Article

Machine Learning-Based Forest Burned Area Detection with Various Input Variables: A Case Study of South Korea

Changhui Lee [1,†], Seonyoung Park [2,†], Taeheon Kim [1], Sicong Liu [3], Mohd Nadzri Md Reba [4], Jaehong Oh [5] and Youkyung Han [1,*]

[1] Department of Civil Engineering, Seoul National University of Science and Technology, 232, Gongneung-ro, Nowon-gu, Seoul 01811, Korea
[2] Department of Applied Artificial Intelligence, Seoul National University of Science and Technology, 232, Gongneung-ro, Nowon-gu, Seoul 01811, Korea
[3] College of Surveying and Geoinformatics, Tongji University, Shanghai 200092, China
[4] Geoscience and Digital Earth Center (INSTeG), Research Institute for Sustainable and Environment (RISE), Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Malaysia
[5] Department of Civil Engineering, Korea Maritime and Ocean University, Busan 49112, Korea
[*] Correspondence: han602@seoultech.ac.kr
[†] These authors contributed equally to this work.

Abstract: Recently, an increase in wildfire incidents has caused significant damage from economical, humanitarian, and environmental perspectives. Wildfires have increased in severity, frequency, and duration because of climate change and rising global temperatures, resulting in the release of massive volumes of greenhouse gases, the destruction of forests and associated habitats, and the damage to infrastructures. Therefore, identifying burned areas is crucial for monitoring wildfire damage. In this study, we aim at detecting forest burned areas occurring in South Korea using optical satellite images. To exploit the advantage of applying machine learning, the present study employs representative three machine learning methods, Light Gradient Boosting Machine (LightGBM), Random Forest (RF), and U-Net, to detect forest burned areas with a combination of input variables, namely Surface Reflectance (SR), Normalized Difference Vegetation Index (NDVI), and Normalized Burn Ratio (NBR). Two study sites of recently occurred forest fire events in South Korea were selected, and Sentinel-2 satellite images were used by considering a small scale of the forest fires. The quantitative and qualitative evaluations according to the machine learning methods and input variables were carried out. In terms of the comparison focusing on machine learning models, the U-Net showed the highest accuracy in both sites amongst the designed variants. The pre and post fire images by SR, NDVI, NBR, and difference of indices as the main inputs showed the best result. We also demonstrated that diverse landcovers may result in a poor burned area detection performance by comparing the results of the two sites.

Keywords: forest fire burned area detection; Sentinel-2; Random Forest; LightGBM; U-Net; input variables analysis; model performance analysis

1. Introduction

In the past decade, increasing wildfire events have become one of the world's most destructive natural hazards, which has resulted in extensive damage from economical, humanitarian, and environmental perspectives [1]. Due to climate change and rising temperatures, wildfires have become more severe, frequent, and long-lasting, which has led to the release of massive amounts of greenhouse gases, the destruction of forests and their related habitats, and the damage to infrastructure and property [2–6]. As a result, identifying burned areas is crucial for monitoring wildfire damage. It aids in the growth of vegetation, the nutrient deposition on the forest floor, and the maintenance of a healthy forest ecosystem [7].

Wildfires frequently occur in areas where several causes interact, making early identification by in situ observations challenging. Remote sensing imaging has been useful in detecting wildfire regions due to its ability to identify changes in large areas over long periods of time that are difficult to notice from the ground [8]. Remote sensing has been a method for wildfire monitoring using the Earth observation satellite data [9–11]. Earth observation systems based on satellites can give consistent and regular readings across large portions of the surface. This provides for quick and cost-effective monitoring of wildfire progression around the world [3].

Satellite sensors like Sentinel-2 and Landsat-8 have multispectral bands sensitive to fire disturbance in the visible, near infrared (NIR), and short-wave infrared (SWIR) [12]. The amount of wildfire-related radiation reflectance can be described in terms of the spectral wavelength [13]. For example, fire-disturbed regions absorb more NIR radiation than unburned areas, whereas burned regions reflect more radiation in the visible and SWIR bands [14]. As a result, several spectral indices for burned area detection have been developed, including the Normalized Difference Vegetation Index (NDVI), the Burned Area Index Modified (BAIM), the Normalized Burn Ratio-SWIR (NBRSWIR), and the Normalized Burn Ratio (NBR) [15–21]. Among them, NBR, constructed by NIR and SWIR bands, has been frequently employed in burned area mapping by utilizing the optical data (e.g., [22,23]). The NBR is primarily sensitive to living chlorophyll and water content of vegetation and soil due to the fact that NIR is sensitive to photosynthetic components and SWIR is sensitive to vegetation moisture. To improve the burned area discrimination, the indices such as dNBR and dNDVI can be differently normalized by the relevant bitemporal pair (i.e., pre-fire and post-fire scenes), which also requires cloud-free pre-fire satellite imagery [24]. Optical-based spectral indices for various vegetation types frequently demonstrate strong sensitivity to low, moderate, and high burn severity, with results that are not site-specific [23,24]. For example, an automatic thresholding chain based on the dNDVI and dNBR was proposed to map the burned areas at a national scale by using Sentinel-2 data [25]. Since the next Earth observation satellites will have onboard data processing and constrained storage for earlier scenes, such bitemporal indices increase preprocessing time and restrict potential uses. When applied to a variety of landscapes, certain automated approaches based on spectral signatures necessitate the time-consuming and difficult fine-tuning of a set of thresholds [26]. Therefore, nonparametric machine learning algorithms have received a lot of attention for outperforming traditional threshold-based approaches using spectral indices in burned area detection [27].

Fortunately, combined with recent remarkable improvements in processing capabilities, the availability of remote sensing data has boosted the ability to use data-centric methods for wildfire detection [27]. Consequently, there has been an increasing interest in using machine learning approaches in wildfire monitoring and management. The ability of machine learning algorithms to automatically detect complicated spatiotemporal patterns in data without the requirement for expert descriptions [28]. Since machine learning algorithms rely on the dispersion and distribution of the training data without making any assumptions, automated burned area detection becomes possible [12]. Various machine learning techniques have been widely applied in wildfire monitoring, such as the Random Forest (RF), the Support Vector Machines (SVM), the Artificial Neural Networks (ANN), and the decision trees [27]. In [29], an automatic burned area mapping system was proposed employing paired Sentinel-2 images, using the SVM for an initial classification and a multiple spectral-spatial classification strategy for delineating the final burned area. Moreover, the RF and the seed-growing techniques were used to construct a worldwide burned area mapping approach [30]. Gibson et al. [31] proposed a study that examined large-scale wildfires in New South Wales, Australia. The researchers performed numerous tests by applying the RF method, which is trained on a variety of spectral indices, such as dNDVI and dNBR.

Recently, deep learning-based techniques were developed to detect wildfire regions using satellite imagery due to their benefits of fusing contextual information and multiscale spatial details [12,31–33]. Pinto et al. [32], for example, suggested a deep learning method for mapping burned areas using MODIS satellite imagery. Langford et al. [33] employed a five-layer deep neural network (DNN) with MODIS derived variables (e.g., NDVI and surface reflectance (SR)) to map fires in Interior Alaska. Several machine learning algorithms, e.g., Light Gradient Boosting Machine (LightGBM), HRNet, U-Net, Fast-SCNN, and DeepLabv3+, were applied to Sentinel-2 and Landsat-8 imagery in wildfire regions [12]. Among the deep-learning approaches, U-Net-based semantic segmentation approaches were commonly used to classify the burned areas [7,34,35]. De Bem et al. [34] recently investigated the performance of U-Net and ResUnet using a bitemporal Landsat image pair. Knopp et al. [35] presented a processing chain based on U-Net for burned area detection using uni-temporal Sentinel-2 data.

This work aims at detecting forest burned areas occurring in South Korea using satellite images. To exploit an advantage of applying machine learning methods, here, we employed three representative machine learning methods, LightGBM, RF, and U-Net, to detect forest burned areas with a combination of input variables including SR, NDVI, and NBR. Three different Schemes while varying the combination of input variables were suggested and compared. Two study sites recently occurring forest fire events in South Korea were selected, and Sentinel-2 satellite images were used by considering a small scale of the forest fires. The quantitative and qualitative evaluations according to the machine learning methods and input variables were carried out.

Main contributions of this study can be summarized as follows. First, a detailed comparison of burned area detection performance between the deep learning-based machine learning model and ensemble learning-based machine learning models was performed. Based on the comparative results, advantages and disadvantages of each model for detecting the burned area were presented, allowing readers to choose the proper model for their detection cases. Second, the burned area detection performance was evaluated while varying the combination of input variables for training the model and obtaining the burned area detection map. It was also included to confirm the possibility of detecting the burned area by using only one satellite image acquired after the forest fire event.

