

Article

Not peer-reviewed version

---

# ResNet152-U-Net: Transfer Learning for the Semantic Segmentation of Forest Restoration Site Using Aerial Images

---

[Eunhee Son](#) , Ui Rin Ha , Hojin Seong , Yoonsun Park , Jungeun Song , [Hyungho Kim](#) \*

Posted Date: 29 February 2024

doi: 10.20944/preprints202402.1732.v1

Keywords: semantic segmentation; forest restoration; aerial images; deep learning; convolutional neural network; U-Net



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Article

# ResNet152-U-Net: Transfer Learning for the Semantic Segmentation of Forest Restoration Site Using Aerial Images

Eunhee Son <sup>1</sup>, Ui Rin Ha <sup>1</sup>, Hojin Seong <sup>1</sup>, Jungeun Song <sup>2</sup>, Yoonsun Park <sup>2</sup> and Hyungho Kim <sup>3,\*</sup>

<sup>1</sup> Department of Forest Environmental Resources, Gyeongsang National University, Jinju, 52828, Korea; hope@gnu.ac.kr (E.S.); har2428@gnu.ac.kr (U. R. H.); jhst1517@gnu.ac.kr (H.S.)

<sup>2</sup> Forest Research Bureau, Korea Forest Conservation Association, Daejeon 35262, Korea; readjean@kfca.re.kr (J. S.); yspark@kfca.re.kr (Y. P.)

<sup>3</sup> Department of Forest Environmental Resources / Institute of Agriculture & Life Science, Gyeongsang National University, Jinju, 52828, Korea; khh@gnu.ac.kr (H.K.)

\* Correspondence: khh@gnu.ac.kr; Tel.: +82-55-772-1857

**Abstract:** Accurate detection of forest restoration sites using deep learning is crucial for effective ecosystem management, optimizing reforestation efforts, and ensuring environmental sustainability. While various deep learning-based systems have been developed for this purpose, the influence of different neural network architectures on model performance during the identification of forest restoration sites remains underexplored. This study aimed to address this gap by exploring the optimal methodology for extracting and classifying candidate sites for forest restoration based on Convolutional Neural Networks (CNNs), specialized for image recognition. Four categories—arable land, road and barren, quarry, and forest—were defined as candidate sites for forest restoration. A dataset comprising 17,043 samples was divided into training (11,929) and validation (5,114) sets at a ratio of 7:3 for model learning. Model accuracy was evaluated using pixel accuracy (PA) and mean intersection over union (mean IoU). The Resnet152-U-Net model demonstrated a pixel accuracy of 95.2% and a mean IoU of 61.3%, indicating excellent performance in extracting candidate sites for forest restoration. This approach offers spatial and temporal advantages over conventional field surveys or aerial image-based assessments, potentially serving as valuable data for selecting future forest restoration sites.

**Keywords:** semantic segmentation; forest restoration; aerial images; deep learning; convolutional neural network; U-Net

## 1. Introduction

Forest conservation is essential for biodiversity preservation, managing air and water quality, and mitigating climate change [1,2]. In recent years, forests have faced significant losses due to climate change and human activities, including deforestation and forest degradation, which have substantial impacts on climate, biodiversity, and society. Various factors contribute to forest loss, including natural disasters such as wildfires and landslides, as well as human-induced land use changes like urbanization and small-scale agriculture [3].

Given the diverse nature and extent of deforestation along with limited financial resources for restoration, prioritizing degraded forest areas for restoration is essential [4]. For effective environmental conservation, it's crucial to classify forest restoration sites based on their types and assess their current conditions to facilitate appropriate restoration efforts.

Traditional methods for forest monitoring, such as ground surveys using the Global Navigation Satellite System (GNSS) and remote sensing techniques, have been widely used. Remote sensing, in particular, enables the detection of various land disturbances over time [5–12]. However, remote sensing methods may be limited in early-stage detection when ground access is restricted [13].

Advancements in high-resolution satellite and aerial imagery, along with deep learning techniques, have enhanced geospatial analysis capabilities for change detection.

The development of computer vision, especially with the rapid advancement of artificial intelligence, has revolutionized various fields, including biometrics, medical imaging, and autonomous driving [14]. Among these, convolutional neural networks (CNNs) have shown promising results, particularly in image recognition tasks. Semantic segmentation, a technique based on CNNs, enables simultaneous segmentation and classification of various objects. In forestry, deep learning analysis has been applied to detect forest fires, landslides, pest infestation, tree species classification, and land cover classification [15–23].

Additionally, deep learning algorithms have been utilized in change detection research, which observes differences in objects or phenomena over time [24]. Change detection methods can be categorized into pixel-based and object-based approaches. While pixel-based change detection is suitable for medium- and low-resolution images, object-based approaches are preferred for high-resolution imagery due to spectral and registration errors [25].

This study evaluates the potential of CNN models in forestry applications and change detection. The methodology involves acquiring a large dataset through annotation and open-source images, preprocessing datasets from forest areas in South Korea, fine-tuning the ResNet152-U-Net model for segmenting five categories, and assessing model performance using time-series aerial images of forest degradation areas.

2. Materials and Methods

2.1. Overview

The overall research system can be divided into three main stages: training data preparation, deep learning model training, and model evaluation (Figure 1). During the training data preparation phase, activities such as research data collection, data pre-processing, and dataset segmentation are conducted to prepare for deep learning model training. In the deep learning model training phase, we perform tasks such as input data pre-processing, data augmentation, U-Net deep learning model training, and adjustment of training settings for efficient learning. During the model evaluation phase, time-series analysis and assessment are carried out to identify deforested regions by semantically segmenting spatial data from two different time periods. This process also includes analyzing the features of potential sites for forest restoration.

