

Article

# Automatic Detection and Mapping of Dolines Using U-Net Model from Orthophoto Images

Ali Polat <sup>1</sup> , İnan Keskin <sup>2</sup>  and Özlem Polat <sup>3,\*</sup> 

<sup>1</sup> Department of Planning and Risk Reduction, Provincial Directorate of Disaster and Emergency, Sivas 58000, Turkey; ali.polat@afad.gov.tr

<sup>2</sup> Department of Civil Engineering, Faculty of Engineering, Karabük University, Karabük 78050, Turkey; inankeskin@karabuk.edu.tr

<sup>3</sup> Department of Mechatronics Engineering, Faculty of Technology, Sivas Cumhuriyet University, Sivas 58140, Turkey

\* Correspondence: ozlem.polat@cumhuriyet.edu.tr

**Abstract:** A doline is a natural closed depression formed as a result of karstification, and it is the most common landform in karst areas. These depressions damage many living areas and various engineering structures, and this type of collapse event has created natural hazards in terms of human safety, agricultural activities, and the economy. Therefore, it is important to detect dolines and reveal their properties. In this study, a solution that automatically detects dolines is proposed. The proposed model was employed in a region where many dolines are found in the northwestern part of Sivas City, Turkey. A U-Net model with transfer learning techniques was applied for this task. DenseNet121 gave the best results for the segmentation of the dolines via ResNet34, and EfficientNetB3 and DenseNet121 were used with the U-Net model. The Intersection over Union (IoU) and F-score were used as model evaluation metrics. The IoU and F-score of the DenseNet121 model were calculated as 0.78 and 0.87 for the test data, respectively. Dolines were successfully predicted for the selected test area. The results were converted into a georeferenced vector file. The doline inventory maps can be easily and quickly created using this method. The results can be used in geomorphology, susceptibility, and site selection studies. In addition, this method can be used to segment other landforms in earth science studies.



**Citation:** Polat, A.; Keskin, İ.; Polat, Ö. Automatic Detection and Mapping of Dolines Using U-Net Model from Orthophoto Images. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 456. <https://doi.org/10.3390/ijgi12110456>

Academic Editors: Wolfgang Kainz, Lili Jiang, Di Zhu and An Zhang

Received: 30 August 2023

Revised: 20 October 2023

Accepted: 31 October 2023

Published: 7 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** doline segmentation; deep learning; orthophoto images; geocomputation

## 1. Introduction

Gypsum karsts have been observed in many countries. These landforms cover an area of approximately 5% of Turkey [1]. The largest part of the distribution of gypsum is located in the Sivas basin in Turkey. This region is one of the rare areas in the world where karstic shapes with unique features are found most prominently.

For the formation and development of karstification, rocks must exist that can dissolve in water. The more easily the rocks are dissolvable in water, the faster the development of the karst topography. The term karst topography is mostly associated with limestones. However, relatively large karstic depressions are not observed in those types of areas. Gypsum is a type of low-durable rock with a hardness of around 2.0, and it melts much more easily than limestone. Therefore, its melting rate is higher than that of limestone. Dolines are the most distinctive and common landforms of karstic lands. They have a mostly circular shape, but sometimes, they can also be in different shapes. Their long axes or diameters can vary from a few meters to a kilometer. In general terms, dolines are formed by the dissolution of gypsum at the bottom and the gradual settling that occurs on the surface. Karstic landforms are formed more quickly and easily in gypsum fields. Determining the characteristics of karst shapes is of great importance in terms of both understanding the formation processes and determining new land-use models.

The collapse of gypsum in most countries results in damage to many living areas and various engineering structures. This event creates natural dangers in terms of human safety, agricultural activities, and the economy [2–7]. In addition, it has been observed that most surface water and groundwater are enriched in sulfate by dissolving gypsum in gypsiferous areas. Therefore, the determination of gypsiferous areas and their karstification characteristics should be taken into account in national- and local-scale planning in terms of addressing natural hazards and water resources.

The rapid detection of karstic shapes and their time-dependent changes are very important parameters in terms of predicting damage in karstic lands. Mapping the dolines in an area is the basis for predicting other dolines that may occur in the future. Creating a doline inventory map is quite difficult, especially in polycarst areas [8], and this is because there are often thousands of very similar dolines on the field.

Image segmentation is commonly used to find objects and borders in images [9–13]. Semantic and instance segmentation methods have been used in image segmentation projects. In this study, we used U-Net, which is one of the semantic segmentation methods that is used to predict doline areas. With this method, the image is automatically classified by assigning all of the pixels in an image to any defined class. This method uses convolutional neural network (CNN) architecture. A U-Net convolutional network [14] is an ideal network for semantic segmentation studies. U-Net consists of two parts: an encoder and decoder. It does not need a large number of labeled data, unlike other methods, and it can learn with a small number of effective data. Additionally, with U-Net, skip connections are used to transfer information from previous convolution layers to deconvolution layers. This increases segmentation performance. For this reason, a U-Net model was preferred in detecting dolines in this paper. The U-Net model requires very few training images and yields more precise results in terms of pixel locations. Therefore, it is often preferred in earth science and other studies [15–21].

The novelty of this paper is that the dolines were detected automatically by U-Net's semantic segmentation via transfer learning. There are segmentation studies in which features such as buildings, roofs, and roads were extracted from orthophoto images. However, there have been no studies on doline segmentation. The study area has thousands of dolines, and it is almost impossible to determine all of them with classical field studies. Our model successfully predicted dolines in a randomly selected area.

#### *Literature Review*

In the literature, there have been many studies conducted with various methods, especially in similar areas using machine learning and deep learning methods on doline distribution such as the following: object detection [22–26], object classification [27], image segmentation [28–30], change detection [31], building objects mapping [32], vegetation mapping [33], and structural controls [34].

In the study conducted by Mochales et al. [22], dolines were located in the Ebro basin in northern Spain. The dolines identified were filled with alluvial sediments, agricultural soils, and urban debris. For this purpose, magnetic susceptibility measurements were used, which revealed a remarkable contrast between the host rocks and cavity fillings. Pueyo-Anchueta et al. [23] proposed a geophysical routine for use in alluvial karst regions to create an integration-based model of the many methodologies in their paper. For the detection of doline areas, the geophysical characteristics involved in each approach, as well as their solutions and uncertainties, were investigated. The suggested sequence of implemented procedures indicates a gradual rise in survey time consumption, ambiguity reduction, and improved resolution. In the study by Nahhas et al. [24], a deep learning (DL)-based building detection approach using a combination of Light Detection and Range (LiDAR) data and orthophotos was investigated. The proposed model was employed on two datasets taken from an urban region with different building types, and it was tested with 21 features and 10 features. The experimental results showed that the accuracy of the tests with size reduction was higher both in the study region and in the test region.

Hussain et al. [25] investigated the Vodose and Fluvial caves in Tarimba (Goias, Brazil) with several geophysical approaches. They determined that the findings obtained from their studies were compatible with real field conditions. They found circular and concave landforms formed by karstic processes, which are known as sinkholes or dolines. In their study, Čarni et al. [26] propose using indicator species to describe sections of several dolines and their common species to determine dolines with significant conservation value for cold-adapted species. The primary goal of their work was to classify the dolines into landform-vegetation units (LVUs) by considering essential geomorphic features and indicative plant species.

Zero-shot learning (ZSL) is a method for identifying unseen objects during the training phase that has been known to be beneficial in real-world situations. In the study performed by Pradhan et al. [27], the integration of CNN and ZSL was used as a classification and feature extraction technique to map the land cover using high-resolution orthophotos. High accuracy values were obtained in the experiments, and the effectiveness of ZSL in land cover mapping based on high-resolution photographs has been proven.