2. Study Area and Data

2.1. Study Area

The study focused on two cases of forest fires that occurred in South Korea (Figure 1). Site A is Goseong, Gangwon-do and Site B is Andong, Gyeongsangbuk-do. As the frequency of dry days in winter and spring has grown, wildfires have become more common in these areas. The forest fire in Site A started at about 19:00 on 4 April 2019 and was put out around 18:00 on 5 April 2019. The damaged area was approximately 1227 ha. A forest fire in Site B started at about 15:00 on 24 April 2020 and was put out at 14:00 on 26 April 2020. The forest fire destroyed a forest area of around 1944 ha. While Site A has heterogeneous land-covers, including forest, cropland, built-up, waterbody, and snow with high-altitude mountainous areas, Site B consists of homogeneous land covers, such as forest and cropland.
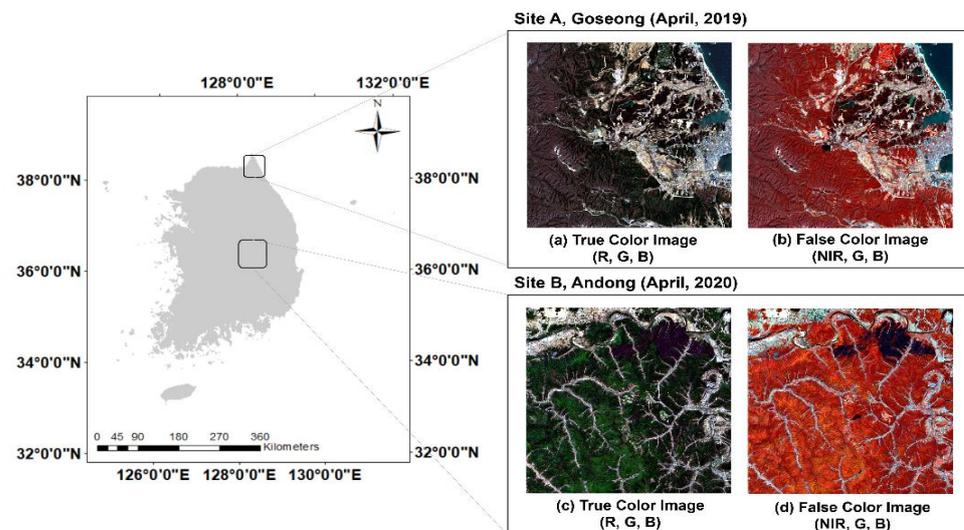
Figure 1. Study area (true color and false color composition of Sentinel-2 image). (**a**) True color image collected on April 2019 of Site A, (**b**) False color image collected on April 2019 of Site A, (**c**) True color image collected on April 2020 of Site B, and (**d**) False color image collected in April 2020 of Site B.

2.2. Data Preparation

Forest fires in Korea are small scale compared to those in Australia or in the United States. Therefore, satellites with low spatial resolution were considered unsuitable due to these characteristics. In this study, the Multi Spectral Instrument (MSI) sensor of the Sentinel-2 satellite operated by the European Space Agency (ESA) was used because it has suitable spatiotemporal resolution to monitor forest fires at a small scale. The Sentinel-2 satellites are made up of two satellites having the same chrematistics (i.e., 2A and 2B) that are moving on the same orbit with 180 degrees apart. The temporal resolution of a single Sentinel-2 satellite is ten days; the combined constellation revisit of the two satellites is thus five days. Table 1 shows the data acquisition date of the study sites. Sentinel-2 satellite images are available at various product levels. Among them, Level 2A products, which have been geometrically and atmospherically corrected by ESA's 'Sen2Cor' process, were used in this study due to their suitability for forest burned area detection studies that require the SR data. There are three spatial resolutions accessible at the Sentinel-2 (10 m, 20 m, and 60 m), and each resolution offers different kind of bands. A total of nine bands with 20 m spatial resolution data—B2 (Blue), B3 (Green), B4 (Red), B5–7 (Red Edge 1–3), B8a (Narrow NIR), B11 (SWIR 1), and B12 (SWIR 2)—were used to classify the forest fire damaged area.

Table 1. Description of satellite image collection.

Sites	Event Date	Acquisition Date		Size	Resolution & Product Level
Site A	4 April 2019	Pre-event	26 March 2019	494 × 655 pixels	
		Post-event	8 April 2019		Spatial Resolution: 20 m
Site B	24 April 2020	Pre-event	14 April 2020	886 × 976 pixels	Product Level: Level-2A
		Post-event	29 April 2020		

To increase the accuracy of the forest fire detection, NDVI, dNDVI, NBR, and dNBR were employed in addition to the nine bands given by Sentinel-2 satellites in this work. NDVI is the most widely used vegetation index in the field of the agriculture and the environment studies. dNDVI was calculated to estimate fire burn damage by subtracting post-fire image from pre-fire image. The NBR is calculated using the NIR band, which is sensitive to vegetation, and the SWIR band, which is sensitive to soil moisture content (Equation (1)) [36]. NBR uses the magnitude of the spectral difference between NIR and SWIR bands normalized by dividing by the sum of the two. A higher NBR value means

healthy vegetation, while a lower NBR value suggests a burned-out area or bare soil. It is possible to distinguish bare soil from the damaged region since it tends to have a value near to 0. In the case of dNBR, it is determined by subtracting the pre-fire and post-fire images in the same way as dNDVI, calculated as Equation (2). The higher the dNBR value in the positive range, the greater the damage; the region where the dNBR value is negative may grow trees and vegetation again after affecting the forest fire. To establish objectivity of the relationship between the dNBR value and the fire burn severity, the United States Geological Survey (USGS) proposed a taxonomy to evaluate forest fire burn severity according to the dNBR value [37].

$$NBR = \frac{NIR - SWIR}{NIR + SWIR} \tag{1}$$

$$dNBR = PrefireNBR - PostfireNBR \tag{2}$$

Ground reference data were collected from satellite images by visual interpretation with the help of the forest map given by the Korea Forest Service. Specifically, the reference data were classified into two classes, i.e., forest and non-forest, by overlaying the forest map. After that, the reference data were refined using Sentienl-2A/B and Google Earth images in 2019 and 2020 through a visual inspection. There are three classes in the ground reference data, i.e., forest, burned area, and non-forest.

3. Methodology

This study aims at detecting small scale forest fire burned areas with various combinations of input variables and machine learning models to compare their performance. To this end, classification of forest fire burned areas was conducted by using nine bands of Sentinel 2A/B satellite images before and after occurring forest fires, together with four vegetation indices and their differences (NBR, NDVI, dNBR, dNDVI) as input variables. Three different machine learning models, LightGBM, RF, and U-Net, were applied.

Of the three types of machine learning models used in this study, the LightGBM and the RF are ensemble learning-based machine learning models that operate on a pixel basis, and the U-Net is a deep learning-based machine learning model that operates based on image patches. The operation process of the pixel-based machine learning models (i.e., Light GBM and RF) and the image patch-based machine learning model (i.e., U-Net) used in this study is shown in Figure 2.
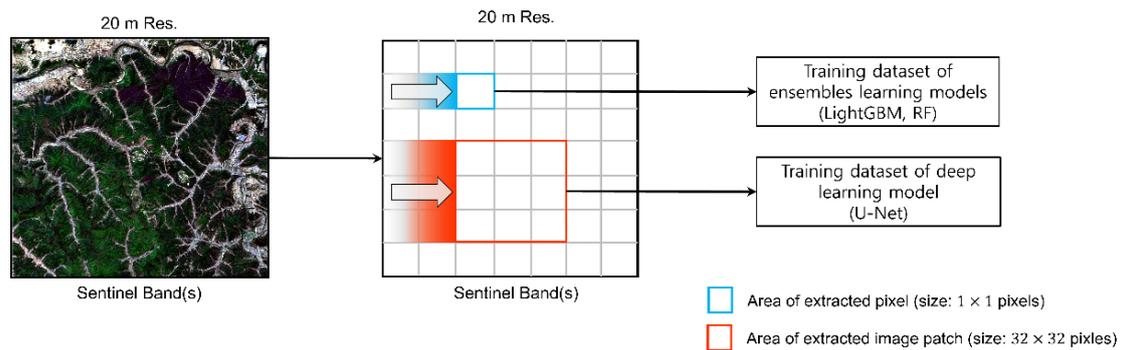

Figure 2. Input data characteristics according to machine learning models.

The flow chart of this study is presented in Figure 3. After conducting the preprocessing to prepare the input variables, three Schemes were developed according to the combination of input variables for performing the machine learning-based burned area detection approaches. More specifically, Scheme 1 used post forest fire image only and two indices, which are NDVI and NBR. Scheme 2 exploited pre and post forest fire images without vegetation index. Finally, using the pre and post forest fire images, four vegetation indices (i.e., two indices for each image) and two difference images of the indices were

