

AUTOMATED UPDATING OF FOREST COVER MAPS FROM CLOUD-FREE SENTINEL-2 MOSAIC IMAGES USING OBJECT-BASED IMAGE ANALYSIS AND MACHINE LEARNING METHODS

M. Debella-Gilo ^{1,*}, B. T. Borchsenius ¹, K. Bjørkelo ¹, J. Breidenbach ¹

¹ Norwegian Institute of Bioeconomy Research (NIBIO), P.O.Box 115, 1431 Ås, Norway - (misganu.debella.gilo, bjorn.borchsenius, knut.bjorkelo, Johannes.breidenbach@nibio.no)

KEY WORDS: Land Cover Mapping, Land Cover Change, Image Classification, Multilayer Perceptron, Random Forest

ABSTRACT:

Planning sustainable use of land resources and environmental monitoring benefit from accurate and detailed forest information. The basis of accurate forest information is data on the spatial extent of forests. In Norway land resource maps have been carefully created by field visits and aerial image interpretation for over four decades with periodic updating. However, due to prioritization of agricultural and built-up areas, and high requirements with respect to the map accuracy, forest areas and outfields have not been frequently updated. Consequently, in some part of the country, the map has not been updated since its first creation in the 1960s. The Sentinel-2 satellite acquires images with high spatial and temporal resolution which provides opportunities for creating cloud-free mosaic images over areas that are often covered with clouds. Here, we combine object-based image analysis with machine learning methods in an automated framework to map forest area in Sentinel-2 mosaic images. The images are segmented using the eCognition™ software. Training data are collected automatically from the existing land resource map and filtered using height and greenness information so that the training samples certainly represent their respective classes. Two machine learning algorithms, namely Random Forest (RF) and the Multilayer Perceptron Neural Network (MLP), are then trained and validated before mapping forest area. The effects of including and excluding some features on the classification accuracy is investigated. The results show that the method produces forest cover map at very high accuracy (up to 97%). The MLP performs better than the RF algorithm both in classification accuracy and in robustness against inclusion and exclusion of features.

1. INTRODUCTION

Detailed forest information is crucial for land use planning, land resource management, biodiversity and climate change monitoring and mitigation (Astrup et al. 2019; Fichtner et al. 2018; Gamfeldt et al. 2013). The basis for forest information is accurate and up-to-date data on the spatial extent of forests. Efficient methods of obtaining accurate forest cover is thus important in up-to-date resource inventory and environmental monitoring, especially with respect to climate change.

Optical satellite remote sensing has long been used in land use land cover mapping including forest cover. However, major parts of countries like Norway are covered by clouds throughout the year with only few days of clear sky. This severely limits the use of optical satellite images. The high temporal frequency of the Sentinel-2 satellite pair (Drusch et al. 2012), every 2 to 3 days for Norway, gives great opportunities to circumvent this challenge. Cloud-free image mosaics over relatively short period of time can be created. The short time spans of one or two months in the middle of the summer, June to July, are important as the phenological conditions of the vegetation are more stable compared to periods of seasonal transitions. Cloud-free mosaic images produced over that period can therefore be considered as representative of the period. Beyond visualisation purposes, such mosaic images can be analysed in a similar approach as single scene images.

Mapping forest cover from remote sensing images is a classification problem where the image characteristics, the

entity to be classified, and the methods used for the classification are important and must be chosen carefully. Sentinel-2 images have already proven to have well suited spectral and spatial characteristics for land cover classification and forest inventory (e.g., Astrup et al. 2019, Puliti et al. 2020). The choice of analysis entity, i.e. pixel, sub-pixel or objects, is nonetheless important. Experiences show that object-based image analysis (OBIA) for land cover classification has a few advantages, including accuracy improvement, compared to pixel-based methods (Blaschke 2010; Myint et al. 2011; Whiteside et al. 2011). Careful implementation is, however, required as OBIA involves some subjective parameter adjustments.

