DEEP LEARNING-BASED HIGH-RESOLUTION FOREST-TYPE CLASSIFICATION USING TRANSFER LEARNING

Seoyeon Cho[1] and Yongil Kim*[2]
[1]Graduate Student, Dept. of Civil & Environmental Engineering, Seoul National University
35-318, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, South Korea
Email: chwh1229@snu.ac.kr

[2]Professor, Dept. of Civil & Environmental Engineering, Seoul National University,
35-410, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, South Korea
Email: yik@snu.ac.kr

KEY WORDS: Deep learning, Transfer learning, Classification, Forest-type map, High-resolution satellite image

ABSTRACT: To enable advanced monitoring in forest areas and maximize the utilization of forest data, the creation of a high-resolution forest-type map becomes imperative. A high-resolution forest map plays a critical role in the early identification of fire-prone zones and contributes significantly to post-fire recovery efforts. Furthermore, it is anticipated that the high-resolution forest-type map will effectively enable the assessment of the forest's carbon sequestration capabilities. As awareness of the significance of forest resources continues to grow, high-resolution forest-type maps draw attention to important spatial information. Especially for countries like South Korea, where approximately 70% of the land is covered by forests, comprehensive information is essential for effective forest management. However, existing methods of collecting information on forests are labor-intensive, time-consuming, and costly to produce, and even that is heavily reliant on expert visual classification. Hence, this study aimed to investigate the feasibility of generating a high-resolution forest-type map, which shows great potential for diverse future applications focused on deep learning and transfer learning techniques. Using the existing training dataset as a foundation, pre-trained models were created for UNet++, DeepLabV3+, and FPN, and their performance and characteristics were compared. Subsequently, these models were applied to real forest environments. As a result, an average improvement of approximately 0.07 in F1-score performance was observed when both transfer learning and CBAM(Channel Attention Module) were applied. This demonstrates the potential for more efficient forest-type classification within the context of forest area analysis by assigning weights to channel characteristics. Additionally, each architecture exhibited distinct response patterns. Therefore, it can be inferred that the choice of architecture should align with specific performance requirements and objectives. Furthermore, the models were applied not only to datasets with the same spatial resolution as the training data but also to datasets with higher spatial resolution for comparison. Although the trends were not consistent, superior performance in terms of F1-score suggests robust utility for transfer learning in various contexts.

1. INTRODUCTION

The expansion of the global satellite market is giving rise to a growing emphasis on the practical application of satellite imagery across various domains. With increased private investment, the remote sensing industry is diversifying, extending beyond government-operated satellites to include commercial ones. Consequently, satellite imagery is finding extensive utility not only in national urban planning, surveillance, and military applications but also in various commercial use cases. These applications range from estimating the number of automobiles and calculating revenue to monitoring petroleum reserves and much more. This trend is further accelerated by the integration of deep learning, leveraging the big data characteristics of satellite imagery, which holds great promise for the future. In contrast, the application of satellite technology combined with deep learning in the field of forestry remains relatively underdeveloped. While urban applications involve object detection and other sophisticated research areas, forestry applications predominantly focus on fundamental tasks such as change detection, wildfire detection, or identifying areas affected by landslides. The limited availability of data, both in terms of various datasets suitable for deep learning and the scarcity of satellite imagery over forested regions, hinders progress. This scarcity not only restricts the supply of data but also impedes its utilization even when needed, thus hindering advancements in the field. However, it is worth noting that in the case of South Korea, forests cover approximately 70% of the land area, making them a crucial resource and subject of monitoring on a global scale. Presently, the production of forest maps remains a time-consuming and labor-intensive process, primarily reliant on expert visual interpretation. Aerial imagery acquisition methods are costly, with budgets running into billions of won and since 2022, there has been a shift towards relying on modification only for areas with observed changes.

When comparing the forest composition between the 5th and 4th forest maps, several changes were observed. Coniferous forests decreased from 42.2% to 37.8%, while deciduous forests (excluding evergreen broadleaf forests) increased from 25.2% to 46.9%. Mixed forests also experienced a decrease, dropping from 29.5% to 11.6% (National Geographic Information Institute, 2021). The increase in the proportion of deciduous forests can be attributed to a more detailed classification of areas previously categorized as mixed forests in the previous forest map at a 1:5,000 scale. Additionally, it is interpreted that the decline of coniferous forests is due to the impact of climate change. This phenomenon emphasizes the changeability of forest composition. Furthermore, the existing classification method relies on the dominant tree type in each region, therefore, it's crucial to recognize that the forest map itself may not be a perfectly accurate reference. As the importance of monitoring forested areas continues to grow, there is a heightened demand for high-resolution satellite imagery. Such imagery not only allows for the creation of detailed forest maps but also facilitates the monitoring of changes in valuable forest resources, aids in wildfire prevention and initial response, supports recovery efforts, and enables effective management of tree diseases. However, given the vast expanse of forested regions and the inherent challenges in generating comprehensive and accurate forest maps based solely on ground surveys, leveraging existing datasets becomes imperative.

This study therefore, aimed to examine the feasibility of generating a comprehensive forest map utilizing high-resolution satellite imagery. The experiment consists of two steps. Initially, a pre-trained model was developed using the existing dataset, followed by its application to the four national parks in South Korea. Subsequently, the most appropriate deep learning and transfer learning techniques for forest-type classification will be identified. This phase involves identifying unique aspects related to forest categorization within the forested areas and the detection of any potential enhancements or challenges that demand attention. Lastly, a consideration of the prerequisites for the future versatile utilization of South Korea's forest satellite data was performed.