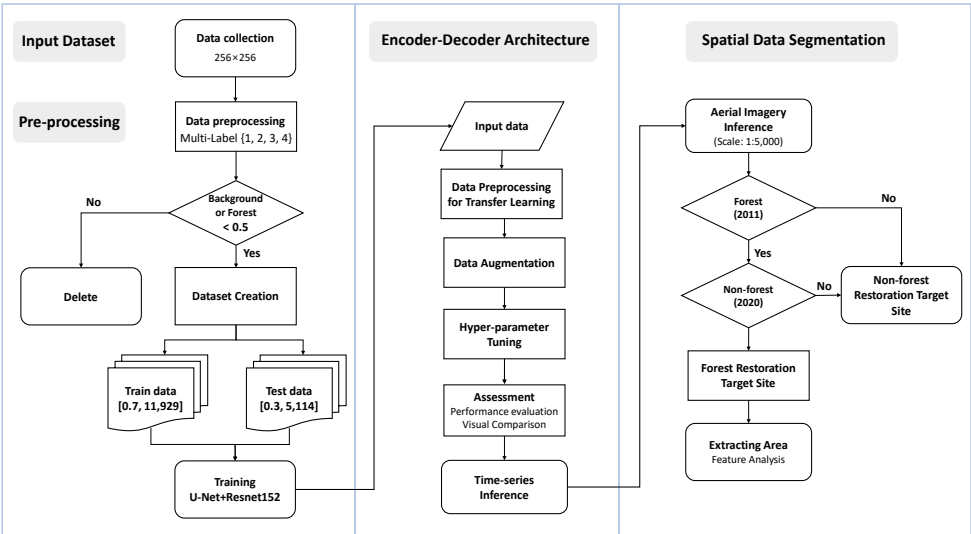


Figure 1. The schema of proposed approach.

2.2. Study Area and Dataset Description

We utilized manually annotated datasets as well as those provided by the National Information Society Agency (NIA) for training our deep learning model. For acquiring annotated datasets, we selected aerial images from 2011 and 2020 of Gangwon Province, which hosts the largest number of candidate sites for forest restoration according to the Korea Forest Service [26]. Figure 2 illustrates an example of the training dataset, which comprises a TIFF file containing an aerial image in RGB bands and corresponding ground truth data in the same area and size.

This study focuses on forests in Gangwon Province, South Korea, chosen for their significant areas of degradation observed in aerial images from the Forest Restoration Site Survey [27]. Specifically, Gochang, Byeonggok, Samcheok, Doam, and Yeongwol were identified as the main degraded areas in Gangwon Province. Consequently, we manually annotated 31 aerial images of these areas for model training.

Moreover, this study utilized the open data source platform AI-Hub, developed by the National Information Agency (NIA), for detecting changes in forest regions using a deep learning model. AI-Hub provides datasets for machine learning research, applications, and education across various fields, including the land environment. Airborne imagery capturing the forest region from 2018 to 2019 with a spatial resolution of 0.51 m and RGB format in the Gangwon area is available in the land environment datasets. Additionally, AI-Hub provides label images with the same location, resolution, and number as the airborne images, encompassing various categories such as roads, paddy fields, fields, forests, and barren lands, which is all pertinent to the forest restoration target area.

To ensure the quality of AI-Hub data, the National Geographic Information Institute (NGII) of South Korea supervised the management and monitoring of image data quality, including radiometric calibration and geometric correction, while KOFPI managed the quality of label image data [28].

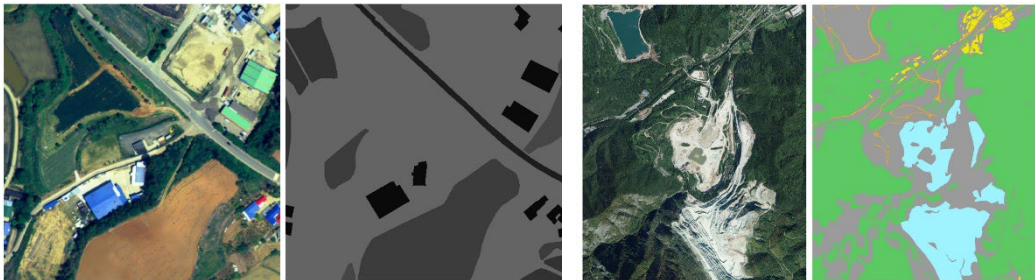


Figure 2. Combined view of NIA dataset and manually annotated data.

The dataset consists of 17,043 images, divided into a 7:3 ratio of training (11,929) and validation (5,114) images. Additionally, we constructed datasets with a resolution of 0.5 meters and a size of 256x256 pixels (Table 1).

Table 1. Dataset composition overview.

Source	Usage	Number of Images	Size	Resolution (m)	Format
Aerial image	Train	11,929	256×256	0.5	Tiff.
	Validation/Test	5,114			
Ground Truth	Train	11,929			
	Validation/Test	5,114			

This study aimed to classify four categories (arable land, road and barren, quarry, and forest) in aerial images using deep learning techniques. To extract features from the aerial images, we employed a convolutional neural network (CNN), a deep learning model specialized in image recognition.

Arable land consists of paddy fields and general fields. Road and barren areas include highways, general roads, forest roads, and barren lands. Quarry areas encompass mining and quarry lands, while forest areas include afforestation lands. Non-target areas comprise various other regions, such as urban areas, shadows, and cemeteries. Roads, especially forest roads in mountainous regions, exhibit similar RGB spectral characteristics to barren lands, hence they are grouped into a single category.

The spectral features of aerial images were compared based on annotation items by analyzing the variations in spectral values throughout the red, green, and blue bands. Comparing spectral values is crucial for training deep learning models as it helps determine the distribution features of each item. The spectral values of arable land, road and barren, quarry, forest, and non-target areas were analyzed to determine the minimum, maximum, and average values (Table 2).

**Table 2.** RGB band detail of annotated area.

Class		Red			Green			Blue		
		Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
	Arable land	13	208	94	22	222	108	19	198	91
	Road-Barren	15	158	80	25	179	102	22	152	85
	Quarry	29	218	135	39	231	157	34	218	138
	Forest	90	127	104	109	152	127	96	133	111
	Background	6	231	100	14	244	122	12	232	113

### 2.3. Pre-Processing

In this section, we conducted normalization of the RGB bands, cropping of images, and data augmentation.