The analysis method is another important issue that affects the outcome. Nowadays, supervised machine learning is widely used in image classification. Some algorithms, such as Random Forest (RF) and Support Vector Machines (SVM), have become state-of-the-art methods for image classification (Khatami et al. 2016). Deep learning based on the Neural Network algorithm are gaining popularity in image classification due to their superior classification accuracy. Multilayer Perceptron Neural Network (MLP), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN) are used depending on the purpose of application. The MLP implements neural networks without considering the spatial and temporal aspects of the data. Detailed description of MLP is given by Panchal et al. (2011). CNNs take the spatial aspect or texture into account. Such an approach is important when object classes have distinct appearances as in the case of animals, fruits, leaf types, etc.

* Corresponding author

However, land cover classes are more complex and satellite images have coarse resolution with a single pixel covering a relatively large area. The spatial aspect is thus better taken care of in a different approach. OBIA is known to be one of the techniques used to include the spatial and neighbourhood context into image analysis (Li et al. 2014). OBIA suppresses noise and enhances uniqueness of the classification entities, i.e. image objects.

In this study, cloud-free sentinel-2 mosaic images over the mid-summer months are used as the main dataset. Image objects created through multiresolution segmentation are used as analysis entities and two machine learning algorithms (i.e. RF and MLP) are trained and implemented in an automated procedure where training data are collected from the existing land resources map. The results of the classification are also automatically integrated with the existing land resources map to create an updated and detailed forest cover map. The effects of including and excluding some important features are also explored.

2. METHODS

2.1 Study area

This study is conducted in the south-eastern part of Norway as shown by the dark area in Figure 1. The area covers about 8000 km². This includes open water areas such as lakes. The region is one of the areas dominated by forest. The whole region is treated as one study area for which the same training dataset is used.

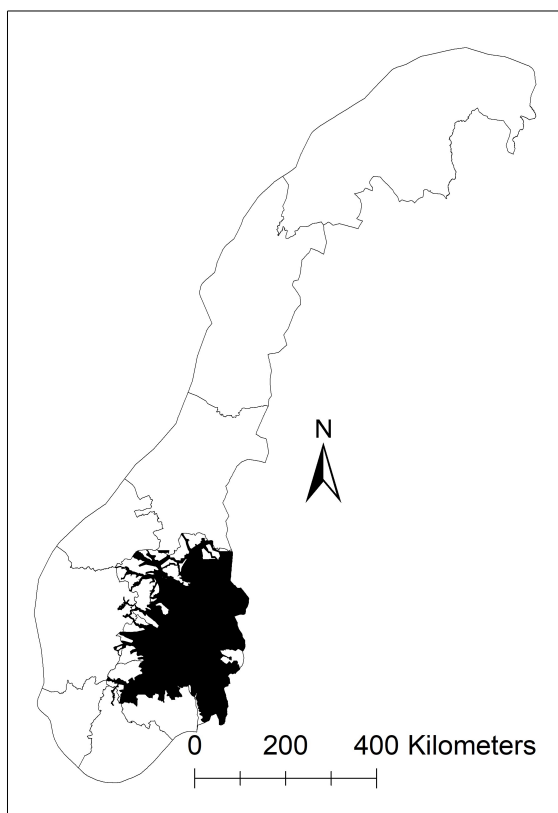


Figure 1. Location of the study area (dark part) in relation to the map of Norway

2.2 Datasets

Sentinel-2 images acquired over the land area of Norway during June and July of 2018 are the image data used in this study. The images were at level 2A, i.e. georeferenced and atmospherically corrected. The Norwegian Mapping Agency produced the cloud-free mosaic images as a national product and provided free public access at a lower bit depth of 8-bit. We are privileged to obtain the mosaics with the full bit depth. The mosaic images are made for all the spectral bands and resampled at 10 m pixel size. All the 10 bands with the original resolutions of 10 m and 20 m are used in this study.

