SSC23-WI-04

Early Wildfire Detection with CubeSat Images Using Single Image Super-Resolution

Jimin Choi
Korea Advanced Institute of Science and Technology (KAIST)
Daejeon 34141, Republic of Korea; 82-10-8411-7209
wlals991215@kaist.ac.kr

Jaemyung Ahn
Korea Advanced Institute of Science and Technology (KAIST)
Daejeon 34141, Republic of Korea; 82-10-4114-2718
jaemyung.ahn@kaist.ac.kr

ABSTRACT

Early detection of wildfires is crucial for preventing the spread of fires and protecting lives and properties. In recent years, satellite-based wildfire detection is becoming popular because of the coverage area and cost limitations of traditional detection methods such as ground-based observation and aerial surveillance. In particular, using CubeSat has the advantage of real-time monitoring and early detection of wildfires in a large area at a low cost. However, the CubeSats have limited image quality due to physical limitations such as size, weight, and power, which reduce detection performance. Therefore, this paper proposes a novel approach for early wildfire detection with CubeSat images using deep learning and super-resolution techniques. Considering the limitations of CubeSat, a dataset of three-channel RGB images was used for binary classification. Landsat-8 images of ten bands were preprocessed into RGB images and enhanced by 4x using Real-ESRGAN. The study utilized transfer learning for wildfire detection using two pre-trained deep learning models, MobileNetV2 and ResNet152V2. The results proved that the super-resolution of the satellite images improved the wildfire detection precision, recall, and f1-score by about 3~5%, depending on the models.

INTRODUCTION

Wildfires can cause catastrophic damage in many areas of life, including transportation, communications, power and gas services, and water supplies. It also causes environmental damage and human casualties. Therefore, early detection and suppression of wildfires are essential because the damage of wildfires increases exponentially with time after they occur. Past approaches for wildfire detection can be divided into three categories which are terrestrial-based, aerial-based, and satellite-based. The previous two approaches have been more widely used than the last method regarding initial cost and technical difficulty. On the other hand, satellite-based wildfire detection research is being actively conducted as the amount of satellite launches has recently increased, and costs have been significantly reduced.

Satellites offer several advantages for wildfire detection. They can cover remote and inaccessible areas where ground-based detection systems are unavailable and continuously monitor fires at night or in bad weather. Moreover, satellite-based systems can be cost-effective in the long run compared to other methods due to reduced maintenance and replacement needs. In particular, wildfire detection using CubeSats maximizes the advantages of satellite-based systems. CubeSat can expand the real-time detection coverage area worldwide at a low cost. Depending on the orbit selection, looking at various sites or continuously observing one place is possible. Additionally, it is possible to address the restrictions of the link budget by employing multiple CubeSats.

However, due to their small size, there are a few obstacles to wildfire detection using CubeSats. Existing satellite-based wildfire detection algorithms are based on comparing multiple bands of satellite images, which are not feasible for CubeSat-based wildfire detection due to the small camera payload capacity. Windfire detection using only limited bandwidth of satellite images can resolve this issue. CubeSat also has a limited payload capacity regarding software memory, making it impossible to use heavy deep-learning models. Therefore, research is needed to maximize the detection performance while satisfying the limited number of bands and memory. Considering the limited payload capacity of CubeSats, this paper proposes an alternative method to improve the CubeSat-based wildfire detection performance using single-image super-resolution.