Abdollahi et al. [28] presented two novel deep convolution models for the segmentation of multiple objects from aerial photos such as buildings and roads based on the U-Net family. Their attention was directed toward buildings and road networks due to their large presence within urban regions. The presented models are called U-Net with multi-level context gate (MCG-UNet) and bidirectional ConvLSTM UNet model (BCL-UNet). Compared to U-Net and BCL-UNet, MCG-UNet increased the average F1 success by 1.85% and 1.19% for road extraction and by 6.67% and 5.11% for building extraction, respectively. Road parts extraction is of great importance in most Geographic Information System (GIS) applications. To derive the road class from orthophoto images, Abdollahi and Pradhan [29] employed an integrated technique that included segmentation and classification methods with connected component analysis. There are three steps to the approach that has been proposed. The fractured pictures were first segmented using the multi-resolution segmentation approach. Using three different classification techniques (support vector machines, decision trees, and k-nearest neighbor), the results were then split into two groups: road and non-road. Finally, connected component labeling was utilized to extract road components, and morphological processing was performed to increase performance by removing off-road parts and noises. The current methods described in the literature for road extraction result in fragmented output due to obstructions like shadows, structures, and vehicles. To address the fragmented results, Abdollahi et al. [30] presented SC-roadDeepNet, which is a deep learning-based architecture that preserves structure shape and connection. In comparison to alternative models like LinkNet, ResUNet, U-Net, and VNet, the proposed approach enhances the F1 score value by 5.49%, 4.03%, 3.42%, and 2.27%, respectively.

In the study conducted by Ghasemkhani et al. [31], the proposed model describes the conversion of bare lands into settled or developed areas. The model consists of a fuzzy system and Ordered Weight Averaging (OWA) methods together. The applied model consists of four parts: physical fitness, accessibility, neighborhood effect, and calculation of general fitness. Experiments have shown that the proposed model predicts changes with high accuracy.

Abdollahi and Pradhan [32] made studies related to the mapping of building objects from aerial images. They proposed a novel deep learning structure called MultiRes-UNet. In their network, they used MultiRes-UNet blocks and convolutional operations with skip connections. They combined semantic edge knowledge with semantic polygons. They tested their model on the roof segmentation dataset and achieved a 0.78% increase in mIoU value.

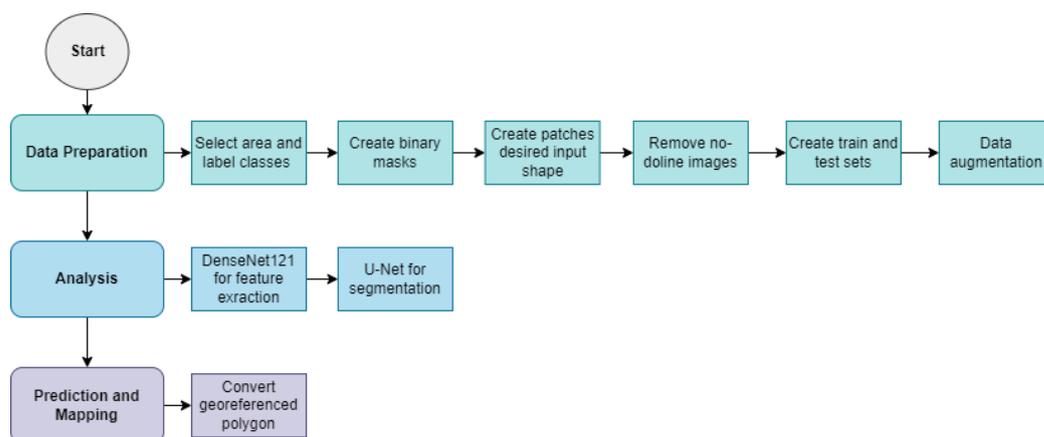
In another study, Abdollahi and Pradhan [33] studied urban vegetation mapping and described how the output of the DNN model used to categorize vegetation can be interpreted using an annotation technique known as Shapley additive explanations (SHAPs). They evaluated the accuracy of their method by mapping vegetation from aerial imagery

using spectral and textural parameters, and their results showed an overall accuracy of 94.44%.

The solution dolines form characteristic landforms that are observed on the high plateaus of the Taurus Mountains. The study by Öztürk et al. [34] concentrated on how the distribution and morphometric characteristics of the dolines in the western portion of the Central Taurus Mountains were influenced by tectonic structures, drainage patterns, and slope conditions. As a result of the research, it has been revealed that the fault and joint systems formed on the thick-bedded limestone between the thrust faults are effective on the change in the intensity of doline.

## 2. Material and Methods

This study consists of three phases: data preparation, analysis, prediction and mapping (Figure 1). First, an area was selected for the analysis. This selection should be balanced for the doline area and the non-doline area. The second process is the labeling of images including dolines. A binary image (mask) is created by labeling all dolines. Later, desired size patches of the image are created, and the images that have no-doline data are removed. Subsequently, classical segmentation processes were performed. These processes are dataset preparation, training, testing, and prediction. In this study, a Segmentation Models library with U-Net architecture [35] was used. This is a Python library with Neural Networks for Image Segmentation based on Keras [36] and TensorFlow [37]. The last task is converting the predicted doline data to a georeferenced vector file. The experiments were completed on a desktop that has an Intel i7 CPU, 32 GB memory and NVIDIA GeForce GTX 1070 graphics card. These processes are explained in detail below.



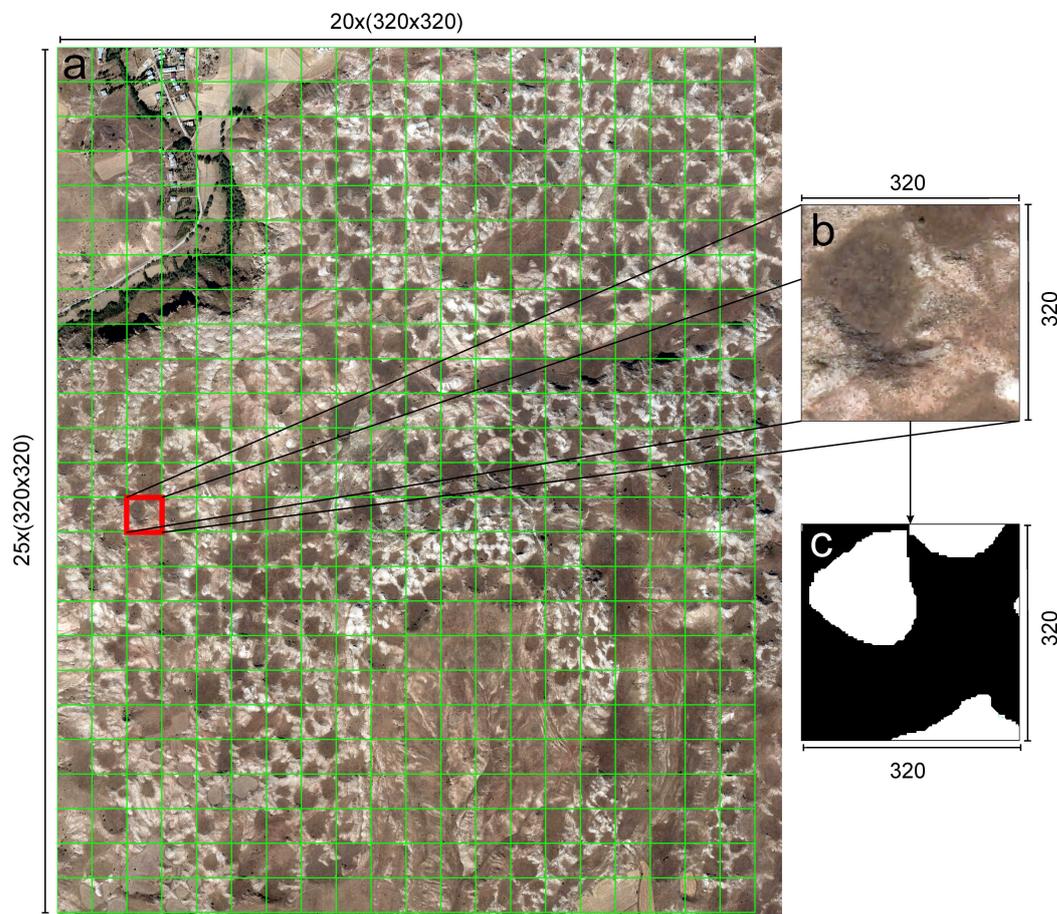
**Figure 1.** Workflow diagram.