The Norwegian land resource map, called AR5, which is a land use/cover map at the scale of 1:5000 (Ahlstrøm et al. 2019), is used as a reference land cover map. The map is originally created as an aggregated product based on the economic map of the country that was created dating back up to four decades. The AR5 has four attributes; namely major land use/cover type; and within forest, site index, tree species and the ground condition. The map is updated continuously and periodically, but some areas are prioritized over others depending on their economic importance.

Norway is conducting a national Aerial Laser Scanning (ALS) campaign since 2015 to produce a fine-resolution digital terrain model. The acquired data is processed and provided as point clouds, digital terrain model (DTM) and digital surface model (DSM), with a resolution of 1 m, freely by the Norwegian Mapping Agency. In this study, the DSM and DTM are used to create the canopy height model (CHM), that is used to filter the training data further.

In addition, the cadastre map of Norway contains point data representing single buildings, among others. The data can be used in post-processing phase to check if houses are constructed in forest areas.

2.3 Segmentation

The Sentinel-2 mosaic is tiled into smaller overlapping tiles. Each tile with all the selected spectral bands is segmented using the eCognition™ software (Trimble 2018). In eCognition, a ruleset that implements the multiresolution segmentation algorithm, with manually adjusted parameters such as the scaling factor, is created. The ruleset implements a few actions before exporting the segments as polygons, i.e. removing sliver polygons, merging segments covering water and missing data areas. The segments are then exported with a long list of attributes. The attributes included are the mean, standard deviation, and percentiles (25 and 75) of the 10 Sentinel-2 bands, and both CHM and NDVI. This means, four features for each of these attributes is included, creating 48 features in total. The attributes of the AR5 map and the number of building points in each polygon are also included for post classification. For automation purposes, eCognition is integrated into a Python script by piping through its command line application.

2.4 Training data collection

Six major classes are defined based on the AR5 map by aggregating similar classes for this particular purpose; for example, built-up areas and roads, grazing and open areas, etc. Table 1 presents the classes with brief description for each. The training data are collected first as the centroids of the polygons in AR5. As the polygons are large and the landcover type at the

centroid area may have changed since the creation of the map, the collected sample points are further filtered using the CHM and NDVI to confirm whether they accurately represent their respective classes. For example, forest samples must have certain height above ground and must be green during the summer, grazing and open areas (outfields) must be green and not significantly above ground, etc. The final training points are used to extract the attribute values from the segmentation dataset hence creating the complete training dataset. The complete training dataset is composed of about 125000 samples with six classes and 48 features.

Code	Class	Description
0	Built-up	Buildings, roads and paved areas
1	Bare land	Cultivated and other bare areas
2	Forest	All high vegetation
3	Outfields	Open areas which are not forest and not bare
4	Mire	Open mire areas that are not cultivated and not covered by forest
5	Water	All open water bodies

Table 1. The six classes and their brief description

2.5 Training and parameter tuning

During the training stage, 70 percent of the training data is used to train the algorithms with a 10-fold cross-validation approach. That means during every iteration one-tenth of the training data is randomly selected for the cross-validation. The final evaluation sample is prepared in such a way that 30 percent of the dataset is randomly selected and kept out of the training process for later evaluation.

Two of the popular machine learning algorithms in image classification are selected, i.e. the RF and the MLP. Both are non-parametric and widely used machine learning algorithms but operate under different principles. RF is an ensemble classifier that uses multiple decision trees through a voting system (Breiman 2001). RF is conceptually easy to understand, computationally fast and parallelizable (Biau and Scornet 2016). Ma et al. (2017) state that RF shows best performance in object-based classification. RF has some hyperparameters that need to be tuned. In this work, we used the randomized search procedure of the hyperparameter tuning from the *Scikit* Python package.

The MLP is a light-weight deep learning algorithm. As stated by Zhang et al. (2018), although mathematically complicated, the MLP architecture is simple to implement. It does not implicitly consider the spatial or temporal context as the CNN and the RNN respectively do. Image segmentation is one of the methods in which spatial-contextual aspects are included in image analysis Li et al. (2014). Here, the MLP was implemented on segmented objects so that the spatial and neighbourhood context is included implicitly through the segmentation process, and explicitly by collecting statistical values such as standard deviation and percentiles. The MLP was trained by changing the hyperparameters, specifically number of hidden layers, drop-outs and number of epochs.