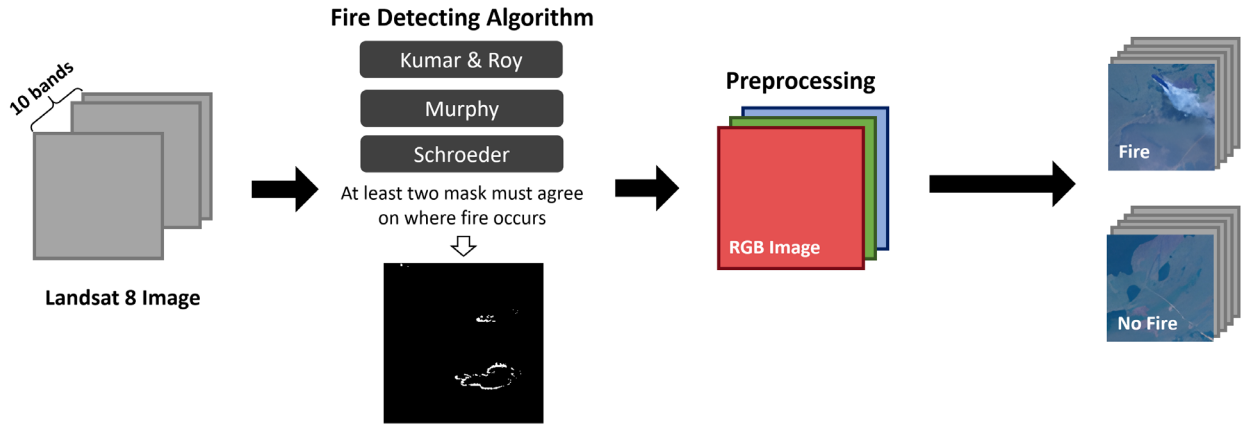

Figure 1: Data Preprocessing for Binary Classification

The super-resolution technique in the ground station can enhance the wildfire detection performance without changing the specification of CubeSat.

This study presents the possibility of worldwide real-time wildfire detection using a CubeSat-based system with super-resolution. This approach provides a solution for maximizing wildfire detection performance while satisfying the limitations of CubeSats, such as their small size, limited bandwidth, and payload capacity. The proposed method applied the super-resolution technique at the ground station to improve wildfire detection performance in RGB images received from CubeSats. The study proved that super-resolution images were superior to the original CubeSat images in both learning speed and performance by applying the same deep-learning models.

MATERIALS

Original Dataset

The data used for learning were preprocessed from wildfire satellite images released as open sources in the previous study [1]. The dataset was created by collecting Landsat-8 imagery from the United States Geological Survey (USGS). Since the Landsat-8 images are multispectral, including 11 bands, they were encoded in TIFF format. The image of the original dataset is composed of ten bands, excluding band 8 (panchromatic). The wavelength and resolution of each band in Landsat-8 images are shown in Table 1.

Table 1: Wavelengths and Resolutions of Bans in Landsat-8 Imagery

Bands	Wavelength [μm]	Resolution [m]
Coastal aerosol	0.43-0.45	30
Blue	0.45-0.51	30
Green	0.53-0.59	30
Red	0.64-0.67	30
Near Infrared	0.85-0.88	30
SWIR 1	1.57-1.65	30
SWIR 2	2.11-2.29	30
Panchromatic	0.50-0.68	15
Cirrus	1.36-1.38	30
Thermal Infrared 1	10.60-11.19	100
Thermal Infrared 2	11.50-12.51	100

The previous study trained a deep learning model on images of size 256x256. Considering the resolution of Landsat-8, which is 30m, each image covers an area of approximately 59 km^2. From a wildfire detection perspective, this is a relatively large area, so in the previous study, a segmentation technique was used to create fire masks based on whether each pixel was on fire. However, much effort is required to create manually annotated fire masks because manual annotation is possible only when there is a large amount of data on the time and point of the actual fire, and satellite images are complicated to process. Therefore, they used three well-established fire detection algorithms of satellite images to create a massive amount of fire masks presented by Schroeder et al. [2], Murphy et al. [3], and Kumar and Roy [4]. Since the three algorithms are not ground truth, they sometimes produce slightly different results. In this paper, the model was trained based on the assumption that a pixel within the mask should be classified as an active fire if there is an agreement between at least two sets of algorithms indicating it is a fire pixel.

