FINAL REPORT

BRASKOG

Bærekraftig **R**essursforvaltning **A**v **S**kogbestand vha mas**K**inlært, aut**O**matisk kartle**G**ging

ALLSKOG
EID AV SKOGEIERE

TABLE OF CONTENT

ABSTRACT / KORTVERSJON

Kjernevirksomhetene til ALLSKOG er basert på detaljerte ressursoversikter over skogen som inneholder informasjon om volum av stående skog, treslagsfordeling, miljø- og vernedata, veinett og driftsbegrensninger. Disse ressursoversiktene er brukt til å forvalte skogen på best mulig måte, både i forhold til bærekraft og økonomisk avkasting til skogeier. Å lage disse ressursoversikter er både arbeids- og tidsintensivt på grunn av at bildetolkning foregår manuelt.

Prosjektet BRASKOG går ut på å utvikle og teste kvaliteten og nytteverdien av automatisert skogbestandskartlegging med utgangspunkt i fjernmålte datasett.

Her undersøkes spesielt potensialet for "deep learning" metoder for automatisert ressurskartlegging. Convolutional neural networks (CNNs) er en type deep learning metode som har endret fagfeltet innenfor kunstig syn i de siste årene. CNN modeller trenger store datamengder for å sette parameterne på modellen (trening), tidligere kartlagte ressursoversikter kan brukes som treningsdata der de tilfredsstiller krav til for eksempel tilgang til ortofoto av en viss kvalitet. I dette prosjektet brukes det CNNs for bildesegmentering for å klassifisere skog, hogstklasse og treslagsfordeling på pixel nivå fra ortofoto og andre fjernmålingsdata. En casestudie er gjennomført i flere kommuner i Trøndelag og Nordland for å sikre representativ variasjon i skogen med tanke på miljø, klima og tilgjengelig datagrunnlag. Resultater er validert ved å sammenligne med de manuell kartlagte dataene. Denne forskningen viser at CNNs kan utnyttes for å klassifisere skog, treslagsfordeling og hogstklasse. Resultatene for treslag og hogstklasse klassifiseringen er lovende og viser potensialet for å erstatte noe av det manuelle arbeidet på sikt.

Prosjektet ønskes å videreføres til en ny fase der prediksjonsmodellen kan forbedres ytterligere, og modellen etableres med brukervennlige grensesnitt som en start til kommersiell utnyttelse.

ALLSKOG
EID AV SKOGEIERE

INTRODUCTION

In 2018, ALLSKOG SA initiated a project to make use of remote sensing data as the availability of datasets increased and their costs are minimal. Therefore, it has become more realistic for smaller companies to utilize this kind of data more extensively.

A major pillar for the business of ALLSKOG SA is the availability of forest inventory data. The forest inventory data usually contains information on forest volume (trevolum), maturity class (hogstklasse), tree species (treslag), information for biodiversity considerations and conservation demands as well as visualisation of forest roads and in some cases expected challenges for logging operations. Forest inventory data is essential for long term sustainable forest management guaranteeing a high level of felling volume throughout the years. Currently the production line for creating forest inventory data takes advantage of both lidar data and aerial imagery. The lidar data is used to retrieve the tree height and volume, whilst aerial imagery is interpreted by humans to derive among others the tree species distribution and maturity class. The manual interpretation and mapping is very labour intensive work and therefore expensive. Firstly, the forest is divided into stands, which are defined as homogeneous parts of the forest that will be treated as units. Secondly, there are certain attributes for each stand that must be set. The most important attributes are maturity class, siteclass and tree species distribution on which the calculation of volume attributes from lidar data depends. Due to the high costs related to creating the forest inventory, data north of Dovrefjell is lacking, extensive areas have never been inventoried or are outdated.

BRASKOG (**B**ærekraftig **R**essursforvaltning **A**v **S**kogbestand vha mas**K**inlært, aut**O**matisk kartle**G**ging) as the project is called is a cooperation with the company Science [&] Technology (S[&]T).

Science & Technology AS was established in 2010 and has since 2015 worked with automatic forest monitoring using Copernicus Sentinel-2 data. By exploiting state of the art machine learning techniques S[&]T aims to increase efficiency of European forestry industry and make it more profitable. In these european projects S[&]T cooperated closely with pilot customers in Portugal and Norway. S[&]T brings in a lot of knowledge and experience in the field of deep learning and remote sensing.

GOAL

The goal of the project is to test the quality and benefit of a stand classification created by machine learning models using remotely sensed data as input. Especially the potential of deep learning methods will be evaluated for automated classification.