The effects of including and excluding the spatial-contextual and other features on the classification accuracy is investigated. The spatial-contextual features are features that represent the local textural and contextual variations, for example the standard deviation, the 25 percentiles and 75 percentiles of the

features. Even the mean values produced through the segmentation process can be considered as spatial-contextual as they are aggregates of the pixel values within the respective segments. Additionally, the effects of including and excluding the precomputed NDVI and CHM are also investigated. Table 2 below presents the different cases of including and excluding features.

Case	Description
All Features	All features including CHM and NDVI are included, i.e. all 48 features.
Excluding spatial features	Only the mean values of the features ignoring the statistics that represent the textural variation in the segments is ignored, i.e. 12 features
Excluding CHM	All the CHM related features are removed, i.e. 44 features
Excluding CHM and NDVI	All the CHM and NDVI related features are removed, i.e. 40 features
Only band means	Only the mean values of the spectral bands are used, i.e. 10 features

Table 2. Different combinations of features in the analysis process

After the trained models are evaluated using the test dataset, the model is implemented on the entire segmentation dataset in order to classify each segment into one of the six classes. The classification results are further processed as explained under section 2.6 below.

2.6 Creating the updated forest cover map

As the aim of the work is to create an up-to-date and accurate forest cover map, the final classification result is evaluated in reference to the existing AR5 land resource map. A few rules are created to include new forests and exclude erroneous forest areas from the existing map. First, all areas that are forest in both maps are kept as forest. Second, new forests on areas previously mapped as open areas or mire or grazing areas are included as regenerated or planted forests. Third, areas that are forest in the land resource map, but are not classified as forest in the new map are evaluated further: if their site index does not show productive forest, it is considered as wrong classification in the old map and removed from the new forest cover map. However, if the site index shows productive forest and there is no building point in the polygon, it is kept as forest (possibly harvested) in the new map. Old forest areas now covered by buildings are not included into the new forest cover map. The procedure is presented in a workflow diagram in Figure 2 below.

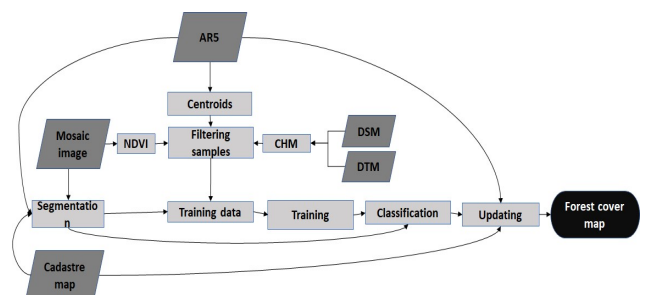


Figure 2. Workflow diagram of the overall automated procedure

3. RESULTS AND DISCUSSION

Table 3 and Table 4 present the normalized confusion matrix of the evaluation result based on the test data for the RF and MLP algorithms respectively. In this analysis all the 48 features are included. The accuracies are computed based on the number, not area, of the training records which are supported by image objects of various sizes. There is very small difference in classification accuracy of forest between the RF and MLP algorithms (96% and 97% respectively). The training data is dominated by forest (well over one-half). Therefore, a better way to look at the performances is the balanced accuracy rather than the overall accuracy and the accuracies of individual classes. As presented in Table 5, MLP resulted in a slightly higher balanced accuracy (91%) compared to the RF (88%).