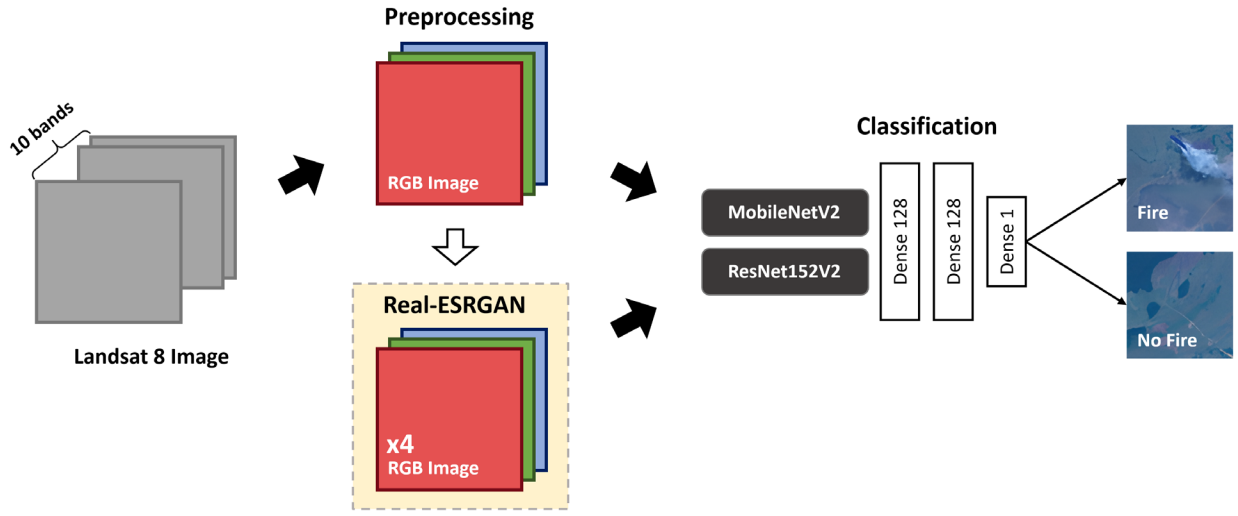

Figure 2: Learning Architecture of Early Wildfire Detection Using Super-Resolution

Preprocessing

The image size was reduced to 64x64 to take advantage of the CubeSat-based system, which can operate multiple units simultaneously with a small capacity. As the image size decreased, the area occupied by one image shrank to 3.7 km^2. Hence, the problem changed to a binary classification of whether the image has fire rather than pixel-by-pixel segmentation. In addition, since camera payloads containing ten bands could not be loaded due to CubeSat's capacity limitations, the original dataset images were converted to images containing only RGB channels. (Band 4, 3, 2 in Landsat-8).

The dataset preprocessing process for the learning is shown in Figure.1. First, ten bands of multispectral images format in TIFF is converted to three-channel RGB images in PNG format. The Python library GDAL was used for this process. When the image format is changed from TIFF to PNG, each pixel value is converted from a float to an integer, losing some information. However, since learning through the model uses normalized information, the effect on the learning result is insignificant. Next, Landsat-8 images of size 256x256 are divided into 16 new images of 64x64. The divided images are newly labeled according to whether or not they contain wildfire areas based on the fire masks of the original dataset. For example, if the fire Mask in Figure 1 is divided into 16 parts, there are areas where an image with white parts (fire) and an image with only black parts (no fire) are generated. This study assumes that the existing fire masks are as the ground truth. A new binary classification dataset was created based on the ground truth.

The original dataset was intended for pixel-level segmentation. Therefore, even if only one pixel corresponded to fire, it was included in the dataset, resulting in a severe class imbalance of about 10:1. When the dataset is imbalanced, the model may be biased toward the majority class and ignore the minority class as it attempts to optimize overall accuracy. Since it is essential to accurately detect fire in wildfire detection, dataset imbalance causes model performance degradation. Among several methods to solve dataset imbalance, undersampling was used in this study. The images for undersampling were randomly selected to avoid losing important information. The final dataset for binary classification deep-learning contains 5,966 images with equal proportions of fire and no fire images, and the total dataset split is presented in Table 2. The identical images were used for training and testing regardless of whether super-resolution was used.