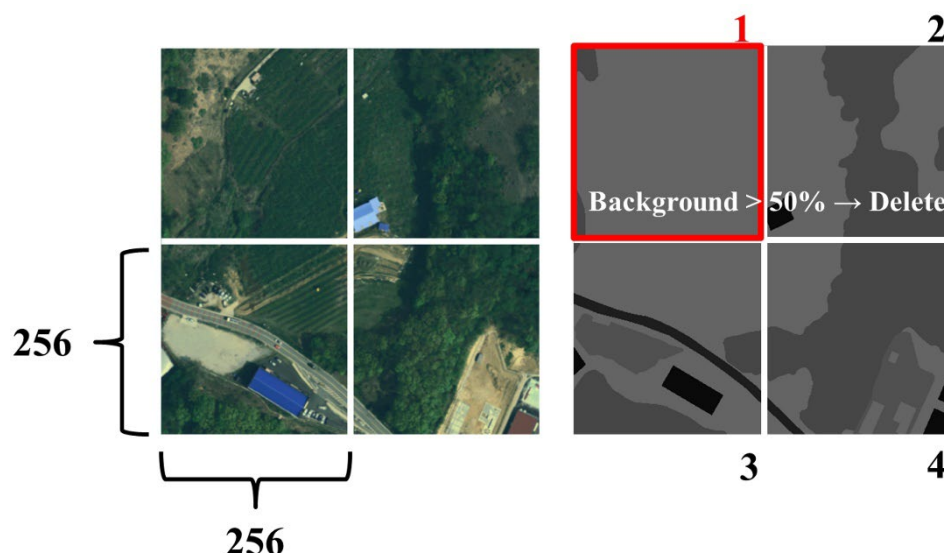
Data pre-processing is the initial stage in machine learning where the data is changed or encoded to a state that allows the machine to efficiently analyze the input. Data pre-processing significantly impacts the generalization performance of a supervised machine learning algorithm. Enhancing data quality is crucial for optimizing performance[29].

#### 2.3.1. Normalizing RGB Band and Cropping Images

In the pre-processing step, the RGB channels of two datasets are normalized using Min-Max normalization. Normalization is a scaling technique, mapping technique or pre-processing stage [30]. Min-Max normalization is a simple technique where the technique can specifically fit the data in a pre-defined boundary with a pre-defined boundary [31]. They are divided into  $256 \times 256$  patches and finally, the patches which contain the higher portion of the target in forest areas are chosen. Moreover, the ground truth is converted into  $256 \times 256$  patches so that it is suitable as an input of the ResNet152-Unet.

The selected aerial images and label images were cropped to the size of  $256 \times 256$  for feeding to the deep learning model because this study adopted a pre-trained deep learning model architecture that allowed for the aforementioned image size in terms of the input. Each image of  $512 \times 512$  size produced 4 cropped images. The labels for arable land, road and barren, quarry, and forest are each assigned by one-hot-encoding. On the other hand, the label for non-target regions was assigned to 0. If more than 50% of the areas in the label data were background, both the aerial and label image data were deleted from the training and validation dataset to prevent ambiguity in the performance of the deep learning model (Figure 3).



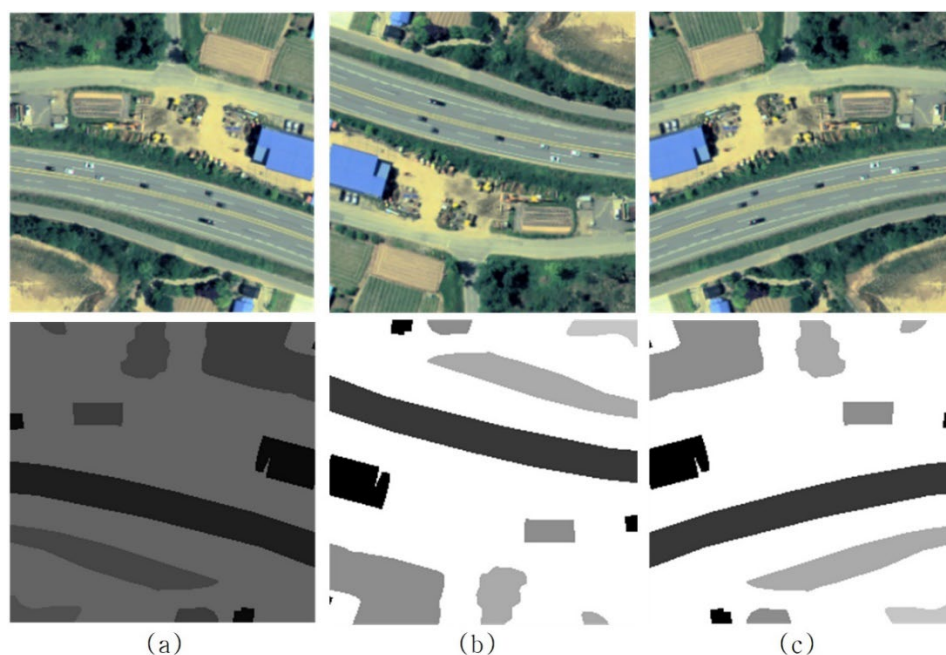


**Figure 3.** Cropping images by 256×256 and removal condition on the dataset.

### 2.3.2. Data Augmentation

When training with a limited training dataset, underfitting or overfitting can occur [32]. Underfitting and overfitting affect the ability to accurately generalize models to both the observed training data and unknown testing data. To solve these problems, pre-training each layer of deep learning, dropout, and data augmentation methods have been proposed to randomly train some nodes [33–35]. Deep artificial neural networks require a large amount of training data to successfully learn a process that is sometimes costly and labor-intensive. Data augmentation addresses this problem by artificially increasing the size of the training dataset [36].

Common data augmentation techniques for image data are Horizontal/Vertical Flips, Random Crops/Scales, and Color Jittering. In this study, we used horizontal flips and vertical flips to increase the amount and diversity of training data in order to preserve the pixel-by-pixel image information as much as possible and apply it to training (Figure 4). ImageDataGenerator from the Python Keras library is used to augment the input data.



**Figure 4.** The schema of data augmentation. (a) indicates the original dataset, (b) present horizontal flip, (c) present vertical flip.