### 2.1. Data Preparation

The data preparation process is one of the very important steps of classification and segmentation studies. Many factors in this process directly affect the results of the model. Therefore, the task of data selection and preparation of the data should be completed carefully.

In this study, orthophoto images with 0.3 m spatial resolution of the study area were used as input data. An area of about 4.835 km<sup>2</sup> was selected randomly to create training and test data. In the next step, boundaries of all dolines in this area were drawn and labeled. The next processes were completed by Python scripts. The first script named “create image and mask” creates two separate images in RGB and binary format. The U-Net model requires mask images related to each class. We used dolines and backgrounds as class labels. Therefore, we prepared a binary image having the values of 0 and 255. Zero values were assigned to the background pixels, and 255 values were assigned to the doline pixels. Later, 320 × 320 sized patches were created from these single images (Figure 2). Images with no doline data were removed. In addition, the number of pixels as labeled background and doline in the images was determined. Then, images containing too small

doline areas were removed. A total of 374 RGB images and the same size of mask images with the same names were created. These data were split into training and test sets. The splitting ratio was selected as 67% for training and 33% for test data. Finally, 251 training images and 123 test images were created. The splitting process was also completed for the mask images.



**Figure 2.** (a) Patched original image, (b) Sample of a patch, (c) Mask image of patch.

### 2.1.1. Study Area

The study area is located about 25 km northwest of the Sivas city (Figure 3). An area of approximately 4.835 km<sup>2</sup> where dolines are observed very densely has been chosen as the study area. This area consists entirely of gypsum units. The main lithology in the study area and its surroundings consists of Lower Miocene–Middle Miocene gypsum with considerable karstification (Figure 4a). There are thousands of dolines with different sizes and types within the formation. Gypsum bedding in the region is mostly not obvious. Gypsums show a massive structure with thickness ranging from 100 to 200 m [38].

The areas where dissolution dolines are frequently seen can be defined as doline karst areas. Dolines in the study region are defined as dissolution dolines with a depth of 5–20 m. These are mostly flat dolines that elongate in one direction and with shapes of circular or nearly circular. These dolines generally do not have water (Figure 4b,c).

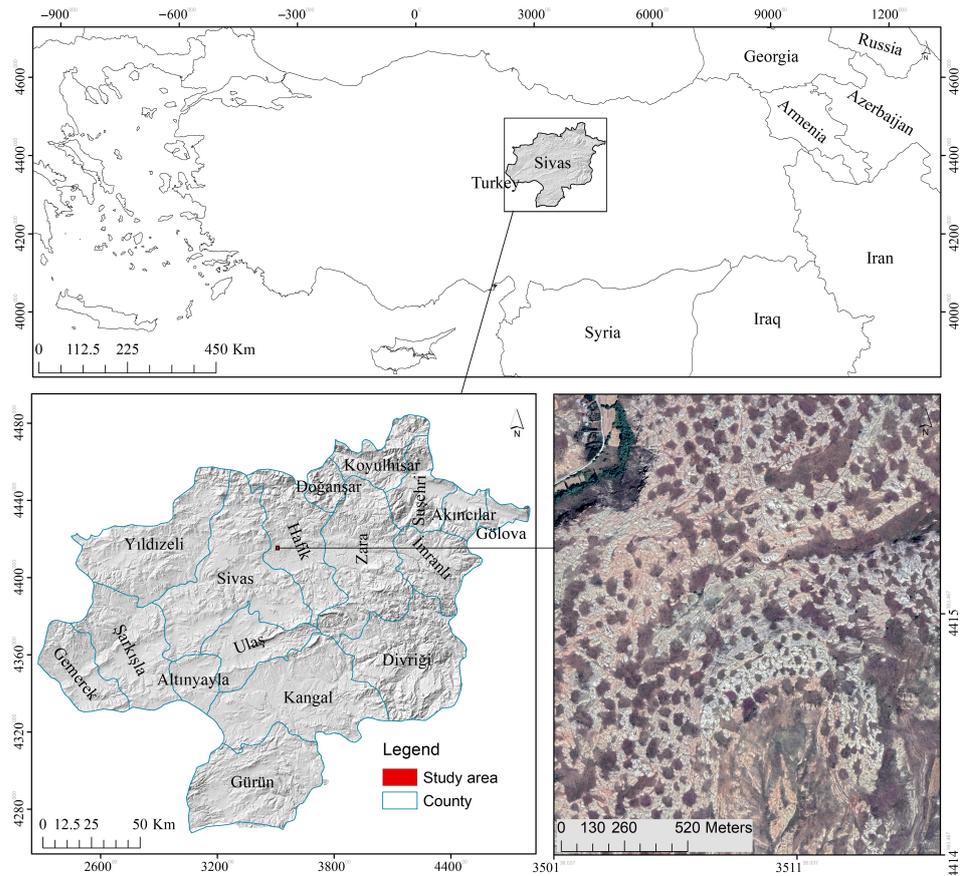


Figure 3. Location map of study area.

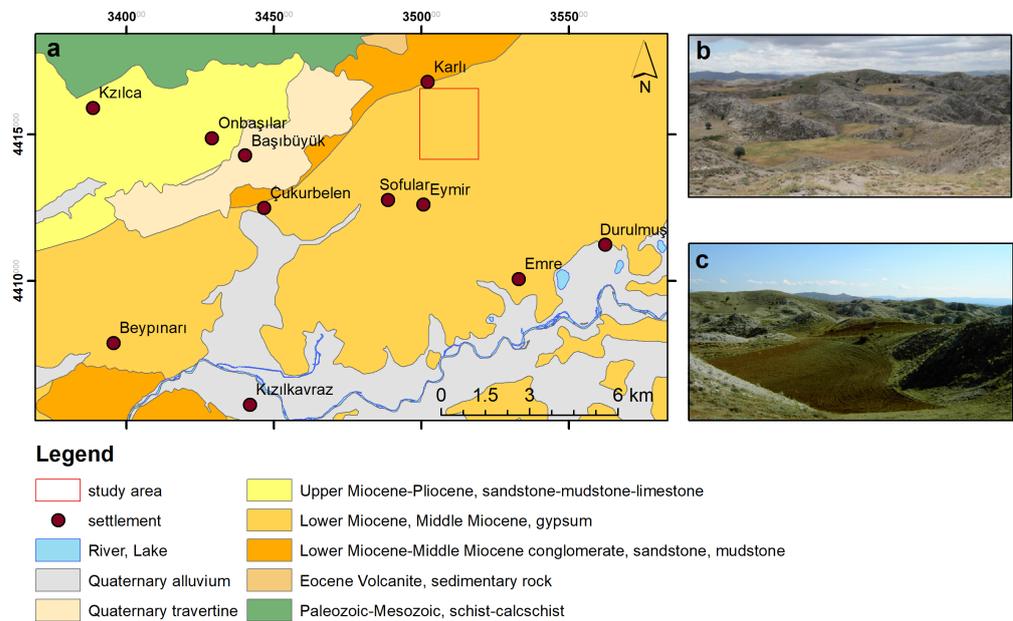
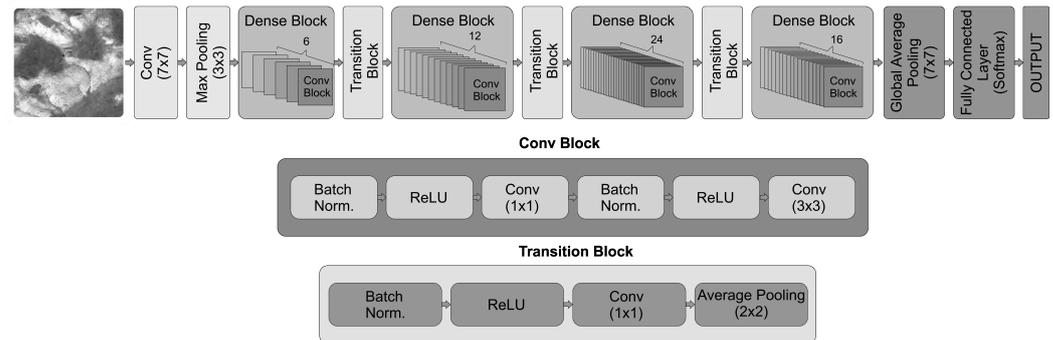


Figure 4. (a) Geological map of study area (1/25,000 scale geological map of General Directorate of Mineral Research and Explorations) (b,c) Doline images.