selected as input variables for Scheme 3. The dataset of the three divided cases was trained in machine learning models, LightGBM, RF, and U-Net. Then, 60% of the dataset was utilized for training, and 40% of the remaining data was used for validation. Models' performance was tested using the label image of post forest fire. As a result, the quantitative and qualitative forest fire burned area detection performances were analyzed using nine difference cases (i.e., three Schemes with three models for each Site A and Site B). The detailed explanation of the machine learning approaches with the tuned hyperparameter values and of the classification Scheme design was given in the following subsections.
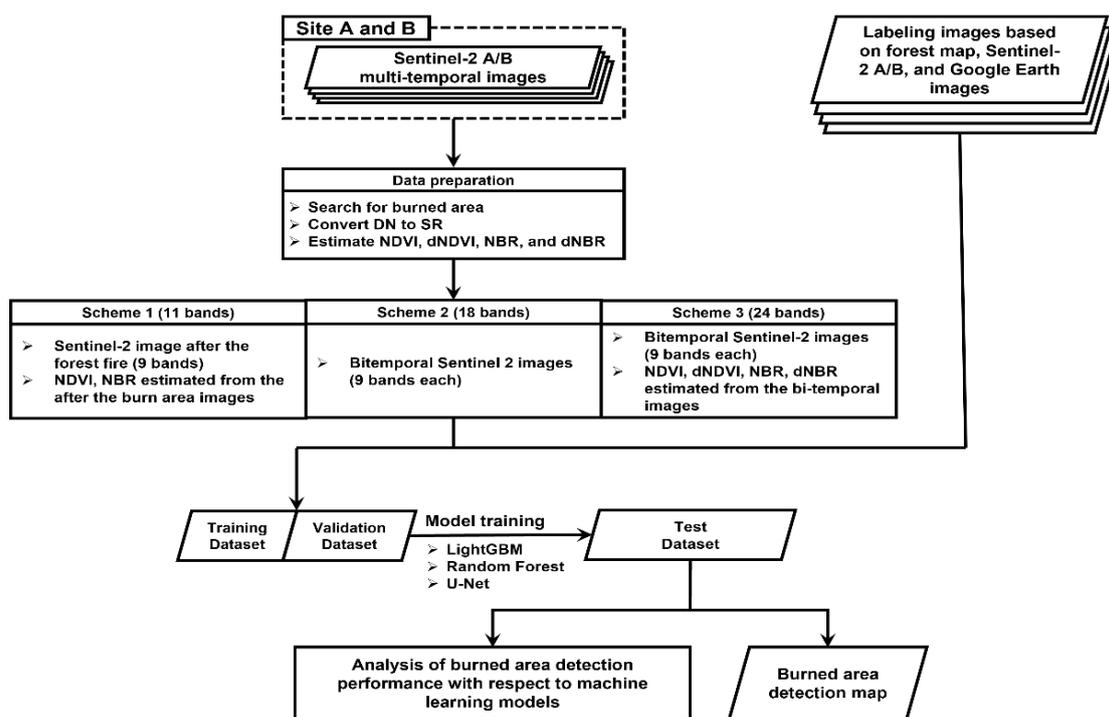

Figure 3. Flow chart of the burned area detection by applying three Schemes.

3.1. Machine Learning Approches

3.1.1. LightGBM

The Gradient Boosting Machine (GBM) is a machine learning model that employs the weights of model as gradients. There are several algorithms developed from existing GBM, including Adaboost, XGboost, and LightGBM. Among them, LightGBM, an ensemble-based machine learning model that has recently been announced, is known to perform much more quickly and accurately [38,39]. Therefore, the LightGBM was employed in this study because of its simplicity but reliable classification performance.

LightGBM is a decision tree-based method that significantly reduces learning time compared to existing GBMs, and it presented a solution to the existing limitation introducing Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB). GOSS is a learning approach that uses only the remaining data with large gradients except for the small gradient element. EFB is a feature reduction technique. In this study, optimal hyperparameters were identified using a grid search based on cross validation. Table 2 shows the optimized parameters for each Scheme.

Table 2. Parameters of LightGBM model.

Model	Parameters		
	Scheme 1	Scheme 2	Scheme 3
Boosting type	GBDT (traditional Gradient Boosting Decision Tree)		
Colsample bytree	0.6	0.8	0.8
Maxdepth	7	1	4
Min. child weight	10	20	10
Num. of estimators	900	300	900
Num. leaves	15	15	15
Subsample	0.6	0.6	0.6
Learning rate	0.005	0.005	0.005

3.1.2. RF

RF is a Classification And Regression Trees (CARTs) based on ensemble approach proposed to overcome the overfitting problem that frequently happens in existing decision trees [40]. It has an advantage of being computationally efficient due to its simple structure [41]. In addition, it has characteristics suitable for training datasets with plenty of variables while showing high performance in multiclass classification of remote sensing data [42]. Considering that many variables are used in this study and performing multiclass classification to detect forest fire burned area, the RF model was judged to be a suitable model for the comparative evaluation.

RF generates multiple decision trees, and each tree predicts the final classification result by aggregating the results of the trees through majority voting. The RF model includes two randomization procedures, random selections of training samples, and input features. These processes alleviate CARTs weaknesses of the overfitting issue. In comparison to the LightGBM model, the RF has less hyperparameter influence on the model's performance. It is also optimized based on the same grid search method as LightGBM. Model parameters optimized were 'Num. of estimators' (50, 70, and 10), 'Max depth' (16, 13, and 10), 'Min. sample leaf' (10, 5, and 5), and 'Min. Sample split' (13, 9, and 9) for Scheme 1, Scheme 2, and Scheme 3, respectively.

3.1.3. U-Net

U-Net is a Fully Convolution Network (FCN)-based model for semantic segmentation introduced in [43]. Despite being intended for biomedical image segmentation, it has been used in many studies in the remote sensing field, such as land cover classification and object detection, because it performs well, even with a small amount of training data [44]. The study sites in South Korea also have limitations in extracting training data because of the small scale of the fire damaged region. For this reason, the U-Net was selected as a deep learning model for performance comparison with ensemble learning-based machine learning models in this study while considering the advantage that high performance can be derived even in the limited training data.

The structure of U-Net is shown in Figure A1 at Appendix A. In U-Net, accuracy was improved by minimizing information loss in the convolution layer using Skip-connection. Furthermore, it is useful since it allows for the usage of previously inaccessible location data [45,46]. In this study, the optimal hyperparameters were determined based on the grid search method. Ten times tests were performed, and different hyperparameter values were applied for each test. Through this process, optimal hyperparameters were determined in consideration of training accuracy, validation accuracy, and loss. Model parameters optimized were Batch Size = 30, Epoch = 100, Learning Rate = 1×10^4, Optimizer = 'Adam', and Loss function = 'Categorical cross entropy'. The 32×32 size patches of input images were extracted and fed into U-Net.

3.2. Classification Scheme Design

A total of 920 sample points were randomly extracted, and then a 32 × 32 image patch was extracted from the points as center. The ratio of samples for each class, which were evenly extracted while considering the area portion of the classes over the study sites, was determined to be 5:2:3 (forest: burned area: non-forest) (Figure 4).
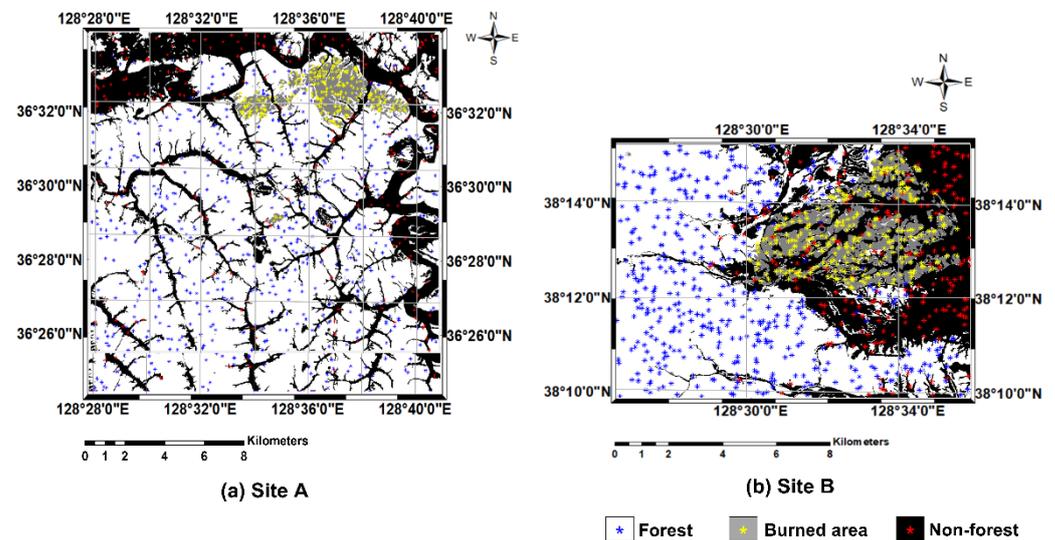

Figure 4. Visualization of sample points extraction for each class. (**a**) Sample points extraction of Site A, (**b**) Sample point extraction of Site B.

In this study, three Schemes were constructed based on different input variables to evaluate the performance of forest fire detection. Scheme 1 is composed of SR, NDVI, and NBR data, which were collected after the forest fire. Scheme 2 utilized data from both before and after the forest fire but did not include the NDVI and NBR. SR, NDVI, and NBR data from both before and after forest fires are included in Scheme 3. The impact of the input data composition on classification accuracy was identified. Table 3 summarizes the data included in each Scheme.

Table 3. Components of dataset by scheme.

Scheme	Num. Spectral Bands	Num. Indices	Total
Scheme 1	9 images (post fire image, 9 bands)	2 images (PostNDVI, PostNBR)	11 variables
Scheme 2	18 images (pre and post fire images, 9 bands each)	-	18 variables
Scheme 3	18 images (pre and post fire images, 9 bands each)	6 images (PreNDVI, PreNBR, PostNDVI, PostNBR, dNDVI, dNBR)	24 variables

3.3. Accuracy Assessment

In this study, the forest fire burned area detection based on machine learning models was conducted by classifying into three classes, i.e., forest, burned area, and non-forest.