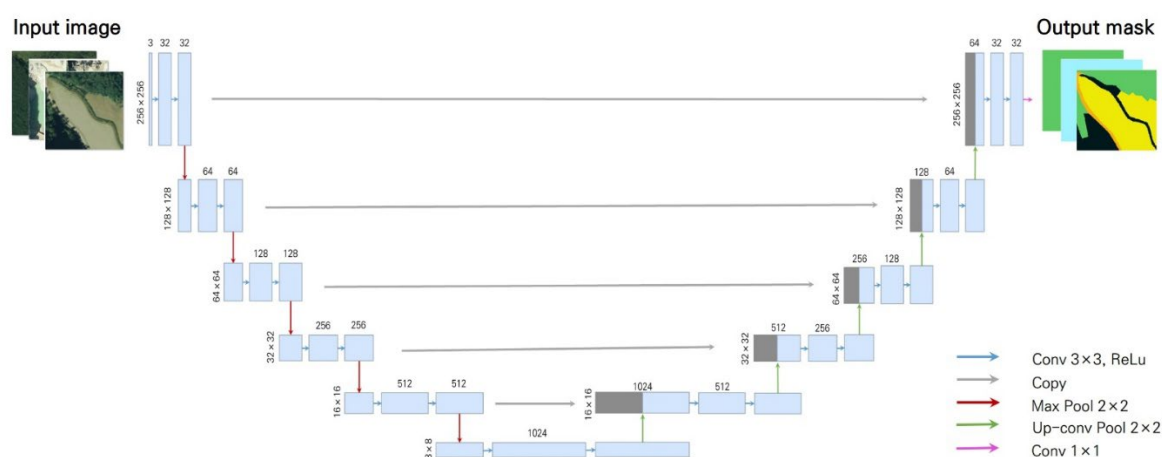
## 2.4. The Architecture Proposed in this Paper ResNet152-UNet

Convolutional Neural Network (CNN) is a deep learning neural network model inspired by the structure of animal visual cortex. It can be divided into the first half, which extracts features by performing convolutional operations, and the second half, which classifies using features. CNN is a model that performs well in recognition problems such as image classification and character recognition [37]. U-Net is a fully convolutional neural network based on FCN (Fully Convolution Network). The original purpose of U-Net's development was to segment biological images [38]. In order to classify the human cells, this model demonstrated strong semantic segmentation performance. As shown in Figure 5, the network model has a U-shape and consists of a total of 23 convolutional layers, which can be divided into an encoder part and a decoder part, i.e., a contracting path to extract image features and an expanding path to reproduce the extracted image features as the original input [39].

For large images, U-Net uses the overlap-tile strategy, which divides the image into tiles instead of using the entire image. For segmenting pixels at the boundary of the image, it uses an extrapolation method using mirroring of the image boundary instead of using arbitrary padding values. With this structure, it performs well on many image segmentation problems using data augmentation even with a small amount of training data. Although it was developed for biomedical image segmentation, it has also been used for land cover classification using various images because it segments images with less information [38,40–42].

In this study, U-Net with a ResNet152 backbone model pre-trained on the ImageNet dataset. Transfer learning is a widely used technique in computer vision for efficiently constructing precise models [43]. Transfer learning in computer vision typically involves utilizing pre-trained models. A pre-trained model is a model that has been trained on a substantial benchmark dataset to address a problem akin to the one we aim to solve. Due to the high computational cost of training these models, it is typical to import and utilize models from published literature such as VGG, Inception, ResNet and MobileNet [44].

ResNet introduces the concept of residual learning. Residual learning solves the problem of gradient vanishing in deep neural networks by using shortcut connections to skip one or more layers and learn to minimize the residual ( $F(x) = H(x) - x$ ). In addition, a  $1 \times 1$  convolutional layer was added before and after the  $3 \times 3$  convolutional layer through a bottleneck structure to reduce and expand the dimensionality of the feature map and reduce computation [45]. This ResNet152-U-Net model comprises approximately 67.2 million parameters.



**Figure 5.** Schematic diagram of the U-Net network structure (Source: Adapted from Figure 5 in Ronneberger et al., 2015).

3. Results and Discussions

3.1. Experimental Settings

The initial learning rate for model training is set to 0.001, with a decay rate of 1e-2. The Adam optimizer is employed with a beta\_1 value of 0.9 and a beta\_2 value of 0.999. Additionally, an exponential decay learning rate schedule is implemented, reducing the learning rate by 5% every ten epochs, and the total number of epochs is set to 200. One of the most renowned optimization algorithms, Adam, is chosen to enhance the model's capabilities. For the loss function, we utilize the categorical focal jaccard loss. The reason is that the foreground has taken a much smaller place than the background, and this disproportion will lead to more missing detection problems. The Focal loss focuses training on a sparse set of hard examples and prevents the vast number of easy negatives from overwhelming the detector during training [46]. The Jaccard loss measures the similarity between finite sample sets A,B as the Intersection over Union (IoU) [47].

In addition, this experiment uses data augmentation methods, which can play an important role in avoiding overfitting, improving the robustness of the model, and improving the expression ability of the model. The overall experiments are conducted on two RTX 3070 GPUs (64 GB memory).

3.2. Accuracy Evaluation

Assessing accuracy is an integral component of every remote sensing task. This study compares the end outcomes of the suggested strategy with the aerial images and ground truth using quantitative and qualitative measures. The quantitative comparison is based on the metrics which are described subsequently.

The test dataset prediction results typically include four categories of pixels: truly positive (TP), false negative (FN), truly negative (TN), and false positive (FP), as illustrated in Table 3. We assess the outcomes by using four well-established accuracy evaluation measures, each derived using distinct methodologies based on the 4 types of pixels mentioned above. The Precision is the ratio of predicted real-change pixels to predicted overall change pixels. The Recall is the ratio of projected real-change pixels to actual overall change pixels. The Pixel Accuracy (PA) is calculated by measuring the percentage of correctly classified pixels in relation to the total number of pixels in the image. The Intersection over Union (IoU) represents the ratio of the intersection to the union of the real-change pixels and the predicted change pixels, which measures the similarity between the predicted situation and the actual situation. Each index may emphasize distinct performances using diverse methodologies, thus requiring a comprehensive evaluation.

Table 3. Formulas for accuracy assessment metrics.

Metric	Formula
Precision	$\frac{TP}{TP + FP}$
	$\frac{TP}{TP + FN}$
Recall	$\frac{TP}{TP + FN}$
	$\frac{TP + TN}{TP + TN + FP + FN}$
Pixel Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
IoU	$\frac{TP}{TP + TN + FP}$
	$\frac{TP}{TP + TN + FN}$

3.3. Segmentation of Forest Region

The model showed a reliable level of performance with an overall accuracy of 95.2% and mean IoU 61.3% (Table 4). The accuracy of each segmentation item showed excellent performance in the arable land, road and barren, quarry, forest. It indicates that the CNN model can efficiently determine the target site for forest restoration using aerial images.



