International Journal of Engineering and Geosciences
https://dergipark.org.tr/en/pub/ijeg
e-ISSN 2548-0960

Performance analysis of rule-based classification and deep learning method for automatic road extraction

Zeynep Bayramoğlu[*1], **Melis Uzar** [1]

[1]*Yıldız Technical University, Department of Geomatics Engineering, Türkiye*

Keywords
Automatic road extraction
Remote Sensing
Deep Learning
Ontology
UAV

Research Article
DOI: 10.26833/ijeg.1062250

Received: 24.01.2022
Accepted: 25.02.2022
Published: 13.04.2022

Abstract

The need for accurate and up-to-date spatial data by decision-makers in public and private administrations is increasing gradually. In recent decades, in the management of disasters and smart cities, fast and accurate extraction of roads, especially in emergencies, is quite important in terms of transportation, logistics planning, and route determination. In this study, automatic road extraction analyses were carried out using the Unmanned Aerial Vehicle (UAV) data set, belonging to the Yildiz Technical University Davutpasa Campus road route. For this purpose, this paper presents a comparison between performance analysis of rule-based classification and U-Net deep learning method for solving automatic road extraction problems. Objects belonging to the road and road network were obtained with the rule-based classification method with overall accuracy of 95%, and with the deep learning method with an overall accuracy of 86%. On the other hand, the performance metrics including accuracy, recall, precision, and F1 score were utilized to evaluate the performance analysis of the two methods. These values were obtained from confusion matrices for 4 target classes consisting of road and road elements namely road, road line, sidewalk, and bicycle road. Finally, integration of classified image objects with ontology was realized. Ontology was developed by defining four target class results obtained as a result of the rule-based classification method, conceptual class definition and properties, rules, and axioms.

1. Introduction

Roads are one of the important parts of the transportation infrastructure in spatial systems, especially for cities. In addition to the general management of urban areas, road maps are basic components of effective and planned management before and after disasters. Especially in emergencies, the management of roads with fast and up-to-date data is vital in solving transportation problems. To solve the requirements, the extraction of roads is carried out by spending a long time with traditional methods (e.g., manual digitizing) depending on the operators. Road extraction by the use of the classical methods such as pixel-based classification carries forward to time and resolution-related errors with misclassification problems. Moreover, most of the results of these methods are full of incomplete and erroneous data. Such errors negatively affect the accuracy of analyses for extracting roads. For that reason, automatic road extraction is aimed to solve these problems. At the stage of processing the data, determining the road network with photogrammetry and remote sensing methods is significant in terms of obtaining up-to-date, fast, and accurate data. For this purpose, in addition to using existing orthophoto and satellite images, Unmanned Aerial Vehicles (UAV) are preferred, which provide the opportunity to obtain fast, economical, and up-to-date data. The method of classifying objects to make inferences from the image is a method frequently

* **Corresponding Author**

*(zynpbayramoglu@gmail.com) ORCID ID 0000-0003-0324-3561
(auzar@yildiz.edu.tr) ORCID ID 0000-0003-0873-3797

Cite this article

Bayramoğlu, Z., & Uzar, M. (2023). Performance analysis of rule-based classification and deep learning method for automatic road extraction. International Journal of Engineering and Geosciences, 8(1), 83-97

preferred by researchers [1]. Generally, object-based and pixel-based classification methods are used for object extraction. In the case of non-homogeneous objects with various spectral values on the image, class confusion problems occur when using pixel-based methods [2]. For that reason, an object-oriented rule-based classification method is preferred to eliminate the problem of class confusion in automatic object extraction. In this method, spectral information with the properties of objects such as size, shape, texture, location was used to develop rules for extracting target classes. The analyses of segmentation and classification steps of the method are utilized to increase the accuracy of the target classes. A summary review of the recent state of the art in road extraction can be found in [3-4]. In addition, [5-7] utilized object-oriented classification methods for road extraction.

Deep learning has become popular in big data analysis and plays a big role in object detection, semantic segmentation and classification. Deep learning can be a bridge filling the gap between object extraction and remote sensing. The convolutional neural network (CNN) model, which can automatically extract objects and make classifications through sequential convolutional and fully connected layers, is used to improve efficiency and accuracy in remote sensing applications with the development of deep learning architectures. U-Net is the one of the most famous CNN architectures that is widely used for object segmentation and extraction. The U-Net architecture is based on fully convolutional networking. The advantage of the U-Net is that it works with fewer training images, providing more precise segmentation and allowing very fast results.

Today, the use of deep learning algorithms for the automatic extraction of roads [8-10] and feature information with machine learning has become prevalent. Particularly, studies conducted with remote sensing images and U-Net architecture are becoming widespread, [11-14]. In addition, there are also scientific studies in the literature that perform the classic methods with deep learning methods [15-16]. On the other hand, [17-21] employ deep learning methods for extraction of roads.

Ontology is used for presenting data and performing logical inference operations, as well as for leveraging and sharing these data by different systems and processing them syntactically and semantically by computers [22]. Ontology has been developed to identify and explain related concepts, establish the logical structure, and create taxonomic relationships. The difference in the result obtained from the image due to the user interpretation of the data creates the semantic gap problem [23-24]. To solve this problem, it is necessary to customize the a priori and extract information from the image by incorporating the conceptualization into image analysis systems [25]. Therefore, there is a need for a structure that can formalize field knowledge and expert knowledge. This structure should be in a state that can be applied to the image which will be classified. The concept of ontology, which helps to conceptualize a particular field, responds to this need by connecting spatial contents and presenting a formalized structure. The availability of geospatial data with ontology integration is increasing in recent years. Sener and Uzar [26] developed ontology by defining conceptual class definitions, object and data properties, rules, and axioms. Another study of geospatial ontology to define multi-scale representations and detailed cardinality relations of the building features was developed by [27].

In this study, automatic road extraction possibilities were investigated by using rule-based classification and U-Net [28] architecture. In addition, an ontology with classified objects was developed. The remainder of this paper is organized as follows. The properties of the study area and the dataset obtained by the UAV system are presented in the next section, followed by an explanation of the methodology adopted. The results of the experiment are reported along with an accuracy assessment of the automatic road extraction results with the comparison of two methods, i.e., rule-based classification and U-Net deep learning analyses, in the Results and Discussion section, and the conclusion is given in the last section.

2. Method

Object-oriented rule-based classification and U-Net deep learning methods were utilized for the automatic road extraction analysis, with the data obtained with UAV. The proposed methodology workflow is given in Fig. 1. In the rule-based classification method, orthophoto, DSM, and DTM were produced during data preparation. For automatic road extraction with deep learning U-Net architecture, data preparation, i.e., labeling, pre-processing, training and test data separation, has been completed. The performance analysis of target classes (road, sidewalk, road line, and bicycle road) was evaluated with both methods and then accuracy analyses were performed. Thus, it is aimed to solve the automatic road extraction problems with performance analysis of two methods using remote sensing and photogrammetric data. In this study, Pix4D software was used for photogrammetric data production processes using UAV images. The e-Cognition Developer software was used to develop a rule set with segmentation and classification analyses. In the deep learning method, the labeling processes of the target classes were created using GNU Image Manipulation Platform (GIMP) software. Accuracy, F1 score, precision, and recall parameters were evaluated for performance analysis of two methods. Finally, an ontology was developed to conceptualize the semantic information of objects and object classes with each other. And then the integration of classified image objects with ontology was carried out. Java-based and open-source Protégé was utilized for the creation of the semantic model and ontology development processes.

2.1. Study area and dataset

In this study, a road route was chosen as a study area found in the Davutpasa Campus of Yildiz Technical University in Istanbul, Turkey (Fig. 2).