However, the purpose of this study is to detect the burned area. Therefore, two classes (forest and non-forest) were combined and expressed as a binary classification image (i.e., burned area and unburned area) for an intuitive understanding of the areas affected by the forest fire burned area. Classification performance was then assessed using five accuracy evaluation metrics: Overall Accuracy (OA), recall, precision, F1-score, and kappa. The OA represents the ratio of the number of correctly predicted samples among all samples (Equation (3)). The recall, also known as sensitivity, is calculated in Equation (4). It represents the ratio between the predicted true value and the actual true value. The

precision (Equation (5)) can be used to assess the accuracy of the recall value because there is a trade-off relationship between these two accuracy assessment metrics. Even though all three metrics are useful, there are limitations to effectively representing the algorithm's performance under the unbalanced dataset. For this reason, quantitative analysis was conducted using both F1-Score and Kappa, which are accuracy evaluation metrics that can sufficiently maintain reliability in an unbalanced dataset. F1-Score is the harmonic average of recall and precision values (Equation (6)), and Cohen's Kappa can evaluate the degree of agreement between the classification results and reference data (Equation (7)).

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

where *TP*, *TN*, *FP*, and *FN* stand for true positive, true negative, false positive, and false negative, respectively.

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{6}$$

$$Kappa = \frac{P_0 - P_e}{1 - P_e}$$
$$P_0 = \frac{TP + TN}{TP + TN + FP + FN} = OA \tag{7}$$
$$P_e = \frac{(TP + FN) \times (TP + FP)}{(TP + TN + FP + FN)^2} + \frac{(FN + TN) \times (FP + TN)}{(TP + TN + FP + FN)^2}$$

4. Result

4.1. Comparison of Forest Burned Area Detection Performance

The classification performance for three Schemes in Site A and B is summarized in Tables 4 and 5, respectively. The values in the tables were averaged with 10 times repetitions of the results. In the case of binary classification of Scheme 3, a simple threshold value was applied to the dNBR image to extract the burned area for the comparison point of view. To this end, the threshold value applied for the extraction of the forest fire burned area was classified into the burned area and the unburned area based on the burn severity standard proposed by the USGS [37]. As a result, the dNBR binary classification map was constructed by mapping the values of low severity (<+0.1) or higher to the wildfire burn area and the values below (>+0.1) to the unburned area. The results from U-Net showed the highest accuracies (OA and Precision) among the three machine learning approaches in both Sites A and B. However, all Schemes in Site B have lower Recall values in U-Net compared to other methods. It is considered that LightGBM and RF tend to over detect burned areas compared to U-Net. As mentioned above, F1-Score and Kappa are much appropriate metrics in the unbalanced dataset. U-Net exhibits an enhanced Kappa and F1-Score in both multiclass and binary classification when compared to LightGBM and RF, indicating that the deep learning model was successful in identifying forest burned regions. The LightGBM, RF, and U-Net classification results also compared with the dNBR result, and showed the better classification performance for three Schemes in site A and B.

Table 4. Summary of the burned area detection results for Site A.

		OA	Recall	Precision	F1-Score	Kappa
Multiclass Classification						
-		OA	Recall	Precision	F1-Score	Kappa
Scheme 1	LightGBM	0.85	0.78	0.60	0.68	0.73
	RF	0.85	0.74	0.62	0.67	0.74
	U-Net	0.93	0.87	0.89	0.88	0.88
Scheme 2	LightGBM	0.86	0.80	0.68	0.73	0.76
	RF	0.86	0.78	0.67	0.72	0.75
	U-Net	0.93	0.88	0.89	0.89	0.88
Scheme 3	LightGBM	0.88	0.81	0.71	0.76	0.78
	RF	0.87	0.81	0.70	0.75	0.77
	U-Net	0.93	0.89	0.88	0.89	0.88
Binary Classification						
-		OA	Recall	Precision	F1-Score	Kappa
Scheme 1	LightGBM	0.92	0.78	0.60	0.68	0.63
	RF	0.92	0.74	0.62	0.67	0.62
	U-Net	0.97	0.87	0.89	0.88	0.86
Scheme 2	LightGBM	0.94	0.80	0.68	0.73	0.69
	RF	0.93	0.78	0.67	0.72	0.68
	U-Net	0.98	0.88	0.89	0.89	0.87
Scheme 3	LightGBM	0.94	0.81	0.71	0.76	0.73
	RF	0.94	0.81	0.70	0.75	0.71
	dNBR	0.92	0.56	0.63	0.59	0.55
	U-Net	0.98	0.89	0.88	0.89	0.88

Table 5. Summary of the burned area detection results for Site B.

		OA	Recall	Precision	F1-Score	Kappa
Multiclass Classification						
-		OA	Recall	Precision	F1-Score	Kappa
Scheme 1	LightGBM	0.90	0.96	0.61	0.74	0.79
	RF	0.91	0.97	0.60	0.74	0.80
	U-Net	0.94	0.90	0.91	0.90	0.87
Scheme 2	LightGBM	0.91	0.97	0.67	0.79	0.81
	RF	0.92	0.97	0.63	0.76	0.81
	U-Net	0.94	0.91	0.92	0.92	0.87
Scheme 3	LightGBM	0.91	0.96	0.67	0.79	0.80
	RF	0.91	0.96	0.68	0.79	0.81
	U-Net	0.94	0.92	0.91	0.92	0.88
Binary Classification						
-		OA	Recall	Precision	F1-Score	Kappa
Scheme 1	LightGBM	0.98	0.96	0.61	0.74	0.74
	RF	0.98	0.97	0.60	0.74	0.73
	U-Net	0.99	0.90	0.91	0.90	0.90
Scheme 2	LightGBM	0.98	0.97	0.67	0.79	0.78
	RF	0.98	0.97	0.63	0.76	0.75
	U-Net	0.99	0.91	0.92	0.92	0.91
Scheme 3	LightGBM	0.98	0.96	0.67	0.79	0.78
	RF	0.98	0.96	0.68	0.79	0.79
	dNBR	0.98	0.93	0.67	0.78	0.77
	U-Net	0.99	0.92	0.91	0.92	0.92

As highlighted in the analysis according to the Schemes, Scheme 3, which used both pre- and post-fire images, showed the highest accuracy, while Scheme 1, which used only

post-fire images, was the lowest accuracy. Because pre- and post-fire images provide change information, it is not surprising that accuracy is high. The use of additional indices (including NDVI, NBR, and dNBR) in Scheme 3 did not significantly improve the detection performance, yielding results comparable to Scheme 2. However, Scheme 3 was more accurate than Scheme 2 in site A, implying that using more indices could be useful in detecting burned areas. The improvement of accuracy is not dramatic in U-Net, showing stable classification accuracy. U-Net conducts image-based learning, while LightGBM and RF conduct pixel-based learning. It demonstrates that image-based learning is more effective in detecting the burned area because it can consider the shape as well as pixel value of the burned area.

4.2. Comparison of Spatial Distributions of Forest Burned Area Detection Results

Figures 5 and 6 illustrate the maps of the burned area for Site A and B, respectively. The area affected by forest fires is expressed as a 'burned area', and the area without forest fire damage is expressed as an 'unburned area'. According to the results in Tables 4 and 5, U-Net shows the best performance among three machine learning approaches in both sites. Pixel-based learning approaches, including LightGBM and RF, showed the noise in burned area detection, while there is little noise in U-Net results. As mentioned above, it is because U-Net could consider both the shape and pixel value of the burned area. The performance of the spatial distribution of burned area detection improved with the use of both pre- and post-fire images (Scheme 2) and additional indices including NDVI, NBR, dNDVI, and dNBR (Scheme 3). It produces less noises of burned area detection as more information is employed.

Different types of land cover can be identified at Sites A and B. While Site A has a heterogeneous distribution of land cover, Site B has a homogeneous distribution. These features explain why Site B's accuracy outperforms than that of Site A (Tables 5 and 6). Similarly, Site B had a less noisy spatial distribution of the burned area. Sites A and B also differ in terms of their topography, which includes elevation and latitude/longitude. Site A has a lower latitude and lower elevation than Site B. This can be confirmed through the True Color Image (TCI) images of Sites A and B. Although the two images were acquired at a similar time, vegetation did not grow in the left area of Site A, and snow was distributed only in some areas at Site A only. On the other hand, it is not possible to check the snow-covered land at all areas of Site B, and it is assumed that vegetation is growing in a balanced manner in all areas by seeing that most of the mountain areas in the image are green. As a result, the forest class in Site A is mixed with snow, resulting to a significantly greater rate of false detection than the forest class in Site B. The spatial distribution of dNBR was compared with machine learning results. It performed poor with lots of noise. False detections in the water body were also identified at both sites when using the dNBR (Figures 5 and 6).

Table 6. Time requirement for training to detect burned area.