Table 4. Overall accuracy of ResNet152-U-Net.

Model	PA	mIoU
ResNet152-U-Net	95.2	61.3

Table 5 shows the classification accuracy for each category calculated as a result of the model's inference. The accuracy is calculated as the mean Intersection over Union (mIoU), and the highest classification accuracy was obtained for forest, arable land, road and barren, and quarry. Among the classification categories, Forest had the highest accuracy of all categories at 94.6%. Arable land was second with 87.4% accuracy, followed by road and barren with 68.3%, and quarry with a mean IoU of 42.8%.

When comparing the ground truth and the inferred results of each model, it can be concluded that all models can predict the outline of the deforested area relatively accurately, although there are differences in the detailed shape (Figure 6). In addition, the model inference results showed good detection results, such as identifying areas with forest but not labeled in the ground truth data as forest, or identifying non-target areas in the background as items if there is a corresponding area. However, for confusing items such as quarry and arable land, the model prediction results were mixed or fragmented.

Table 5. mIoU (%) of ResNet152-U-Net for 4 class.

Class	mIoU
Arable land	87.4
Road-Barren	68.3
Quarry	42.8
Forest	94.6

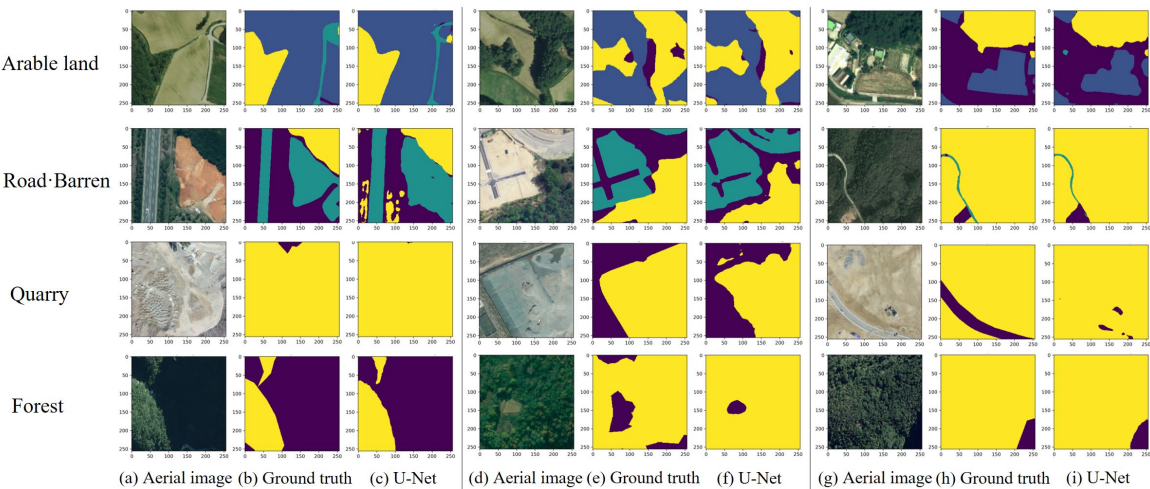


Figure 6. (a, d, g) indicate Aerial image, (b,e,h) indicate Ground truth, (e, f, i) indicate U-Net prediction result.

Table 6 displays a confusion matrix generated by assessing a ResNet152-U-Net model using test data. The confusion matrix arranges the model's predictions in a tabular style, comparing the actual labels with the anticipated labels. Every column in the matrix represents the number of occurrences categorized into a certain combination of expected and actual labels. The classes being considered include arable land, road and barren, quarry, and forest.

The model accurately predicts arable land 80% of the time based on these metrics. The model has a high recall rate of 95.5% for identifying actual arable land cases. In the road and barren class, the model exhibited similar precision but a poorer recall compared to arable land. The model has a precision of 76.3% for predicting road and barren pixels but a recall of 81.8% indicates that it may miss some actual instances. The model demonstrates excellent predictive accuracy (96.0%) for

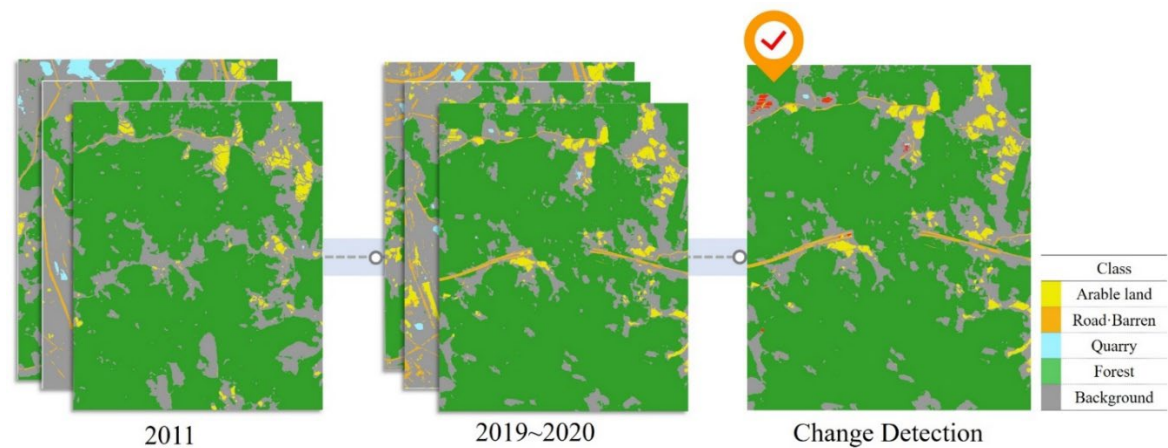
identifying quarries. The model's lower recall rate of 61.9% indicates that it fails to identify a considerable amount of real quarry events. The model's precision of 99.2% and recall of 93.8% indicate its strong ability to identify forested regions. This could be beneficial for tasks such as land cover mapping, environmental monitoring, and related areas.