Table 2. Size of Dataset for Learning

Train	Validation	Test
3818	954	1194

METHODS

Early Wildfire Detection Framework

The proposed approach involves utilizing two deep learning models for wildfire detection - one deployed on an onboard satellite and the other deployed on the ground. This approach is adopted to address the inefficiency of using many CubeSats for real-time wide-range surveillance, resulting in a large volume of captured images. Thus, a lightweight model has to be utilized for wildfire detection onboard the CubeSat.

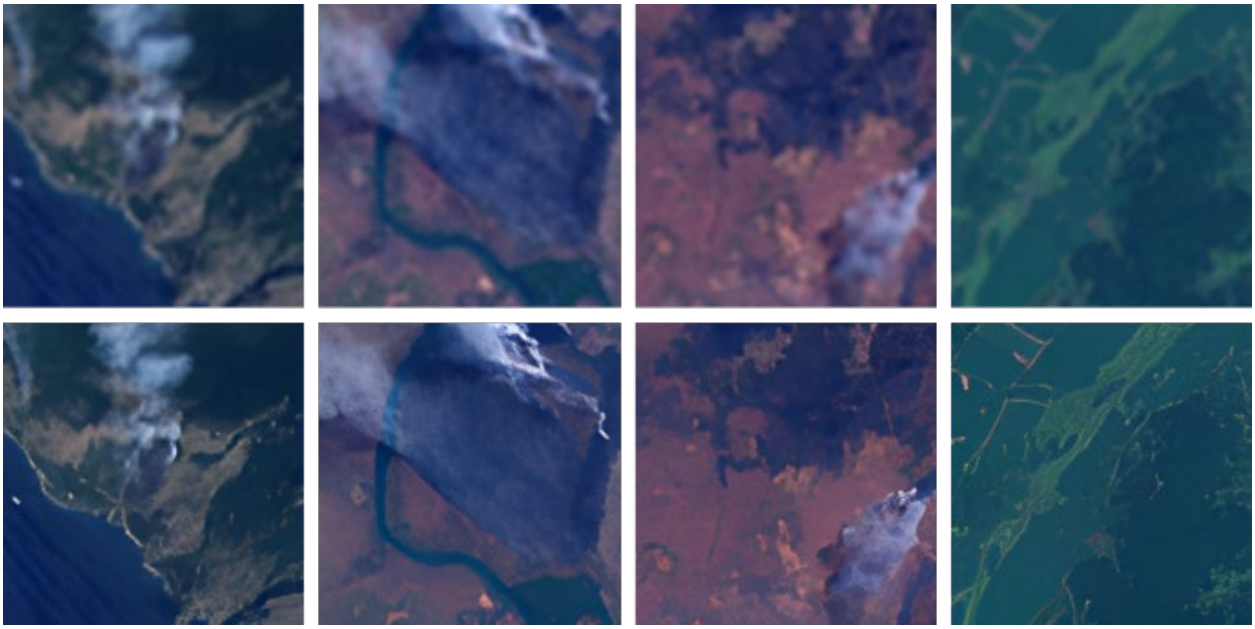

Figure 3. Examples of Training Dataset Super-Resolved by Real-ESRGAN: The upper row is the original images (spatial resolution 30m), and the lower row is the images after x4 super-resolution (spatial resolution 7.5m).

However, the limited payload capacity of the CubeSat poses a challenge for performing wildfire detection onboard. Several studies have consistently shown that the detection performance of a deep learning model increases as the number of parameters increases. Therefore, it is beneficial to perform this task on the ground by leveraging adequate computing power.