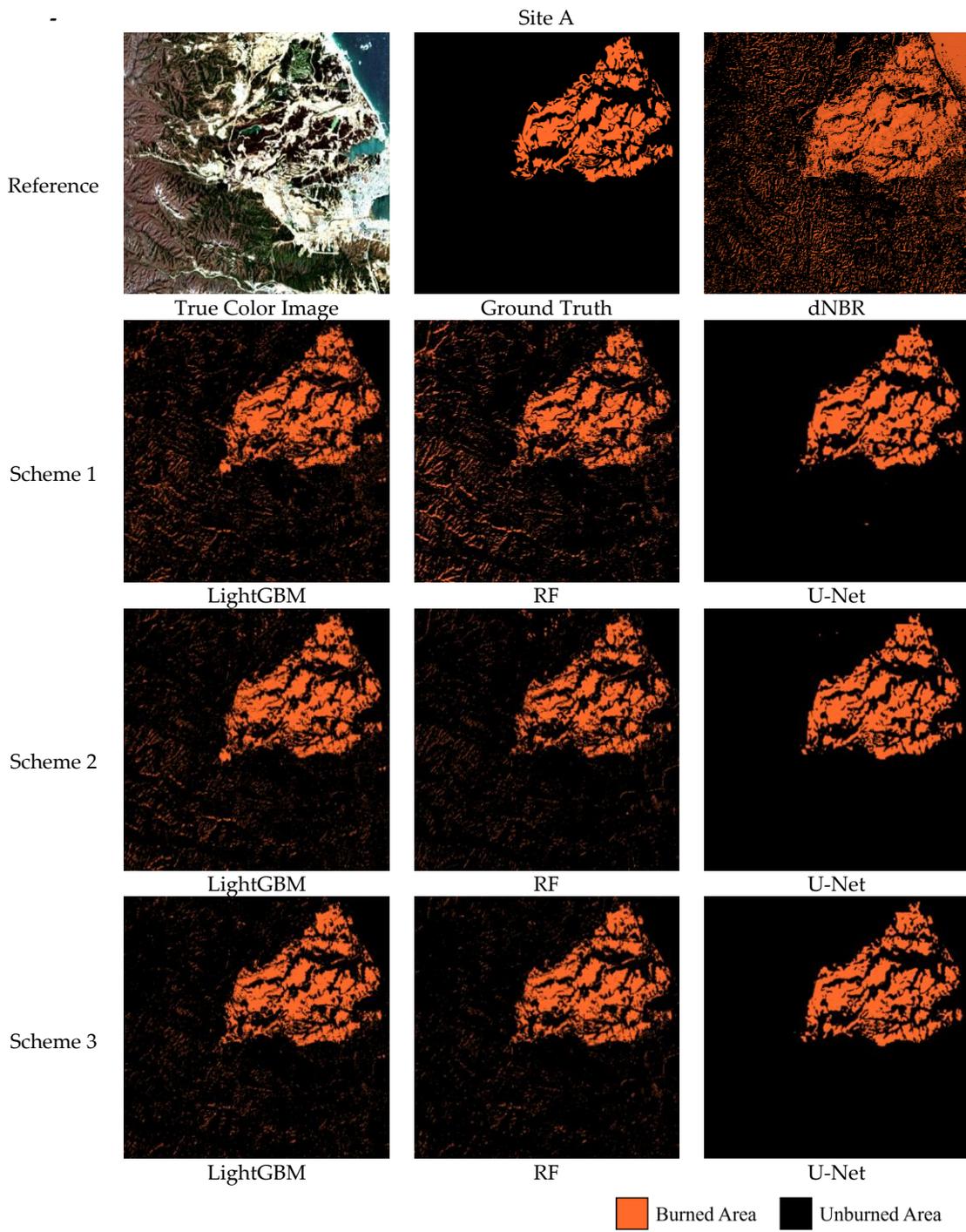| -		Site A		Site B	
-		Accuracy	Training Time	Accuracy	Training Time
Scheme 1	LightGBM	0.88	175.5 s	0.92	143.8 s
	RF	0.78	17.5 s	0.95	17.6 s
	U-Net	0.91	173.2 s	0.94	175.4 s
Scheme 2	LightGBM	0.89	190.4 s	0.93	163.9 s
	RF	0.84	18.2 s	0.94	17.3 s
	U-Net	0.90	177.2 s	0.97	176.8 s
Scheme 3	LightGBM	0.90	196.2 s	0.94	175.6 s
	RF	0.88	18.5 s	0.90	17.3 s
	U-Net	0.91	174.1 s	0.97	179.2 s

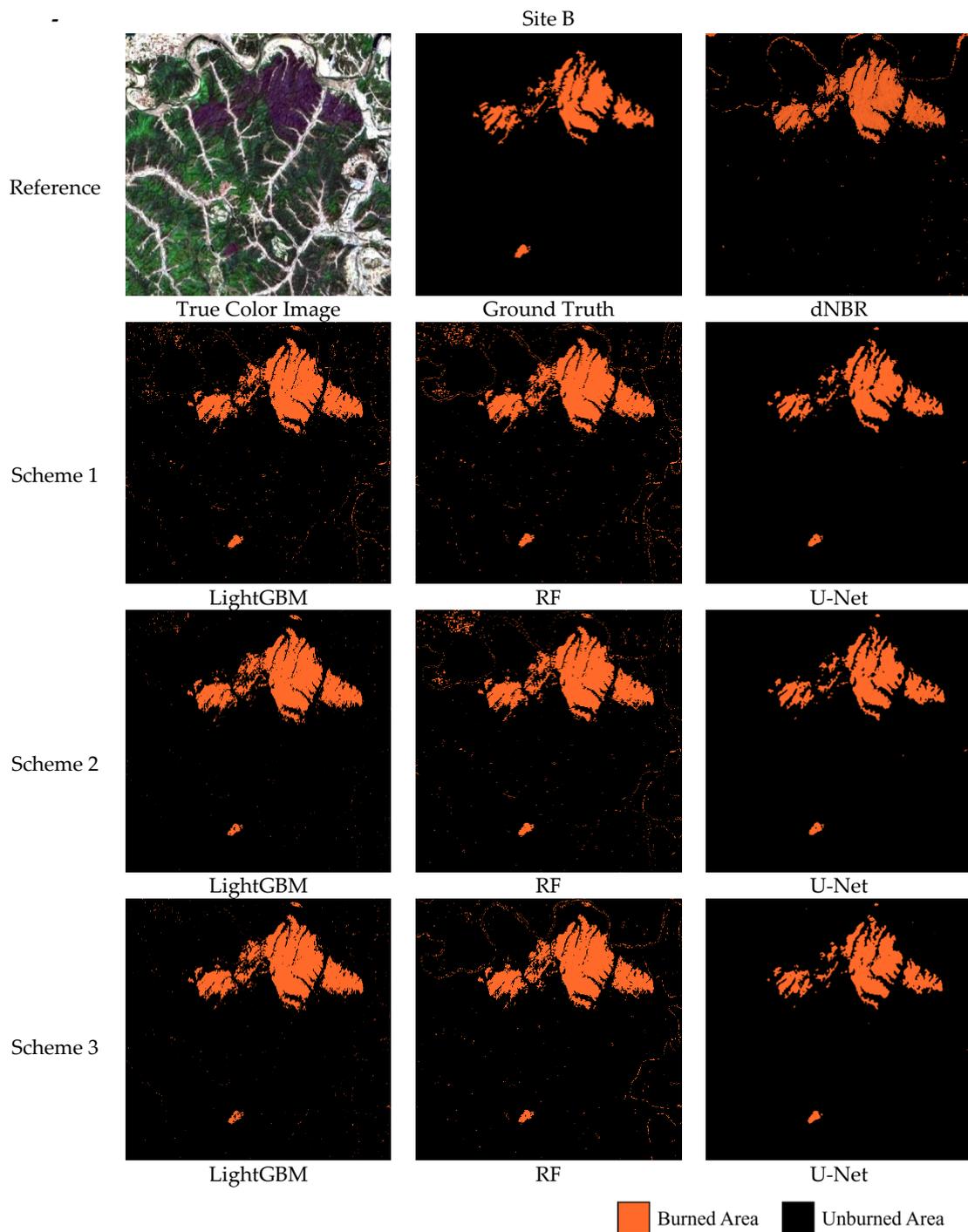**Figure 5.** Spatial distribution of burned area detection results in Site A.

Figure 6. Spatial distribution of burned area detection results in Site B.