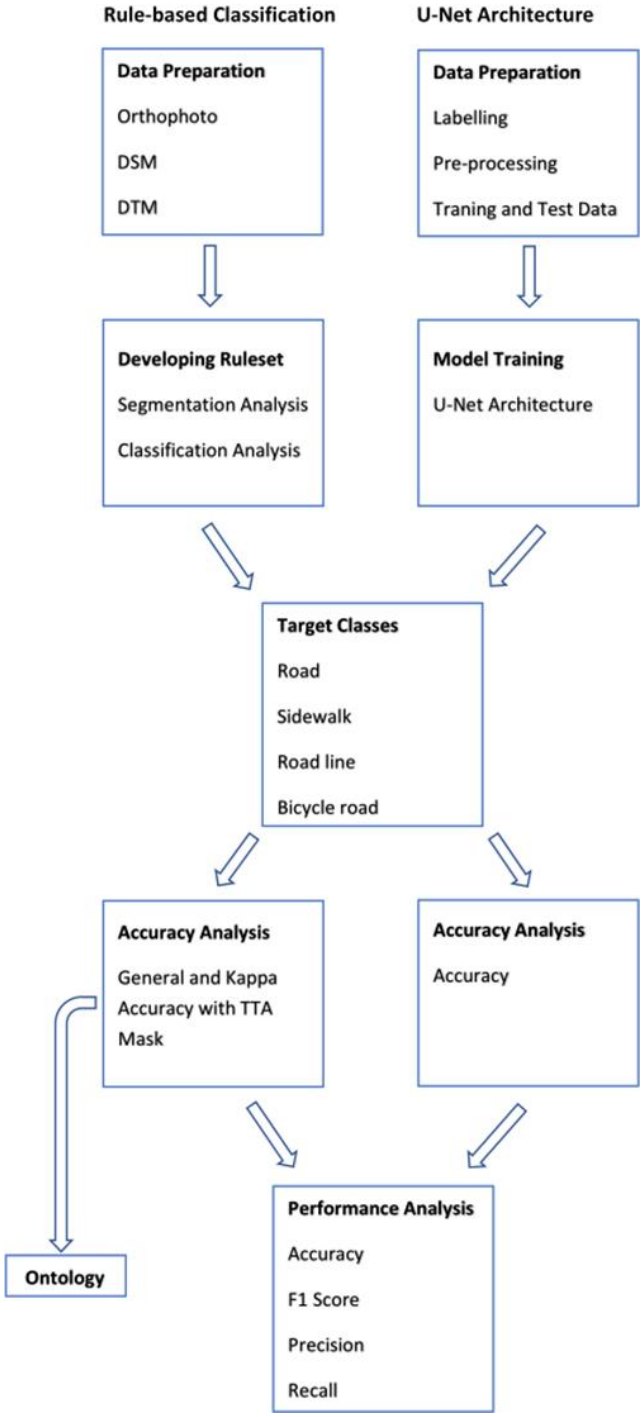

Figure 1. Workflow for the proposed methodology

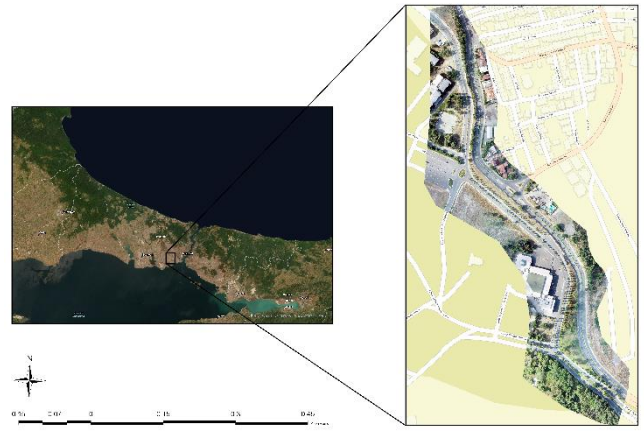

Figure 2. Study area

The data set was obtained with the UAV system, including DJI Phantom 4 Pro, GPS/IMU, and a digital camera positioned at the same platform. The UAV flight was carried out on 25.07.2019 with a flight altitude of 103 m. The images collected with the UAV were examined and the images with high error values and missing orientation parameters were eliminated. At this point of view, 157 images out of 182 images were used for method analysis. The FC6310R digital camera used in the UAV system was calibrated and Ground Control Points (GCP) were also measured. GCP and Root Mean Square (RMS) Errors are shown in Table 1.

Table 1. GCP and RMS Error

GCP	Error X(m)	Error Y(m)	Error Z(m)	Error (pixel)
10	0.004	0.001	-0.000	0.567
11	0.013	0.010	0.014	0.544
12	0.006	-0.013	0.038	0.340
RMS Error(m)	0.008	0.0094	0.0236	

After the measurement of the GCPs, the three-dimensional dense point cloud data was automatically generated in the TUREF/TM30 coordinate system. It was determined that the parameter obtained in the camera calibration report, camera distortion errors, and standard deviations were suitable for photogrammetric data generation [29]. A three-dimensional model was produced by a mesh model and texture from the point cloud data. A digital surface model (DSM), a digital terrain model (DTM) and an orthophoto with 2.55 cm ground sampling interval were produced from point cloud (Fig. 3).

2.2. Rule-Based classification method

In this study, by developing a rule set for automatic road extraction with object-based image analysis method, target classes were defined as road, road line, sidewalk, and bicycle road. And also additional classes, namely, green area, shadow, ground, sidewalk, and non-road were created to avoid class confusion. At the first stage, segmentation and classification analyses were performed to solve the class confusion problem, and rules were created as a result of the appropriate parameter estimation. In the second stage, a rule set was developed using brightness, texture analysis with a gray level co-occurrence matrix (GLCM) entropy, normalized digital surface model (NDSM) analysis, band combinations, and digital image processing techniques. Band combinations used to develop the ruleset are shown in Table 2 (Eq. 1-6). At the final stage, the error matrix based on the Test and Training Area (TTA) mask method was performed to test the accuracy of the proposed approach. In the TTA mask method, kappa, the overall, producer's and user's accuracies were computed. As a result, target classes are extracted using an original ruleset after the accuracy analysis for the performance analysis of two methods.

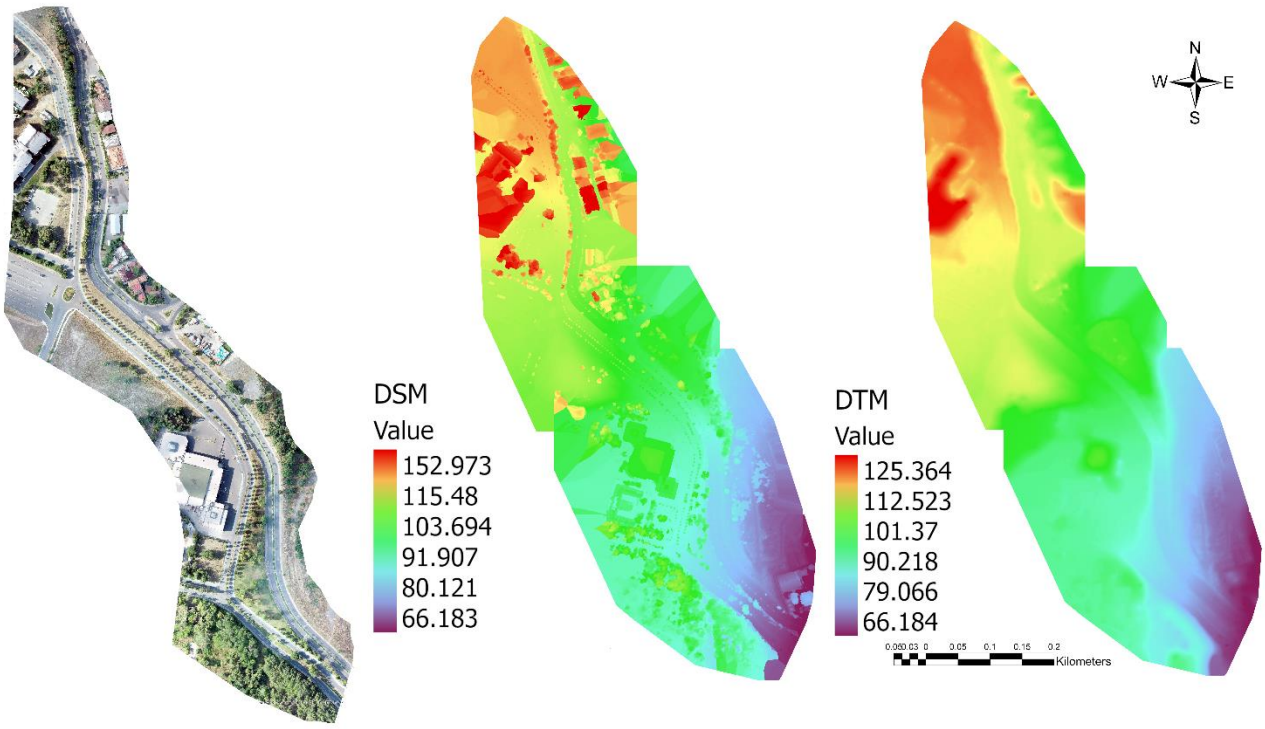

Figure 3. Orthophoto, DSM, and DTM

Table 2. Band combinations

Method	Abbr.	Formula	Eq. No
Brightness	B	$\frac{R+G+B}{3}$	1
Red Green Band Difference	R-G	$R-G$	2
Excess Index	RGI	$(2xG)-B-R$	3
Green Leaf Area Index	GLI	$\frac{((2\times G)\ -\ B-R)}{((2\times G)\ +\ B+R)}$	4
The Triangular Greenness Index	TGI	$G–(0.39R–0.69B)$	5
Haralick	H	$\frac{(0.5*(R-G)+(R-B))}{((R-G)^2+(R-B)(G*B))^{0.5})}$	6