2. MATERIALS

2.1 Training Dataset : AI-Hub

As a pre-training dataset, a dataset from AI-Hub which was provided by the NIA (National Information Society Agency) was selected. AI Hub is an integrated AI platform designed to facilitate AI technology development, product creation, and service deployment by providing essential AI infrastructure. This infrastructure includes AI data, AI software APIs, and computing resources, making it accessible to a wide range of users and participants. The AI-Hub’s forest tree-type image dataset is South Korea's most extensive collection of forest tree species images, established in 2021 and made available for distribution from July 2022. This dataset is constructed using high-resolution aerial imagery obtained from two reputable sources: the National Geographic Information Institute and the Korea Aerial Photography. While primarily built upon 0.25-meter aerial images, the dataset also includes simulated 5-meter satellite images(Rapideye) to complement the aerial data.

For training purposes, the dataset offers both fine and coarse annotations, which involve segmenting and labelling the data for four forest tree species categories: Coniferous, Decicuous, Other, and Unidentifiable. Each category has its dedicated dataset. The image size for training is standardized at 512×512 pixels. The dataset comprises a total of 1,010 images, distributed across four regions in South Korea: Gyeongsang (570 images), Jeolla (380 images), and Jeju (60 images). The regions for dataset selection were chosen based on a criterion that considered the proportion of each tree species within the target area. Specifically, regions with a mixed forest ratio of 7% or less among the forest area depicted on 1:5000 topographic maps were chosen for inclusion in the dataset. As a result, the dataset accurately represents the forest composition in the selected regions, with Coniferous and Deciduous trees accounting for nearly equal proportions at 41.71% and 40.58%, respectively, while Mixed forests constitute 12.29%.

The dataset is composed of images and masks. Figure 1 shows the example data of fine annotated 5m resolution image data. Those annotation labels are made up of 4 categories, coniferous, deciduous, others, and illegible. During preliminary experimentation, both Coarse annotation data and Fine annotation data were considered for training. However, due to a significant drop in accuracy observed when incorporating Coarse data, the decision was made to proceed with Fine annotation data exclusively for training purposes.

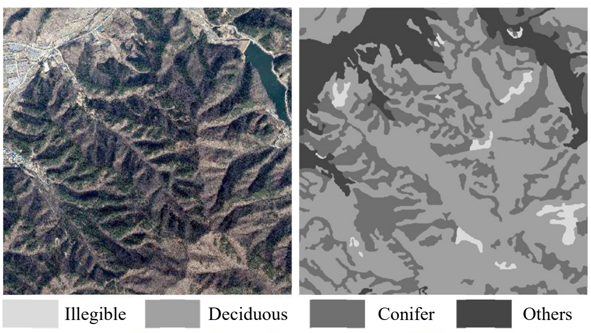
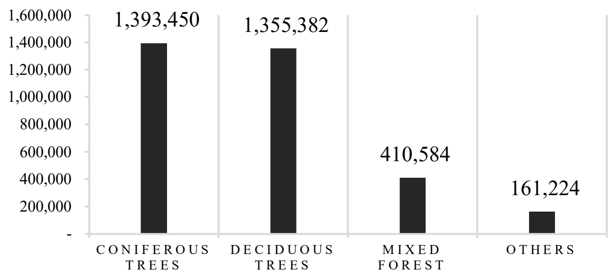


Figure 1. Annotation (Left : Image, Right : Mask)

Table 1. Forest distribution of dataset areas (ha)
(Gyeongsang, Jeolla, Jeju)



2.2 Test Data : PlanetScope

Due to the extensive coverage of forested regions, obtaining a high-resolution satellite dataset that encompasses the entire area can be quite challenging. Moreover, these regions often contend with cloud cover and exhibit varying vegetation characteristics across different seasons, making it complicated to acquire consistent, high-quality datasets. With the research objective of comprehending and identifying an appropriate learning process, our focus shifted towards evaluating the effectiveness of transfer learning. As a result, we made the decision to work with a limited amount of data.

For our data source, we selected PlanetScope satellite imagery due to its high resolution, frequent revisits, and exceptional quality(Acharki, S., 2022). Additionally, prior research suggested that spectral separability in May was superior for vegetation classification in satellite imagery(Lim et al., 2019). After analyzing the similarities between various months in the same area, we concluded that early May imagery would best align with our research objectives. However, despite the generally high quality of PlanetScope imagery, we observed significant spectral variations between adjacent areas captured on the same day. To ensure consistency, we visually compared cloud-free images that had similar characteristics with the training data and finally chose the date of acquisition. After the selection of the date, we excluded the regions that were already included in the AI-Hub training dataset to prevent areas that have already been learned from being used for testing. The process referenced the AI-Hub training data guidelines and tried to minimize the inclusion of mixed forests.

Ultimately, four regions were selected, as outlined in Table 3, consisting of images captured within approximately 40 seconds by the same sensor. The dataset consisted of a total of 100 patches, each measuring 512×512 pixels, along with the corresponding forest cover masks derived from the region's forest map.