### 2.1.2. DenseNet121

Transfer learning is a machine learning technique. In this technique, pre-trained model weights are used for a new task. The new model is trained faster by fine-tuning using the knowledge from the previously trained models. In addition, this method provides high performance by using fewer data.

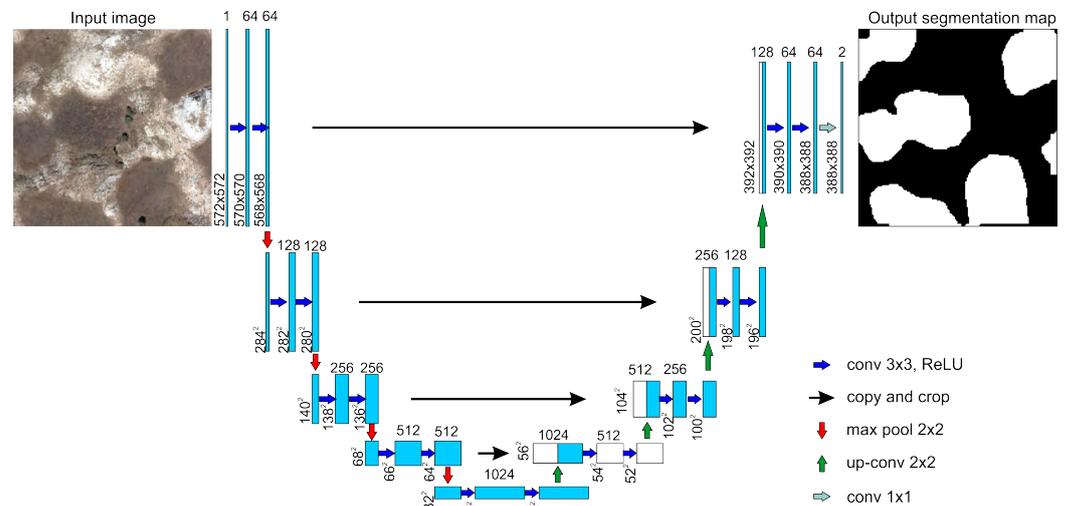
In this study, we used DenseNet121 architecture as a pre-trained model. This model consists of four dense blocks. These blocks include 6, 12, 24, and 16 convolution blocks, and each convolution block also has two convolution layers, Conv( $1 \times 1$ ) and Conv( $3 \times 3$ ), respectively. Also, there is a transition block between dense blocks. This block has a convolution layer, Conv( $1 \times 1$ ), and a  $2 \times 2$  average pooling layer. The last block is the classification block. This block has a global average pooling followed by the fully connected layer. There are 121 convolution layers in DenseNet architecture (Figure 5). Therefore, it is called DenseNet121. For more information, see [39].



**Figure 5.** DenseNet121 architecture [40].

### 2.1.3. U-Net Model

In this study, U-Net architecture [14] developed for biomedical image segmentation was used as the segmentation model. This architecture is a type of fully convolutional networks. It consists of two parts. It is called U-Net because its architecture looks like a U-shape (Figure 6).



**Figure 6.** U-net architecture.

Conventional CNN architectures require too much training data. The U-Net model performs effectively with a limited amount of training data and produces more accurate results. The network architecture of U-Net is shown in Figure 6. It consists of a contracting path (left side) and an expanding path (right side). The contracting path captures context, and the expanding path enables precise localization. The contracting path is like the typical architecture of a convolutional network. Each standard convolution process is activated by the Rectified Linear Unit (ReLU). In the expanding path of the model, the upsampling process is applied. Thanks to the layers in the expanding path, it is aimed to augment the output resolution. The sampled output is combined with high-resolution features throughout the model for localization. Cropping is required because the boundary pixels

disappear with each convolution. In the last layer, a  $1 \times 1$  convolution process is applied to transfer each feature vector to the intended number of classes.

#### 2.1.4. Data Augmentation

The DNNs require a large amount of training data. The more data, the better the model performance. In addition, it prevents overfitting of the model. Although the U-Net architecture requires less data, we implemented a data augmentation method for our model. We used the Albumentations method [41] for data augmentation. Albumentations offers an interface for various computer vision tasks, including classification, segmentation, detection, estimation, and more. The library is extensively utilized in industry, deep learning research, machine learning competitions, and open-source projects.

Albumentations includes many transform methods, and all these transforms can be easily applied. Since we have a small dataset, we used numerous augmentation methods including:

- Horizontal flip;
- Affine and perspective transforms;
- Brightness, contrast and color adjustments;
- Sharpening and blurring process;
- Adding gaussian noise.

#### 2.2. Evaluation Metrics and Loss Functions

In this study, IoU (Intersection over Union), also known as the Jaccard index, and F-score metrics were used to evaluate the model performance. The Jaccard index [42], also called the Jaccard similarity coefficient, is a statistical metric used to compare the similarities and differences between sample sets (1). In our study, the relationship between ground truth (real dolines) and predicted dolines was compared. The higher the Jaccard index, the higher the accuracy of the classifier.

$$J(A, B) = \frac{A \cap B}{A \cup B} \quad (1)$$

We used the F-score (Dice coefficient) as the second metric for the model evaluation. F-score is a metric that integrates both precision and recall into a single value. Precision is the proportion between the True Positives and all the Positives (2). Recall calculates what proportion of actual positives was identified correctly (3). The F-score considers the importance of precision and recall equally. The  $\beta$  value in Equation (4) is the F-score coefficient utilized for balancing precision and recall. The F-score value is between 0 and 1. 1 represents the highest score, and 0 represents the lowest score. The calculation of F-score is shown in Equation (4).

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (2)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (3)$$

$$F_{\beta}(\text{precision}, \text{recall}) = (1 + \beta^2) \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}} \quad (4)$$

Segmentation model losses can be combined. We used Binary Focal Loss (BFL) and Dice loss (DL) together. BFL measures the Binary Focal Loss between the ground truth and the prediction by the following equation.

$$L(gt, pr) = -gt\alpha(1 - pr)^{\gamma}\log(pr) - (1 - gt)\alpha pr^{\gamma}\log(1 - pr) \quad (5)$$

where  $gt$  is ground truth,  $pr$  is prediction,  $\alpha$  is the same as a weighting factor in balanced cross-entropy, and  $\gamma$  is the focusing parameter for the modulating factor.