Study area and ground truth data

The study area comprises seven municipalities in the counties Trøndelag and Nordland (see figure 1).

The selection of the study area was done based on available recent forest inventory data, the availability for aerial imagery/ lidar data on a certain criteria and last but not least distribution to cover different areas where ALLSKOG is located and has their major impact area.

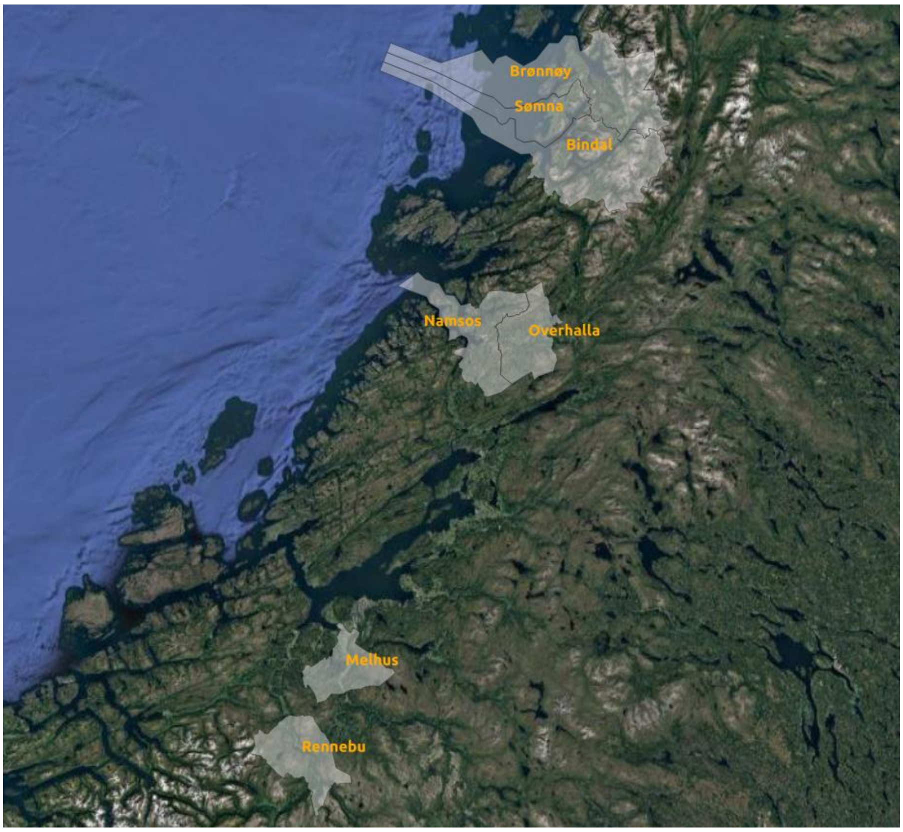

Figure 1: Overview of study area showing the particular involved municipalities.

METHODOLOGY

Background DL/CNN

Deep learning is a subset of methods within the large domain of Machine learning. All the deep learning methods are based on multilayer artificial neural networks (NNs), hence the name deep. In recent years, these techniques have improved the state of the art in speech recognition, computer vision and many other domains (Lecun et al., 2015). Although artificial neural networks have been around for several decades, research and applications have exploded in recent years. Mainly because advances in hardware enabled training of more resource demanding complex/deeper networks. The strength of these types of models is that complex non-linear relationships can be mapped/learned from the data, achieving very high accuracy on many different kinds of classification tasks. Another trade-off is that increased model complexity also reduces the explain ability and risk of overfitting. The downside is, however, that this comes at the cost of models that are computationally resource expensive to train and correspondingly require extensive training data sets.

Convolution neural networks (CNNs) are a particular type of neural network, specifically designed for working on image data. These networks make use of image kernels which are applied on the image, this is called convolution. Major advantages of these image kernels is that weights are shared across the image and connectivity is local. In practise this means that the network becomes less complex and resource demanding and is therefore easier to train and use.

CNNs have changed the field of computer vision and are specifically used for image classification, object detection and image segmentation. Image classification is the process of classifying an entire image into a single class. Object detection is the detection of one or more objects belonging to a certain class within the image. Image segmentation is the partitioning of an image into segments with similar characteristics. In essence the image segmentation can be seen as classification on pixel level. By determining the class of every pixel, segments naturally form by connected pixels with the same class. Popular CNN architectures for image segmentation are U-net and segnet (Ronneberger et al., 2015) (Badrinarayanan et al., 2017).