5. Discussion

5.1. Analysis of Forest Fire Burned Area Detection Errors

In this study, LightGBM and RF, which are ensemble learning-based machine learning approaches, and U-Net, which is a deep learning-based machine learning approach, were used to detect the burned area caused by forest fires. Sentinel-2A/B satellite images before and after occurring forest fires were used as input data (e.g., SR, NDVI, and NBR). Three Schemes, each of which was composed of different combinations of the aforementioned data, were evaluated. Based on accuracy assessment, the optimal model for detecting burned areas was identified. In both sites, the U-Net showed greater overall accuracy than

the LightGBM and RF (Figures 5 and 6). Figure 6 also demonstrates this tendency. The diverse patterns of land cover at Site A revealed noise, resulting in misclassification in the burned area [47]. Although the ensemble-based machine learning models generated numerous misclassified pixels in Schemes 1 and 2, the rate of misclassified pixels was reduced in Scheme 3 (Figures 7 and 8). This is because of the incorporation of vegetation indices as input data in Scheme 3. It indicates that use of temporal data (before and after the forest fire) as well as additional input data (vegetation indices) contributed to improve the burned area detection accuracy.
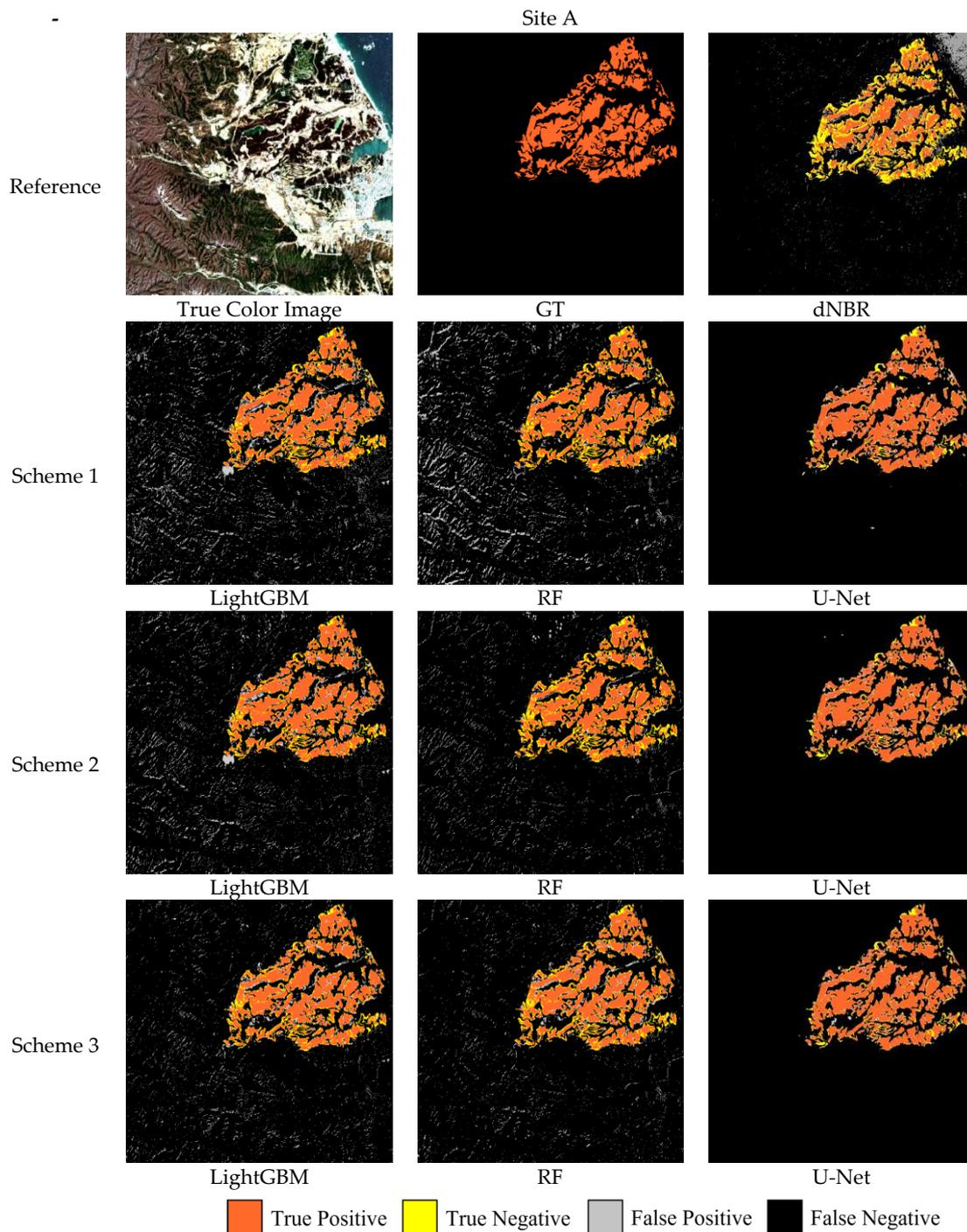
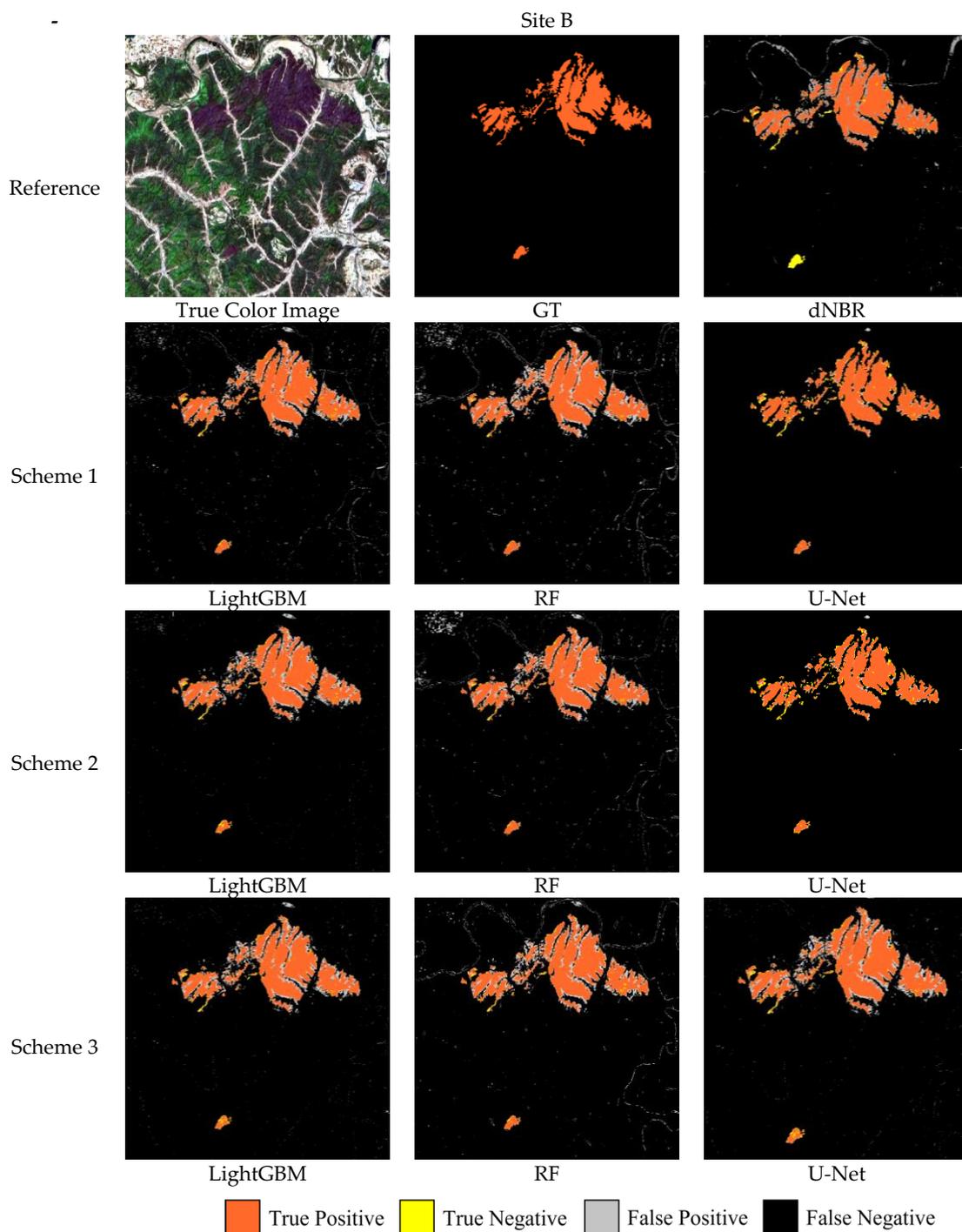**Figure 7.** Visualization of burned area detection errors in Site A.

Figure 8. Visualization of burned area detection errors in Site A.

The U-Net obtained robust modeling performance in all Schemes. Although misclassified pixels were detected, most of the pixels were identified close to the burned area's boundary. This is because the degree of damage weakens toward the boundary [48]. The ensemble learning-based machine learning models generally exhibits a strong sensitivity to the properties of the training data and weak performance in noise [49]. In contrast, the U-Net model is more resilient to noise and less sensitive to the properties of the training data than the ensemble-based machine learning models [50–52]. It is therefore appropriate to use the U-Net to identify burned areas from forest fires.

In Site B, the ensemble learning-based machine learning models outperformed U-Net in terms of recall value (Table 4). It is because the ensemble-based models over-detected the burned areas. This tendency is clearly shown in Figure 8. The recall value errors are due to the FP pixels, which are presented as gray colors in Figure 8. Detecting more pixels than the actual number of pixels indicates false alarms, and the recall value could be abnormally high [53]. In this case, using the F1-Score and Kappa coefficient can be suitable for evaluating the detection performance rather than the recall error. [54]. The highest F1-Score and Kappa values were consistently achieved by the U-Net for all Schemes in Site B.

5.2. Assessment of Training Efficiency According to the Models

The training time and accuracy of the models were tested for evaluating the efficiency. The identical hardware environment was used to train the machine learning models (i.e., AMD Ryzen 95900X 12-Core CPU, NVIDIA Geforce RTX 3060, and 64GB RAM). The results are summarized in Table 6. There was a notable difference in training times between the ensemble learning-based machine learning models. The LightGBM model, compared to the RF model, has more parameters to be tuned [55,56] so that the longer computation time is required. The training accuracy for each Scheme for LightGBM and RF models is also shown to be unstable. It exhibited low accuracy in Site A and high accuracy in Site B. Although the computation speed of U-Net for burned area detection was not the fastest, U-Net demonstrated consistent accuracy regardless of Sites and Schemes. In order to detect burned areas, deep learning approaches like U-Net could produce reliable results.

