WEAKLY SUPERVISED LEARNING FOR TREELINE ECOTONE CLASSIFICATION BASED ON AERIAL ORTHOIMAGES AND AN ANCILLARY DSM

Jakub Dvořák[1,*], Markéta Potůčková[1], Václav Treml[2]

[1] Department of Applied Geoinformatics and Cartography, Faculty of Science, Charles University, Prague, Czechia –
jakub.dvorak@natur.cuni.cz, marketa.potuckova@natur.cuni.cz
[2] Department of Physical Geography and Geoecology, Faculty of Science, Charles University, Prague, Czechia –
vaclav.treml@natur.cuni.cz

ABSTRACT:

Convolutional neural networks (CNNs) effectively classify standard datasets in remote sensing (RS). Yet, real-world data are more difficult to classify using CNNs because these networks require relatively large amounts of training data. To reduce training data requirements, two approaches can be followed – either pretraining models on larger datasets or augmenting the available training data. However, these commonly used strategies do not fully resolve the lack of training data for land cover classification in RS. Our goal is to classify trees and shrubs from aerial orthoimages in the treeline ecotone of the Krkonoše Mountains, Czechia. Instead of training a model on a smaller, human-labelled dataset, we semiautomatically created training data using an ancillary normalised Digital Surface Model (nDSM) and image spectral information. This approach can complement existing techniques, trading accuracy for a larger labelled dataset while assuming that the classifier can handle the training data noise. Weakly supervised learning on a CNN led to 68.99% mean Intersection over Union (IoU) and 81.65% mean F1-score for U-Net and 72.94% IoU and 84.35% mean F1-score for our modified U-Net on a test set comprising over 1000 manually labelled points. Notwithstanding the bias resulting from the noise in training data (especially in the least occurring tree class), our data show that standard semantic segmentation networks can be used for weakly supervised learning for local-scale land cover mapping.

1. INTRODUCTION

CNNs are currently one of the most commonly used classifiers in remote sensing, but their effective application requires large amounts of training data. Because acquiring more training data is not always feasible, many authors focus on approaches for reducing training data requirements. The most common strategies are transfer learning and data augmentation (Kattenborn et al., 2021). In this paper, we instead focus on a third approach – weakly supervised learning. To explain the rationale underlying our choice, we will describe these three approaches and their applications, discussing their advantages and disadvantages.

Transfer learning involves pretraining the network or a part of it on a larger dataset. These datasets are often unrelated to the specific application, and most pretrained networks are based on three-band RGB images, for example ImageNet. As a result, remote sensing practitioners must either transform their data into three bands (Kattenborn et al., 2021) or use more than one encoder, as proposed by Audebert et al. (2018). Their approach can be utilised with multiple pretrained backbones, which can be particularly helpful when analysing multimodal data. Overall, this approach has been successful, but more significant advantages for remote sensing may derive from network backbones pretrained on large multispectral datasets, such as SEN12MS (Schmitt et al., 2019).

Data augmentation is another effective technique that compensates for insufficient training data. This approach augments existing data by changing them slightly, which allows the network to generalise more successfully. The most common approaches to data augmentation consist of rotating the existing

training data and introducing random noise into the imagery. This technique has been exceptionally successful; in fact, almost half of all recent studies on CNNs for vegetation remote sensing augment their training data (Kattenborn et al., 2021).

Weakly supervised learning stands out among the less commonly used strategies for overcoming the lack of training data, such as synthetic data creation and semi-supervised learning (Kattenborn et al., 2021). Weakly supervised learning consists of training a network using low-quality labels (Kattenborn et al. (2021)). As noted by Schmitt et al. (2020), most remote sensing studies involving weak supervision learning have focused on object detection, whereas only a couple of studies have used this approach for semantic segmentation. These semantic segmentation studies trained models on sparse or image-level annotations, but Schmitt et al. (2020) worked with dense, noisy labels. More precisely, these authors attempted to classify global land cover from high resolution Sentinel-1/-2 data using CNNs trained on a MODIS-based land cover map. Despite highlighting that this complex classification scheme remains a challenge, they also indicated that weakly supervised learning may be useful in less complex scenarios. For example, Weinstein et al. (2019) successfully detected trees in RGB imagery by first training a network on a LiDAR-based unsupervised annotation and subsequently refining the model on a small number of hand-annotated images. For this reason, we focus on weakly supervised learning for semantic segmentation in this contribution.