However, using CubeSats for image acquisition presents a challenge due to the fixed resolution of the images. We cannot simply apply more complex models to improve performance because of the fixed resolution. This study proposes a super-resolution technology to address this challenge, which can enhance image quality and achieve more accurate wildfire detection. The super-resolution techniques can identify image features that are not readily discernible at lower resolutions. The study investigates the impact of super-resolution on wildfire detection performance compared to using the same image and training model without super-resolution. The overall flow of wildfire detection using super-resolution is illustrated in Figure 2.

Super-Resolution of Satellite Images

Deep learning-based super-resolution techniques have become increasingly popular due to their capacity to learn complex image features and generate high-quality results. However, applying super-resolution to satellite imagery poses several challenges. Since super-resolution research has developed in the field of computer vision, the majority of the research has focused on processing RGB images. On the other hand, satellite imagery typically comprises multiple spectrums, including infrared bands which are frequently used for wildfire detection. As a result, most pre-trained models for super-resolution are unsuitable for satellite imagery. In addition, there is a requirement for wildfire image datasets that are appropriate for super-resolution training that considers atmospheric conditions, clouds, and cloud shadows. For this reason, despite the advantages of super-resolution technology, super-resolution is not being used for CubeSat-based wildfire detection. However, the existing super-resolution algorithm was easily applied in this study because the satellite image was preprocessed and converted into an image with three RGB channels.

Furthermore, there are several reasons why the decision was made to utilize single-image super-resolution instead of multi-image super-resolution. Firstly, given the limited bandwidth allocated to satellite communications, transmitting multiple photos of a specific area to the ground proved challenging, as satellites are constantly in orbit. The amount of data that could be transmitted is restricted, making capturing images at multiple time intervals difficult. Second, creating a dataset of satellite images of wildfires at consecutive time intervals was virtually impossible. It is the same reason for using the algorithms rather than

creating manually annotated masks explained earlier. Therefore, single-image super-resolution was adopted as a super-resolution technique.

The pre-trained "Real-ESRGAN" model [5] was used to super-resolve the satellite image dataset. Real-ESRGAN is a deep learning-based super-resolution model using a generative adversarial network (GAN) architecture to generate realistic images of high quality. The model is trained on real-world images and incorporates perceptual loss to enhance the super-resolution performance. The application of super-resolution of the model increases the image resolution by a factor of 4, resulting in an image size of 256x256. As the image size increases, the resolution has improved up to 7.5m, while one pixel has the same spatial resolution of 30m. The generated satellite images using the Real-ESRGAN model were of significantly higher quality, as visually evident in Figure 3, with a few examples of super-resolution images shown.

Wildfire Detection Learning Models

Two models (MobileNetV2 [6] and ResNet152V2 [7]) were used as deep-learning models for wildfire detection. The two models explored the impact of the model size on wildfire detection capability. Both models were the pre-trained model provided by the Keras application. The fully connected layers were added to change both models' roles from multi-class classification to binary classification. The model size and the number of parameters are summarized in Table 3.

Table 3: Model Size and Number of Parameters of MobileNetV2 and ResNet152V2

	MobileNetV2	ResNet152V2
No. of Parameters	2,438,593	58,610,561
Model Size (MB)	9.75	234.44

Two separate iterations of training and testing were conducted on each model. One of these rounds utilized the original dataset, while the other employed the super-resolved dataset. Both datasets were trained to leverage an equivalent model architecture, with only the input image size as the distinguishing factor.

Both models were trained for up to 100 epochs, and the model that achieved the minimum validation loss was selected as the best model and applied to the test dataset. The input images were preprocessed by normalizing and fed into a model as three-channel inputs. In addition, in this study, data augmentation was performed by allowing horizontal and vertical flips of the input image. However, care was taken to avoid enlargement techniques that could result in loss of information from the original image, such as cropping or rotating the original image. Multiple tests for hyperparameter selection were conducted using varying values of the learning rate, batch size, and other relevant parameters. Table 4 exhibits the resulting hyperparameters used in the training process.