**Table 3.** Confusion matrix for U-Net model using test data.

Class		True labels				Total	Precision
		Arable land	Road-Barren	Quarry	Forest		
Predicted labels	Arable land	13,577,664	1,280,209	232,831	1,872,599	16,963,303	80.0
	Road-Barren	476,340	6,909,885	325,833	1,348,608	9,060,666	76.3
	Quarry	63	30,296	925,855	8,520	964,734	96.0
	Forest	165,390	224,255	12,302	48,867,163	49,269,110	99.2
	Total	14,219,457	8,444,645	1,496,821	52,096,890	Total Accuracy	
	Recall	95.5	81.8	61.9	93.8	92.2	

3.4. Change Detection of Forest Region

To detect forest degradation, we utilized aerial photographs not employed during the training phase. Comparisons between images from 2011 and those from 2019-2020 facilitated the identification of changes in forested regions. Model prediction outputs stored as raster files were vectorized using Raster to Polygon Vectorizing for quantitative analysis, enabling the utilization of additional spatial analysis and Geographic Information System (GIS) functionalities. Additionally, the Clip function was employed to analyze change areas between the two time periods. To ensure alignment in the time-series imagery, a uniform coordinate system of Korea 2000 / Central Belt 2010 (EPSG:5186) was adopted. For instance, in the case of map index number 377112047, visually discernible changes such as newly constructed roads and land use conversions were observed. This analysis was conducted using the open-source software QGIS (Figure 7).



**Figure 7.** Time series analysis to identify forest restoration targets (index number 377112047).

Table 7 compares the 2011 and 2020 classes and their changes. The model's projections for 50 aerial images from 2011 indicated that there were 744 ha (2.2%) of arable land, 412 ha (1.2%) of road and barren, 585 ha (1.7%) of quarry, 21,394 ha (62.7%) of forest, and 11,010 ha (32.2%) of non-target regions. Model forecasts for 50 aerial images taken between 2019 and 2020 1,468 hectares (4.3%) of arable land, 1,524 hectares (4.5%) of road and barren land, 380 hectares (1.1%) of quarry, 21,799 hectares (64.0%) of woodland, and 8,912 hectares (26.1%) of non-target areas were seen.

Area changes from 2011 to 2020: Arable land increased by 97%, road and barren areas increased by 270%, quarry decreased by 35%, woodland increased by 2%, and non-targeted areas decreased by 19%. The quarry area was reduced while the forest, arable land, road and barren areas rose. The reduction in Non-target area was owing to the prediction that forest shadows will appear as forest in 2020 because of variations in time-of-year and seasonal spectral characteristics.

**Table 4.** Result of Change detection in Gangwon province (2011, 2020).

Class	2011		2019~2020		2011~2020	
	Area (ha)	Rate (%)	Area (ha)	Rate (%)	ΔArea (ha)	ΔRate (%)
Arable land	744	2.2	1,468	4.3	724	97%
Road-Barren	412	1.2	1,524	4.5	1,112	270%
Quarry	585	1.7	380	1.1	(205)	-35%
Forest	21,394	62.7	21,799	64.0	405	2%
Non-target	11,010	32.2	8,912	26.1	(2,098)	-19%

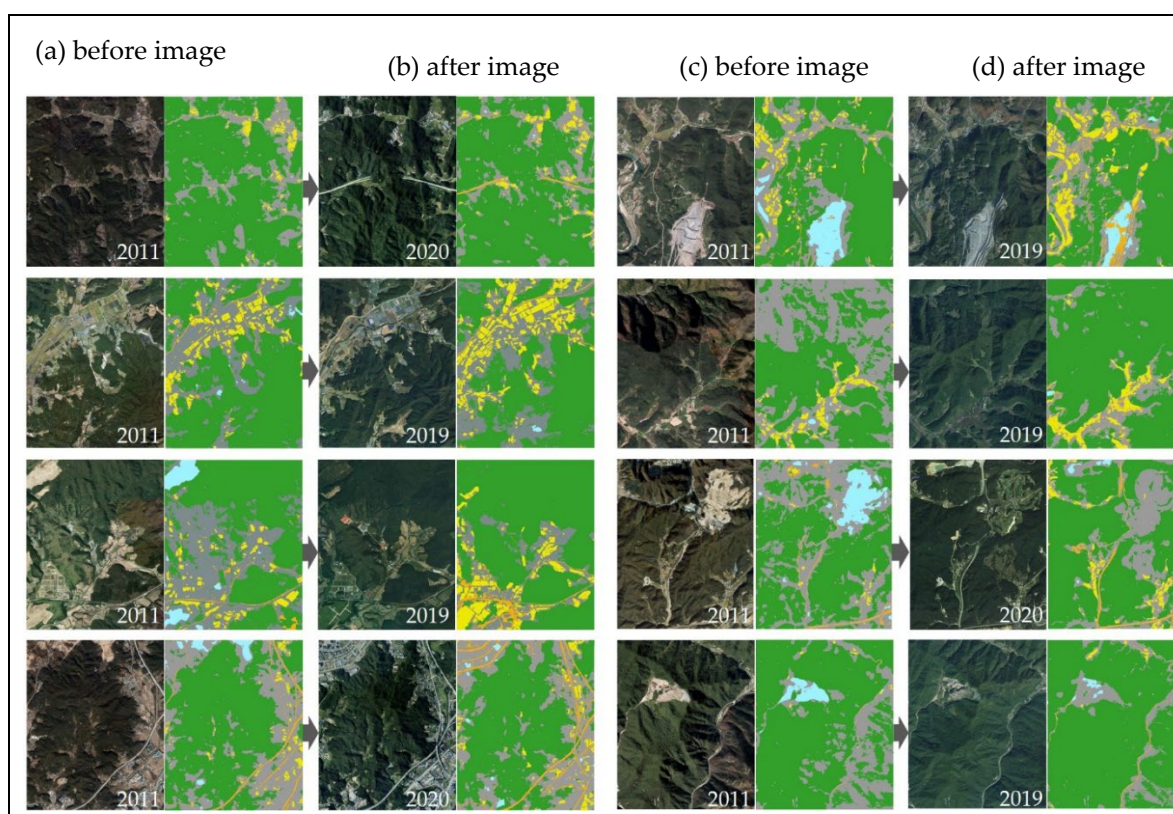
Figure 8 shows the model's prediction findings on a picture of Gangwon Province. The presentation includes a sequence of time-series aerial photos together with forecasts produced by ResNet152-U-Net. The display features eight pairs of photographs, each showing a comparison between a photograph from 2011 and images from 2019 and 2020. This setup clarifies how the model analyzes and forecasts land use changes over time, offering insights into the evolving landscape of the area.