Our goal is to discriminate Norway spruce (*Picea abies*) trees from dwarf pine (*Pinus mugo*) shrubs in multispectral (visible and near infrared) aerial images. Conventional machine learning approaches (support vector machine, random forest, etc.) were
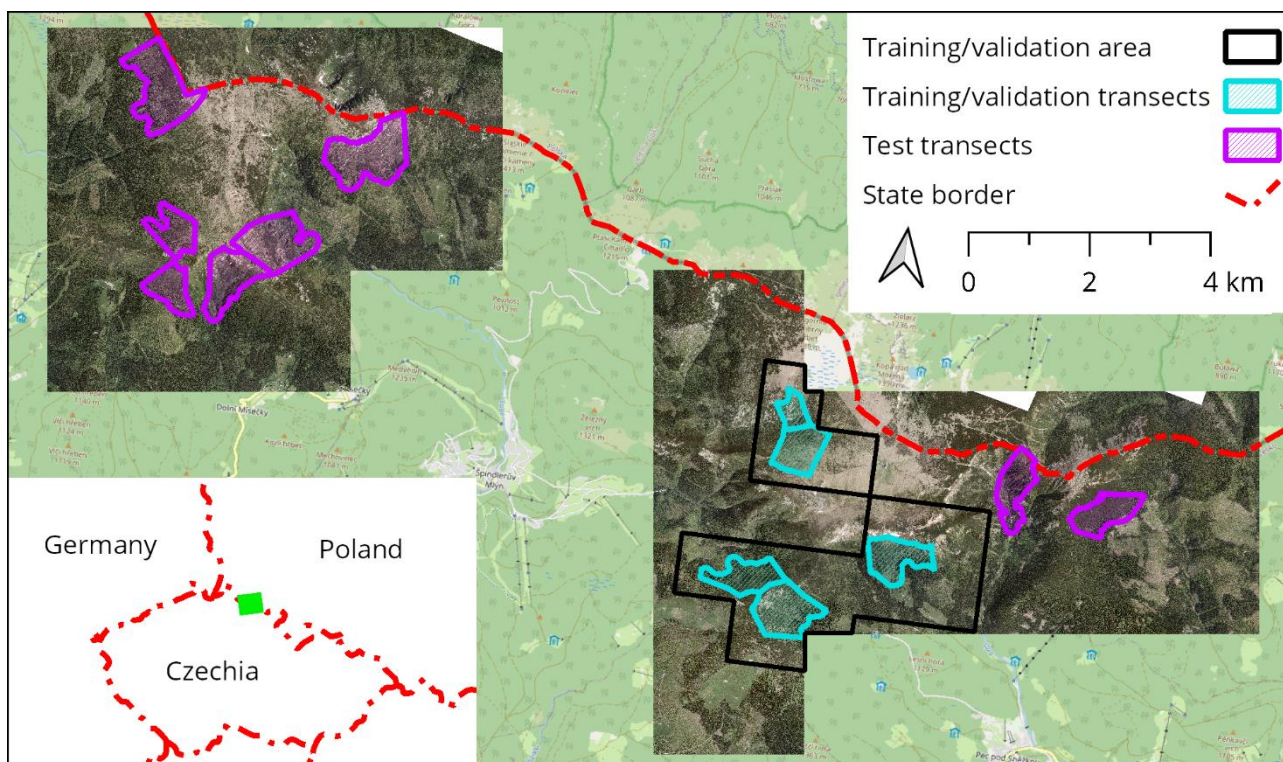
---

* Corresponding author

Figure 1. Area of Interest in the treeline ecotone of Krkonoše mountains. Depicts both the western and eastern tundra. Made with Natural Earth data, © OpenStreetMap contributors

first considered. However, preliminary data exploration using conventional machine learning and image segmentation has shown that these classifiers are unsuitable for the given dataset as was further confirmed by Dvorak (2020). The complex spatial structure of spruces, especially their shadows, and dwarf pines showed to be difficult to address using object based as well as pixelwise classification.