	Built-up	Arable land	Forest	Outfields	Mire	Water
Built-up	0.96	0.01	0.03	0.01	0.00	0.00
Arable land	0.00	0.93	0.01	0.01	0.05	0.00
Forest	0.00	0.00	0.96	0.01	0.02	0.00
Outfields	0.00	0.04	0.03	0.77	0.15	0.00
Mire	0.00	0.01	0.17	0.01	0.81	0.00
Water	0.00	0.00	0.04	0.00	0.01	0.94

Table 3. The normalized confusion matrix for the RF algorithm

	Built-up	Arable land	Forest	Outfields	Mire	Water
Built-up	0.92	0.00	0.05	0.02	0.00	0.00
Arable land	0.00	0.95	0.01	0.01	0.02	0.00
Forest	0.00	0.00	0.97	0.01	0.02	0.00
Outfields	0.00	0.04	0.03	0.83	0.10	0.00
Mire	0.00	0.00	0.13	0.01	0.86	0.00
Water	0.00	0.00	0.03	0.00	0.01	0.96

Table 4. The normalized confusion matrix for the MLP algorithm

Table 5 presents the effects of including and excluding the different features on the accuracy levels. The values in the parenthesis are from the RF algorithm while the others are from the MLP. As the table depicts, the highest overall accuracy, balanced accuracy, and accuracy of forest class is obtained when all the features are included. The removal of statistical attributes such as standard deviation and the percentiles affects the overall performance of the classification in both RF and MLP almost equally. RF is more sensitive to the removal of CHM and NDVI than the MLP is. The MLP is not at all affected by the removal NDVI. This is in line with the theoretical claim that there is no need of manual extraction of features for deep learning algorithms. The lowest accuracy levels in both algorithms is registered when the spatial-contextual features and the CHM features are excluded. The balanced accuracy in this case is reduced by about 6% and 9%, from that of the case where all the features are used, for RF and MLP respectively. However, in this study, the classification accuracy of forest is almost insensitive to the inclusion or exclusion of these variables. Another important point to note is that RF was more sensitive to the class imbalance than MLP as can be observed by the difference between their respective overall and balanced accuracies.

The automated workflow produced an up-to-date and detailed forest cover map by removing previously erroneous forest areas in the land resources map (Figure 3) and adding new forests (Figure 4). Up-to-date and accurate forest cover map improves the accuracy of the estimation of other forest properties as pointed out in Breidenbach et al. (2020). A closer look at the

new forests show that much of it is regeneration in previously open outfields. However, there are also some areas of mire that are either planted or regenerated in the years after the map is originally created. Some grazing fields have also grown into forest possibly due to abandonment.

Accuracy measure/Case	All features	Excluding spatial features	Excluding CHM features	Excluding NDVI and CHM features	Only band means
Balanced accuracy	91 (88)	91 (88)	89 (81)	87 (80)	84 (79)
Overall accuracy	94 (92)	93 (91)	92 (90)	92 (90)	91 (90)
Forest accuracy (producer)	97 (96)	96 (96)	96 (96)	96 (96)	96 (95)

Table 5. The effects of including and excluding some features on the different measures of accuracy. Values for RF are the ones in the parentheses.

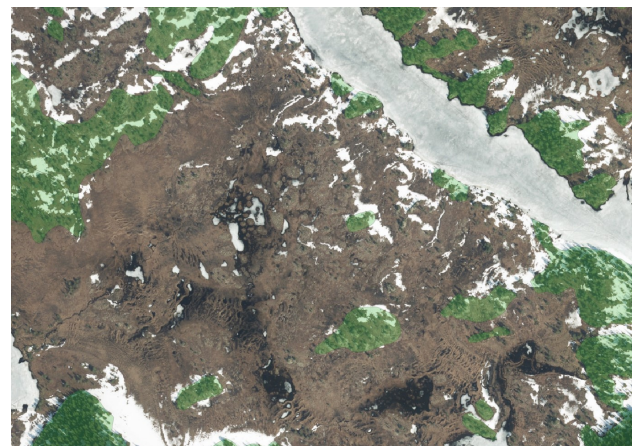


Figure 3. Example of open areas erroneously recorded in the AR5 as forest and removed from the updated forest cover map. Only the green areas are kept in the new forest cover map.