The other loss function is DL, which is developed based on the dice coefficient for binary data and can be calculated as follows:

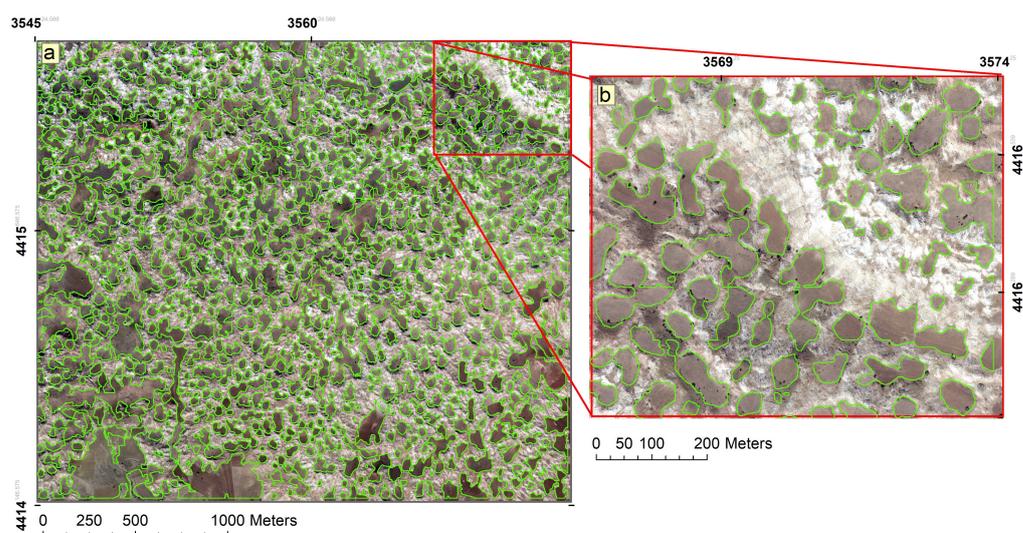
$$L(\text{precision}, \text{recall}) = 1 - (1 + \beta^2) \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}} \quad (6)$$

where  $\beta$  is the coefficient for precision and recall balance.

### 3. Results and Discussion

In this study, the Segmentation Models library was used for the segmentation process with transfer learning techniques. The transfer learning allows you to take pre-trained model weights for any task and reuse them for another task. In this technique, the layers and weights of a pre-trained model for classification purposes are used in the first part (encoder) of the U-Net architecture. Then, the layers in the second part of the U-Net (decoder) are trained with the augmented dataset. After preparing the data, the model was trained in U-Net architecture with three different pre-trained models. These are ResNet34, EfficientNetB3 and DenseNet121. Each model was trained and evaluated using the same parameters. Model results are close to each other. However, the results of DenseNet121 have higher values than the others. Therefore, the DenseNet121 model was used for the prediction of dolines.

A new area (not including the train and test data) was selected from the study area, and the model was run. This area is about 7.289 km<sup>2</sup> (Figure 7). A Python script was written for the prediction task. This script loads the model, patches the image, makes a prediction, and creates a polygon file including dolines. Our script works successfully and is completed in a short time. This process was completed in approximately 30 s. A total of 808 dolines were predicted in the selected area. The total area of the dolines is 2.689 km<sup>2</sup>. This area corresponds to 37% of the total area. The smallest doline area is 1.5 m<sup>2</sup>, and the largest doline area is 115,775.2 m<sup>2</sup>. All the predicted doline areas were almost the same as the real shapes of dolines. The boundaries of dolines are correctly predicted as seen in Figure 7b. Figure 7a shows the predicted dolines, and Figure 7b shows a zoomed view of an area of the image.



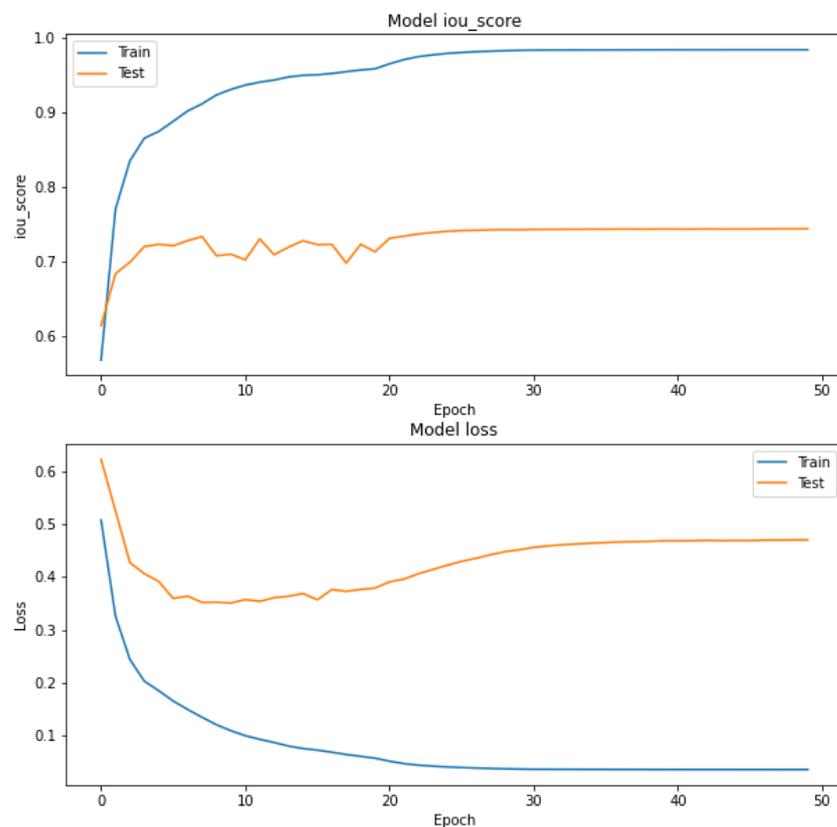
**Figure 7.** (a) Predicted dolines on test area, (b) Zoomed view.

In this study, the Adam optimizer with a learning rate value of 0.0001 was used as an optimization algorithm. In addition, ReduceLROnPlateau is used, which reduces the learning rate when the metric stops developing. The combination of Binary Focal Loss and

Dice loss was used as the loss function. The batch size value was selected as 8. Higher batch sizes caused memory errors. The epoch value was selected as 50. Too many epochs can cause the model to overfit, and very few epochs can cause the model to underfit. The optimal epoch value can be determined by assessing training and test results.

The image size is a factor that affects the performance of the model. All classes must be visible in the images. If the image size is small, the number of data increases, but images could be produced in which all classes are not visible. If the image size is large, the number of data decreases. Datasets of  $128 \times 128$ ,  $256 \times 256$ , and  $320 \times 320$  dimensions were created to determine the optimum size. The best dimension was determined as  $320 \times 320$  for our data.

Image augmentation is a method to overcome overfitting trouble in deep CNNs and is widely utilized to enhance performance on a variety of applications [41]. The DenseNet121 model was tested with non-augmented data. The results show an overfitting problem and an unstable model (Figure 8). The mean IoU value of the training set is 98%, while that of the test set is 70%. In addition, the loss value is not as low as expected. Therefore, data augmentation was applied in this study and the results were improved.

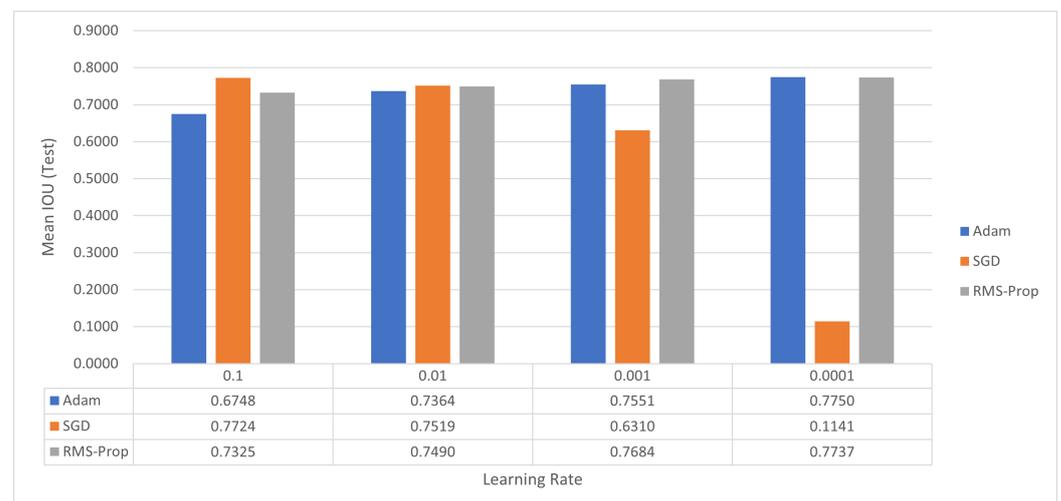


**Figure 8.** Model results without augmentation.

To create a robust model, it is necessary to determine the appropriate parameters. For this purpose, some parameters should be tuned. Tuning the parameters of the model requires a significant amount of time and computational resources, so researchers often decide on settings from previous experience. The settings used in similar studies can also be preferred. Also, some methods such as random search, and grid search automatically select the parameters [43,44]. In this study, loss functions, optimizers, and learning rates were tuned. Loss functions are one of the factors affecting model performance. They do not show the same performance in every model. For complex objectives such as segmentation, it is not possible to decide on a universal loss function [45]. In our study, Jaccard loss, Dice loss, and Focal Loss were tested with the DenseNet121 model. In addition, Jaccard loss–Dice loss, Binary Focal Loss–Dice loss, and Binary Cross-Entropy Loss were combined