Our results provide for alternative restoration site detection approaches that have a Spatiotemporal advantage over the existing field survey or manual detection of aerial images. These results can contribute to appropriate decisions on forest restoration. Hence, the integration of open-source data, manually annotated data, and deep learning models not only bolstered the viability of change detection but also extended its applicability to diverse domains such as land use classification, tree species classification and biomass estimation [19,48–52].

Furthermore, our study demonstrated the robust efficacy of open-source data, manually annotated data, and deep learning methodologies in forest change detection. However, it also underscored the potential drawback of delayed image acquisition, which could lead to overlooking crucial change occurrences. Image acquisition is time-consuming and costly because it requires the labor of forest experts to create a dataset of forest restoration sites [53,54].

Consequently, further research is imperative to enable real-time detection of forest area transformations induced by phenomena such as floods, landslides, and earthquakes, possibly facilitated by promptly acquiring image data from aerial platforms like drones.





**Figure 8.** Time series aerial images and ResNet152-U-Net prediction. (a, c) indicates before image (b, d) indicates after image.

#### 4. Conclusions

This study evaluated the practicality of utilizing open-source aerial pictures and a manually annotated dataset along with the U-Net model to analyze forest change detection in Gangwon Province, South Korea. Using open-source data helped overcome the limitations of remote sensing acquisition and improved the overall effectiveness of the data-driven approach. The enhancement of the pre-trained U-Net model further optimized the performance of the deep learning model. The ResNet152-U-Net model did a great job of separating forest and arable land, roads and barren, and quarry at the pixel level in images, both during training and validation. The training stages and the exclusion of more than 50 percent of the background label were crucial for improving the semantic segmentation performance of U-Net, specifically in forest regions. U-Net and RGB-based aerial images were used to perform forest restoration site segmentation based on four classes in Gangwon Province, South Korea, with an accuracy value of 95.2%. The accuracy values for 4 classes forest had the highest accuracy of all categories, at 94.6%. arable land came in second with 87.4% accuracy, then road and barren came in third with 68.3%, and quarry came in fourth with a mean IoU of 42.8%. It has been established that the ResNet152-U-Net model is effective in identifying potential places for forest restoration. This will likely be better than both the field survey, which can be expensive, and the visual reading method, which can make mistakes due to the reader's experience and personal thoughts. In the future, if the diversity of data secured by collecting aerial images by various seasonal and types of forest restoration target sites and technical conditions is further supplemented through model optimization research, it is believed that it will be able to contribute to more precise and efficient classification and change monitoring of candidate sites for forest restoration.

**Author Contributions:** Conceptualization, E.S. and H.K.; methodology, E.S. and Y.P.; investigation, Y.P., E.S., U. R. H. and H.S.; resource, J. S. and Y.P.; writing—original draft preparation, E.S.; visualization, E.S.; writing—review and editing, H.K.; project administration, H.K. and J. S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the R&D Program for Forest Science Technology (2021358B10- 2323-BD01) provided by the Korea Forest Service (Korea Forestry Promotion Institute).

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

## References

1. Foley, J.A., et al., *Global consequences of land use*. science, 2005. **309** (5734): p. 570-574.
2. Foley, J.A., et al., *Solutions for a cultivated planet*. Nature, 2011. **478** (7369): p. 337-42.
3. Hosonuma, N., et al., *An assessment of deforestation and forest degradation drivers in developing countries*. Environmental Research Letters, 2012. **7** (4): p. 044009.
4. Marjokorpi, A. and R. Otsamo, *Prioritization of Target Areas for Rehabilitation: A Case Study from West Kalimantan, Indonesia*. Restoration Ecology, 2006. **14** (4): p. 662-673.
5. Wilson, N.R., et al., *Comparison of remote sensing indices for monitoring of desert cienegas*. Arid Land Research and Management, 2016. **30** (4): p. 460-478.
6. Schug, F., et al., *Mapping patterns of urban development in Ouagadougou, Burkina Faso, using machine learning regression modeling with bi-seasonal Landsat time series*. Remote Sensing of Environment, 2018. **210**: p. 217-228.
7. Hermosilla, T., et al., *Prevalence of multiple forest disturbances and impact on vegetation regrowth from interannual Landsat time series (1985–2015)*. Remote Sensing of Environment, 2019. **233**.
8. Deng, Z. and B. Quan, *Intensity Analysis to Communicate Detailed Detection of Land Use and Land Cover Change in Chang-Zhu-Tan Metropolitan Region, China*. Forests, 2023. **14** (5).
9. Yin, X., et al., *Tropical Forest Disturbance Monitoring Based on Multi-Source Time Series Satellite Images and the LandTrendr Algorithm*. Forests, 2022. **13** (12).
10. Abdikan, S., et al., *Burned Area Detection Using Multi-Sensor SAR, Optical, and Thermal Data in Mediterranean Pine Forest*. Forests, 2022. **13** (2).
11. Qarallah, B., et al., *Assessment of Small-Extent Forest Fires in Semi-Arid Environment in Jordan Using Sentinel-2 and Landsat Sensors Data*. Forests, 2022. **14** (1).
12. Lim, S.V., et al., *Attention-Based Semantic Segmentation Networks for Forest Applications*. Forests, 2023. **14** (12).
13. Hamdi, Z.M., M. Brandmeier, and C. Straub, *Forest Damage Assessment Using Deep Learning on High Resolution Remote Sensing Data*. Remote Sensing, 2019. **11** (17).
14. Leo, M., et al., *Computer vision for assistive technologies*. Computer Vision and Image Understanding, 2017. **154**: p. 1-15.
15. Kussul, N., et al., *Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data*. IEEE Geoscience and Remote Sensing Letters, 2017. **14** (5): p. 778-782.
16. Ienco, D., et al., *Weakly Supervised Learning for Land Cover Mapping of Satellite Image Time Series via Attention-Based CNN*. IEEE Access, 2020. **8**: p. 179547-179560.
17. Pan, S., et al., *Land-cover classification of multispectral LiDAR data using CNN with optimized hyper-parameters*. ISPRS Journal of Photogrammetry and Remote Sensing, 2020. **166**: p. 241-254.
18. Brand, A.K. and A. Manandhar, *Semantic Segmentation of Burned Areas in Satellite Images Using a U-Net-Based Convolutional Neural Network*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2021. **XLIII-B3-2021**: p. 47-53.
19. Nuijten, R.J.G., et al., *Monitoring the Structure of Regenerating Vegetation Using Drone-Based Digital Aerial Photogrammetry*. Remote Sensing, 2021. **13** (10).
20. He, Y., et al., *Land Cover and Land Use Mapping of the East Asian Summer Monsoon Region from 1982 to 2015*. Land, 2022. **11** (3).
21. Gomroki, M., M. Hasanlou, and P. Reinartz, *STCD-EffV2T Unet: Semi Transfer Learning EfficientNetV2 T-Unet Network for Urban/Land Cover Change Detection Using Sentinel-2 Satellite Images*. Remote Sensing, 2023. **15** (5).
22. Lee, M.-G., et al., *Detection of Pine Wilt Disease Using Time Series UAV Imagery and Deep Learning Semantic Segmentation*. Forests, 2023. **14** (8).
23. Shahid, M., et al., *Forest Fire Segmentation via Temporal Transformer from Aerial Images*. Forests, 2023. **14** (3).
24. Schmidt, M., et al., *Multi-resolution time series imagery for forest disturbance and regrowth monitoring in Queensland, Australia*. Remote Sensing of Environment, 2015. **158**: p. 156-168.
25. Hussain, M., et al., *Change detection from remotely sensed images: From pixel-based to object-based approaches*. ISPRS Journal of Photogrammetry and Remote Sensing, 2013. **80**: p. 91-106.
26. Service (KFS), K.F., *Forest Restoration Implementation Plan*. 2021. p. 52.
27. Service (KFS), K.F., *Survey of actual conditions in areas subject to forest restoration*. 2016.
28. Pyo, J., et al., *Generalization of U-Net Semantic Segmentation for Forest Change Detection in South Korea Using Airborne Imagery*. Forests, 2022. **13** (12).
29. Maharana, K., S. Mondal, and B. Nemade, *A review: Data pre-processing and data augmentation techniques*. Global Transitions Proceedings, 2022. **3** (1): p. 91-99.