Within the field of remote sensing CNNs have been applied on all types of imagery, from different sensors and platforms. CNNs are exploited for image classification, object detection and segmentation as well (Zhu et al., 2017). Image classification of entire images is uncommon as images are typically extremely large, comprising large areas with many different objects/features. Examples do exist for example determining whether a satellite image is clouded or not. There are many applications for object detection, examples include detection of vehicles (boats, cars, planes), buildings, roads and trees. Finally, image segmentation is also exploited for many different use cases such as, cloud masking and land cover detection. Some applications where CNNs are exploited within the forestry sector specifically are: clearcut detection, forest fire detection, biomass estimation, tree species determination, forest health monitoring, detection of disease outbreaks and drought.

CNN Classification

In this project the main focus will be on the classification of forest, tree species (treslag) and maturity class (hogstklasse) of the forest. Treslag and hogstklasse are the two main characteristics of the forest that determine the outline of stand (bestand) data. Since we have an extensive dataset of manually collected stand data at our hand, CNNs can be exploited. Without such a dataset the CNN models would be impossible to train. The mapping from some pixel values on the aerial image to the treslag or hogstklasse of the forest is very complex. As an example, there is no simple way to distinguish between a spruce and a pine forest on an aerial image. The appearance of the forest does not only depend on the tree species but on many factors. From which positions and at what angle was the image taken? What is the lighting/shading conditions in the image? How big, old, healthy are the trees in the forest? What is the climate, topography, soil conditions in the area? Even individual trees within the same species are different. As a result we can't simply use, for example, the average color of some pixels to distinguish between spruce and pine. Even if we extract multiple/many features and try to classify based on those, any additional relevant information that is captured within the raw image is ignored. In order to get the best possible classification results we therefore want to exploit all information that is present within the image. CNN models take a window of raw pixel values as input rather than only some features that are extracted from the raw pixel values.

The type of problem we are dealing with within the project is a segmentation problem because, the goal is to detect forest, determine tree species and growth stage throughout the entire image on pixel level. The u-net is selected as a base architecture and modified to optimize for the specific data we are dealing with. According to standard machine learning practices the available data is split into a train and test set. The training set is used for training the model and fine-tuning hyper parameters. The test set is used exclusively for model validation.

AGGREGATION AND VALIDATION

The classification results are on pixel level, however the manually derived reference data is at stand level. In order to be able to validate the model, the classification results are aggregated per stand. This is done by taking the average value of the pixels within each stand.

RESULTS AND DISCUSSION

The results and discussion chapter is subdivided into three sections relating to the main classification tasks that have been investigated during this project. 1) the classification of forest 2) the classification of treslag 3) the classification of hogstklasse.

Forest classification:

Figure 2 shows an example of forest classification for one of the test areas in Melhus. Already in this image alone it can be seen that the color variation within the forest is large. For example, a shaded part of the forest appears very similar to a dark lake, whilst a forest in sunlight might appear very similar to a shaded agricultural field. By exploiting the context, texture and color within the image the forested areas can be extracted.