and tested. The model was tested three times for every loss function. The mean IoU values of Dice loss, Binary Cross-Entropy Loss, Jaccard loss, Binary Focal Loss, and Focal Loss were calculated as 0.7728%, 0.7755%, 0.7750%, 0.7258%, and 0.7206%, respectively. Also, some loss metrics were combined. The mean IoU values of Jaccard loss + Dice loss and Binary Focal Loss + Dice loss were calculated as 0.7494% and 0.7762%, respectively. The mean IoU value of Binary Focal Loss+Dice loss is higher than the others.

Optimizers are methods used to minimize loss function and to maximize performance. They are mathematical functions dependent on the model's weights and biases. In this study, the three most used optimizers were performed. These are Adaptive Moment Estimation (Adam), Root Mean Square Propagation (RMS-Prop), and Stochastic Gradient Descent (SGD). The learning rate is a hyperparameter that controls how much the model weights are changed with each update. This is one of the most important parameters when configuring the model. The learning rate values of 0.1, 0.01, 0.001, and 0.0001 were tested for each optimizer. The tests were performed on the DenseNet121 model, and the batch size was chosen as 8. Test IoU values were used to evaluate the results. Results show us the best optimizer is Adam with a learning rate value of 0.0001 (Figure 9). The result of the evaluation of RMS-Prop with a learning rate value of 0.0001 is very close to Adam's. The best learning rate value for the SGD optimizer was determined as 0.1.



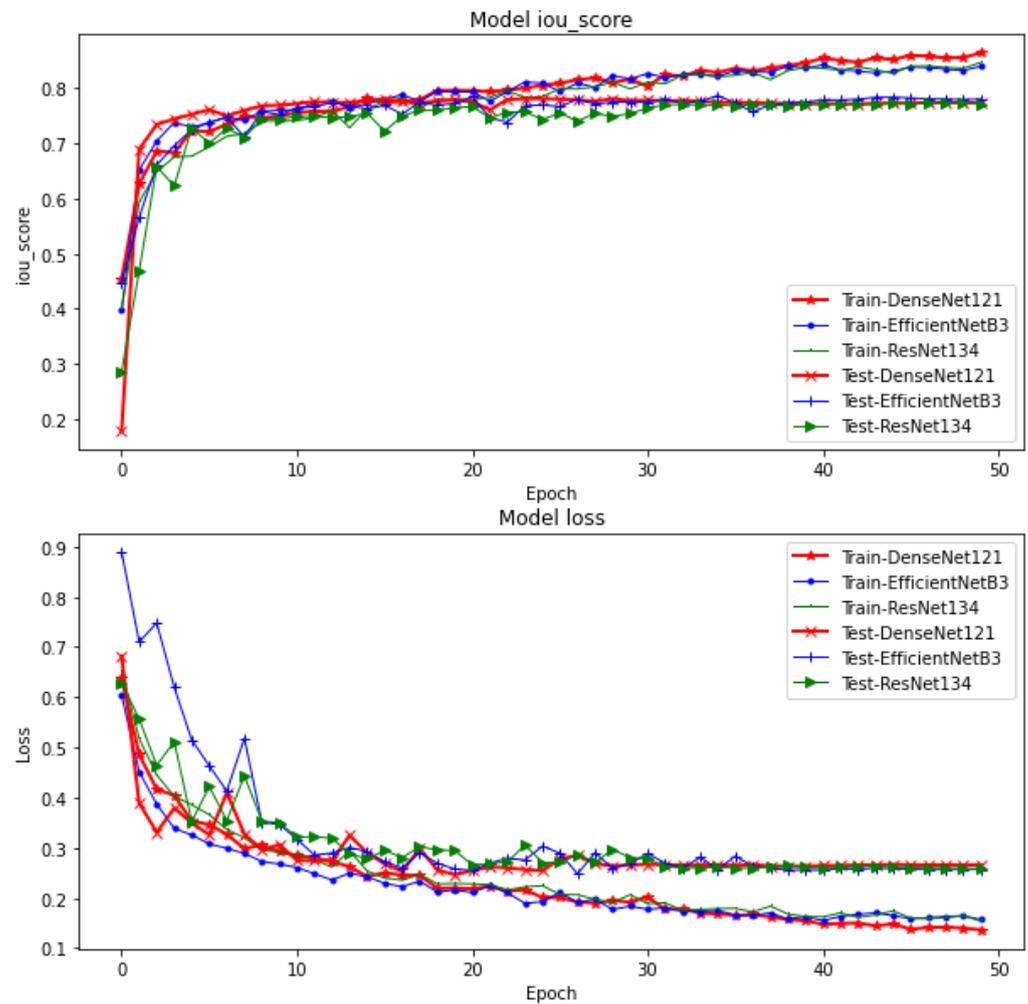
**Figure 9.** The performance of learning rates and optimizers.

We trained our model three times with the same dataset and the same parameters for each transfer learning method. Mean IoU and mean F-score were calculated for training and test data, and the results are shown in Table 1.

**Table 1.** The results of pre-trained models.

Pre-Trained Model	Train IoU	Train F-Score	Test IoU	Test F-Score
ResNet34-1	0.8433	0.9146	0.7634	0.8537
ResNet34-2	0.8533	0.9203	0.7668	0.8575
ResNet34-3	0.8386	0.9117	0.7586	0.8500
Mean	0.8451	0.9155	0.7630	0.8538
EfficientNetB3-1	0.8513	0.9191	0.7741	0.8634
EfficientNetB3-2	0.8392	0.9121	0.7728	0.8624
EfficientNetB3-3	0.8334	0.9082	0.7710	0.8598
Mean	0.8413	0.9131	0.7726	0.8619
DenseNet121-1	0.8323	0.9079	0.7712	0.8622
DenseNet121-2	0.8634	0.9265	0.7800	0.8689
DenseNet121-3	0.8490	0.9279	0.7773	0.8672
Mean	0.8482	0.9180	0.7762	0.8661