Thus, our approach is similar to that of Weinstein et al. (2019), albeit aimed at semantic segmentation of two vegetation classes while avoiding training on hand-annotated data. For this purpose, we trained two CNNs on a noisy classification, which was independently derived using an ancillary normalised Digital Surface Model (nDSM) and spectral information. The nDSM is available only for a part of our study area and, therefore, cannot be used as a classification feature. Despite the inherent training data noise, we hypothesize that CNNs can identify the spatial structure of trees in multispectral images.

2. KRKONOŠE MOUNTAINS

The Krkonoše Mountains in the Sudetes Mountain system form a natural border between Czechia and Poland. The area is designated as a UNESCO biosphere reserve and as a national park on both sides of the border – known as *Krkonošský Národní Park* (KRNAP) in Czechia and as *Karkonoski Park Narodowy* in Poland (UNESCO, 2016). This area contains an arctic-alpine tundra, a relict ecosystem unique for its flora. The Krkonoše tundra consists of two large, detached sections: Western (1 284 ha) and Eastern (2284 ha). Connecting the tundra to Spruce-dominated montane forests below is the treeline ecotone (Treml and Chuman, 2015).

The alpine tree line ecotone is characterised by an extensive dwarf pine shrub cover and by spruce stands sparsening with altitude. Local dynamics of the ecotone were analysed by Treml and Chuman (2015), who found that the timberline has advanced

upwards by 0.43 m per year since the 1930s. The authors noted that this process is comparatively slower in in the Sudetes Mountain system than in other European mountain ranges, possibly due to its high pine shrub cover.

These and potential other changes in treeline ecotone dynamics are monitored by the KRNAP administration. In total, 13 transects have been delineated for this purpose, as shown in Figure 1. The transects were defined with different spruce and dwarf pine densities and distributions throughout the area. In total, the transects cover 7.79 km^2, with six of the transects located in the western tundra and the remaining seven in the eastern tundra.

3. METHODOLOGY

3.1 Data

3.1.1 Aerial imagery: Our study uses multispectral aerial imagery captured on the 18th and 27th of June 2012. The imagery covers the whole study area and is divided into 2×2.5 km tiles with 8bit radiometric resolution. As is relatively common in real-world tasks, we had no access to the original data, only to derived products – a Red, Green, Blue (RGB) and Near Infrared, Red, Green (CIR) orthorectified composites with ground sampling distance (GSD) of 0.2 m.

These products share two spectral bands (Red and Green), but their simple combination into a four-band dataset (NIR, R, G, B) is impractical because both composites were radiometrically enhanced separately. Therefore, the red and green bands do not share values between products, so the two datasets were used separately.

3.1.2 nDSM: A LiDAR-based surface model covering a part of the area of interest (training/validation area in Figure 1) forms

a basis for creating training and validation labels. The data for this model were gathered with a point cloud density of 5 points/m^2 using a RIEGL LMS Q-680i full waveform scanner in spring of 2013 (Puchrik and Nýdrle, 2013). We had access to the nDSM as a preprocessed raster with a GSD of 0.125 m. This original raster was subsequently resampled using Nearest Neighbour interpolation to match the multispectral imagery (0.2 m × 0.2 m). While both the datasets were captured a year apart, we don't consider this to be an issue as there were no observable changes in the area.

3.2 Experiment setup

Our goal was to classify spruce trees and dwarf pine shrubs in the areas of interest of the Krkonoše treeline ecotone. Consequently, we defined three semantic classes for classification – spruce trees (*picea abies*), dwarf pines (*pinus mugo*) and background. The complex spatial structure of these classes is extractable using a deep Convolutional Encoder-Decoder network. Accordingly, we trained two deep models, as described in section 3.4.