Table 2. PlanetScope Sensor Specifications (PLANET.COM, 2018)

Product Attribute	Description
Sensor Type	Four-band frame Imager or four-band frame Imager with a split-frame VIS+NIR filter (DOVE-C)
Analytic Bands	4-band multispectral image (blue, green, red, near-infrared)
GSD (nadir)	3.7 m - Delivered at 3.0m
Frame Size	24 km x 8 km (approximate)

Table 3. Information on acquisition datas (Date : 2019-05-02)

	1. Seolraksan National Park (01:52:30 UTC)	2. Chiaksan National Park (01:52:40 UTC)	3. Woraksan National Park (01:52:50 UTC)	3. Jirisan National Park (01:53:10 UTC)
Ground sample distance	3.9m	3.9m	3.9m	3.9m
Off-nadir angle	3.2°	3.2°	3.2°	3.1°
Sun elevation	59.4°	59.7°	60.1°	60.8°
Sun azimuth	131.3°	130.5°	129.1°	126.7°

2.3 Reference Data

The Korea Forest Service's forest type map, representing the distribution and composition of forests in Korea, was chosen for comparison with the satellite image classification results. We selected the most recent 1:25,000 scale map, which is a product of the Fifth National Forest Resources Investigation (2006~2010). This map provides valuable attributes such as tree species and crown closure information. The criteria for classifying each forest type were based on factors like crown area and stand proportion, such as more than 75% coniferous, deciduous, larch, etc.

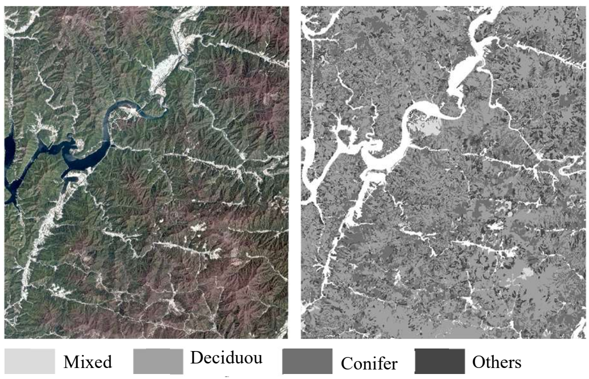


Figure 2. Example dataset of Seolraksan National (Left : Image, Right : Mask)

3. METHODOLOGY

3.1 Overall Process

This study aimed to perform forest-type classification, beginning with the assessment of the performance of three deep learning models (FPN, DeepLabV3+, UNet++). Subsequently, to incorporate vegetation characteristics, the CABM (Channel Attention Module) was applied, and its effects were examined. Finally, the pre-trained model trained in this manner was applied to real-world cases for transfer learning. The application of the case was tested not only on data of the same resolution as the one used for pretraining (5m) but also on finer resolution data (3m) for comparison.

During this process, we observed under which conditions the highest accuracy was achieved and investigated any instances of accuracy reduction, along with the reasons behind them. Ultimately, we contemplated what information would be necessary for future forest-type classification and considered directions for improving deep-learning architectures.

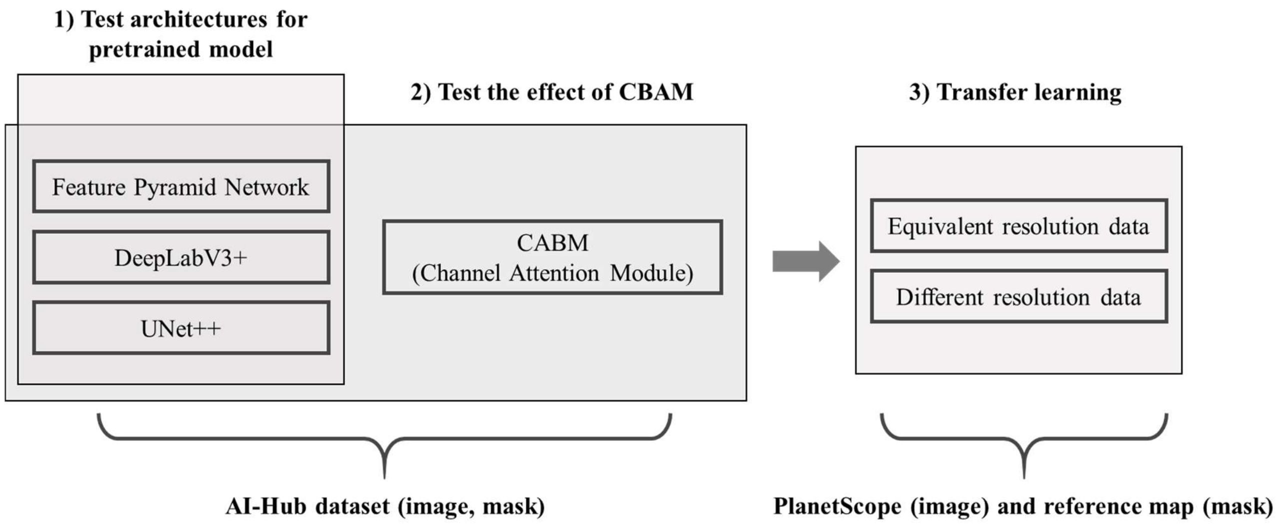


Figure 3. Flowchart

3.2 Deep Learning Details and Preprocessing

Several deep learning architectures were assessed, and their results were compared based on various characteristics. Ultimately, three representative segmentation models, renowned for their strong performance, were selected. The training process was executed using PyTorch, and data augmentation techniques such as horizontal flipping, rotation, shifting, and random cropping were applied. Methods related to color or brightness adjustments were accepted, considering the distinct nature of forest data. In the preprocessing stage, histogram equalization and matching were carried out. Pretraining experiments were conducted for 100 epochs, while for transfer learning, experiments were performed for 30 epochs.