2.3. *U-Net Architecture*

A Fully Convolutional Neural Network (FCNN) based U-Net architecture, which allows provides the use of global location and global context simultaneously with a few training samples, consists of the contracting path and the expansive path, also known as the encoder and decoder for analysis and synthesis path, respectively. The contracting path utilizes a typical CNN. It consists of two consecutive 3x3 convolutions followed by a batch normalization layer, a Rectified Linear Unit (ReLU) activation unit and a 2x2 max-pooling layer. This order is repeated as the number of downsampling operations. It should be noted that after each downsampling, the number of feature channels to be extracted is doubled [30] (Bayrak 2020). The expansive path includes an upsampling of the feature map by 2x2 transpose convolution to halve the number of feature channels and concatenation with the corresponding feature map from the contracting path. The final layer employs a 1x1 convolutional layer to classify each pixel. U-Net architecture for road extraction is given in Fig. 4.

2.3.1. *Training Details of U-Net Architecture*

Target classes (bicycle road, sidewalk, road, and road line) were annotated in GIMP software. U-Net implementation was performed using Keras 2.3.1 and the train/test process was made on a single NVIDIA RTX 2080 GPU. Firstly, the study area and corresponding annotated image were split into the sub-images with the size of 512×512 pixels and the numbers of train/test images and batch size were set as 320/140 and 8, respectively. Secondly, minimum maximum (min-max) normalization technique was applied to images due to data imbalance between classes. No pre-trained models were loaded; hence, the network was trained from scratch. Finally, the initial number of feature channels was set as 64 and after each downsampling, it is doubled and iterated 4 times. Adam optimizer with a learning rate of 1e-3 and exponential learning rate decay of 0.90 applied after each epoch (Table 3).

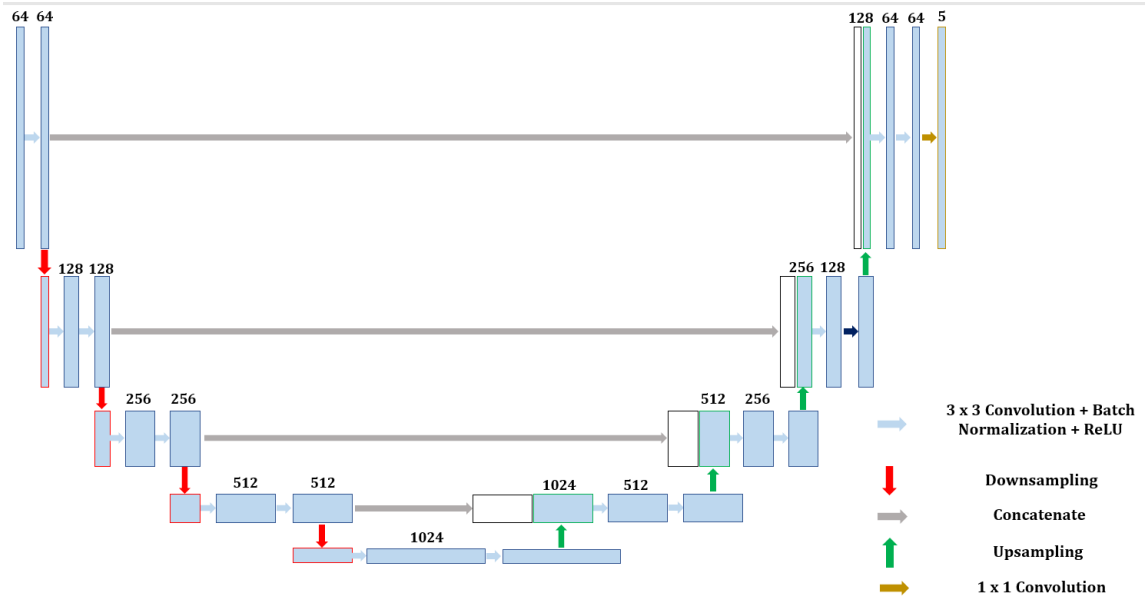

Figure 4. U-Net architecture [28]

Table 3. GCP and RMS Error

Train/Test Ratio	Optimizer	Learning Rate	Epoch Number
320/140	Adam	0.001	30

The model hyperparameters are given in Table 3. Loss calculation was performed by Dice Similarity Coefficient (DSC) [31] loss function in 30 epochs (Fig. 5). To evaluate the predictions, final feature maps were thresholded at 0.5 to obtain the masks, before comparison with ground truth annotations.

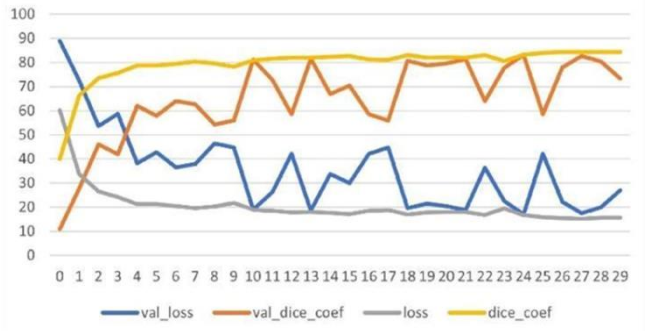

Figure 5. U-Net results

2.4. Developing ontology with classified image objects

In this study, shapefile data format was used for exporting data. The main reason for this approach is to preserve the properties of the classified image objects and to use them in information representation and ontology. This stage includes class definitions, object and data relations definitions, axiom, and rule definitions. The defined classes are shown in Fig. 6.

For the integration of classified object images with ontology, the data in shapefile format was converted into Web Ontology Language (OWL) format using the Comma Separated Value (CSV) as an intermediate format. After that step, the semantic model definition was completed. The QGIS program was used for conversion to CSV data format. The properties of image objects classified in CSV format are preserved. The data in CSV format was converted to an .xlsx file with excel and used as input in the ontology development platform. The steps required to create a file in OWL format are shown in Fig. 7.

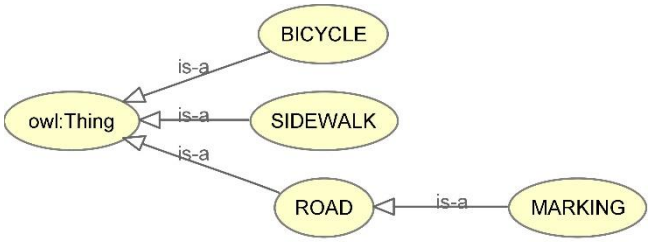

Figure 6. Classes in Ontology

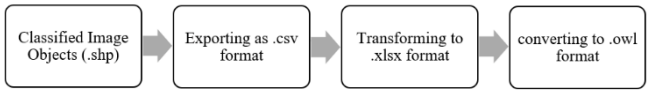

Figure 7. Converting shapefile to .owl format

After developing an ontology and integrating object instances into the ontology, the first requirement is the definition of data properties. For this purpose, the data properties defined for the integration of image objects into the ontology are given in Fig. 8.

These data features were determined by analysis after image segmentation. After defining data properties, axioms and rules are defined and ontology is developed. The rules used in transferring the data to the Protege program are shown in Fig. 9. An example representation of the developed ontology is given in Fig. 10.