To confirm the results, we generated 1086 points in the 13 evaluated transects by stratified random sampling; 543 points for the eastern and 543 points for the western part of the territory. The number of points was derived according to (Foody, 2009) for a two-sided confidence interval α = 2%. The points were consequently labelled based on visual interpretation of the multispectral imagery, essentially creating a testing dataset.

As the available nDSM covered only a portion of the treeline ecotone, we generated a semiautomatic annotation in the area where imagery and nDSM overlap. This semiautomatically labelled region was subsequently used for training and validating our Encoder-Decoder networks. Five of the eastern transects were also selected to validate our training / validation area. The Encoder-Decoder results found in these five transects were also compared with results in the remaining transects using the aforementioned hand-labelled validation points. This comparison allows us to gauge the generalizability of our models.

All Encoder-Decoders were trained on a Windows workstation equipped with an Intel i9-7940X CPU, 64 GB of RAM and a Nvidia GeForce GTX 1070 GPU (8GB VRAM). The models were written for PyTorch 1.4 with CUDA 10.2.

3.3 Semiautomatic annotation

Creating training data for our task manually is impractical given the overall size of the study area. To simplify this task, we took advantage of spruce tree and pine shrub aboveground heights in the area. As we were operating in the Krkonoše tundra, the only other objects with similar heights were either manmade structures or dead vegetation (Treml and Chuman, 2015). This fact allowed us to use aboveground heights (where nDSM is available) as a feature to create the training and validation dataset.

Thus, when creating our labels, we assumed that:
1) green trees and shrubs have normalised difference vegetation index (NDVI) higher than zero
2) trees and shrubs can be identified from their height

We computed NDVI from CIR imagery and assigned all pixels with NDVI below 0 to the background class. To differentiate vegetation types, we assigned pixels with nDSM value above 2 m as spruce trees and pixels with nDSM value between 0.12 m and 2 m as dwarf pine shrubs. These arbitrary values were selected after visual comparison of multiple potential height values.

Our semiautomatic annotation considerably reduced the time required to create enough training data. However, this method introduces a degree of uncertainty. For this reason, we only used the labels for model training and validation, whereas the reported accuracy of the test set was based on visual interpretation of the original imagery.

3.4 Encoder-Decoder networks

A family of CNNs sometimes known as Convolutional Encoder-Decoders includes networks such as SegNet or U-Net, which are well suited to pixelwise classification/semantic segmentation. U-Net was selected as the basis for this task, given its recent popularity for solving comparable remote sensing tasks (Brandt et al., 2020; Huang et al., 2018; Kattenborn et al., 2019).

We performed no data augmentation as it may not be needed given the relatively large size of our training and validation dataset. Rotating the input patches may even have been harmful to the classification, given that this would rotate tree shadows, while their location in relation to the trees is a potentially valuable feature for classification.

Spruce trees cover a relatively small area in comparison to the other two classes. We therefore assign weights when computing the loss function with relation to the individual classes, a similar scheme was previously used by Audebert et al. (2018). The weights were assigned as: background – 0.1, pine shrubs – 0.2 and spruce trees – 0.7. This setup will likely improve the results for spruce trees at the expense of overall accuracy.

Data is fed into the networks in the form of 512×512-pixel tiles with 50% overlap and six-band inputs. Using these relatively large tiles results in 1005 tiles for training and 252 tiles for validation. Both networks were trained using Adam (Kingma and Ba, 2015) as the optimiser and cross entropy as the loss function. While training, batch size of 2 was selected for both our models, due to memory constraints of our GPU. Suitable values for other hyperparameters were identified using three-fold cross validation, and they are detailed in Table 1.

	# of epochs	Initial learning rate	Learning rate halved after epochs #
U-Net	80	1.25e-5	50, 70
KrakonosNet	100	1e-3	60, 80, 95

Table 1. Hyperparameter settings for training

3.4.1 U-Net: Our U-Net implementation pads convolutions, similarly to Kattenborn et al. (2019) and Wagner et al. (2019). This adjustment results in network outputs with the same spatial resolution as inputs. As in the original U-Net (Ronneberger et al., 2015) the weights and biases of the network were randomly initialized from a Gaussian distribution.