Table 4: Hyperparameters for Learning

Hyperparameter	Value
Optimizer	Adam
Learning rate	1e-5
Batch size	128
Epochs	100
Loss function	Binary Cross-entropy

RESULTS

Evaluation Metrics

The models' performances are evaluated by computing precision, recall, and f1-score. Precision (P) evaluates the accuracy of the model's positive predictions by dividing true positives (TP) by the sum of TP and false positives (FP). Recall (R), also known as sensitivity, measures the model's capacity to identify positive instances by dividing TP by the sum of TP and false negatives (FN). Finally, the F1-score (F), which balances precision and recall, is calculated as the harmonic mean of these two measures.

$$P = \frac{TP}{TP + FP} \tag{1}$$

$$R = \frac{TP}{TP + FN} \tag{2}$$

$$F = \frac{2}{1/P + 1/R} \tag{3}$$

In addition, graphical representations of the models' binary classification performances were generated by plotting a ROC curve. The ROC curve plots the true positive rate or sensitivity against the false positive rate or 1-specificity at different threshold settings as the discrimination threshold changes. AUC is a measure that quantifies the overall performance of a binary classifier by computing the area under the ROC curve, which ranges from 0 to 1. For example, a perfect classifier is represented by an AUC of 1, while an AUC of 0.5 denotes a random classifier.