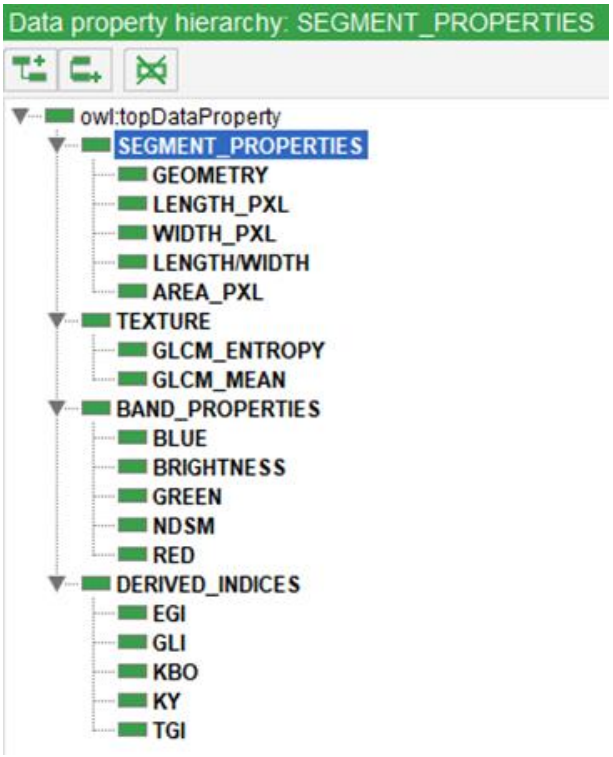

Figure 8. Defined data properties for Ontology

Figure 9. The image of the transfer process of rules

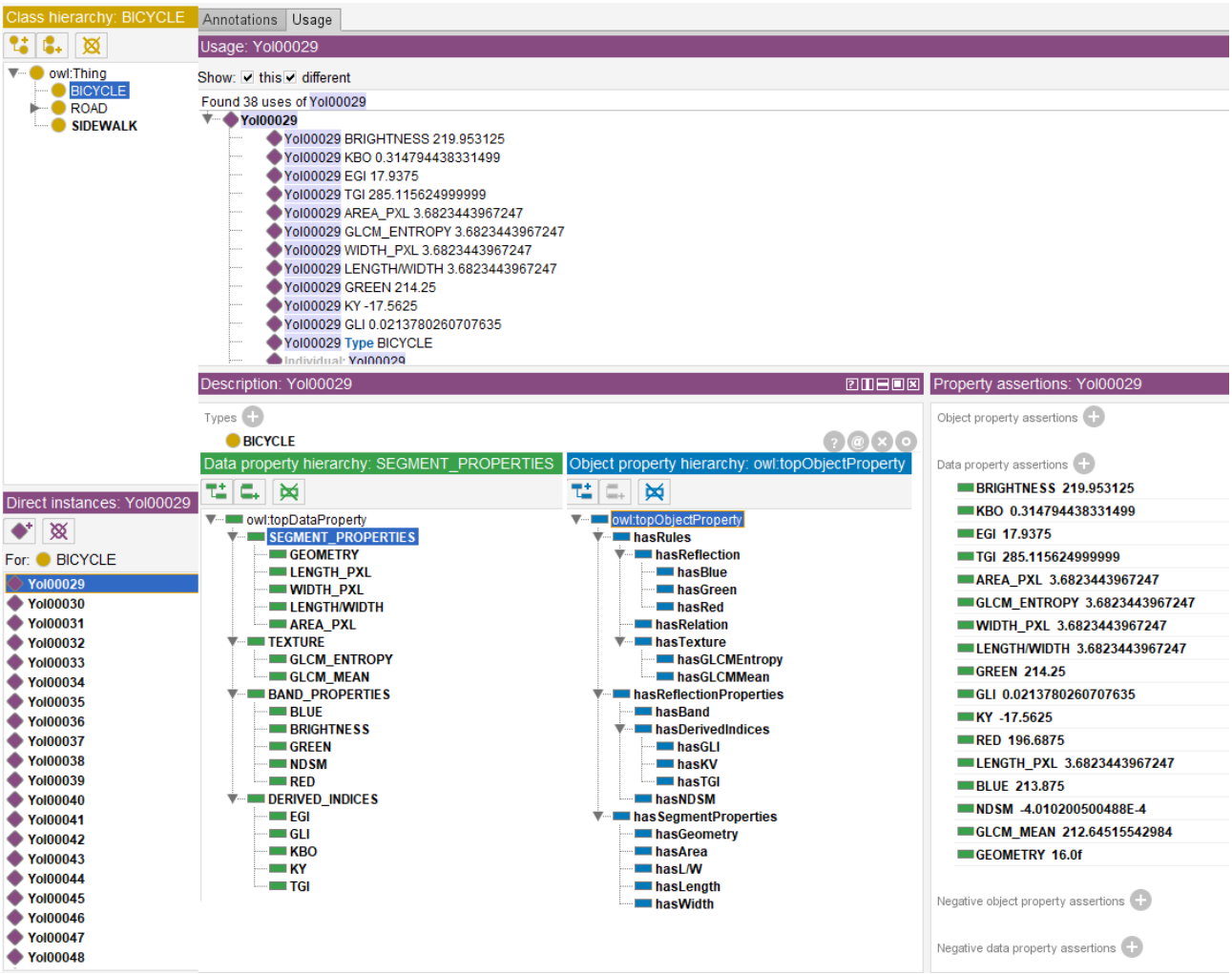

Figure 10. Ontology developed for the target classes

2.4.1. Evaluation parameters for performance analysis

A confusion matrix was used to perform the performance analysis in this study (Table 4). A confusion matrix is made up of four components, namely, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [32-33]. The assessment metrics engaged here are accuracy, precision, recall, and F1 score (Eq.7-10). True Positive (TP) refers to the number of predictions that correctly predict the object in reality. True Negative (TN) refers to the number of predictions that correctly predict the object that does not exist in reality. False Positive (FP) refers to the number of predictions in which it incorrectly predicts the object that does not exist in reality. And finally, False Negative (FN) refers to the number of predictions that incorrectly predict the object in reality. Precision shows how many of the positively predicted objects are Positive. The sensitivity indicates how many of the transactions that should be predicted as positive are predicted as positive. F1 Score value shows the harmonic mean of Precision and Sensitivity values. Accuracy is used to find the rate of correct predictions.

Table 4. Confusion matrix and evaluation metrics

		Prediction	
		Positive	Negative
Reference	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

$$Precision=TP/ (TP+FP) \quad (7)$$

$$Recall= TP/ (TP+FN) \quad (8)$$

$$F1\ Score = 2*((P*R)/ (P+R)) \quad (9)$$

$$Accuracy= (TP+TN)/ (TP+FP+FN+TN) \quad (10)$$

3. Results and Discussion

3.1. Automatic extraction of road using the rule-based classification method

Segmentation, which is the first stage of the rule-based classification method, is the process of dividing the image into smaller parts by creating meaningful segments. These segments are created according to scale, shape, compactness, contrast difference, brightness, and statistical parameter values. A multi-resolution segmentation algorithm using shape, compactness, and scale factor was used. Analyses were carried out to determine the appropriate parameters and as a result, it was decided to prefer shape, integrity, and scale factors as 0.3, 0.8, and 20, respectively. The classes, parameters and analysis results created for the rule set developed for automatic road extraction are given in Table 5.

Table 5. Band combinations

Parameter	Analysis	Execution Class	Target Class
NDSM data	NDSM≥1		Non-road
GLI combination	threshold value ≥0.1		Green area
Length Width analysis	threshold value≥4		Line
GLCM Entropy	2≥threshold value≤6		Road
GLCM Mean	threshold value ≤170		Shadow
Classification		Unclassified	Ground
R-G	threshold value ≤-16	Road	Bicycle road
EGI	-8≥ threshold value ≤9	Bicycle road	Road
TGI	285≤ threshold value≥ 350	Road	Sidewalk
Brightness	230≤ threshold value ≥ 350	Road, Ground	Sidewalk
Hue	0.56≤ threshold value≥ 0.6	Shadow	Road
EGI/Road relation	value ≤-4 / 0.4≥ value	Shadow	Road
Area	100000	Road	Road
Boundary relationship with road	threshold value ≥0.1	Line	Road line
Area	600	Bicycle road	Ground

In this study, a rule set for automatic road extraction was developed by using the appropriate parameters obtained as a result of the analysis to prevent class confusion. Initially, the normalized digital surface model (NDSM) was obtained to distinguish between ground and non-ground objects. In this analysis, segments with an NDSM value greater than 1 were assigned as a non-road class. The proposed method improves the target class with rules using brightness, texture analysis with a gray level co-occurrence matrix (GLCM) entropy, NDSM analysis, band combinations, and digital image processing techniques. To classify the segments of the road, Haralick GLCM entropy texture analysis was performed and values between 2≥ value ≤ 6 were assigned as road class. The bicycle roads in the study area were obtained by taking the difference of the red and green bands. And the green leaf index (GLI) band combination was used to classify the green areas. Excess

green index (EGI) and triangular greenness index (TGI) band combinations were used to generate the road class and improve the classification.

