Remote sensing-based Plantation Forest Mapping in the Central Highlands of Vietnam: A Deep Learning Approach

CORNELIA ZYGAR[1] & JULIANE HUTH[2]

Abstract: For Vietnam, valuable up-to-date information on plantation forest areas and their species is not publicly available. Therefore, in this thesis, a Long Short-Term Memory (LSTM) approach for mapping rubber and acacia plantation forests was conducted for the Central Highlands of Vietnam. The input data for this time series-based classification were 12 Sentinel-2 monthly median composites for 2020. This LSTM model was compared to a random forest baseline. Accuracies were higher for the LSTM-based classification than for the random forest-based one. Also, rubber F1-values generally were higher than the F1-values for the acacia class. This can be explained by rubber being better suited for a time series classification because of its characteristic phenology compared to the evergreen acacia plant.

1 Introduction

An ongoing loss of forest area can be observed globally. However, this trend does not exist in Vietnam. Between 2010 and 2020, Vietnam was amongst the top five countries with the largest yearly increase in forest area of 126,000 hectares, an annual change of 0.9%. But not all forests are the same. A fundamental distinction must be made between “natural regenerating forests” and “plantation forests”. In Vietnam, the area of natural regenerating forests stayed almost constant, while the area covered by plantation forests increased. Plantation forests provide fewer ecosystem services than natural forests and are part of complex ecological and socioeconomic relationships (FAO 2020). Thus, there is the need for detailed knowledge about the locations of plantations and their species. However, high-resolution, large-scale, up-to-date information about the location and species of plantation forests in Vietnam is not publicly accessible which shows the need for further research on this topic. At the same time, deep learning methods (i.e. neural networks) have recently been intensively applied in the context of land cover classifications and have been shown to produce better results than “classical” machine learning methods (REN et al. 2020; REUß et al. 2021). One type of neural network that was previously used for land cover classifications is a Long Short-Term Memory network (LSTM) (RUßWURM & KÖRNER 2020). In this thesis, it is tested whether the application of a such an LSTM architecture is suitable for detecting plantation forests in the Central Highlands of Vietnam. This is assessed based on rubber and acacia plantations and is compared with a baseline random forest classification.

[1] Julius-Maximilians-Universität Würzburg, Institute of Geography and Geology, Department of Remote Sensing, Am Hubland, D-97074 Würzburg,
c/o German Aerospace Center (DLR), Earth Observation Center (EOC), German Remote Sensing Data Center (DFD), Münchener Str. 20, D-82234 Weßling, E-Mail: corneliazygar@gmx.de

[2] German Aerospace Center (DLR), Earth Observation Center (EOC), German Remote Sensing Data Center (DFD), Münchener Str. 20, D-82234 Weßling, E-Mail: Juliane.Huth@dlr.de

2 Materials and methods

2.1 Data

The data used for the analysis were monthly median composites of cloud-masked Sentinel-2 imagery acquired in 2020. For the presented analysis, the atmospherically corrected level 2A data was used. The imagery consists of 12 bands from which the 10 bands with resolutions of 10 and 20 metres have been used for the presented analysis.

Label data was sampled based on Sentinel-2 and high-resolution Google background images. The data was sampled in point format in the provinces Dak Lak and Dak Nong, located in the Central Highlands of Vietnam, and split into training, validation, and testing partitions with a ratio of 60:20:20. The classes that were sampled are the two plantation classes (acacia and rubber), natural evergreen forest, natural deciduous forest and a background class, covering all remaining plantation types and other landcover classes like agricultural areas. Acacia is an evergreen plant mainly used for pulp and paper production, while rubber is a deciduous tree which, in Vietnam, defoliates in January or February and refoliates shortly afterwards. Its main purpose is the production of latex.

In this thesis, neural networks with three different kernel sizes (1, 3, and 5) were tested. The kernel size describes the distance of surrounding pixels that were considered in the classification. I.e. a kernel size 3 describes a 3 × 3 pixel area centred by the originally sampled pixel. For the random forest, only the kernel 1 data was used as input data.