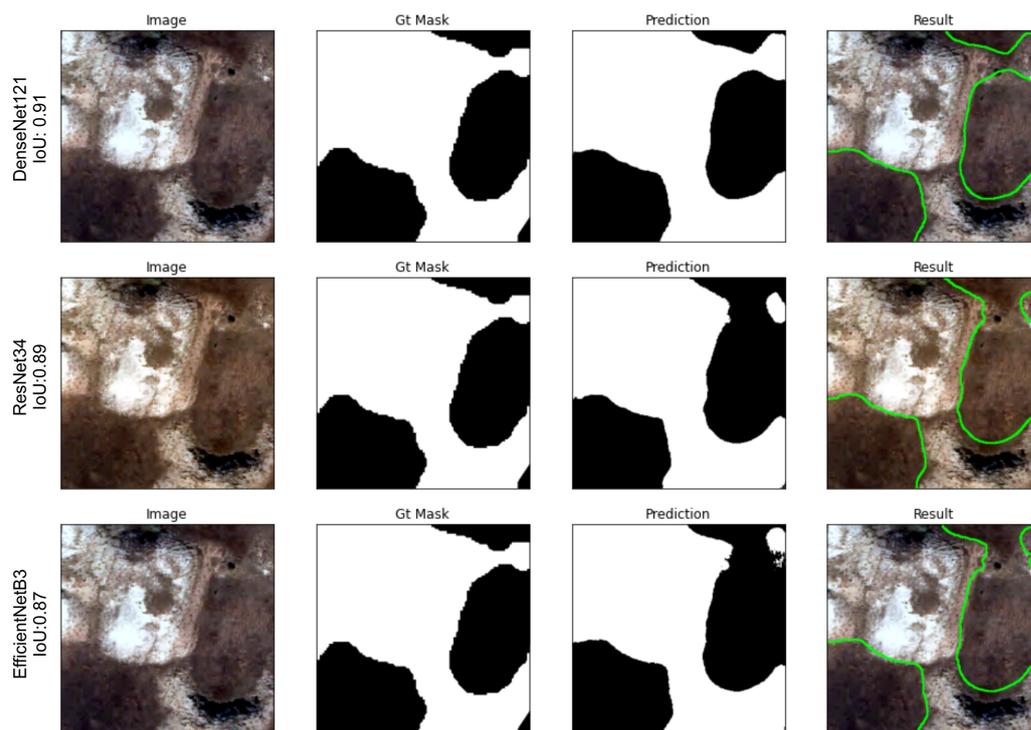
Experiments show that the results of each method are very close to each other. The best results were obtained by using DenseNet121. These are 0.8482, 0.9180, 0.7762, and 0.8661 for train IoU, train F-score, test IoU, and test F-score, respectively. These results are satisfactory for a segmentation task. IoU scores and loss graphs of the models are shown in Figure 10.



**Figure 10.** IoU scores and losses of models.

After the models were created, the performance of each model was examined by selecting a single image from the test dataset. For this process, images in which dolines can be observed well were selected. The IoU values of the models and the predicted masks were compared. Performances are different in each image. ResNet34 was successful on some images, DenseNet121 was successful on others, and EfficientNet B3 was successful on others. In brief, no single model is successful on all images. However, borders of predicted masks of DenseNet121 are observed more appropriately in visual control (Figure 11). These results can be combined in future study topics to benefit from all model results.

In the literature, various studies [22,23,25] have been conducted on the detection of dolines. Their common feature is that they identify areas containing dolines using geotechnical methods. Geotechnical methods are also more costly and time consuming than deep learning methods. To the best of our knowledge, there is no study in the literature that detects dolines using deep learning and creates a doline inventory map.



**Figure 11.** The performances of models for a single image.

The main objective of our article is to successfully detect and inventory uncovered collapse and dissolution dolines. To achieve this, our study utilizes the U-Net model and deep learning techniques. The proposed model gives successful results in large areas containing many dolines and where the depressions are not filled. However, distinct approaches may be necessary for areas where the dolin count is low, depressions are filled, and dolines are covered by vegetation. We recognize the need to develop specialized image processing techniques and employ appropriate suitable satellite imagery and filtering for the detection of covered dolines.

At this point, in future studies, a more specific method will be necessary for the detection and inventorying of covered dolines. This should particularly focus on which image processing techniques can be employed for the concealment and accurate detection of such dolines and which satellite images and filters should be preferred. This can be organized as part of a more comprehensive study on karst areas.

#### 4. Conclusions

Karstic field studies are carried out to generally include measurements and evaluations made in the field. However, karstic areas can not be fully evaluated by field studies due to their characteristics and complex features. One of the reasons is their abundance in the field. Researching in these areas takes a lot of time and effort.

This study proposes a model for the fast and easy detection of dolines in the karstic areas. Deep CNN techniques were used for the segmentation of dolines. U-Net architecture, which requires less data and gives good results, is preferred for this task. All processes are completed automatically by Python scripts.

The results obtained from our model clearly reflect almost the entire karst structure in the field with all its morphological features. All predicted data were georeferenced and can be used in any GIS software. Thus, morphological measurements regarding dolines can be made easily. These results can be used in inventory studies, risk assessment studies, and geomorphological studies. Moreover, our model can be used to detect different types of landforms in future works.

The applied method introduces a novel approach to doline detection and mapping, which is not commonly found in existing methods. The results indicate that the applied method achieves a higher level of accuracy in identifying and mapping dolines, making it superior to previous techniques. Also, the method was designed to be more efficient and had allowed for faster doline detection and mapping, which can be particularly valuable for large-scale applications.

In conclusion, while this article focuses on its achievements in detecting uncovered dolines, it lays the groundwork for guiding future research. Specialized studies will be required for the detection and inventorying of covered dolines, which will be an essential step for a more comprehensive analysis of karst areas.

Based on the findings and discussions presented in the paper, the main recommendation for further research would be to expand the application of the proposed method to different geographic regions and diverse geological conditions. This would help assess the method's robustness and adaptability in various contexts and improve its generalizability for doline detection and mapping. Additionally, exploring the potential integration of other data sources, such as LiDAR or hyperspectral imagery with the existing method could enhance its accuracy and applicability. Further investigations into the scalability of the approach for larger areas and its potential for automation would also be valuable for practical applications in environmental and geological studies.

**Author Contributions:** Ali Polat: Conceptualization, Methodology, Software, Data curation, Formal analysis, Writing—original draft. İnan Keskin: Conceptualization, Methodology, Investigation, Writing—reviewing and editing. Özlem Polat: Conceptualization, Methodology, Investigation, Writing—reviewing and editing, Supervision. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Computer codes related to this article can be found at GitHub: <https://github.com/apolat2018/Dolin-Segmentation-UNet.git> (accessed on 29 August 2023).

**Acknowledgments:** The authors want to thank the Prime Ministry Disaster and Emergency Management Authority for supplying orthophoto images and geological map.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Alagöz, C. Türkiye’de karst olayları hakkında bir araştırma. *Türk Coğrafya Dergisi* **2014**, *1*, 86–92.
2. Thomas, B.; Roth, M. Evaluation of site characterization methods for sinkholes in Pennsylvania and New Jersey. *Eng. Geol.* **1999**, *52*, 147–152. [[CrossRef](#)]
3. Arkin, Y.; Gilat, A. Dead Sea sinkholes—an ever-developing hazard. *Environ. Geol.* **2000**, *39*, 711–722. [[CrossRef](#)]
4. Hu, R.; Yeung, M.; Lee, C.; Wang, S.; Xiang, J. Regional risk assessment of karst collapse in Tangshan, China. *Environ. Geol.* **2001**, *40*, 1377–1389.
5. Cooley, T. Geological and geotechnical context of cover collapse and subsidence in mid-continent US clay-mantled karst. *Environ. Geol.* **2002**, *42*, 469–475.
6. Doğan, U. Dipsiz Göl Kapalı Havzası’ndaki Çökme ve Sübsidans Dolinleri (Batı Toroslar). *Fırat Üniversitesi Sosyal Bilimler Dergisi* **2003**, *13*, 1–21.
7. Zhou, G.; Yan, H.; Chen, K.; Zhang, R. Spatial analysis for susceptibility of second-time karst sinkholes: A case study of Jili Village in Guangxi, China. *Comput. Geosci.* **2016**, *89*, 144–160. [[CrossRef](#)]
8. Keskin, I.; Yılmaz, I. Morphometric and geological features of karstic depressions in gypsum (Sivas, Turkey). *Environ. Earth Sci.* **2016**, *75*, 1040. [[CrossRef](#)]
9. Karimpouli, S.; Tahmasebi, P. Segmentation of digital rock images using deep convolutional autoencoder networks. *Comput. Geosci.* **2019**, *126*, 142–150. [[CrossRef](#)]
10. Rahmani, H.; Scanlan, C.; Nadeem, U.; Bennamoun, M.; Bowles, R. Automated segmentation of gravel particles from depth images of gravel-soil mixtures. *Comput. Geosci.* **2019**, *128*, 1–10. [[CrossRef](#)]
11. Chen, Z.; Liu, X.; Yang, J.; Little, E.; Zhou, Y. Deep learning-based method for SEM image segmentation in mineral characterization, an example from Duvernay Shale samples in Western Canada Sedimentary Basin. *Comput. Geosci.* **2020**, *138*, 104450. [[CrossRef](#)]
12. Kotaridis, I.; Lazaridou, M. Remote sensing image segmentation advances: A meta-analysis. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 309–322. [[CrossRef](#)]