One of the most common problems in object classification analysis is a shadow in the study area. Hue analysis was used to separate the road class from the shadow class. As a result, the classification was performed with the defined rules, and the results were improved with additional classes; Green area, Line, Shadow, Ground, Non-road to accurately represent the target classes. The processing step of target classes' improvement was implemented in Table 8 with several special indices and methods, such as boundary relationship, area determination, rectangular fitting, and merge classes. The image of classification results of additional and target classes is given in Fig. 11. Finally, automatic road extraction was performed using the defined rules with the proposed approach as shown in Fig. 12. Subsequently, TTA Mask analysis was used to evaluate the accuracy. For the selection of sample areas to be used in the test, multi-resolution segmentation was performed with the same parameters at a different level. TTA Mask was created with selected sample segments and the classes obtained in the study were compared with the reference data set. As a result of the accuracy analysis, 89% and overall accuracy of 95% were yielded for the target classes (Table 6).

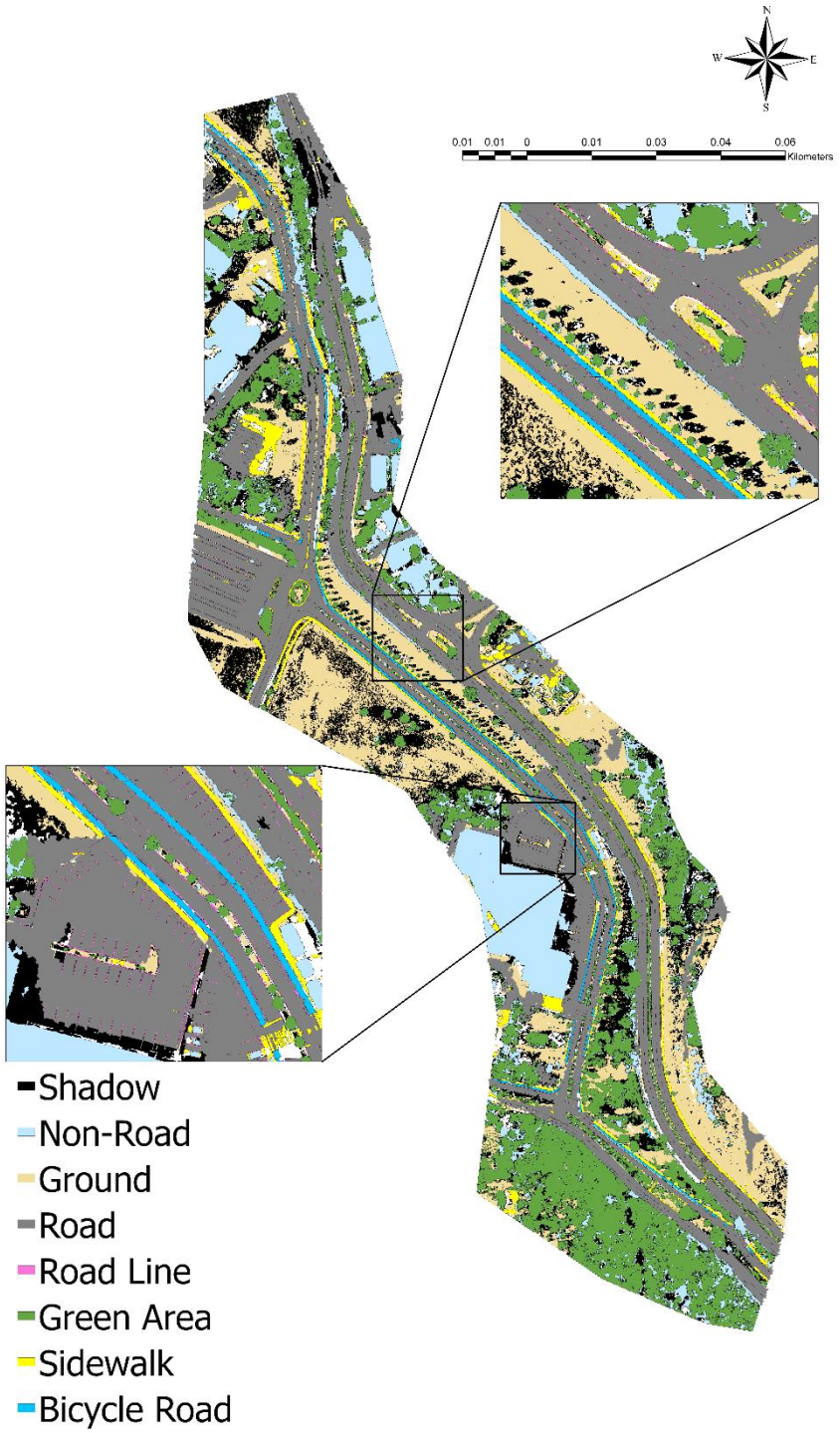

Figure 11. The image of classification results of additional and target classes

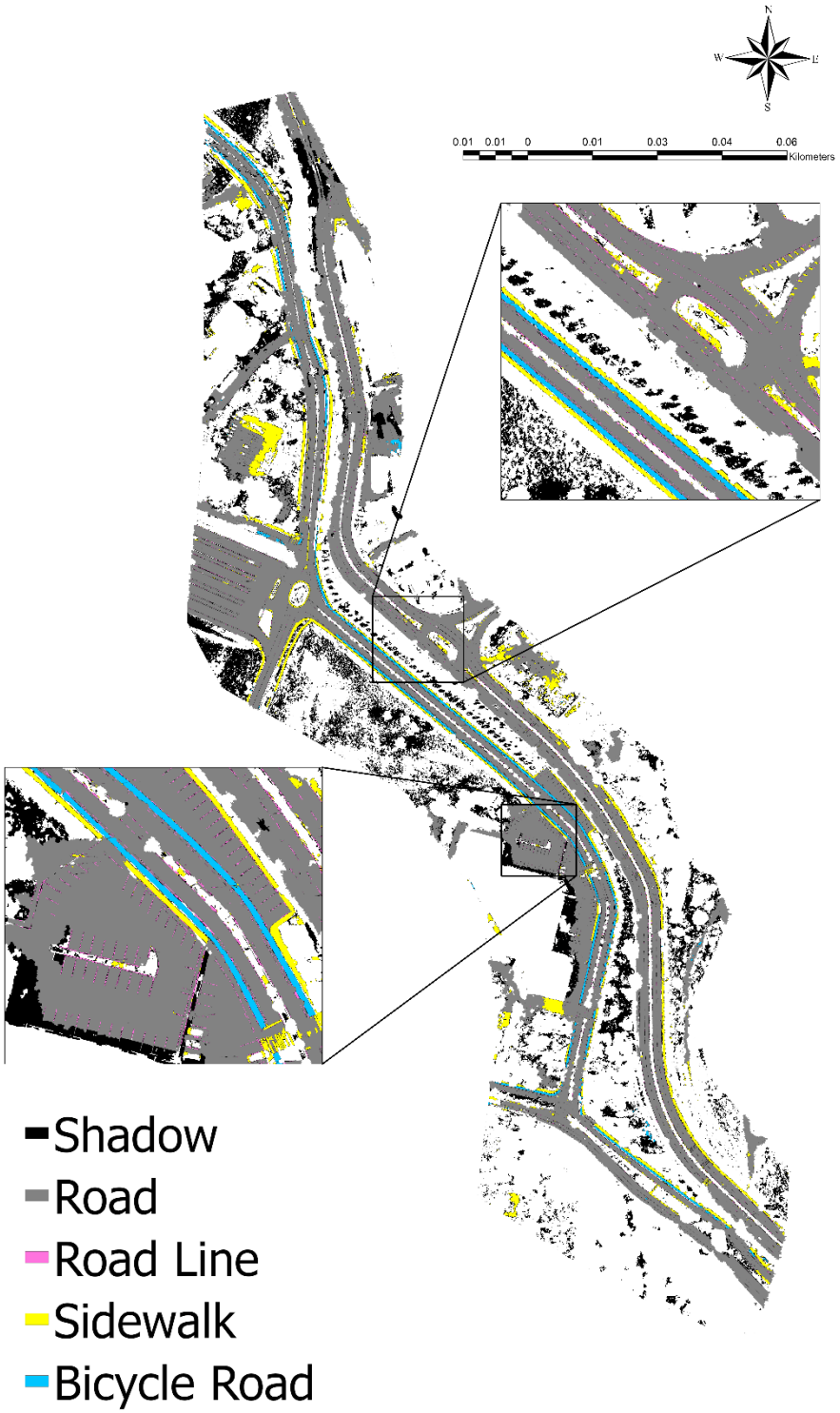

Figure 12. The image of the target classes and shadow class with the rule-based classification method