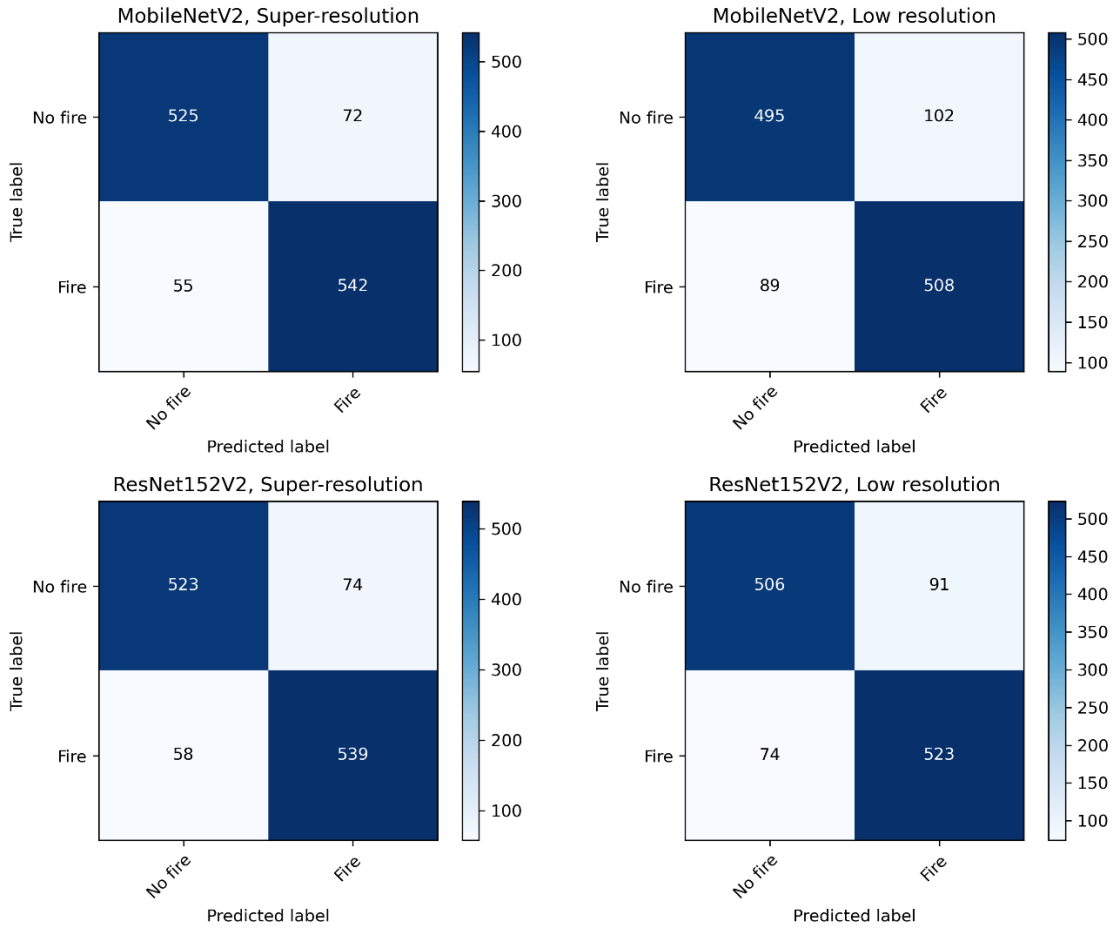

Figure 4: Confusion Matrices of Trained Models for Test Dataset

Performance of Trained Models

The four trained models were evaluated against 1,194 test datasets, and the results are summarized in Table 5. The models using super-resolution images demonstrated superior performance compared to those not using super-resolution. In the case of MobileNetV2, using super-resolution images resulted in a 5% increase in precision, recall, and f1-score, while ResNet152V2 showed a 3% increase. The learning performance of the models trained without super-resolution images improved as the model's size increased. However, in the case of the model using the super-resolution images, the performance difference between ResNet152V2 and MobileNetV2 observed during the training process did not occur during the test process. This suggests that super-resolution images contain more explicit information, allowing the model to learn all the features in the training dataset regardless of the model size. It implies that model performance can be improved by increasing the size of the training dataset.

Table 5: Comparison of Precision, Recall, and F1-Score of Trained Models

Models	P	R	F1
Super-resolution, MobileNetV2 (HMob)	0.88	0.91	0.90
Super-resolution, ResNet152V2 (HRes)	0.88	0.90	0.89
Low resolution, MobileNetV2 (LMob)	0.83	0.85	0.84
Low resolution, ResNet152V2 (LRes)	0.85	0.88	0.86

The confusion matrices in Figure 4 display the classification results for the test data. As it is a binary classification problem, the confusion matrix has a size of 2x2, with fire representing negative and no fire representing positive. Consistent with the evaluation metrics results, the true positive and true negative ratio was higher in the model using a super-resolution image. A significant difference occurred due to model size only in the model without using super-resolution. Since all four confusion matrices show higher values for TP and TN, it can be inferred that the models have learned the classification task well. However, there is still room for improvement as the test accuracies of the models are

lower than the training accuracies. Notably, all four models exhibited more FP than FN. This can be solved by changing the decision threshold value of the model or increasing the size of the data used for learning. A higher threshold would make it more difficult for the model to predict a positive outcome, potentially reducing the number of FP. Increasing the size and diversity of the dataset used to train the model can help reduce FP.

The ROC curves in Figure 5 clearly show the performance differences between models trained with and without super-resolution images. Despite using the same model architecture, super-resolution images increased the AUC score by 0.04. Unlike the previous results, the AUC scores were high when the model size was large in both cases, regardless of the image resolution. This seems to contradict previous results, but it is not. A model can have a high AUC score but low P, R, and F when the threshold used to make predictions may not be optimal for maximizing precision or recall, even if the model can distinguish between positive and negative cases effectively. ResNet152V2 can outperform MobileNetV2 in other evaluation metrics if the Decision Threshold value is adjusted accordingly.