13. Taghanaki, S.; Abhishek, K.; Cohen, J.; Cohen-Adad, J.; Hamarneh, G. Deep semantic segmentation of natural and medical images: A review. *Artif. Intell. Rev.* **2021**, *54*, 137–178. [[CrossRef](#)]
14. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
15. Bai, Y.; Mas, E.; Koshimura, S. Towards operational satellite-based damage-mapping using u-net convolutional network: A case study of 2011 tohoku earthquake-tsunami. *Remote Sens.* **2018**, *10*, 1626. [[CrossRef](#)]
16. Iglovikov, V.; Shvets, A. Ternaunet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. *arXiv* **2018**, arXiv:1801.05746.
17. Hordiiuk, D.; Oliinyk, I.; Hnatushenko, V.; Maksymov, K. Semantic segmentation for ships detection from satellite imagery. In Proceedings of the 2019 IEEE 39th International Conference On Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine, 16–18 April 2019; pp. 454–457.
18. Soares, L.; Dias, H.; Grohmann, C. Landslide Segmentation with U-Net: Evaluating Different Sampling Methods and Patch Sizes. *arXiv* **2020**, arXiv:2007.06672.
19. Pan, Z.; Xu, J.; Guo, Y.; Hu, Y.; Wang, G. Deep learning segmentation and classification for urban village using a worldview satellite image based on U-Net. *Remote Sens.* **2020**, *12*, 1574. [[CrossRef](#)]
20. Khryashchev, V.; Larionov, R. Wildfire Segmentation on Satellite Images using Deep Learning. In Proceedings of the 2020 Moscow Workshop On Electronic and Networking Technologies (MWENT), Moscow, Russia, 11–13 March 2020; pp. 1–5.
21. Marangio, P.; Christodoulou, V.; Filgueira, R.; Rogers, H.; Beggan, C. Automatic detection of Ionospheric Alfvén Resonances in magnetic spectrograms using U-net. *Comput. Geosci.* **2020**, *145*, 104598. [[CrossRef](#)]
22. Mochales, T.; Pueyo, E.L.; Casas, A.M.; Soriano, M.A. Magnetic prospecting as an efficient tool for doline detection: A case study in the central Ebro Basin (northern Spain). *Geol. Soc.* **2007**, *279*, 73–84. [[CrossRef](#)]
23. Pueyo-Anchuela, O.; Casas-Sainz, A.; Soriano, M.; Pocovi-Juan, A. A geophysical survey routine for the detection of doline areas in the surroundings of Zaragoza (NE Spain). *Eng. Geol.* **2010**, *114*, 382–396. [[CrossRef](#)]
24. Nahhas, F.; Shafri, H.; Sameen, M.; Pradhan, B.; Mansor, S. Deep learning approach for building detection using lidar–orthophoto fusion. *J. Sens.* **2018**, *2018*, 7212307:1–7212307:12. [[CrossRef](#)]
25. Hussain, Y.; Uagoda, R.; Borges, W.; Prado, R.; Hamza, O.; Cárdenas-Soto, M.; Havenith, H.; Dou, J. Detection of cover collapse doline and other Epikarst features by multiple geophysical techniques, case study of Tarimba cave, Brazil. *Water* **2020**, *12*, 2835. [[CrossRef](#)]
26. Čarni, A.; Čonč, Š.; Valjavec, M. Landform-vegetation units in karstic depressions (dolines) evaluated by indicator plant species and Ellenberg indicator values. *Ecol. Indic.* **2022**, *135*, 108572. [[CrossRef](#)]
27. Pradhan, B.; Al-Najjar, H.; Sameen, M.; Tsang, I.; Alamri, A. Unseen land cover classification from high-resolution orthophotos using integration of zero-shot learning and convolutional neural networks. *Remote Sens.* **2020**, *12*, 1676. [[CrossRef](#)]
28. Abdollahi, A.; Pradhan, B.; Shukla, N.; Chakraborty, S.; Alamri, A. Multi-object segmentation in complex urban scenes from high-resolution remote sensing data. *Remote Sens.* **2021**, *13*, 3710. [[CrossRef](#)]
29. Abdollahi, A.; Pradhan, B. Integrated technique of segmentation and classification methods with connected components analysis for road extraction from orthophoto images. *Expert Syst. Appl.* **2021**, *176*, 114908. [[CrossRef](#)]
30. Abdollahi, A.; Pradhan, B.; Alamri, A. SC-RoadDeepNet: A New Shape and Connectivity-preserving Road Extraction Deep Learning-based Network from Remote Sensing Data. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5617815. [[CrossRef](#)]
31. Ghasemkhani, N.; Vayghan, S.; Abdollahi, A.; Pradhan, B.; Alamri, A. Urban development modeling using integrated fuzzy systems, ordered weighted averaging (owa), and geospatial techniques. *Sustainability* **2020**, *12*, 809. [[CrossRef](#)]
32. Abdollahi, A.; Pradhan, B. Integrating semantic edges and segmentation information for building extraction from aerial images using UNet. *Mach. Learn. Appl.* **2021**, *6*, 100194. [[CrossRef](#)]
33. Abdollahi, A.; Pradhan, B. Urban vegetation mapping from aerial imagery using Explainable AI (XAI). *Sensors* **2021**, *21*, 4738. [[CrossRef](#)]
34. Öztürk, M.; Şener, M.; Şener, M.; Şimşek, M. Structural controls on distribution of dolines on Mount Anamas (Taurus Mountains, Turkey). *Geomorphology* **2018**, *317*, 107–116. [[CrossRef](#)]
35. Yakubovskiy, P. Segmentation Models. GitHub Repository. 2019. Available online: [https://github.com/qubvel/segmentation\\_models](https://github.com/qubvel/segmentation_models) (accessed on 9 February 2021).
36. Chollet, F. Others Keras: Deep Learning Library for Theano and Tensorflow. 2015. Available online: <https://keras.io> (accessed on 12 June 2023).
37. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. 2015. Available online: <https://www.tensorflow.org> (accessed on 10 June 2023).
38. Poisson, A.; Guezou, J.; Ozturk, A.; Inan, S.; Temiz, H.; Gürsöy, H.; Kavak, K.S.; Özden, S. Tectonic setting and evolution of the Sivas Basin, central Anatolia, Turkey. *Int. Geol. Rev.* **1996**, *38*, 838–853. [[CrossRef](#)]
39. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

40. Polat, Ö.; Güngen, C. Classification of brain tumors from MR images using deep transfer learning. *J. Supercomput.* **2021**, *77*, 7236–7252. [[CrossRef](#)]
41. Buslaev, A.; Iglovikov, V.; Khvedchenya, E.; Parinov, A.; Druzhinin, M.; Kalinin, A. Albumentations: Fast and Flexible Image Augmentations. *Information* **2020**, *11*, 125. [[CrossRef](#)]
42. Jaccard, P. The Distribution of the Flora in the Alpine Zone. *New Phytol.* **1912**, *11*, 37–50. [[CrossRef](#)]
43. Elgeldawi, E.; Sayed, A.; Galal, A.; Zaki, A. Hyperparameter Tuning for Machine Learning Algorithms Used for Arabic Sentiment Analysis. *Informatics* **2021**, *8*, 79. [[CrossRef](#)]
44. Polat, A. An innovative, fast method for landslide susceptibility mapping using GIS-based LSAT toolbox. *Environ. Earth Sci.* **2021**, *80*, 217. [[CrossRef](#)]
45. Jadon, S. A survey of loss functions for semantic segmentation. In Proceedings of the 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Virtual, 27–29 October 2020; pp. 1–7.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.