Table 6. Accuracy analysis results of rule-based classification

User/Reference Class	Bicycle Road	Side walk	Road	Road Line	Sum	User
Bicycle Road	65903	0	0	0	65903	1.00
Side walk	0	187654	4152	2425	194231	0.97
Road	14194	27847	907900	3364	953305	0.95
Road Line	371	0	2091	35819	38281	0.94
Sum	80468	215501	914143	41608	1251720	
Producer	0.82	0.87	0.99	0.86	Overall Acc.	0.95
					Kappa	0.89

3.2. Automatic extraction of road using U-Net architecture

To obtain the U-Net architecture's performance, a quantitative comparison between ground truth annotation and segmented images was employed by confusion matrix elements. Table 7 which covers the results of U-Net architecture includes Accuracy, Sensitivity, Specificity, and F1-Score metrics for each target class. According to Table 7, U-Net architecture achieved 92.4%, 97.9%, 70.5%, and 97.9% accuracy for sidewalk, road line, road, and bicycle road classes, respectively. However, U-Net architecture obtained 0 value for road line class in terms of Sensitivity and F1-Score value, due to lack of road line class in the study area. It means that U-Net architecture was under-fitted for the road line class. Even though the accuracy value of sidewalk and bicycle road classes' accuracy is greater than 90%, the Sensitivity and F1-Score values (0.320 and 0.639) of these classes show that U-Net architecture was under-fitted for both classes, too. Besides that, 61.7% and 50.3% for Sensitivity and F1-Score values for road class proved the same reason, so the lack of data in the small study area. The image of U-Net's classification results of target classes is given in Fig. 13.

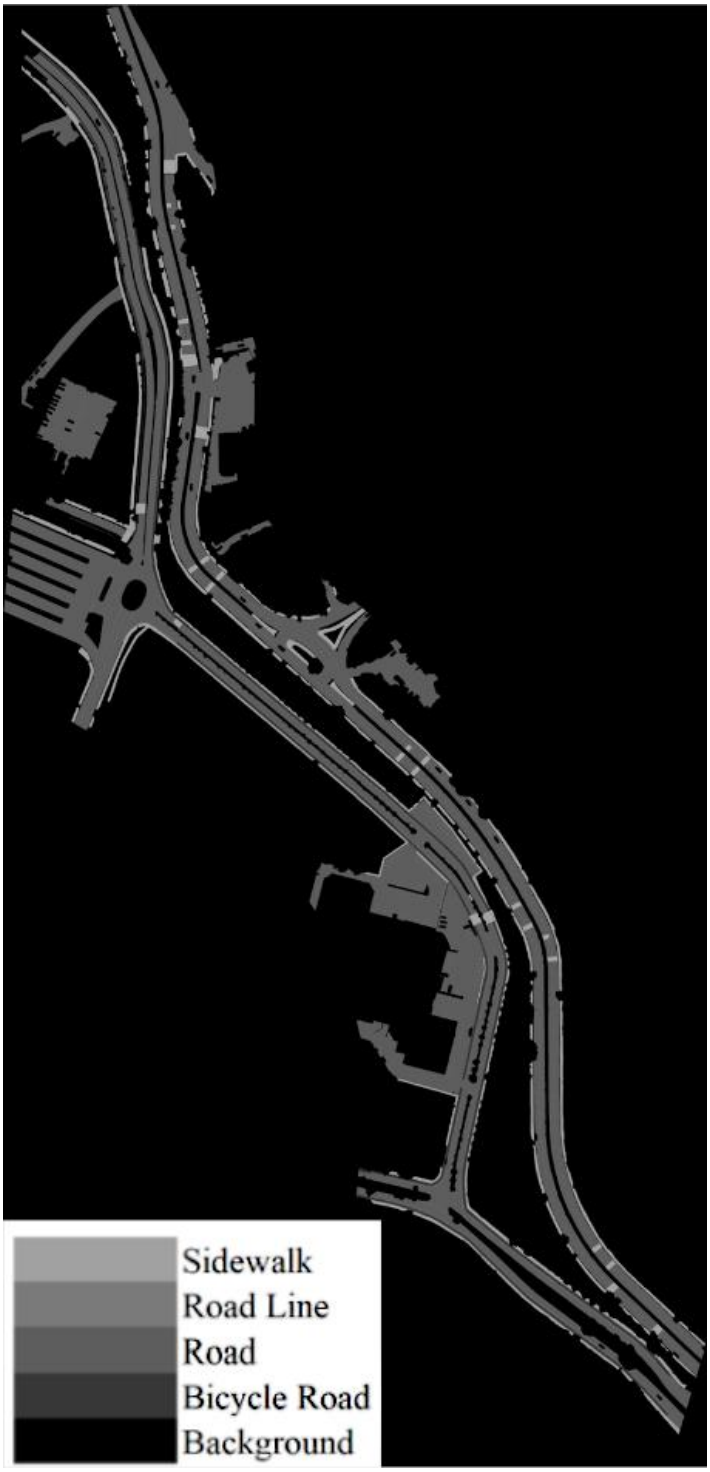

Figure 13. The image of the target classes with U-Net architecture

3.2.1. Performance analysis of target classes extraction with rule-based classification and U-Net architecture

In this study, performance analysis of object-oriented rule-based classification and U-Net architecture was produced by a confusion matrix according to accuracy, recall, precision, and F1 score values. As can be seen in Table 8, the accuracies of the rule-based classification method and U-Net architecture were 96% and 71% for the road class, 99% and 98% for the road line class, 97% and 92% for the sidewalk class, and 92% for the bicycle road class 99% and 98%, respectively.

Inaccuracy analysis of object extraction, calculating the only parameter of accuracy value does not always give the correct class accuracy in terms of completeness and quality. For that reason, precision, recall, accuracy, and F1 score parameters were calculated for the performance analysis of the target classes obtained according to the performance of both methods. The details of accuracy analysis according to two method performance are given in Table 8.

Table 7. Quantitative results of U-Net architecture

	Accuracy	Sensitivity	Specificity	F1-Score
Bicycle Road	0.979	0.639	0.985	0.517
Side walk	0.924	0.320	0.953	0.281
Road	0.705	0.617	0.733	0.503
Road Line	0.979	0.000	0.979	0.000

Table 8. Accuracy analysis of target classes according to two method's performance

	Confusion Matrix Elements	Bicycle Road	Side walk	Road	Road Line
Rule-Based Classification	Accuracy	0.988	0.972	0.959	0.993
	Precision	0.819	0.871	0.993	0.861
	Recall	1	0.966	0.952	0.936
	F1-Score	0.9	0.916	0.972	0.897
U-Net Architecture	Accuracy	0.979	0.924	0.705	0.979
	Precision	0.434	0.251	0.424	0
	Recall	0.639	0.32	0.617	-
	F1-Score	0.517	0.281	0.503	0

4. Discussion

In this study, it is aimed to obtain fast and accurate information in regard to the current spatial attributes of roads in case of any emergency, at the Davutpasa Campus of Yildiz Technical University. Therefore, an area with 244000m^2 was chosen as the study area. The advantages and disadvantages of rule-based classification and U-Net architecture were determined by comparing the performance analysis with target classes; a road, road line, sidewalk, bicycle road. The accuracy of target classes obtained from the rule-based classification method has outperformed U-Net architecture, because of the study area's size, since the rule-based classification method utilizes task-specific defined rules whereas U-Net architecture requires a large number of training images. Despite the lack of training data, Deep networks have proven to outperform at extracting mid-and high-level abstract and discriminative semantic features from images [34]. In such a study area, the similar or same textures which belong to different classes may provide an advantage for the rule-based classification method, whereas it became a disadvantage for U-Net in multiclass segmentation tasks.

In the rule-based classification method, a mixture of classes was prevented by creating auxiliary classes such as green area, shadow, ground, sidewalk, and non-road with scale, shape, and compactness analyses. In addition, with the defined rules, the class features were improved and classification accuracy was increased. More accurate results were obtained with the rule-based classification method by developing a rule set as follows: height, NDSM, pattern, color, texture, band combination, length, width, brightness, etc. to eliminate class confusion between target classes. All of these processes were performed to obtain a high accuracy for the target classes.