30. Al Shalabi, L., Z. Shaaban, and B. Kasasbeh, *Data mining: A preprocessing engine*. Journal of Computer Science, 2006. **2** (9): p. 735-739.
31. Patro, S. and K.K. Sahu, *Normalization: A preprocessing stage*. arXiv preprint arXiv:1503.06462, 2015.
32. Caruana, R., S. Lawrence, and C. Giles, *Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping*. Advances in neural information processing systems, 2000. **13**.
33. Hinton, G.E., S. Osindero, and Y.-W. Teh, *A fast learning algorithm for deep belief nets*. Neural computation, 2006. **18** (7): p. 1527-1554.
34. Gal, Y. and Z. Ghahramani, *Dropout as a bayesian approximation: Representing model uncertainty in deep learning*. in *international conference on machine learning*. 2016. PMLR.
35. Shorten, C. and T.M. Khoshgoftaar, *A survey on image data augmentation for deep learning*. Journal of big data, 2019. **6** (1): p. 1-48.
36. Taylor, L. and G. Nitschke, *Improving deep learning with generic data augmentation*. in *2018 IEEE symposium series on computational intelligence (SSCI)*. 2018. IEEE.
37. LeCun, Y., et al., *Backpropagation applied to handwritten zip code recognition*. Neural computation, 1989. **1** (4): p. 541-551.
38. Ronneberger, O., P. Fischer, and T. Brox, *U-net: Convolutional networks for biomedical image segmentation*. in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18. 2015. Springer.
39. Siddique, N., et al., *U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications*. IEEE Access, 2021. **9**: p. 82031-82057.
40. Ulmas, P. and I. Liiv, *Segmentation of satellite imagery using u-net models for land cover classification*. arXiv preprint arXiv:2003.02899, 2020.
41. Son, S., et al., *Land-Cover-Change Detection with Aerial Orthoimagery Using SegNet-Based Semantic Segmentation in Namyangju City, South Korea*. Sustainability, 2022. **14** (19).
42. Garioud, A., et al., *FLAIR: a Country-Scale Land Cover Semantic Segmentation Dataset From Multi-Source Optical Imagery*. Advances in Neural Information Processing Systems, 2024. **36**.
43. Rawat, W. and Z. Wang, *Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review*. Neural Comput, 2017. **29** (9): p. 2352-2449.
44. Marcelino, P., *Transfer learning from pre-trained models*. Towards data science, 2018. **10**: p. 23.
45. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
46. Lin, T.-Y., et al. *Focal loss for dense object detection*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
47. Duque-Arias, D., et al. *On power Jaccard losses for semantic segmentation*. in *VISAPP 2021: 16th International Conference on Computer Vision Theory and Applications*. 2021.
48. Potapov, P., et al., *Mapping global forest canopy height through integration of GEDI and Landsat data*. Remote Sensing of Environment, 2021. **253**.
49. Chen, C., et al., *A New Individual Tree Species Classification Method Based on the ResU-Net Model*. Forests, 2021. **12** (9).
50. Korznikov, K.A., et al., *Using U-Net-Like Deep Convolutional Neural Networks for Precise Tree Recognition in Very High Resolution RGB (Red, Green, Blue) Satellite Images*. Forests, 2021. **12** (1).
51. Zhang, B., et al., *A Novel Multi-Scale Attention PFE-UNet for Forest Image Segmentation*. Forests, 2021. **12** (7).
52. Lee, Y., et al., *Evaluation of Hyperparameter Combinations of the U-Net Model for Land Cover Classification*. Forests, 2022. **13** (11).
53. Jia, X., et al., *Small sample-based disease diagnosis model acquisition in medical human-centered computing*. EURASIP Journal on Wireless Communications and Networking, 2019. **2019**: p. 1-12.
54. Do, S., K.D. Song, and J.W. Chung, *Basics of deep learning: a radiologist's guide to understanding published radiology articles on deep learning*. Korean journal of radiology, 2020. **21** (1): p. 33-41.

**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.