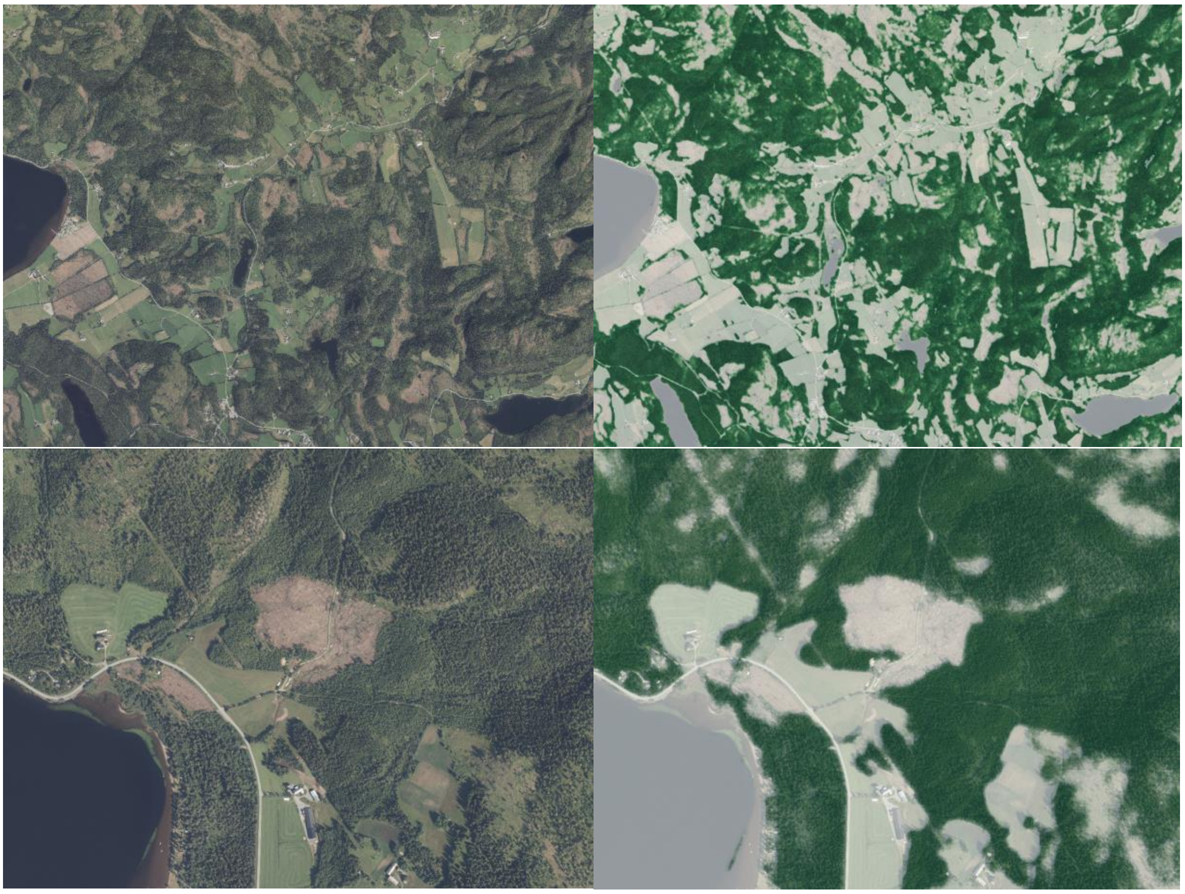

Figure 2: Aerial image of production forest in Melhus, Norway (left-top). Forest classification result, dark-green areas are forested (right-top). A closeup of clearcut area (left-bottom) with classification result (right-bottom).

Currently, no database is present over when and where clearcuts have occurred. Therefore, it is impossible to keep stand data up to date in an automated way. In figure 2 we see a closeup of a clearcut area and the corresponding classification result. The clearcut area is classified as non-forest. By aggregating all pixels within a stand we can determine the percentage of a stand that is forested. In this way we can automatically check the stand data for stands that have been cleared. If hogstklasse is set to 1 or 2 this is in accordance with the manual mapping. On the other hand, if hogstklasse is 4 or 5 this usually means the manual mapping is outdated. Typically, the manual mapping was done on an image before the clearcut happened. Figure 3 shows the stand data over the same clearcut area with respective hogstklasse and percentage of forest. The stands covering the clearcut can be filtered out or updated using this method.

Figure 3: Stand data labelled with respective hogstklasse and colored by percentage of forest within stand according to forest classifier. 0% forest cover (white) to 100% forested (dark-green).

Treslag classification:

Figure 4 shows an example of treslag classification for one of the test areas in Melhus. The classification indicates a mix of gran and furu forest with some patches of lauv forest in the upper right part. The closeup is of an area with some very sharp transitions between patches of gran, furu and lauv forest. More commonly, the transition and difference between gran, furu and lauv is much more gradual and difficult to see on the aerial image.