3.2.1 Feature Pyramid Network (FPN): FPN is a convolutional neural network that is used for object detection and segmentation. It uses a top-down architecture with lateral connections to combine low-level and high-level features to improve the accuracy of object detection and segmentation(C. Wang & Zhong, 2021).

3.2.2 DeepLabV3+: DeepLabV3+ is a convolutional neural network that is used for semantic image segmentation. It uses atrous convolution to capture multi-scale context information and employs a decoder module to refine the segmentation results(Z. Wang et al., 2022).

3.2.3 UNet++: UNet++ is a convolutional neural network that is used for biomedical image segmentation. It uses a nested architecture with skip connections to combine low-level and high-level features to improve the accuracy of segmentation (Siddique et al., 2021).

3.3 Attention module

Traditional forest research has primarily relied on low to moderate-resolution spectral information rather than high-resolution satellite imagery. This preference for spectral information is due to its significance in vegetation classification tasks. However, conventional deep learning methods applied to high-resolution satellite data have been specialized in shape-based feature recognition. Consequently, in the construction of deep learning networks, it is deemed essential to maximize the utilization of spectral information. Therefore, we applied the CBAM (Channel Attention Module) module with the aim of harnessing spectral information to the fullest extent possible and subsequently evaluated its effectiveness for our research. The attention mechanism computes weights for different segments of the input data, aiding the model in concentrating on the crucial aspects of the input information.

3.4 Transfer Learning

The most effective way to create a high-performing deep learning model is to acquire a large volume of data. However, there are cases where data may be limited or acquiring it comes at a significant cost. Transfer learning is a method that address those problems, where the knowledge gained from training a neural network in a specific domain is utilized in the learning process of a neural network used in a similar or entirely new domain. Transfer Learning is highly effective when the training dataset is small, offering significantly higher accuracy and faster learning compared to training from scratch.

The neural networks that have been trained and are used in Transfer Learning are referred to as pretrained models, and some well-known examples include ImageNet, GoogleNet, and VGGNet. These pretrained models, which have been effectively trained on large datasets, are brought in, and their model weights are slightly adjusted to match the specific problem scenario the user is addressing (considering factors such as dataset size, relevance of the pretrained model's dataset to the problem scenario, etc.).

However, for forest-type classification, the aforementioned pretrained models could not be utilized due to domain differences. Therefore, in this study, we attempted to construct a pretrained model using the AI-Hub dataset and validate its applicability.

4. RESULTS AND DISCUSSION

Table 4. F1-Scores of test results on four of Korea's national parks (5m resolution)

Architectures	Pretraining		Transfer learning			
	DL dataset		5m resolution data (train : 80, test : 20)			
	w/o CBAM	w/ CBAM	w/o CBAM		w/ CBAM	
			w/o transfer	w/ transfer	w/o transfer	w/ transfer
Unet++	0.7416	0.751	0.707	0.6648	0.5899	0.657
DeepLabV3+	0.7235	0.7254	0.627	0.6409	0.6042	0.6707
FPN	0.732	0.7255	0.62	0.6766	0.603	0.6726

4.1 Performance Comparison of Deep Learning Architectures

In comparing three different architectures, UNet++, DeepLabV3+, and FPN, it was observed that UNet++ achieved the highest average F1-Score of 0.6852 among the comparison group. Following closely, FPN scored 0.6716, while DeepLabV3+ scored 0.6653. On the other hand, when examining the averages of experiments involving transfer learning, FPN achieved the highest score with 0.6746, followed by UNet++ with 0.6609, and DeepLabV3+ with 0.6558. Without transfer learning, UNet++ scored 0.64845, DeepLabV3+ scored 0.6156, and FPN scored 0.6115. This implies the presence of architectures suitable for utilizing transfer learning when striving to do so.

The figure below, labeled as Figure 4, illustrates an example of how each architecture classifies the same image when CBAM is uniformly applied. Firstly, for the (a) UNet++ model, the F1-Score is 0.7932, with a precision of 0.7866 and a recall of 0.7999. This model exhibits a relatively balanced trade-off between precision and recall. Secondly, for the (b) DeepLabV3+ model, the F1-Score is 0.7994, with a precision of 0.8371 and a recall of 0.7650. This model prioritizes precision, resulting in a relatively higher precision but lower recall. In other words, it focuses more on obtaining accurate positive predictions, leading to a higher ratio of true positives among the predicted positives. Lastly, for the FPN model, the F1-Score is 0.8382, with a precision of 0.8677 and a recall of 0.8107. This model has the highest precision among the models and a significantly high recall. It is confident in obtaining accurate results and has the highest ratio of true positives among the predicted positives.

These results indicate that the choice of model should depend on the performance requirements and objectives. The UNet++ model can be useful when a balanced consideration of accuracy and recall is required. The DeepLabV3+ model may be suitable when precise positive results are crucial, and the FPN model is appropriate when high precision and a substantial recall are needed.