3.4.2 KrakonosNet: A modification of U-Net was introduced by (Dvorak, 2020). Its goal is to reduce overfitting of the network for tree classification in aerial imagery. The author proposes four modifications to the original U-Net structure:
1) Halving the number of feature maps in each convolution layer – inspired by (Wagner et al., 2019).
2) Batch normalization after each convolution – inspired by Zhao et al., (2019)

3) Dropout after the last convolutional layer (50 % probability)
4) Parametric Rectified Linear Unit (ReLU) activation function instead of traditional ReLU – inspired by Nwankpa et al. (2018).

Other implementation specifics are the same as for our U-Net.

4. RESULTS

To evaluate our noisy training data, we created a confusion matrix which compares the training data to 382 overlapping hand-labelled points (Figure 2). This analysis showed a relatively high confusion between the background and pine shrubs, especially for pines, which were misclassified as background. Even more notable is the underrepresentation of spruce trees, with only over a third (38%) of spruces being correctly represented in the weak training labels.

Figure 3 shows patches from the five training / validation transects, highlighting the underrepresentation of spruce trees. The automatically labelled patches also show a characteristic "salt and pepper" effect, which is especially apparent in a confusion between the background and dwarf pines. The resulting classifications have however filtered out most of this noise. Differences between individual results obtained with U-Net and KrakonosNet appear minor based on visual interpretation, but the importance of these slight differences may show during quantitative analysis.

Figure 2. Confusion matrix describing the agreement between our noisy labels and the hand-labelled points

Analysis of confusion matrices presented in Figure 4 reveals that both models performed quite similarly, with KrakonosNet slightly outperforming U-Net. However, both models significantly underestimate spruce trees, and the producers' accuracy for this class was only 55.46% (U-Net) and 63.03% (KrakonosNet). This result is nevertheless considerably better than that of the original training data (38%). In other words, the networks learned some properties of trees from noisy annotations and delineated the trees more accurately than the training data. A notable improvement over the noisy labels has also been achieved in confusion between pine shrubs and the background.
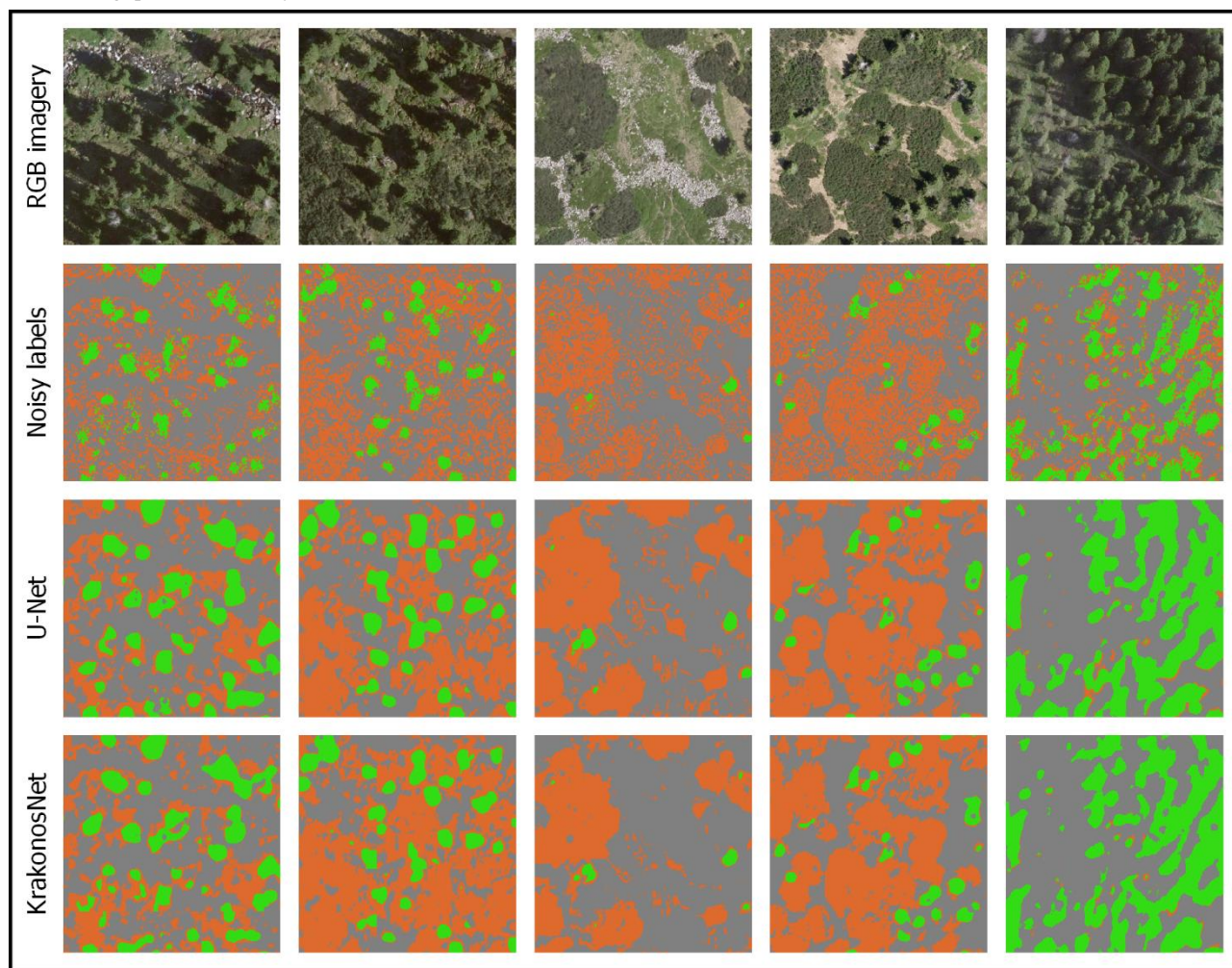