When we analyzed the properties of the study area, data distribution was found as follows: (i) 54% of the study area belong to the background, (ii) 6% to the sidewalk, (iii) 2% to the road line, (iv) 35% to the road and (v) 3% to the bicycle road classes. Therefore, the difference in classification results between target classes arose owing to a class imbalance in the study area (Table 8). It is thought that this problem may be solved by expanding the study area, namely the percentages of the sidewalk, road line, and bicycle road class. Despite these circumstances, U-Net architecture achieves more accurate segments on images with noise and variable details. From literature, it is known that deep learning networks are capable of classifying the objects under different conditions, although the spectral values of the target class vary, through medium and high-level feature extraction. Qualitative and quantitative analysis results showed that target classes could not be segmented since road lines, sidewalks and pedestrian crossings have variable characteristics and there are no adequate land-cover structures for training. In Fig. 14, the sample images of U-Net and Rule-based classification results for target classes are analyzed. When Fig. 14 is examined sequentially, bicycle road (Fig. 14a), shadow effect (Fig.

14b), road line and sidewalk (Fig. 14c), road and road line (Fig. 14d), and road (Fig. 14e) extraction results were given as examples. Two different situations were encountered in U-Net architecture when the extraction results of the bicycle road were analyzed. The bicycle road could not be extracted because it was not labeled as seen in Fig. 14 (a), while in Fig. 14 (b), the U-Net architecture was able to detect the bicycle road correctly. One of the main problems encountered in the automatic extraction of roads is the shadow effect (Fig. 14b). In the rule-based classification method, the objects representing shadow on the road are assigned to the road class by creating auxiliary classes using band combinations and texture properties. Also, the objects representing shadows on the sidewalk and road lines caused mix-classification. This situation was resolved by improving the classes created by length, width, and brightness analysis. In the U-Net architecture, it has been determined that the existence of shadows in the images does not pose a problem, especially in the segmentation of roads. For example, although there is a shadow on the bicycle road in the image, all the details were captured with the U-Net architecture, and classes were created correctly. In Fig.14 (c), the target class of road line, sidewalk, and road were analyzed. When the road lines are clear and have equal marking in aerial photos, the rule-based classification method detects and classifies the objects correctly with the analysis of pattern, color, and texture. However, in the U-Net architecture, the road line class could not be extracted correctly because the model could not fit due to insufficient segmentation of the road line training data. Contrary to this, the extraction of the road and sidewalk were obtained with successful results in U-Net architecture.

Another important condition in the extraction of roads that affects the accuracy is that roads have different textures in the aerial images. The structural material of roads; such as asphalt, ground fill, paint, patch, etc. have different textures. Because of that, different reflection values have emerged and this has prevented the model from learning in U-Net architecture. Qualitative results prove this as seen in Fig.14 (d) and Fig.14 (e). In Fig.14 (d) the road could not be extracted even though the road class was labeled in U-Net architecture. The result of a similar faulty road extraction is seen in Fig.14 (e). Despite these erroneous road extraction results, rule-based classification results were quite successful. The main reason for this is that the developed rule set increases the accuracy and improves the classes. On the other hand, it should also be noted that the U-Net architecture also gave successful results for this test area for the study.

5. Conclusion

In this study, object-oriented rule-based classification and U-Net deep learning architecture methods were utilized for automatic road extraction performance analysis using aerial images obtained by UAVs. Firstly, the object-oriented rule-based classification method was applied and the target classes of roads, road lines, sidewalks, and bicycle roads were extracted. After determining the appropriate parameter values obtained as a result of segmentation and classification analyses with the proposed method, the developed ruleset was created. With this ruleset, class confusions were removed and target classes were improved. In addition, auxiliary classes such as green area, shadow, and ground class were also extracted to eliminate class confusion and increase the accuracy of target classes. As a result of the original ruleset developed, 85% kappa accuracy and 90% overall accuracy were obtained for the accuracy analysis of the extracted target classes. Secondly, automatic road extraction was performed with the U-Net algorithm, which is one of the deep learning architectures, using data from the same study area. Along with the road class, road lines, sidewalks, bicycle roads, and background were labeled. In the deep learning method, an overall accuracy of 86% was obtained. Thirdly, for the performance analysis in this study, a confusion matrix was created for the target classes obtained for both methods and the results were compared with accuracy, recall, precision, and F1 score. In this context, the advantages and disadvantages of two different methods on the road route belonging to the same study area were investigated in the performance analysis and were given with details in the discussion section.

An important work step within the scope of this study is ontology. After the object extraction process, studies were carried out for the development of the ontology. Developing the ontology is aimed to define the high-level semantic concept of the image objects, automate the process of classifying image objects and discover the implicit knowledge in the classified image objects. After the conceptual design, the properties of image objects are defined in the ontology. By defining rules between data properties and segmented objects, classified objects are integrated with ontology. In order to use spatial data in operations regarding roads, various heterogeneity in the image object data should be solved. For this purpose, in this study, ontology was developed to conceptualize the semantic knowledge of image objects and target classes for road networks. As a result of associating the ontology with target classes, the spatial data has become reusable, fully semantic and interoperable.

As a result of this study, two different methods were utilized for the performance evaluation of the proposed automatic road approach. The obtained accuracy analysis as a result of the two different performance evaluations confirms the success of the approach for automatic road extraction due to its high accuracy, reliability, speed, and automation. Alternatively, the classification performance of road unit classes that have a low percentage in the study area may increase by utilizing object detection deep neural networks such as You Look Only Once (YOLO) and R-CNN. Considering this study, the rapid and accurate extraction of roads belonging to small study areas and its applicability for the whole region of the same area will be a source for future studies.

In particular, the correct use of road data in disaster and city administrations and the rapid transfer of these updated data in cooperation with different institutions are important. In this study, it has been determined that developed ruleset with rule-based classification and U-Net architecture deep learning methods can be used in

automatic road extraction according to the performance analysis results. It has been observed that the products obtained as a result of both methods can be used as reference data in the determination of post-disaster damages and inventory transportation to the needy, in route creation applications, in smart city systems, in the creation of city plans, in the follow-up of rapid construction, in the working areas of public and private administrations.