6. Conclusions

Forest fires are occurring frequently not just in Korea but all around the world. For the rapid restoration of forests, a more accurate burned area detection map is required. Satellite remote sensing effectively provides the accurate extent of burned area and damaged map from forest fire because it can provide spatiotemporal images with a low cost. In this study, three machine learning models (LightGBM, RF, and U-Net) were used to detect burned areas and three Schemes were tested to investigate how the input data composition affects the performance of each model. Through the qualitative and quantitative assessment, a model adequate for detecting burned areas was identified. Among the three machine learning approaches, U-Net, a deep learning-based model, showed the robust and high accuracy in detecting burned areas. Ensemble-based approaches, such as LightGBM and RF, performed poorly and showed noise in Site A showing the heterogeneous land covers. When using satellite images before and after forest fires as well as NDVI and NBR indices, the burned area detection performance was the best regardless of the input variables in both Sites A and B. It was also found that obtaining the accurate burned area detection results was difficult by using one image acquiring after the forest fire event. Although machine learning approaches performed significantly worse due to a lack of information, the U-Net had the highest accuracy. It is considered that the U-Net can be used even with limited data, i.e., only satellite images after occurring the forest fires. Post-disaster images are generally easy to collect, while prior images are more challenging. This benefit is regarded as an important feature in monitoring disasters, such as the detection of burned areas from forest fires. The results of the qualitative and quantitative studies lead to the conclusion that the deep learning-based model outperforms the ensemble learning-based machine learning models in identifying burned regions from forest fires.

This study tested three Schemes using different combinations of SR and two indices. In addition to used indices, various satellite-based factors that are related to forest fires, such as the soil index or water index, can be used to detect the burned area. It is expected that by using more information, it will be feasible to detect burned areas with greater accuracy. Additionally, the distribution of land covers was not properly accounted for the training data, especially in Site A where the land covers are heterogeneous. Future research should thus consider topographical variables such as land covers and broaden the variety of data by including information directly impacted by forest fires. Furthermore, beyond using

the well-known model like U-Net itself, future works will develop the model to improve the performance of the forest fire burned area detection. Finally, study sites occurring in large-scale forest fires, e.g., in the United States or Australia, will be considered to confirm whether the demonstrated results from this study are applicable, such as in wide areas as well.

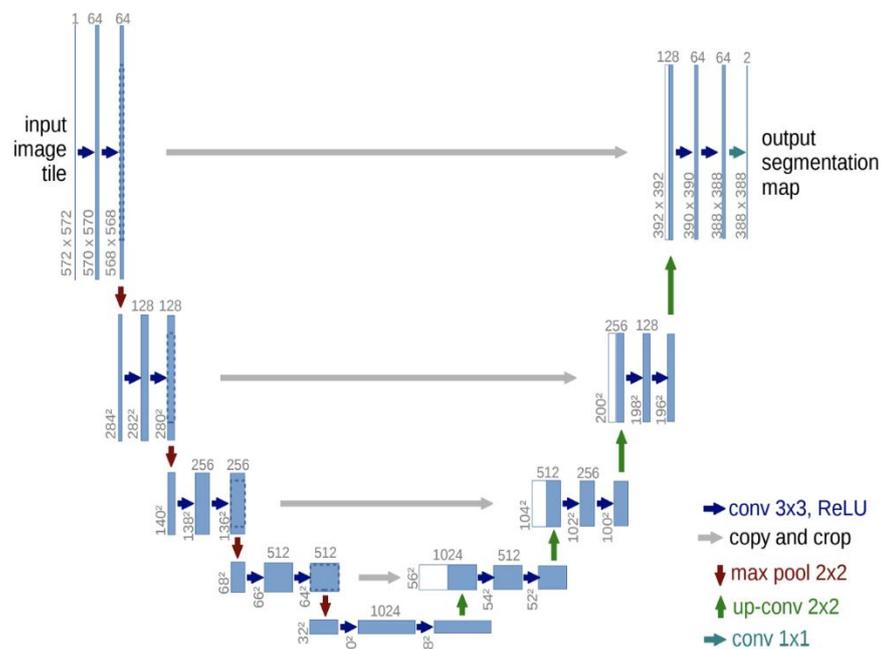**Appendix A. Structure of U-Net**

Figure A1. Structure of U-Net.

Appendix A shows the structure and characteristics of U-Net. Figure A1 shows the U-Net structure. In the structure of the encoder and decoder, this model has 23 convolutional layers. U-Net's encoder is called a contracting path, and it converts a high-dimensional image into a low-dimensional image and extracts features. After that, it is composed of a decoder structure called an expanding path that changes the encoded image back to its original form. Unlike other semantic segmentation models, U-Net adapts to the skip connection mechanism to connect compressed data from the contracting path to the expanding path directly.

References

1. Farasin, A.; Colomba, L.; Garza, P. Double-step u-net: A deep learning-based approach for the estimation of wildfire damage severity through sentinel-2 satellite data. *Appl. Sci.* **2020**, *10*, 4332. [CrossRef]
2. Rashkovetsky, D.; Mauracher, F.; Langer, M.; Schmitt, M. Wildfire Detection From Multisensor Satellite Imagery Using Deep Semantic Segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7001–7016. [CrossRef]
3. Coogan, S.C.; Robinne, F.N.; Jain, P.; Flannigan, M.D. Scientists' warning on wildfire—A Canadian perspective. *Can. J. Forest Res.* **2019**, *49*, 1015–1023. [CrossRef]
4. Zhang, P.; Ban, Y.; Nascetti, A. Learning U-Net without forgetting for near real-time wildfire monitoring by the fusion of SAR and optical time series. *Remote Sens. Environ.* **2021**, *261*, 112467. [CrossRef]
5. Lozano, O.M.; Salis, M.; Ager, A.A.; Arca, B.; Alcasena, F.J.; Monteiro, A.T.; Finney, M.A.; Del Giudice, L.; Scoccimarro, E.; Spano, D. Assessing climate change impacts on wildfire exposure in Mediterranean areas. *Risk Anal.* **2017**, *37*, 1898–1916. [CrossRef]
6. Littell, J.S.; McKenzie, D.; Wan, H.Y.; Cushman, S.A. Climate change and future wildfire in the western United States: An ecological approach to nonstationarity. *Earth's Future* **2018**, *6*, 1097–1111. [CrossRef]
7. Vil'a-Vilardell, L.; Keeton, W.S.; Thom, D.; Gyeltshen, C.; Tshering, K.; Gratzer, G. Climate change effects on wildfire hazards in the wildland-urban-interface–blue pine forests of Bhutan. *For. Ecol. Manag.* **2020**, *461*, 117927. [CrossRef]
8. Liu, C.C.; Chen, Y.H.; Wu, M.H.M.; Wei, C.; Ko, M.H. Assessment of forest restoration with multitemporal remote sensing imagery. *Sci. Rep.* **2019**, *9*, 7219. [CrossRef]
9. Chuvieco, E.; Mouillot, F.; van der Werf, G.R.; San Miguel, J.; Tanase, M.; Koutsias, N.; García, M.; Yebra, M.; Padilla, M.; Gitas, I.; et al. Historical background and current developments for mapping burned area from satellite Earth observation. *Remote Sens. Environ.* **2019**, *225*, 45–64. [CrossRef]
10. Fornacca, D.; Ren, G.; Xiao, W. Performance of three MODIS fire products (MCD45A1, MCD64A1, MCD14ML), and ESA Fire_CCI in a mountainous area of Northwest Yunnan, China, characterized by frequent small fires. *Remote Sens.* **2017**, *9*, 1131. [CrossRef]
11. Wang, J.; Sammis, T.W.; Gutschick, V.P.; Gebremichael, M.; Dennis, S.O.; Harrison, R.E. Review of satellite remote sensing use in forest health studies. *Open Geogr. J.* **2010**, *3*, 28–42. [CrossRef]
12. Hu, X.; Ban, Y.; Nascetti, A. Uni-Temporal Multispectral Imagery for Burned Area Mapping with Deep Learning. *Remote Sens.* **2021**, *13*, 1509. [CrossRef]
13. Kontoes, C.C.; Poilvé, H.; Florsch, G.; Keramitsoglou, I.; Paralikidis, S. A comparative analysis of a fixed thresholding vs. a classification tree approach for operational burn scar detection and mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 299–316. [CrossRef]
14. Quintano, C.; Fernández-Manso, A.; Stein, A.; Bijker, W. Estimation of area burned by forest fires in Mediterranean countries: A remote sensing data mining perspective. *For. Ecol. Manag.* **2011**, *262*, 1597–1607. [CrossRef]
15. Roteta, E.; Bastarrika, A.; Padilla, M.; Storm, T.; Chuvieco, E. Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa. *Remote Sens. Environ.* **2019**, *222*, 1–17. [CrossRef]
16. Trigg, S.; Flasse, S. An evaluation of different bi-spectral spaces for discriminating burned shrub-savannah. *Int. J. Remote Sens.* **2001**, *22*, 2641–2647. [CrossRef]
17. Chu, T.; Guo, X. Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: A review. *Remote Sens.* **2013**, *6*, 470–520. [CrossRef]
18. Huang, H.; Roy, D.P.; Boschetti, L.; Zhang, H.K.; Yan, L.; Kumar, S.S.; Gomez-Dans, J.; Li, J.; Huang, H.; Roy, D.P.; et al. Separability analysis of Sentinel-2A Multi-Spectral Instrument (MSI) data for burned area discrimination. *Remote Sens.* **2016**, *8*, 873. [CrossRef]
19. Navarro, G.; Caballero, I.; Silva, G.; Parra, P.C.; Vázquez, Á.; Caldeira, R. Evaluation of forest fire on Madeira Island using Sentinel-2A MSI imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 97–106. [CrossRef]
20. Quintano, C.; Fernández-Manso, A.; Fernández-Manso, O. Combination of Landsat and Sentinel-2 MSI data for initial assessing of burn severity. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *64*, 221–225. [CrossRef]
21. Liu, S.; Zheng, Y.; Dalponte, M.; Tong, X. A novel fire index based burned area change detection approach using Landsat-8 OLI data. *Eur. J. Remote Sens.* **2020**, *53*, 104–112. [CrossRef]
22. Escuin, S.; Navarro, R.; Fernandez, P. Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *Int. J. Remote Sens.* **2008**, *29*, 1053–1073. [CrossRef]
23. Cardil, A.; Mola-Yudego, B.; Blázquez-Casado, Á.; González-Olabarria, J.R. Fire and burn severity assessment: Calibration of Relative Differenced Normalized Burn Ratio (RdNBR) with field data. *J. Environ. Manag.* **2019**, *235*, 342–349. [CrossRef] [PubMed]
24. Miller, J.D.; Thode, A.E. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sens. Environ.* **2007**, *109*, 66–80. [CrossRef]
25. Pulvirenti, L.; Squicciarino, G.; Fiori, E.; Fiorucci, P.; Ferraris, L.; Negro, D.; Gollini, A.; Severino, M.; Puca, S. An automatic processing chain for near real-time mapping of burned forest areas using sentinel-2 data. *Remote Sens.* **2020**, *12*, 674. [CrossRef]
26. Smith, A.M.; Drake, N.A.; Wooster, M.J.; Hudak, A.T.; Holden, Z.A.; Gibbons, C.J. Production of Landsat ETM+ reference imagery of burned areas within Southern African savannahs: Comparison of methods and application to MODIS. *Int. J. Remote Sens.* **2007**, *28*, 2753–2775. [CrossRef]
27. Jain, P.; Coogan, S.C.; Subramanian, S.G.; Crowley, M.; Taylor, S.; Flannigan, M.D. A review of machine learning applications in wildfire science and management. *Environ. Rev.* **2020**, *28*, 478–505. [CrossRef]

28. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [CrossRef]

29. Stavrakoudis, D.; Katagis, T.; Minakou, C.; Gitas, I.Z.; Stavrakoudis, D.; Katagis, T.; Minakou, C.; Gitas, I.Z. Automated Burned Scar Mapping Using Sentinel-2 Imagery. *J. Geogr. Inf. Syst.* **2020**, *12*, 221–240. [CrossRef]

30. Long, T.; Zhang, Z.; He, G.; Jiao, W.; Tang, C.; Wu, B.; Zhang, X.; Wang, G.; Yin, R. 30 m resolution global annual burned area mapping based on landsat images and Google Earth Engine. *Remote Sens.* **2019**, *11*, 489. [CrossRef]

31. Gibson, R.; Danaher, T.; Hehir, W.; Collins, L. A remote sensing approach to mapping fire severity in south-eastern Australia using sentinel 2 and random forest. *Remote Sens. Environ.* **2020**, *240*, 111702. [CrossRef]

32. Pinto, M.M.; Libonati, R.; Trigo, R.M.; Trigo, I.F.; DaCamara, C.C. A deep learning approach for mapping and dating burned areas using temporal sequences of satellite images. *ISPRS J. Photogramm. Remote Sens.* **2020**, *160*, 260–274. [CrossRef]

33. Arruda, V.L.; Piontekowski, V.J.; Alencar, A.; Pereira, R.S.; Matricardi, E.A. An alternative approach for mapping burn scars using Landsat imagery, Google Earth Engine, and Deep Learning in the Brazilian Savanna. *Remote Sens. Appl. Soc. Environ.* **2021**, *22*, 100472. [CrossRef]

34. De Bem, P.P.; de Carvalho Júnior, O.A.; de Carvalho, O.L.F.; Gomes, R.A.T.; Fontes Guimarães, R. Performance analysis of deep convolutional autoencoders with different patch sizes for change detection from burnt areas. *Remote Sens.* **2020**, *12*, 2576. [CrossRef]

35. Knopp, L.; Wieland, M.; Rättich, M.; Martinis, S. A deep learning approach for burned area segmentation with Sentinel-2 data. *Remote Sens.* **2020**, *12*, 2422. [CrossRef]

36. Key, C.H.; Benson, N.C. Measuring and remote sensing of burn severity. In Proceedings of the Joint Fire Science Conference and Workshop, Boise, ID, USA, 15–17 June 1999; Neuenschwander, L.F., Ryan, K.C., Eds.; Univeristy of Idaho: Moscow, ID, USA, 1999; Volume II, p. 284.

37. United Nations. Office for Outer Space Affairs UN-SPIDER Knowledge Portal. Normalized Burn Ratio (NBR). Available online: https://www.un-spider.org/advisory-support/recommended-practices/recommended-practice-burn-severity/in-detail/normalized-burn-ratio (accessed on 19 September 2022).

38. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. Lightgbm: A highly efficient gradient boosting decision tree. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.

39. Park, J.; Chung, Y.R.; Nose, A. Comparative analysis of high-and low-level deep learning approaches in microsatellite instability prediction. *Sci. Rep.* **2022**, *12*, 12218. [CrossRef]

40. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

41. Shadman Roodposhti, M.; Aryal, J.; Lucieer, A.; Bryan, B.A. Uncertainty assessment of hyperspectral image classification: Deep learning vs. random forest. *Entropy* **2019**, *21*, 78. [CrossRef]

42. Mahapatra, D. Analyzing training information from random forests for improved image segmentation. *IEEE Trans. Image Process.* **2014**, *23*, 1504–1512. [CrossRef]

43. 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 1999; Springer: Cham, Switzerland, 2015; pp. 234–241.

44. Solórzano, J.V.; Mas, J.F.; Gao, Y.; Gallardo-Cruz, J.A. Land use land cover classification with U-net: Advantages of combining sentinel-1 and sentinel-2 imagery. *Remote Sens.* **2021**, *13*, 3600. [CrossRef]

45. Zhang, J.; Du, J.; Liu, H.; Hou, X.; Zhao, Y.; Ding, M. LU-NET: An improved U-Net for ventricular segmentation. *IEEE Access* **2019**, *7*, 92539–92546. [CrossRef]

46. Wonho, J.; Park, K.H. Deep Learning Based Land Cover Change Detection Using U-Net. *J. Korean Geogr. Soc.* **2022**, *57*, 297–306.

47. Karpatne, A.; Jiang, Z.; Vatsavai, R.R.; Shekhar, S.; Kumar, V. Monitoring land-cover changes: A machine-learning perspective. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 8–21. [CrossRef]

48. FuenTes-sAnTos, I.; Marey-Pérez, M.F.; González-Manteiga, W. Forest fire spatial pattern analysis in Galicia (NW Spain). *J. Environ. Manag.* **2013**, *128*, 30–42. [CrossRef] [PubMed]

49. Duan, S.; Huang, S.; Bu, W.; Ge, X.; Chen, H.; Liu, J.; Luo, J. LightGBM low-temperature prediction model based on LassoCV feature selection. *Math. Probl. Eng.* **2021**, *2021*, 1776805. [CrossRef]

50. Oreski, S.; Oreski, D.; Oreski, G. Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessment. *Expert Syst. Appl.* **2012**, *39*, 12605–12617. [CrossRef]

51. Yeh, I.C.; Lien, C.H. The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Syst. Appl.* **2009**, *36*, 2473–2480. [CrossRef]

52. Bui, D.T.; Tsangaratos, P.; Nguyen, V.T.; Van Liem, N.; Trinh, P.T. Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. *Catena* **2020**, *188*, 104426. [CrossRef]

53. Park, S.; Im, J.; Park, S.; Yoo, C.; Han, H.; Rhee, J. Classification and mapping of paddy rice by combining Landsat and SAR time series data. *Remote Sens.* **2018**, *10*, 447. [CrossRef]

54. Yang, X.; Sun, H.; Fu, K.; Yang, J.; Sun, X.; Yan, M.; Guo, Z. Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks. *Remote Sens.* **2018**, *10*, 132. [CrossRef]

55.	Minastireanu, E.A.; Mesnita, G. Light gbm machine learning algorithm to online click fraud detection. *J. Inform. Assur. Cybersecur.* **2019**, *2019*, 263928. [CrossRef]
56.	LightGBM, Parameters Tuning. Available online: https://lightgbm.readthedocs.io/en/latest/Parameters-Tuning.html (accessed on 19 September 2022).