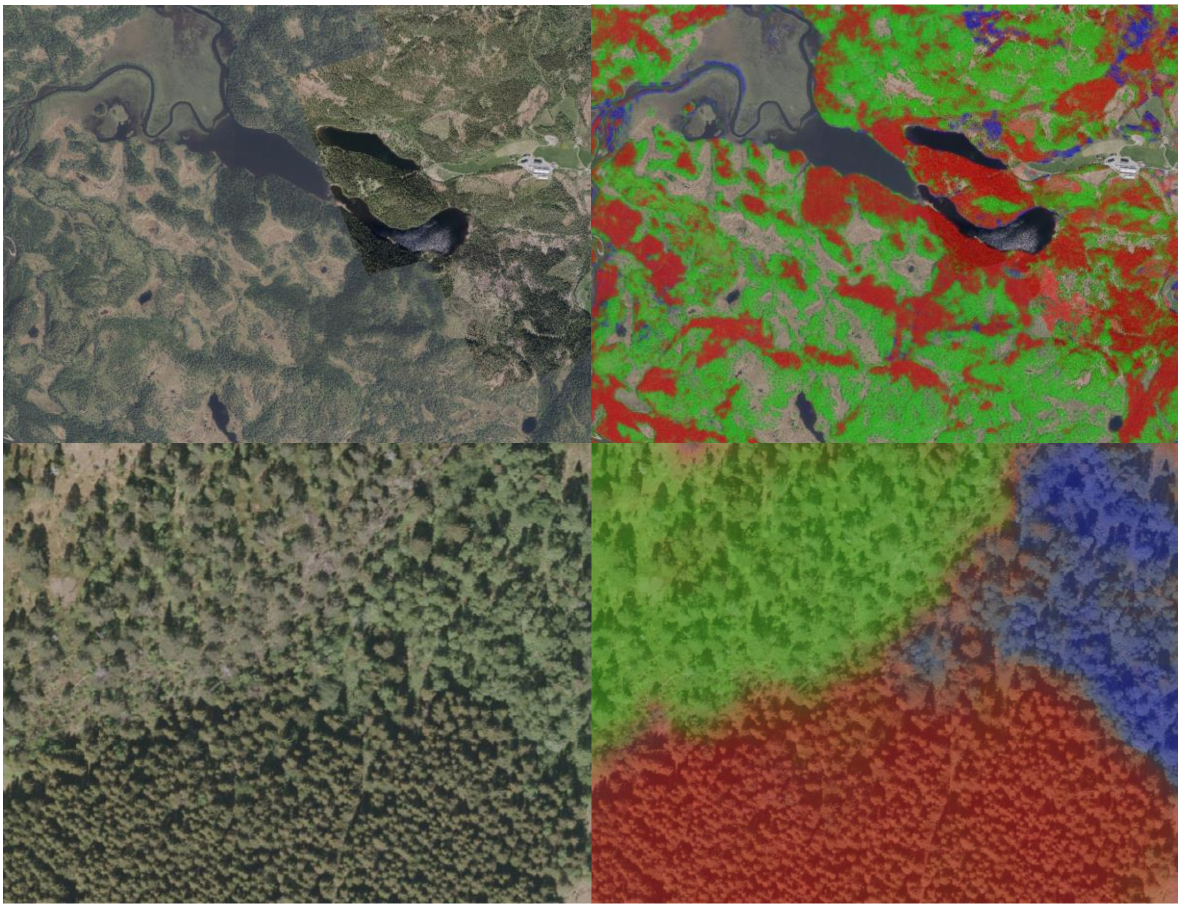

Figure 4: Aerial image of production forest in Melhus, Norway (left-top). Treslag classification (right-top): gran (red), furu (green) and lauv (blue). A closeup of an area with gran, furu and lauv forest (left-bottom) with classification (right-bottom).

For treslag both local (single municipality) and regional (multiple municipalities) models are trained. As explained in the methodology these models are validated by aggregating the classification results per stand and comparing them to the manually derived reference data. Figure 5 shows the confusion matrix for one of the validated treslag models. The accuracy of determining the dominant tree species within a stand is dependent on the treslag. For this particular model the accuracy for gran, furu and lauv are 74%, 79% and 90% respectively. Keep in mind that these values only show a comparison with the manually mapped data, which also contains errors. As expected, the main difficulty is in distinguishing gran and furu. The example in figure 6 shows some common challenges. These presence of low contrast shaded areas, the gradual transition between tree species and corresponding difficulty of delineating stands.

		y_pred		
		gran	furu	lauv
	gran	74%	19%	7%
	furu	19%	79%	0%
y_true	lauv	10%	0%	90%

Figure 5: Confusion matrix showing the performance of a treslag classification model based on 4 test areas within Melhus. The class is determined by the dominant tree species within a stand.