Figure 3. Comparison of noisy training labels and resulting classifications, **grey: background**, **orange: dwarf pine**, **green: spruce**
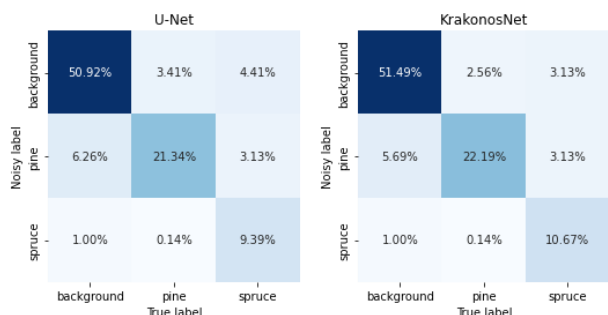
Figure 4. Confusion matrices for the test set.

Estimations of how well a given model generalised are shown in Table 2. A lower difference between values for the test and training/validation set indicates that the model has learned truly valuable features for the given task, while a larger difference means that the model is potentially overfitted.

	Test set (8 transects)		Train/val set (5 transects)	
	U-Net	KrakonosNet	U-Net	KrakonosNet
Background				
F_1	87.10	89.27	88.16	87.72
IoU	77.16	80.62	78.83	78.13
Pine shrubs				
F_1	76.73	79.39	86.03	86.45
IoU	62.24	65.82	75.48	76.13
Spruce trees				
F_1	68.39	74.26	82.11	80.43
IoU	51.97	59.06	69.64	67.27
Overall				
OA	81.65	84.35	86.65	86.39
Mean F1	77.41	80.97	85.43	84.87
Mean IoU	63.79	68.50	74.65	73.84

Table 2. Results validated on hand-labelled points [%]. Comparison of values obtained within the eight test set transects and the five transects that overlap the training / validation set. F_1: F_1 score, IoU: intersection over union, OA: overall accuracy

Both our models had more than 8% difference in F_1 score and IoU for pine shrubs and spruce trees. This difference is especially notable given the relatively low agreement between noisy labels and our validation points (Figure 2). KrakonosNet also has lower differences between results for the training/validation and test dataset than U-Net across all classes and metrics while reaching better overall results on the test set.

5. DISCUSSION

Our results show that our Encoder-Decoders have learned useful features of spruce trees and pine shrubs from aerial imagery while being trained on noisy labels. Learning these features allows both models to outperform original labels by filtering out the noise.