Figure 1 shows a schematic representation of how the input data for the 3 × 3 kernel version was extracted from the Sentinel-2 median images. Each image consists of 10 bands and the total time series consists of 12 images. For the kernel 1 version, this leads to an array shape of (12,10). For the kernel 3 and 5 versions, the values of the neighbouring pixels are cascaded one after the other. For kernel size 3 × 3, this results in an array shape of (12,90) and for kernel size 5 × 5 in a shape of (12,250).

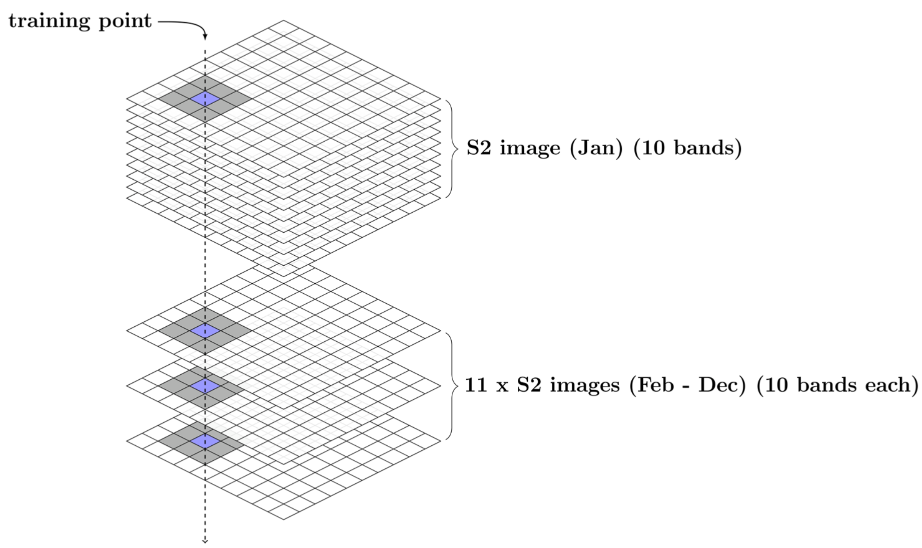

Fig. 1: Schematic representation of input data extraction from Sentinel-2 (S2) monthly median composite time series; own graphic

2.2 Classification methods

LSTMs are a specific type of neural network that can be used for time series forecasting as well as for the classification of time series. They were first introduced by HOCHREITER & SCHMIDHUBER (1997). The architecture that was used in the presented analysis is shown in Fig. 2. It is a modified version of the LSTM used by KERNER et al. (2020). In Fig. 2, blue colours indicate optional parts of the network, that were included in case the hyperparameters were defined accordingly. The input of the network are the 12 time steps of the reflection curves (x1 − x12) with their corresponding band values. The network consists of a variable number of up to three stacked LSTM layers. Optionally, dropout can be applied between these layers. If only one LSTM layer is used, the optional dropout (DO) is applied between the different time steps that are analysed. The output of the last LSTM layer is put into a fully connected layer to reduce the number of network outputs to five (same as the number of classes in the classification). As a final step, a Softmax layer is applied, which returns probabilities for each of the five classes.
To find the model with optimal parameters, a hyperparameter tuning for some of the parameters was conducted. For this, the Ray Tune python library (LIAW et al. 2018) was used in combination with PyTorch lightning (FALCON et al. 2019). Tuning was done for each of the three kernel versions of the LSTM, i.e., kernel 1, 3, and 5. This results in a total of 972 tested hyperparameter combinations.
Random Forest is a decision tree-based machine learning algorithm developed by BREIMAN (2001). It is commonly used for classification tasks in remote sensing and was used as a baseline classification approach.