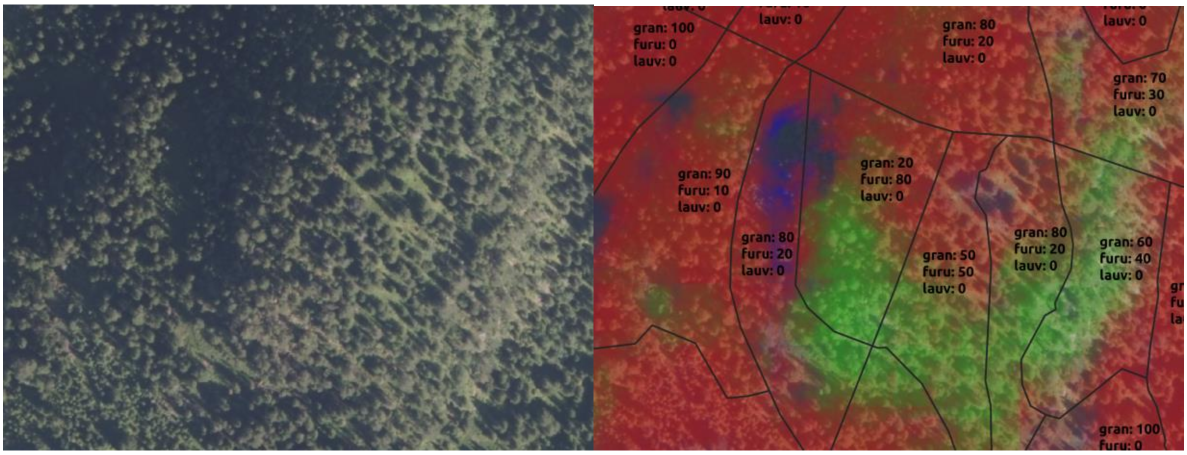

Figure 6: Aerial image (left) with treslag classification and stand delineation with manual derived values (right).

Hogstklasse classification:

Within the CNN, hogstklasse is represented by a single class. The raw output of the hogstklasse model is a single percentage per pixel. The value 0 can be interpreted as certainly hogstklasse 2 and the value 100 can interpreted as certainly hogstklasse 5. After the model is finished training the probability density functions are created for hogstklasse 2-5 using the test data (figure 7). This allows to determine the boundary value between the different hogstklasse, such that the continuous value can be converted to one of the hogstklasse (2-5). The probability density functions also give information about the certainty of a prediction. As an example, a model output value of 10 has a very high likelihood of being hogstklasse 2, however there is a small likelihood of it being a hogstklasse 3 stand. The likelihood of it being a hogstklasse 4 or 5 stand is near 0.

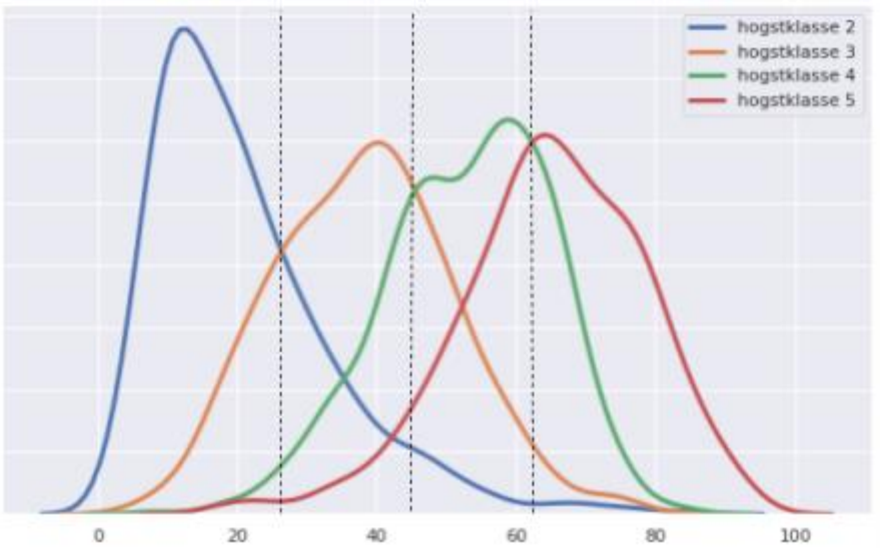

Figure 7: Probability density functions of y_pred for the different y_true hogstklasse classes (2-5).

Figure 8 shows an example of hogstklasse classification for one of the test areas in Melhus. The forest varies from newly planted forest (hogstklasse 2) to fully grown forest (hogstklasse 5). The closeup shows an area with forest in different stages, with fully grown forest at the bottom and some newly planted forest at the top-left.