(a)
(b)
(c)
(d)

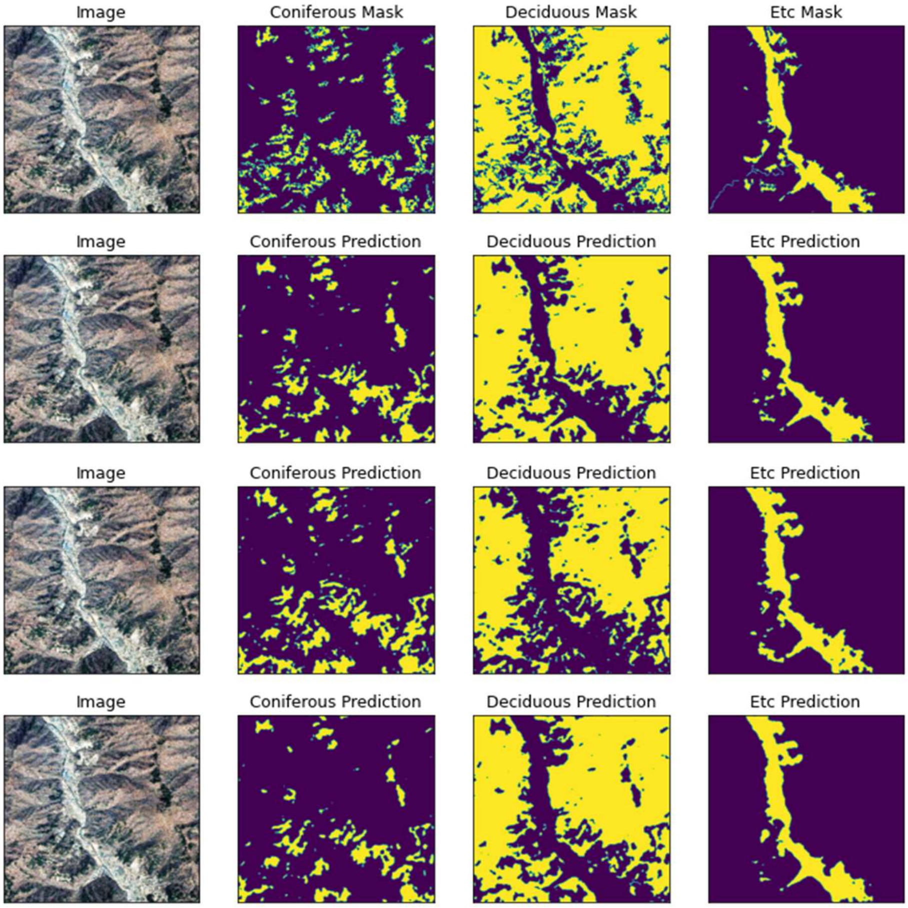


Figure 4. Effect of architecture with CBAM (a) Ground truth,
(b) Unet++ : F1-Score - 0.7932, Precision - 0.7866, Recall - 0.7999
(c) DeepLabV3+ : F1-Score – 0.7994, Precision - 0.8371, Recall - 0.7650
(d) FPN : F1-Score - 0.8382, Precision - 0.8677, Recall - 0.8107

4.2 Analysis on the Attention Module

The first application involved a test on 5m resolution data, which is identical to the resolution of the data used for pretraining models. An important observation is that, when transfer learning was conducted without using CBAM, there was little difference on average compared to when transfer learning was not performed. In the absence of transfer learning, it showed an average F1-score of 0.6513, while with transfer learning, it averaged 0.6608. Furthermore, when considering individual results, the highest score was achieved when training was conducted solely with PlanetScope images without the inclusion of CBAM and transfer learning. However, conversely, when CBAM was used alongside transfer learning, it showed an average improvement of 0.0677. Without transfer learning, it scored 0.5990, and with transfer learning, it demonstrated a substantial improvement to 0.6668 on average. These results indicate the potential for acquiring and utilizing species classification data effectively, even with data acquired through various sensors, by harnessing spectral information. This highlights the feasibility and potential advantages of adopting transfer learning with an attention module in this context. Specific improvements will be visually demonstrated in the subsequent application steps, as will be shown later.

4.3 Transfer Learning Results on Equivalent Resolution Data

4.4 The first application involved a test on 5m resolution data, which is identical to the resolution of the data used for pretraining models. Firstly, in Figure 5, when transfer learning is applied, it starts with a high F1-Score of over 0.6 and a loss below 0.55, demonstrating a very rapid learning speed. An important observation is that, when transfer learning was conducted without using CBAM, there was little difference on average compared to when transfer learning was not performed. In the absence of transfer learning, it showed an average F1-score of 0.6513, while with transfer learning, it averaged 0.6608. Furthermore, when considering individual results, the highest score was achieved when training was conducted solely with PlanetScope images without the inclusion of CBAM and transfer learning. However, conversely, when CBAM was used alongside transfer learning, it showed an average improvement of 0.0677. Without transfer learning, it scored 0.5990, and with transfer learning, it demonstrated a substantial improvement to 0.6668 on average. These results indicate the potential for acquiring and utilizing species classification data effectively, even with data acquired through various sensors, by harnessing spectral information. This highlights the feasibility and potential advantages of adopting transfer learning with an attention module in this context.