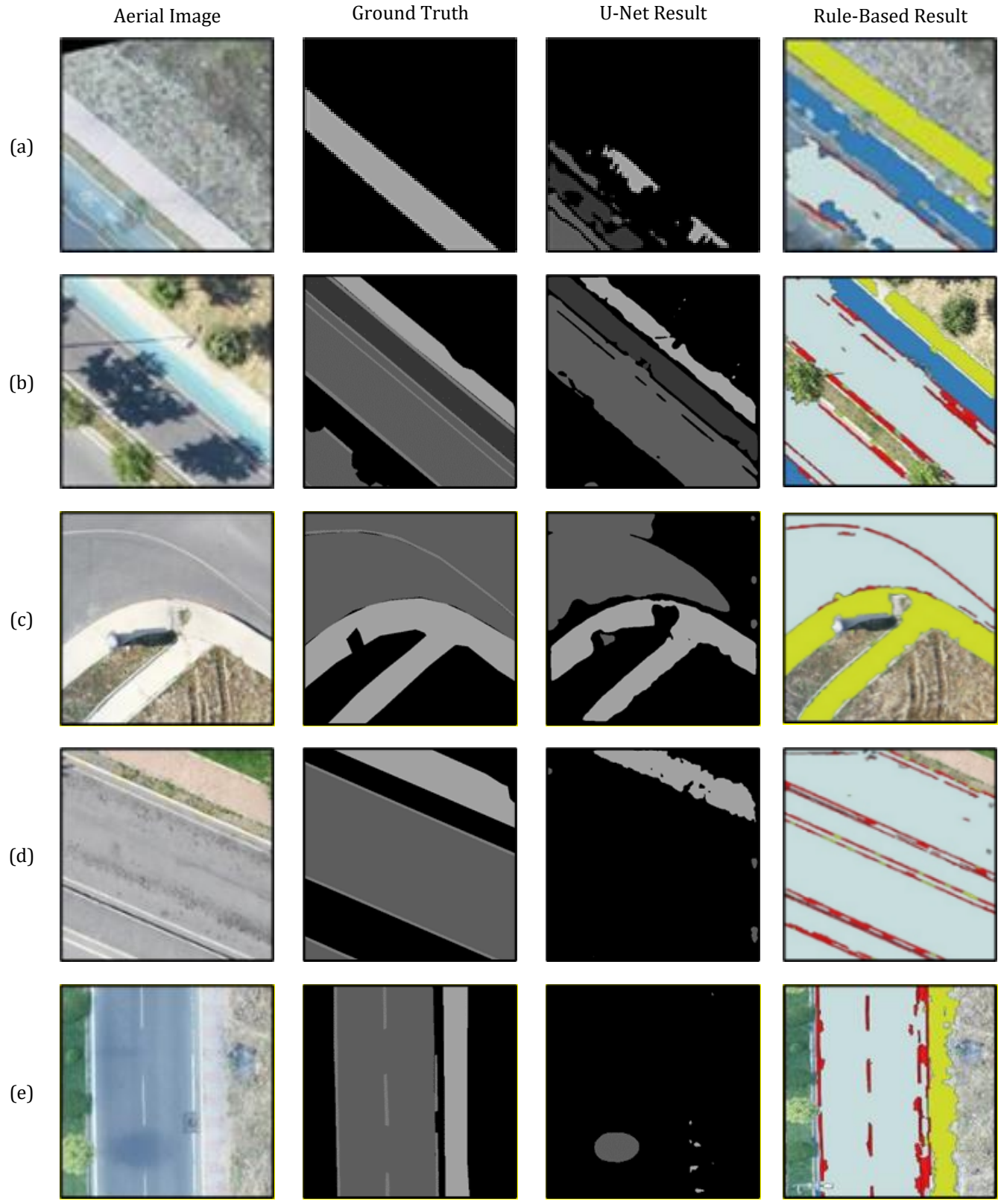

Figure 14. The sample images of U-Net and Rule-based classification results for target classes; a: bicycle road, b: shadow effect, c: road line and sidewalk, d: road and road line, and e: road

Author contributions

Zeynep Bayramoğlu: Methodology, Software, Data curation, Visualization, Writing-Reviewing and Editing.
Melis Uzar: Conceptualization, Methodology, Software, Investigation, Validation, Writing-Original draft preparation.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Fetai, B., Ostir, K., Kosmatin, F. M. & Lisec, A. (2019). Extraction of visible boundaries for cadastral mapping based on UAV ımagery. Remote Sensing, 11(13), 2-20.
2. Kavzoğlu, T., & Tombul, H. (2017). Nesne Tabanlı Sınıflandırmada Segmentasyon Kalitesinin Sınıflandırma Doğruluğu Üzerine Etkisinin İncelenmesi. Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, 17(1), 118-125.
3. Abdollahi, A., Pradhan, B., Shukla, N., Chakraborty, S. & Alamri, A. (2020). Deep Learning Approaches applied to remote sensing datasets for road extraction: A state of the art review. Remote Sensing, 12(9), 4-22.
4. Lian, R., Wang, W., Mustafa, N., & Huang, L. (2020). Road Extraction Methods in High-Resolution Remote Sensing Images: A Comprehensive Review. IEEE journal of selected topics in applıed earth observations and remote sensing, 11(5), 552, 2-16.
5. Yadav, D. P., Nagarajan, K., Pande, H., Tiwari, P., & Narawade, R. (2020). Automatic urban road extraction from high resolution satellite data using object based ımage analysis: a fuzzy classification approach. Journal of Remote Sensing & GIS, 9(1), 279, 1-8.
6. Yiğit, A. Y., & Uysal, M. (2020). Automatic road detection from orthophoto images. Mersin Photogrammetry Journal, 2(1), 10-17.
7. Zhang, X., Han, L., & Zhu, L. (2020). How well do deep learning based methods for land cover classification and object detection perform on high resolution remote sensing imagery. Remote Sensing, 12(3), 2-29.
8. Senthilnath, J., Varia, N., Dokania, A., Anand, G., & Benediktsson, J. A. (2020). Deep TEC: Deep transfer learning with ensemble classifier for road extraction from UAV imagery. Remote Sensing, 12(2), 245-264.
9. Zhang, Z., & Wang, Y. (2019). JointNet: A common neural network for road and building extraction. Remote Sensing, 11(6), 696-718.
10. Gao, L., Song, W., Dai, J., & Chen, Y. (2019). Road extraction from high-resolution remote sensing imagery using refined deep residual convolutional neural network. Remote Sensing, 11(5), 552-568.
11. Zhang, Z., Liu, Q., & Wang, Y. (2017). Road extraction by deep residual U-Net. IEEE Geoscience and Remote Sensing Letters, 5(15), 749-753.
12. Huang, Z., Cheng, G., Wang, H., Li, H., Shi, L., & Pan, C. (2016). Building extraction from multi-source remote sensing images via deep deconvolution neural networks. IEEE International Geoscience and Remote Sensing Symposium, 1835-1838, Beijing, China.
13. Abderrahim, N. Y. Q, Abderrahim, S. & Rida, A. (2020). Road segmentation using U-Net architecture. IEEE International Conference of Moroccan Geomatics, 1-4, Casablanca, Morocco.
14. Emek, R., & Demir, N. (2020). Building detection from SAR images using U-Net deep learnıng method. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIV-4/W3-2020, 5th International Conference on Smart City Applications, 215-218, Virtual Safranbolu, Turkey.
15. Xiaoqiang, Lu, Gong, T., & Zheng, X. (2020). Multisource compensation network for remote sensing cross-domain scene classification, IEEE Trans. Geoscience Remote Sensing, 58(4), 2504-2515.
16. Sarıturk, B., Bayram, B., Duran, Z., & Seker, D. Z. (2020). Feature extraction from satellite images using SEGNET and fully convolutional networks (FCN) International Journal of Engineering and Geosciences, 5(3), 138 - 143.
17. Cheng, G, Xie, X., Han, J., Guo, L., & Xia, G. (2020). Remote sensing ımage scene classification meets deep learning: challenges, methods, benchmarks, and opportunities. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13(1), 3735-3756.
18. Cira, C. I., Alcarria, R., Manso-Callejo, M. Á., & Serradilla, F. (2020). A deep learning-based solution for large-scale extraction of the secondary road network from high-resolution aerial orthoimagery. Applied Sciences, 10(20), 2-18.
19. Filin, O., Zapara, A. & Panchenko, S. (2018). Road detection with EOSResUNet and post vectorizing algorithm. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 211-215, Salt Lake City, UT, USA.
20. Zhang, L., Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: a technical tutorial on the state of the art. IEEE Geoscience and Remote Sensing Magazine, 4(2), 22-40.
21. Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual U-net. IEEE Geoscience and Remote Sensing Letters, 15(5), 749-753.
22. Sener, Z. (2020). Ontology use and evaluation in spatial object extraction from multi sensor system data. Doctoral Thesis, Yıldız Technical University, Institute of Science, Istanbul, 159p
23. Bouyerbou, H., Bechkoum, K., Benblidia, N., & Lepage, R. (2014). Ontology-based semantic classification of satellite images: Case of major disasters. IEEE Geoscience and Remote Sensing Symposium, 2347-2350, Quebec City, Canada.
24. Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A. & Jain, R. (2009). Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), 1349-1380

25. Belgiu, M., & Thomas, J. (2013). Ontology based interpretation of very high-resolution imageries-grounding ontologies on visual interpretation keys AGILE Conference, 14-17, Leuven.
26. Sener, Z., & Uzar, M. (2020). New trend in object oriented image analysis - ontology. Journal of the Faculty of Engineering and Architecture of Gazi University, 35(1), 479-493.
27. Memduhoglu, A., & Basaraner, M. (2022). An approach for multi-scale urban building data integration and enrichment through geometric matching and semantic web. Cartography and Geographic Information Science, 49(1), 1-17.
28. Ronneberger, O., Fischer, P., & Brox, T. (2015) U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention MICCAI 2015. Lecture Notes in Computer Science, 234-241, Springer, Cham. ISBN 978-3-319-24574-4
29. https://www.tmmob.org.tr/en
30. Bayrak, O. C. (2020). Segmentation of liver and brain lesions by deep learning approach from medical images. Master's Thesis, Yıldız Technical University, Institute of Science, Istanbul, 54p.
31. Sørensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species, and its application to analyses of the vegetation on Danish Commons, Kongelige Danske Videnskabernes Selskab, Biologiske Skrifter, 5, 1-34.
32. Savoy, J., Gaussier, E., Savoy, J., & Gaussier, E. (2010). Information retrieval. In N. Indurkhya & F. Damerau (Eds.), Handbook of natural language processing. Boca Raton Chapman; Hall/CRC, 455-484. ISBN: 1420085921
33. Rutzinger, M., Rottensteiner, F., & Pfeifer, N. (2009). A comparison of evaluation techniques for building extraction from airborne laser scanning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2(1), 11-20.
34. Benbahria, Z., Sebari, I., Hajji, H., & Smiej, M. F. (2021). Intelligent mapping of irrigated areas from Landsat 8 images using transfer learning. International Journal of Engineering and Geosciences, 6(1), 40 - 50.