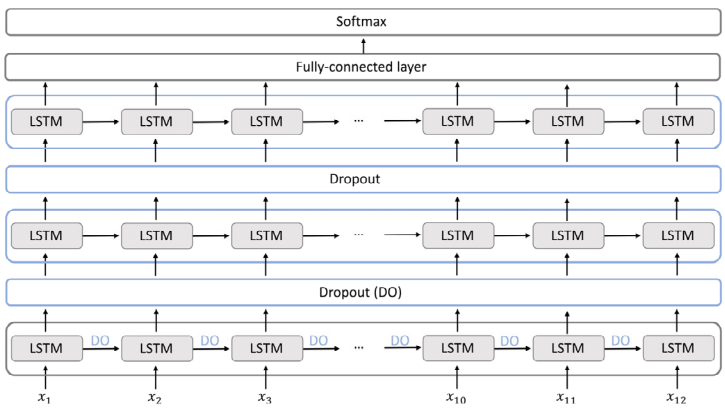

Fig. 2: Adjustable neural network architecture used in this thesis

2.3 Transferability

The data used for the model training was acquired in two provinces of the Central Highlands in Vietnam (Dak Lak and Dak Nong). A transferability analysis was conducted to assess whether the method is applicable to the entire Central Highlands. For this purpose, data was collected in another part of the Central Highlands. This test region is located in the northern half of Gia Lai province. In addition to the spatial transferability, the spatiotemporal transferability of the classification approach was analysed. Here, the models were evaluated on 2021 data from Gia Lai province.

3 Results

3.1 LSTM accuracies and model selection

For each kernel version of the LSTM network, an accuracy assessment was conducted based on a set of testing points located in Dak Lak and Dak Nong. The corresponding accuracy values for the kernel 1, 3, and 5 versions can be found in Table 1. As can be seen here, the overall accuracy and kappa for the three LSTM kernel versions are very similar, with a slight tendency of higher accuracy values with increasing kernel size.

Tab. 1: Test accuracies for kernel 1, 3, and 5

	Kernel 1	Kernel 3	Kernel 5
Overall accuracy	0.943	0.9465	0.9483
Kappa	0.9233	0.9279	0.9302

Despite the similarity of the accuracy values, a visual comparison of the classifications reveals a reduced speckle and “denser” target class areas with increasing kernel size (Fig. 3).

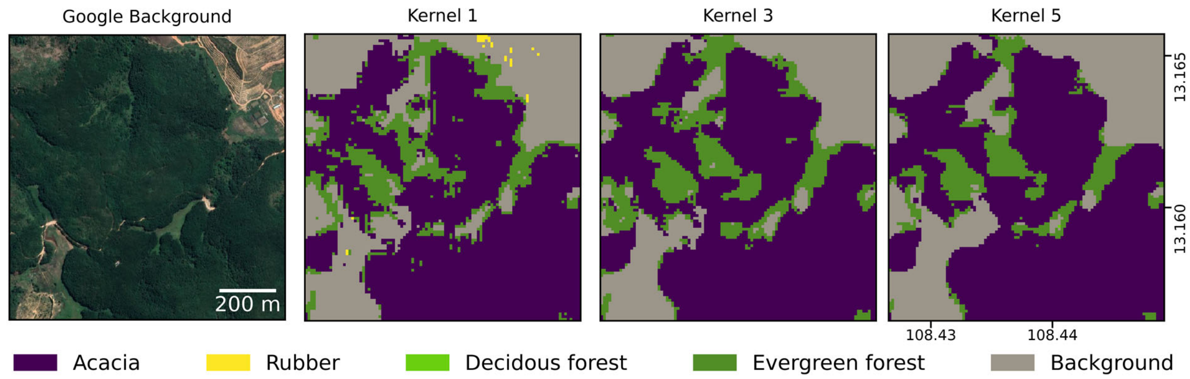

Fig. 3: Effect of different kernel sizes on classification results

Additionally, a visual comparison of detected plantations and the high-resolution Google background image suggests that a larger kernel size leads to an underestimation of plantation extents, as it can be seen in Fig. 4. This might be caused by missing training data at the edges of plantations.

For the final classification, the kernel 3 version was used. It is a suitable compromise between a reduced amount of speckle in the classification and an area estimation that is as correct as possible. The F1 values for the acacia and rubber class in the kernel 3 classification are 0.9653 and 0.9692 respectively (see Table 2).