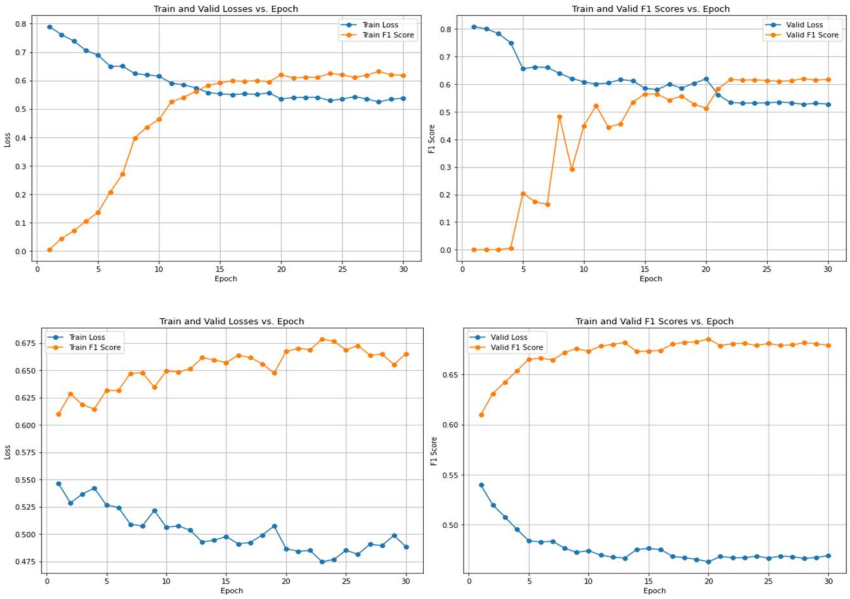


Figure 5. Train and Valid Losses & F1-Scores (Up : without transfer, Down : with transfer)

(a)

(b)

(c)

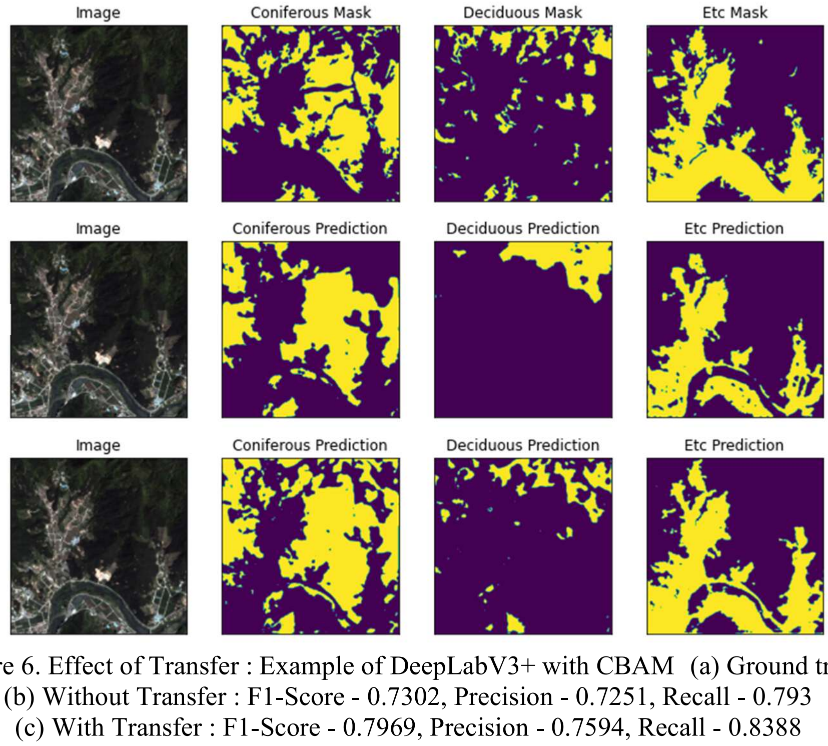


Figure 6. Effect of Transfer : Example of DeepLabV3+ with CBAM (a) Ground truth,
(b) Without Transfer : F1-Score - 0.7302, Precision - 0.7251, Recall - 0.793
(c) With Transfer : F1-Score - 0.7969, Precision - 0.7594, Recall - 0.8388

In the case of performing without transfer learning, the F1-Score is 0.7302, with a precision of 0.7251 and a recall of 0.793. This indicates that the model maintains a certain level of accuracy but falls short in precisely predicting among the actual positives. On the other hand, when transfer learning is applied, the F1-Score increases to 0.7969, with a precision of 0.7594 and a recall of 0.8388. Although the precision is relatively lower in the transfer learning scenario, it significantly improves the recall, meaning it captures more of the actual positives among the predicted positives. This can also be visually observed as the model exhibits a higher level of spatial detail in its predictions.

4.5 Transfer Learning Results on Different Resolution Data

Table 5. F1-Scores of test results on four of Korea's national parks (3m resolution)

Architectures	Application / Test							
	3m resolution data (train : 133, test : 33)				3m resolution data (train : 234, test : 58)			
	w/o CBAM		w/ CBAM		w/o CBAM		w/ CBAM	
	w/o transfer	w/ transfer	w/o transfer	w/ transfer	w/o transfer	w/ transfer	w/o transfer	w/ transfer
Unet++	0.728	0.6454	0.6615	0.6518	0.7453	0.6621	0.7343	0.6548
DeepLabV3+	0.675	0.6652	0.6814	0.6753	0.7079	0.6959	0.7129	0.7002
FPN	0.6037	0.6818	0.6729	0.68	0.6972	0.7126	0.6976	0.705

Lastly, the models pretrained on one resolution were applied to data of a different resolution. Since each pixel in 3m resolution images covers a different area compared to 5m resolution images, two different approaches were employed to compare and determine the amount of data used for training. The first approach aimed to keep the total area covered by pixels in the training dataset equal, selecting 133 images. The second approach involved using even more data to ensure that more information about the target area was incorporated. As a result, both approaches did not yield the same transfer learning and CBAM application effects as observed in the 5m resolution case. While there was no consistent pattern in the overall effects, a closer examination of each patch revealed some tendencies.