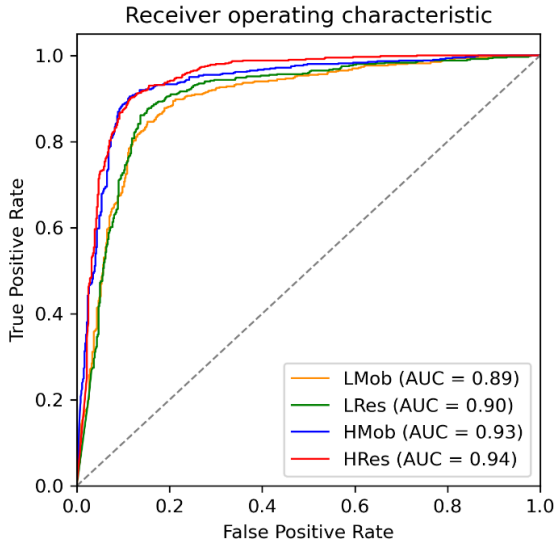

Figure 5: ROC-Curves of Trained Models

CONCLUSION

In conclusion, this paper presents a novel approach for early wildfire detection with CubeSat images, which utilizes deep learning and super-resolution techniques. This approach addresses the challenges posed by the limitations of CubeSats, such as their small size, limited bandwidth, and payload capacity while maximizing wildfire detection performance. The proposed method has the potential to enable worldwide real-time wildfire detection using a CubeSat-based system, which would provide a more cost-effective and efficient alternative to traditional ground-based observation and aerial surveillance methods. The study employed transfer learning for wildfire detection, utilizing two pre-trained deep-learning models, MobileNetV2 and ResNet152V2. Preprocessing and converting the Landsat-8 imagery into an image with three RGB channels enabled the easy application of the existing super-resolution algorithm. The results demonstrated that training with super-resolution images was superior to using original CubeSat images in both learning efficiency and performance, with an improvement in precision, recall, and f1 score by about 3~5%, depending on the model used.

We can explore several future research directions to improve the proposed method. First, increasing the dataset size can improve the test accuracy of the model. The current study utilized a dataset of 1,194 test images. More images of different types of fires, weather conditions, and terrain can be added to the dataset for better performance. Second, the performance of the proposed method is investigated when the super-resolution factor is different. In the current study, Landsat-8 images were enhanced 4x using the Real-ESRGAN algorithm. Changing super-resolution coefficients can affect wildfire detection accuracy. Third, exploring trends in decision thresholds for wildfires can be further studied. Although this study has shown that super-resolution techniques increase the AUC score, the fire decision threshold has yet to be optimized. Performance can be improved by analyzing trends in decision threshold models. Finally, reducing the model size and false negative rate are also future research topics. The current study assumes a situation where the CubeSat makes its first detection onboard. Given the model capacity, reducing the model size and the false negative rate can improve the system's overall performance.

Acknowledgments

This work was supported by Korea Research Institute for Defense Technology Planning and Advancement (KRIT) grant funded by the Korean government (DAPA (Defense Acquisition Program Administration)) (KRIT-CT-22-040, Heterogeneous Satellite constellation based ISR Research Center, 2022)

References

1. G. H. de Almeida Pereira, A. M. Fusioka, B. T. Nassu, and R. Minetto, "Active fire detection in Landsat-8 imagery: A large-scale dataset and a deep-learning study," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 178, pp. 171–186, Aug. 2021.

2. W. Schroeder, P. Oliva, L. Giglio, B. Quayle, E. Lorenz, and F. Morelli, “Active fire detection using Landsat-8/OLI data,” *Remote Sens Environ*, vol. 185, pp. 210–220, Nov. 2016.

3. S. W. Murphy, C. R. de Souza Filho, R. Wright, G. Sabatino, and R. Correa Pabon, “HOTMAP: Global hot target detection at moderate spatial resolution,” *Remote Sens Environ*, vol. 177, pp. 78–88, 2016.

4. S. S. Kumar and D. P. Roy, “Global operational land imager Landsat-8 reflectance-based active fire detection algorithm,” *Int J Digit Earth*, vol. 11, no. 2, pp. 154–178, Feb. 2018.

5. X. Wang, L. Xie, C. Dong, and Y. Shan, “Real-esrgan: Training real-world blind super-resolution with pure synthetic data,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 1905–1914.

6. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

7. K. He, X. Zhang, S. Ren, and J. Sun, “Identity mappings in deep residual networks,” in Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14, 2016, pp. 630–645.