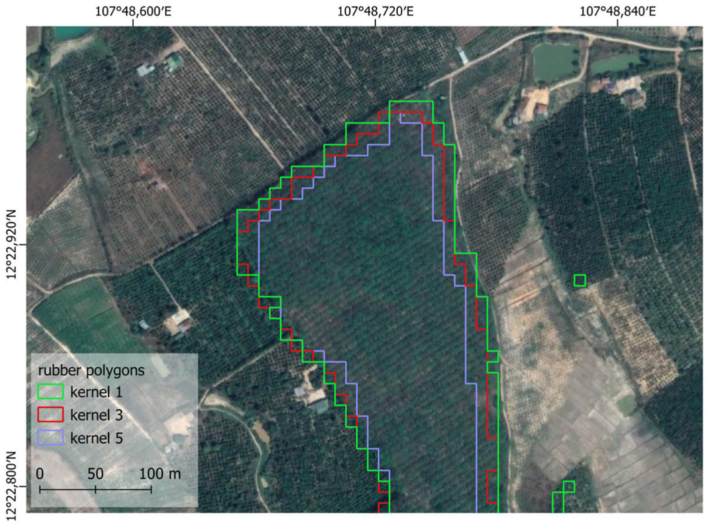

Fig. 4: High resolution Google image in Dak Nong with borders of the detected rubber plantation areas for the different LSTM kernel versions

The accuracy values for the spatial and spatiotemporal transferability analyses are shown in Table 2 as well. There are large differences in the class-wise accuracies. Rubber has high F1 values of 0.9751 and 0.9786 for the spatial and spatiotemporal transferability tests respectively, while the F1 values of the acacia class are 0.8245 for the spatial transferability and 0.6305 for the spatiotemporal transferability. The largest confusion in both tests is acacia being wrongly classified as evergreen forest. This results in acacia recall values of 74.26% for the spatial and 47.06% for the spatiotemporal transferability test.

Tab. 2: Accuracies for kernel 3-based classifications

	Test	Spatial transferability	Spatiotemporal transf.
F1 acacia	0.9653	0.8245	0.6305
F1 rubber	0.9692	0.9751	0.9786
Overall accuracy	0.9465	0.9381	0.9188
Kappa	0.9279	0.9125	0.8844

3.2 Random forest accuracies

Hyperparameter tuning of the number of trees in the random forest revealed that the lowest out-of-bag error was produced with a model consisting of 300 trees. It was used for all further analysis. The corresponding test accuracy values can be found in Table 3. Amongst all five classes, the highest F1 values can be found for rubber (0.9582) and acacia (0.9433).

For the spatial transferability test, the F1 value for the acacia class is 0.7884 and for the rubber class it is 0.9585. For the spatiotemporal transferability analysis, the class-wise F1 scores are 0.56 for acacia and 0.9595 for rubber. Similar to the LSTM-based transferability analyses, the biggest confusion in both random forest transferability tests is found for acacia pixels being classified as evergreen forest. The respective accuracy values are shown in Table 3.

Tab. 3: Accuracies for random forest-based classifications

	Test	Spatial transferability	Spatiotemporal transf.
F1 acacia	0.9433	0.7884	0.56
F1 rubber	0.9582	0.9585	0.9595
Overall accuracy	0.9264	0.9208	0.8951
Kappa	0.9001	0.888	0.8506

4 Discussion

Especially in the context of the transferability analysis, rubber and acacia plantations are predicted with different accuracies. For rubber, the accuracies in both transferability tests are almost 100%. The reason for the good detectability of rubber plantations with the LSTM model might be its distinctive phenology which makes it highly suitable for an LSTM-based classification. Acacia, on the other hand, does not have such a distinct phenology but is very similar to natural evergreen forest. This has two consequences: Firstly, the similar phenology of evergreen forest and acacia plantations makes them generally hard to distinguish. This might be the reason why the biggest confusion between two classes during both transferability tests is between acacia and evergreen forest. And secondly, because acacia is evergreen, the LSTM has the tendency to learn in a similar amount from all timesteps. This missing “focus” towards specific timesteps might lead to a higher sensitivity towards noise in the data (i.e., clouds and cloud shadows). This could also be the reason for the decreased accuracies in the transferability analysis as noise in evergreen forest and acacia plantation points might easily lead to misclassifications.