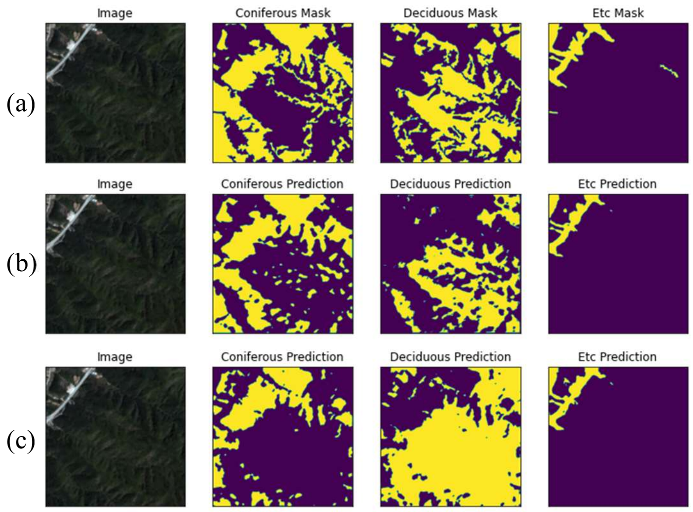


Figure 7. Effect of Transfer : Example of UNet++ with CBAM (a) Ground truth, (b) Without Transfer : F1-Score -0.7445, Precision - 0.8180, Recall - 0.6832 (c) With Transfer : F1-Score - 0.6433, Precision - 0.6147, Recall - 0.6747

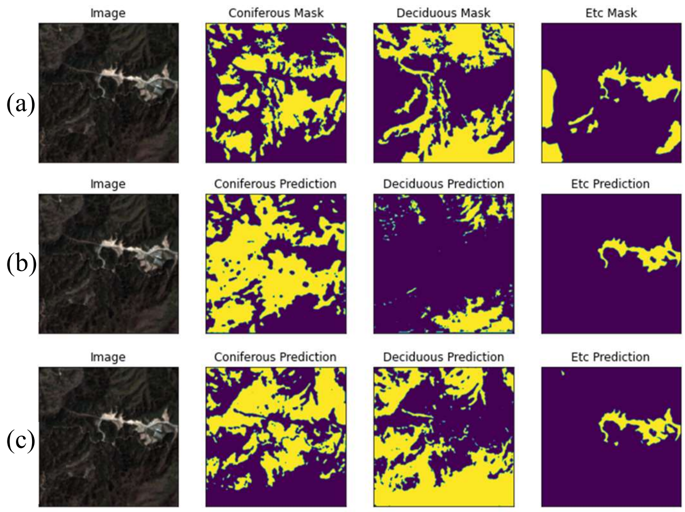


Figure 8. Effect of Transfer : Example of UNet++ with CBAM (a) Ground truth, (b) Without Transfer : F1-Score -0.5832, Precision - 0.6725, Recall - 0.5149 (c) With Transfer : F1-Score - 0.6353, Precision - 0.6172, Recall - 0.6546

Figure 7 and Figure 8 both explore the impact of transfer learning on the CBAM-equipped Unet++ model. Figure 7 reveals a tendency to over-detect deciduous trees, leading to lower F1-Score when transfer learning is applied (0.6433) compared to when it is not (0.7445). However, in scenarios where deciduous trees are prevalent and intricately interspersed with coniferous trees in narrow areas, the model shows improved performance. Figure 8 illustrates an instance where the model excels in distinguishing finely mixed areas between deciduous and coniferous trees while still over-detecting deciduous trees. Consequently, in this specific image, the F1-Score increases with transfer learning applied (0.6353) compared to when it is not (0.5832). These cases demonstrate that applying the model to different resolutions, such as 3m resolution, introduces various variables. Additionally, it is anticipated that the suitable approach may vary based on the distribution patterns and structures of forest in different regions.

4.6 Discussion

When reviewing existing research, the forest-type classification task achieved an accuracy of approximately 81% when using LiDAR data. Additionally, in a simulated evaluation of forest tree species classification based on RapidEye satellite data, using the Random Forest technique with only spectral(RGB) information yielded an overall accuracy of 39.41%, while incorporating both spectral and texture information resulted in an overall accuracy of 69.29%(Park et al., 2011; Kwon et al., 2021) . These findings illustrate that classifying forest types is a challenging task. It is essential to note that the objectives of these previous studies may not be entirely identical to the present research. However, the results of this study are based on the limited spectral information of RGB data. Therefore, it is expected that more detailed classification beyond these previous studies can be achieved in the future by incorporating near-infrared (NIR) and GLCM (Gray Level Co-occurrence Matrix) data, potentially leading to higher accuracy.

Furthermore, this study aimed to verify to what extent improvement could be achieved with different deep learning architectures and transfer learning. However, for practical application in the future, some enhancements are necessary. Firstly, it is crucial to determine how much of the area to be covered by the forest map should be obtained through field surveys and identify the minimum percentage of coverage required. As the area to be classified increases, the amount of data used for training must be adjusted proportionally.