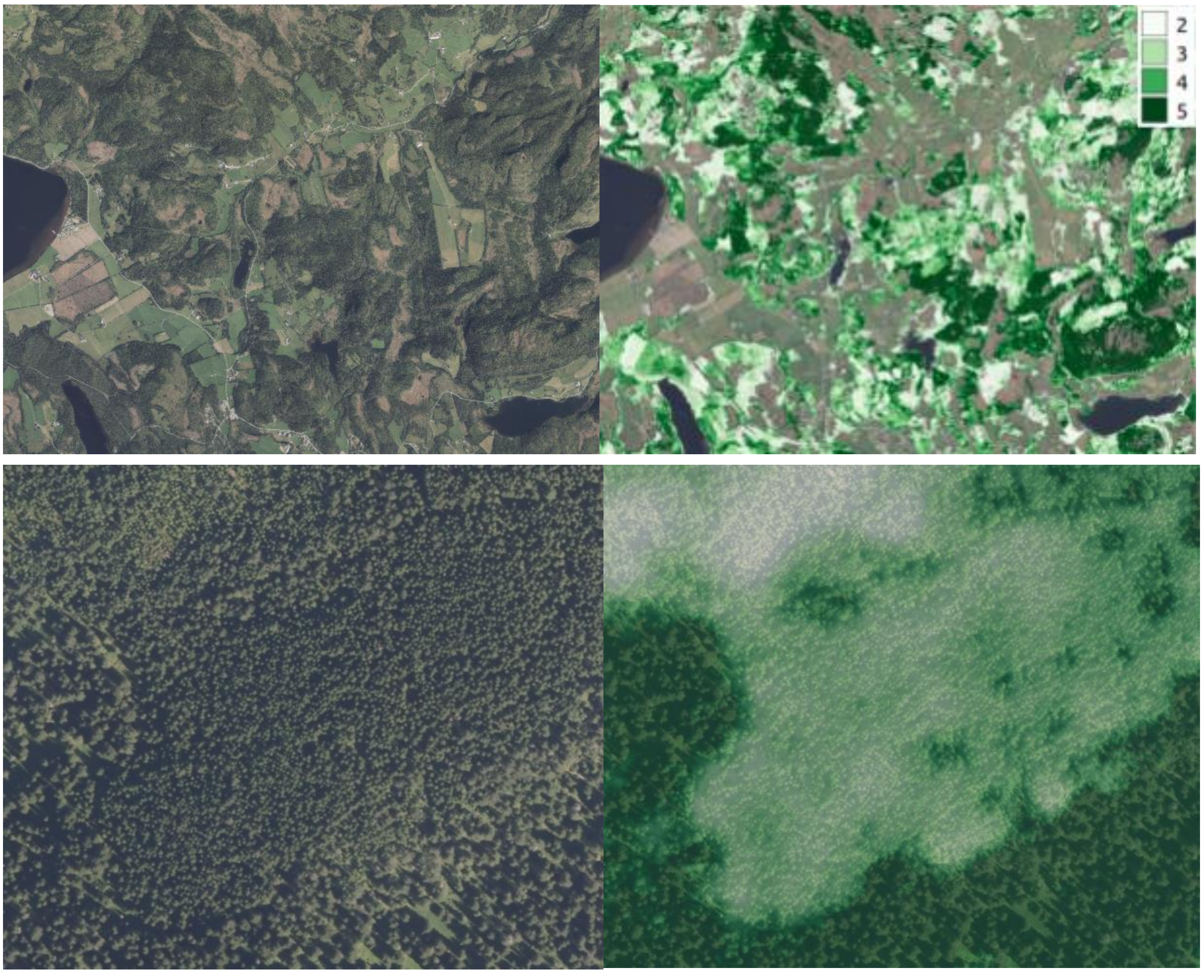

Figure 8: RGB ortho image of production forest in Melhus, Norway (left-top). Hogstklasse (2-5) classification (right-top). A closeup of an area with forest with various hogstklasse (left-bottom) with classification (right-bottom).

Figure 9 shows the validation result for one of the hogstklasse models. Again, values are aggregated per stand for validation purposes. Accuracy of hogstklasse 2 is highest, this is reflected in figure 7, the curve for hogstklasse 2 is most separated from the others. Similarly, the way we pick the thresholds causes accuracy for the classes in the centre (3 and 4) to be lower than for 2 and 5. Overall more than 50% of the stands are classified correctly. Around 90%-95% are not more than 1 class off. The other 5-10% are outliers that are more than 1 hogstklasse off. This can typically be explained by the presence of old/large trees within hogstklasse 2-3 stands or the presence of open spaces or young/small trees with hogstklasse 4-5. This can happen for many reasons such as: when a stand is cleared larger trees are left for seeding or protection of the soil (see figure 10). Natural boundaries are often gradual making it impossible to delineate homogeneous stands perfectly.

		y_pred			
		2	**3**	**4**	**5**
	2	77%	16%	5%	2%
	3	26%	41%	28%	6%
	4	3%	22%	43%	32%
y_true	**5**	1%	6%	26%	67%

Figure 9: Confusion matrix showing the performance of a hogstklasse classification model based on 4 test areas within Melhus.