With increasing kernel sizes, only a slight improvement in accuracy values could be observed for the LSTM networks. However, in combination with a visual analysis it could be shown that including neighbourhood information (surrounding pixels) in the classification improves its result. The improvement of classification results with increasing kernel size suggests that the added information about the neighbourhood of a certain pixel is valuable for the classifier.

The tested random forest classifier received an overall accuracy of 92.64% and a kappa value of 0.9001 on the test dataset. Although these values can be considered very good accuracies, they are worse than the testing accuracies of the LSTM models. The same is true for the transferability analyses. One possible explanation for the lower accuracies of the random forest classification is that the input data used is not optimal for this specific classifier. LSTMs are able to extract relationships between input features. Therefore, they do not need aggregated input data like indices or variance information. For random forests, such input features are usually helpful to provide additional information to the classifier. To keep both classification approaches as comparable as possible, it was decided to use the same input data for both classifiers. However, this might have led to a disadvantage for the random forest.

5 Conclusion and outlook

The LSTM-based classifications were found to achieve slightly higher accuracies than the random forest classification in all tests. Comparing the three LSTM models, it could be shown that the inclusion of neighbourhood information in the LSTM classification improves its quality. However, this improvement was mainly observed through visual assessment and only resulted in a minor improvement in classification accuracy. The transferability tests applied to the selected LSTM model revealed large differences in classification accuracies between both plantation types. For rubber, high accuracies were achieved for the test datasets. Acacia plantations were detected with lower accuracies. A possible explanation for the difference is the characteristic defoliation phase of rubber compared to an evergreen phenology of acacia. Consequently, rubber is more suitable for an LSTM-based classification than acacia.
In further research, other network architectures like convolutional neural networks could be tested in order to further improve the plantation classification, especially of acacia. Also, a random forest-based classification with optimized input features should be considered. Lastly, a country-wide upscaling of the classification approach is particularly promising for the rubber plantation class and could be subject to further research.

6 Publication bibliography

BREIMAN, L., 2001: Random Forests. Machine Learning, (45), 5-32.

FALCON, W. & THE PYTORCH LIGHTNING TEAM, 2019: PyTorch Lightning (Version 1.4) [Computer software]. https://doi.org/10.5281/zenodo.3828935.

FAO, 2020: Global Forest Resources Assessment 2020. Rome: FAO. fao.org/3/ca9825en/ca9825en.pdf, last access 15.09.2023.

HOCHREITER, S. & SCHMIDHUBER, J., 1997: Long short-term memory. Neural computation, 9.8, 1735-1780.

KERNER, H., TSENG, G., BECKER-RESHEF, I., NAKALEMBE, C., BARKER, B., MUNSHELL, B. et al., 2020: Rapid Response Crop Maps in Data Sparse Regions. KDD '20: ACM SIGKDD Conference on Knowledge Discovery and Data Mining, August 22–27 2020, San Diego.

LIAW, R., LIANG, E., NISHIHARA, R., MORITZ, P., GONZALEZ, J. & STOICA, I., 2018: Tune: A Research Platform for Distributed Model Selection and Training. arXiv preprint arXiv:1807.05118.

REN, T., LIU, Z., ZHANG, L., LIU, D., XI, X., KANG, Y., ZHAO, Y., ZHANG, C., LI, S. & ZHANG, X., 2020: Early Identification of Seed Maize and Common Maize Production Fields Using Sentinel-2 Images. Remote Sensing, **12**(13), 2140, https://doi.org/10.3390/rs12132140.

REUß, F., GREIMEISTER-PFEIL, I., VREUGDENHIL, M. & WAGNER, W., 2021: Comparison of Long Short-Term Memory Networks and Random Forest for Sentinel-1 Time Series Based Large Scale Crop Classification. Remote Sensing, **13**(24), 5000, https://doi.org/10.3390/rs13245000.

RUßWURM, M. & KÖRNER, M., 2020: Self-attention for raw optical Satellite Time Series Classification. ISPRS Journal of Photogrammetry and Remote Sensing, **169**(12), 421-435, https://doi.org/10.1016/j.isprsjprs.2020.06.006.