KrakonosNet slightly outperformed U-Net across all metrics in the test set. One notable difference between the networks lies in the producers' accuracy for spruce trees, which also led to the higher F_1 and IoU scores for KrakonosNet for this class. We can therefore confirm that the network structure adaptations prove useful for this task, but a more in-depth analysis of the usefulness of individual network adaptations should be conducted in further studies. KrakonosNet is also considerably faster to train because it only has half of the convolutional filters of U-Net. A considerable improvement in classification accuracy may have been achieved by using a state-of-the-art Encoder-Decoder, but we believe that there is value in first experimenting with a relatively well-understood architecture.

Contrary to Schmitt et al. (2020), we conclude that standard models such as U-Net (perhaps with a few simple adaptations) may be sufficient for our task given that mapping two vegetation classes on a local scale is considerably easier than the global land cover mapping presented in their contribution.

Although our overall results were satisfactory, a more careful analysis of accuracy metrics and the visualised results identifies several shortcomings of this approach. Firstly, shadows are difficult to successfully classify in this context as the appropriate label cannot always be assigned even based on a visual interpretation. We observe (as illustrated in Figure 3) that lower parts of the spruce trees are often misclassified as dwarf pine shrubs, especially on the shadowed half of the trees. This misclassification is likely caused by a systemic issue within our algorithm for generating training data.

Our training data generation technique assigns parts of spruce trees below two meters to the pine shrub class, thus misclassifying tree edges. This height restriction also taught the network to mistake newly planted forests for dwarf pine shrubs. To correct this issue, the existing training data may be improved by a simple procedure. For example, the existing labels for spruce trees could be extended by a buffer of 50 cm in all directions, which would help to properly cover the spruce trees. However, this systematic error may also be resolved by creating training data through a different strategy.

One such strategy for automatically creating an annotation includes finding local maxima of the nDSM raster and creating circular buffers with a set radius around them, thus delineating the spruce tree canopies in the process. The radius of buffers may even be tied to the treetop height, thereby creating larger buffers for higher trees. This relationship between tree height and width would however need to be calibrated each time when using imagery captured by different sensors due to differences commonly found between aerial orthoimages caused by the image geometry and rectification process. These strategies for generating training data should be considered as part of a sensitivity study. Such study should also consider introducing shaded classes.

Our goal was to assess Convolutional Encoder-Decoders for weakly supervised learning, and the results show that these networks can overcome the noise present in the training imagery. However, we believe that a production-focused solution should also involve data augmentation, transfer learning or a small amount of precisely labelled training data. In particular, data augmentation through patch rotation or the introduction of random noise can be easily implemented, while typically improving the generalization capabilities of models. Transfer learning with remote sensing data is often challenging due to the high dimensionality of the data, but Audebert et al. (2018) have presented a potential avenue for transfer learning in the context of our data. We could essentially run both our three-band datasets through separate pretrained encoders and then combine both signals for the decoder part of our network.

Besides already mentioned directions for future research in the context of this task, we aim to adapt weakly supervised learning in the same area to data captured in different time periods and using different sensors. For example, a land cover map created based on 2012 imagery may be used as training data for models training on 2001 imagery to assess historical changes in the area. To achieve satisfactory results, a model must consider specifics related to tree growth and the quality of available aerial imagery.

In summary, the weak labels proved sufficient to teach useful classification features to standard Encoder-Decoder networks, but the labels contain a systematic error which ultimately limited the performance of the model. Nevertheless, the limitations of our results may be overcome by creating training data in a different manner and by improving the training process with data augmentation or transfer learning.

6. CONCLUSIONS

Weakly supervised learning proved viable on our task of classifying land cover in the treeline ecotone using labels based on an ancillary nDSM. The networks could overcome noise present in the training data and learn better representations than the labels. Our results still reflect systematic errors introduced when creating training data, but we show that standard semantic segmentation networks may be sufficient for weakly supervised learning for local-scale land cover mapping.

