teknoloji Dergisi 2020; 95-102 (in Turkish).

[16] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference; Munich, Germany 2015; 234-241.

[17] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. pp. 770-778.

[18] Bendiabdallah MH, Settouti N. A comparison of U-net backbone architectures for the automatic white blood cells segmentation. WAS Science Nature (WASSN) ISSN 2021; 2766-7715: 4 (1).

[19] Jadon S. A survey of loss functions for semantic segmentation. In 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB); 2020. pp. 1-7.

[20] Zhao R, Qian B, Zhang X, Li Y, Wei R et al. Rethinking dice loss for medical image segmentation. In 2020 IEEE International Conference on Data Mining (ICDM); 2020. pp. 851-860.

[21] Chang J, Zhang X, Ye M, Huang D, Wang P et al. Brain tumor segmentation based on 3D Unet with multi-class focal loss. In 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI) 2018; 1-5.

[22] Kang D, Benipal SS, Gopal DL, Cha YJ. Hybrid pixel-level concrete crack segmentation and quantification across complex backgrounds using deep learning. Automation in Construction 2020; 118: 103291.

[23] Zendel O, Murschitz M, Zeilinger M, Steininger D, Abbasi S et al. Railsem19: A dataset for semantic rail scene understanding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops 2019.

[24] Cira CI, Manso-Callejo MÁ, Alcarria R, Fernandez Pareja T, Bordel Sanchez B et al. Generative learning for postprocessing semantic segmentation predictions: a lightweight conditional generative adversarial network based on Pix2pix to improve the extraction of road surface areas. Land 2021; 10 (1): 79.

[25] Katar O, Duman E. Automated Semantic Segmentation for Autonomous Railway Vehicles. Tehnički glasnik 2022; 16 (4): 484-490.

[26] Daimary D, Bora MB, Amitab K, Kandar D. Brain tumor segmentation from MRI images using hybrid convolutional neural networks. Procedia Computer Science 2020; 167: 2419-2428.

[27] Fu X, Li K, Liu J, Li K, Zeng Z et al. A two-stage attention aware method for train bearing shed oil inspection based on convolutional neural networks. Neurocomputing 2020; 380: 212-224.