There is also a need to examine attention modules other than CBAM, as well as conduct additional research on handling data with different resolutions. In this study, when comparing the analysis results of 5m resolution and 3m resolution images, it was observed that accuracy decreased when transfer learning was performed on 3m resolution images. However, the absolute performance of the F1-Score, on the other hand, was higher. This can be interpreted as transfer learning being applicable even to datasets with different resolutions. Therefore, it is worth investigating the causes and consequences for each case more thoroughly, and considering the feasibility of using pretrained models on data with varying resolutions by increasing the dataset size or augmenting data through zooming in and out. Lastly, the mask in the AI-Hub dataset contains "Illegible" and reference map ground truth contains "Mixed" masks. By using transfer learning, "Mixed" area could be classified into coniferous or deciduous trees. Additionally, in most cases of "Illegible", these are shadows, which can potentially be trained more specifically. During the experiment, when CBAM was applied to transfer learning, the ability to distinguish between coniferous and deciduous trees in shadowed areas was superior compared to when transfer learning was not used. In the absence of transfer learning, all shadowed areas were classified as "other," but when CBAM and transfer learning were used simultaneously, more information was extracted from these pixels. Utilizing this feature to finely categorize information, such as distinguishing between coniferous and deciduous areas with shadows, could lead to more accurate classification.

5. CONCLUSION

This research involves pretraining models followed by transfer learning to assess the feasibility of producing high-resolution forest maps. The process involves the utilization of various deep learning techniques, conducting comparative analyses, and the pursuit of identifying the most effective approach. Furthermore, exploring the applicability of existing datasets and validating the effects of pretraining provide insights into the potential utilization and enhancement of deep learning datasets in future forestry-related research endeavours.

As a result, this study was able to achieve a classification accuracy of up to approximately 66% using only RGB color information. It should be noted that the ground truth used in this study is in polygon format not pixel, and there are areas in the training dataset's masks marked as "Illegible," "Mixed" which, if improved, could potentially enhance the actual accuracy further. There are ongoing challenges regarding how ground truth should be established for future work and how satellite image training datasets can be used to build nationwide forest maps in conjunction with field surveys. Additionally, considerations should be made on how to use this data for forest monitoring in the future.

Moreover, aside from the CBAM module, there is a need to explore more active utilization of spectral information. Notably, high-performance pretrained models like ImageNet have demonstrated significant accuracy improvements in object detection tasks, including the classification of objects such as cars, buildings, and roads in remote sensing fields. However, the benefits of high-resolution imagery, with both excellent spatial resolution and spectral characteristics, have not been fully exploited. Future research should focus on making high-resolution satellite images not only spatially precise but also fully accessible in terms of spectral information.

In particular, during these processes, it is crucial to understand the characteristics of different architectures and the specific attributes of the target region to adapt accordingly. Previous experiments have led to the conclusion that the choice of architecture should be based on specific performance requirements and objectives. It is necessary to assess whether a balanced compromise between different aspects, such as accuracy and recall, is needed or if obtaining precise positive results is of utmost importance. Additionally, the selection should align with the forest patterns that are particularly well distinguished by a specific model and approach.

Lastly, the dataset used in this study is labeled as a remote sensing dataset, but it lacks information on NIR (near-infrared), and sensor-related details are limited. In this regard, there is room for improvement in the acquisition process and methodology for creating remote sensing datasets. To harness the full potential of remote sensing, additional information about sensors and location data should be provided, along with the incorporation of NIR and texture information into deep learning datasets. This would enable the creation of multiple pretrained models that can be used effectively even with minimal data acquisition in disaster situations.

Furthermore, this research sheds light on the potential utility of forest satellite imagery of South Korea which is planned in the future, and identifies elements for further enhancement. The creation of high-resolution forest information maps, which go beyond simple forest maps by incorporating infrastructure details such as tunnels and roads, holds promise for diverse applications. By advancing from conventional forest maps to comprehensive spatial information maps of forested areas, a wide range of uses can be expected in the future.

6. ACKNOWLEDGEMENTS

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00155763).

This research was financially supported by the Institute for Peace and Unification Studies (IPUS) at Seoul National University under the project of "Laying the Groundwork for Unification and Peace."

7. REFERENCES

Acharki, S., 2022. PlanetScope contributions compared to Sentinel-2, and Landsat-8 for LULC mapping. Remote Sensing Applications: Society and Environment, 27, 100774.

Kwon, S. K., Kim, K. M., & Lim, J.,2021. A study on pre-evaluation of tree species classification possibility of CAS500-4 using rapideye satellite imageries. Korean Journal of Remote Sensing, 37(2). https://doi.org/10.7780/kjrs.2021.37.2.9

Lim, J., Kim, K. M., & Kim, M. K., 2019. The Development of Major Tree Species Classification Model using Different Satellite Images and Machine Learning in Gwangneung Area. Korean Journal of Remote Sensing, 35(6–2). https://doi.org/10.7780/kjrs.2019.35.6.2.2

National Geographic Information Institute, 2021. THE NATIONAL ATLAS OF KOREA

Park, T., Kwak, D. A., Lee, W. K., & Lee, J. Y., 2011. Stand species classification and biomass estimation using LiDAR data. Proceedings of Korean Society for Geospatial Information Science, pp. 159-162.

Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V., 2021. U-net and its variants for medical image segmentation: A review of theory and applications. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3086020

PLANET.COM, 2018. PLANET IMAGERY PRODUCT SPECIFICATIONS

Wang, C., & Zhong, C., 2021. Adaptive Feature Pyramid Networks for Object Detection. IEEE Access, 9. https://doi.org/10.1109/ACCESS.2021.3100369

Wang, Z., Wang, J., Yang, K., Wang, L., Su, F., & Chen, X., 2022. Semantic segmentation of high-resolution remote sensing images based on a class feature attention mechanism fused with Deeplabv3+. Computers and Geosciences, 158. https://doi.org/10.1016/j.cageo.2021.104969