Figure 4. The open forest seen in this figure is an example of forest areas absent in the land resource map but included in the new forest cover map

4. CONCLUSION

This work has demonstrated that cloud-free Sentinel-2 mosaic images created over relatively short time periods can be used beyond visualization. The proposed automated procedure produced an accurate forest cover map. The newly created

forest cover map can be used as a base map for various purposes including forest inventory, climate change monitoring, land use planning and related purposes.

The two algorithms resulted in similar accuracies for forest cover mapping. If the purpose is, however, to make complete land cover map, the MLP performed better in both classification accuracies and sensitivities to some features. The inclusion of spatial-contextual features and CHM improve the classification accuracy in both the RF and the MLP. While manually computed features such as NDVI did not affect the performance of MLP, it slightly affected that of RF.

Such fully automated can even get more accurate if time series satellite images are used instead of the image mosaics as in the case of this work. The temporal dimension can help capture the temporal differences between the land cover classes enhancing the discriminability of the classes.

ACKNOWLEDGEMENTS

The authors are very grateful to the Norwegian Institute of Bioeconomy Research (NIBIO) for financing the study and the Norwegian Mapping Agency for producing the cloud-free Sentinel-2 mosaic images.

REFERENCES

- Ahlström, A., Bjørkelo, K., & Fadnes, K.D. (2019). *AR5 Klassifikasjonssystem*. Ås, Norway
- Astrup, R., Rahlf, J., Bjørkelo, K., Debella-Gilo, M., Gjertsen, A.-K., & Breidenbach, J. (2019). Forest information at multiple scales: development, evaluation and application of the Norwegian forest resources map SR16. *Scandinavian Journal of Forest Research*, 34, 484-496
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, 25, 197-227
- Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, 2-16
- Breidenbach, J., Waser, L.T., Debella-Gilo, M., Schumacher, J., Rahlf, J., Hauglin, M., Puliti, S., & Astrup, R. (2020). National mapping and estimation of forest area by dominant tree species using Sentinel-2 data. *Canadian Journal of Forest Research*, In Press
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25-36
- Fichtner, A., Härdtle, W., Bruelheide, H., Kunz, M., Li, Y., & von Oheimb, G. (2018). Neighbourhood interactions drive overyielding in mixed-species tree communities. *Nature Communications*, 9, 1144
- Gamfeldt, L., Snäll, T., Bagchi, R., Jonsson, M., Gustafsson, L., Kjellander, P., Ruiz-Jaen, M.C., Fröberg, M., Stendahl, J., Philipson, C.D., Mikusiński, G., Andersson, E., Westerlund, B., Andrén, H., Moberg, F., Moen, J., & Bengtsson, J. (2013). Higher levels of multiple ecosystem services are found in forests with more tree species. *Nature Communications*, 4, 1340
- Khatami, R., Mountrakis, G., & Stehman, S.V. (2016). A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sensing of Environment*, 177, 89-100
- Li, M., Zang, S., Zhang, B., Li, S., & Wu, C. (2014). A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information. *European Journal of Remote Sensing*, 47, 389-411
- Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017). A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 277-293
- Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, 115, 1145-1161
- Panchal, G., Ganatra, A., Kosta, Y., & Panchal, D. (2011). Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. *International Journal of Computer Theory and Engineering*, 3, 332-337
- Puliti, S., Hauglin, M., Breidenbach, J., Montesano, P., Neigh, C.S.R., Rahlf, J., Solberg, S., Klingenberg, T.F., & Astrup, R. (2020). Modelling above-ground biomass stock over Norway using national forest inventory data with ArcticDEM and Sentinel-2 data. *Remote Sensing of Environment*, 236, 111501
- Trimble Geospatial (2018). eCognition developer: Version 9.5. Munich, Germany.
- Whiteside, T.G., Boggs, G.S., & Maier, S.W. (2011). Comparing object-based and pixel-based classifications for mapping savannas. *International Journal of Applied Earth Observation and Geoinformation*, 13, 884-893
- Zhang, C., Pan, X., Li, H., Gardiner, A., Sargent, I., Hare, J., & Atkinson, P.M. (2018). A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 140, 133-144