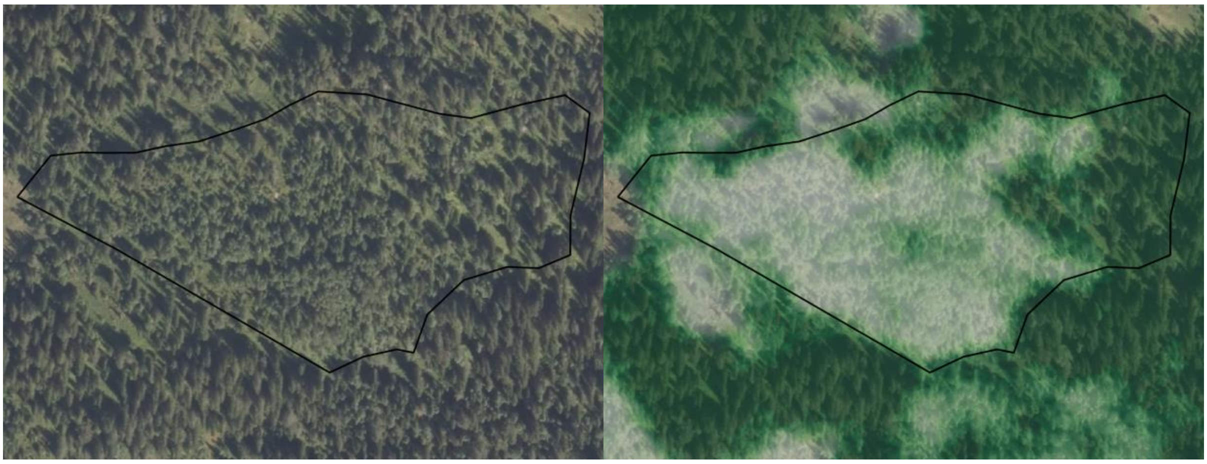

Figure 10: An example of a stand where larger trees are present on the right side of the stand (left). The hogstklasse classification correspondingly shows dark-green in these areas (right).

ALLSKOG
EID AV SKOGEIERE

CONCLUSIONS

A proof of concept is given that CNNs can be exploited for the classification of forest, treslag and hogstklasse specifically. The results are promising and seem to have the potential of substituting some of the manual labour in the long term. The forest classification is already used to clean up stand data, by detecting stands that are clearcut. The treslag classification reaches an average accuracy of around 80% for determining the dominant tree species within a stand. Distinguishing between gran and furu proves to be the most difficult. For hogstklasse over 50% is classified correctly and around 90-95% is at a maximum one class off. Typically, the predictions that are more than 1 class off, can be related to problems with the process of manual delineation and classification of stands. The presented performance values are derived by comparing the model prediction with the manually derived labels. The manual labels are treated as ground truth but in reality, there will be some error. Even though the quality of the manual labels is high, the mapping is done based on aerial imagery only which as well sometimes makes it impossible to determine treslag and hogstklasse.

FUTURE WORK

We see overall promising results of the automatic classifications for the predictions we run in this project. This makes it very interesting for ALLSKOG to continue research on this topic. By implementing results from BRASKOG in a usual production line the man hours could not only be lowered but it is also expected to have a very positive impact on the employees as the work with the forest inventory is very static and requires extreme focus.

First, we see the need to systemize the data very thoroughly if more projects and models should be able to run in the future. For the next phase there are several ideas how to make the tree species classification even more robust. Future work will also aim at developing approaches for automating volume estimation and stand delineation. In order to operationalize the trained models within pilot forest inventory projects we would need to develop an application where non-developers can run the given models with a simple user interface.

Dissemination of results

The main project report in this document is delivered in English language since this have been the main language in the project group and it seemed to be the best way to describe the work that has been done. The report tries to give a good balance of both technical and professional explanations to the audience. An abstract in Norwegian is included in the report for use on the homepage of Skogbrukets verdiskapingsfond. The results will also be published at both S[&]T and ALLSKOGs homepage.

REFERENCES

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, *39*(12), 2481-2495.

Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8-36.