REFERENCES

Audebert, N., Le Saux, B., Lefèvre, S., 2018. Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks. ISPRS Journal of Photogrammetry and Remote Sensing 140, 20–32.
 https://doi.org/10.1016/j.isprsjprs.2017.11.011

Brandt, M., Tucker, C.J., Kariryaa, A., Rasmussen, K., Abel, C., Small, J., Chave, J., Rasmussen, L.V., Hiernaux, P., Diouf, A.A., Kergoat, L., Mertz, O., Igel, C., Gieseke, F., Schöning, J., Li, S., Melocik, K., Meyer, J., Sinno, S., Romero, E., Glennie, E., Montagu, A., Dendoncker, M., Fensholt, R., 2020. An unexpectedly large count of trees in the West African Sahara and Sahel. Nature. https://doi.org/10.1038/s41586-020-2824-5

Dvorak, J., 2020. Deep learning for tree line ecotone mapping from remote sensing data (Bachelor's Thesis). Charles University, Praha. https://hdl.handle.net/20.500.11956/120388

Foody, G.M., 2009. Sample size determination for image classification accuracy assessment and comparison. International Journal of Remote Sensing 30, 5273–5291. https://doi.org/10.1080/01431160903130937

Huang, B., Lu, K., Audebert, N., Khalel, A., Tarabalka, Y., Malof, J., Boulch, A., Le Saux, B., Collins, L., Bradbury, K., Lefèvre, S., El-Saban, M., 2018. Large-Scale Semantic Classification: Outcome of the First Year of Inria Aerial Image Labeling Benchmark, in: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Valencia, pp. 6947–6950. https://doi.org/10.1109/IGARSS.2018.8518525

Kattenborn, T., Eichel, J., Fassnacht, F.E., 2019. Convolutional Neural Networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery. Sci Rep 9, 17656. https://doi.org/10.1038/s41598-019-53797-9

Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S., 2021. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 173, 24–49.
 https://doi.org/10.1016/j.isprsjprs.2020.12.010

Kingma, D.P., Ba, J., 2015. Adam: A Method for Stochastic Optimization, in: Bengio, Y., LeCun, Y. (Eds.), 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings. https://arxiv.org/abs/1412.6980

Nwankpa, C., Ijomah, W., Gachagan, A., Marshall, S., 2018. Activation Functions: Comparison of trends in Practice and Research for Deep Learning. https://arxiv.org/abs/1811.03378.

Puchrik, L., Nýdrle, J., 2013. Snímkování a letecké laserové skenování Krkonošského národního parku (in Czech). Geodis news 12, page 24.
 https://issuu.com/geodis/docs/geodis__news_stranky/24
(16 January 2022)

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Schmitt, M., Hughes, L.H., Qiu, C., Zhu, X.X., 2019. SEN12MS – a curated dataset of georeferenced multi-spectral Sentinel-1/2 imagery for deep learning and data fusion. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. IV-2/W7, 153–160. https://doi.org/10.5194/isprs-annals-IV-2-W7-153-2019

Schmitt, M., Prexl, J., Ebel, P., Liebel, L., Zhu, X.X., 2020. Weakly Supervised Semantic Segmentation of Satellite Images for Land Cover Mapping -- Challenges and Opportunities. https://arxiv.org/abs/2002.08254.

Treml, V., Chuman, T., 2015. Ecotonal Dynamics of the Altitudinal Forest Limit are Affected by Terrain and Vegetation Structure Variables: An Example from the Sudetes Mountains in Central Europe. Arctic, Antarctic, and Alpine Research 47, 133–146. https://doi.org/10.1657/AAAR0013-108

UNESCO, 2016. Krkonose/Karkonosze [WWW Document]. United Nations Educational, Scientific and Cultural Organization. https://www.unesco.org/new/en/natural-sciences/environment/ecological-sciences/biosphere-reserves/europe-north-america/czech-reppoland/krkonosekarkonosze/ (16 January 2022).

Wagner, F.H., Sanchez, A., Tarabalka, Y., Lotte, R.G., Ferreira, M.P., Aidar, M.P.M., Gloor, E., Phillips, O.L., Aragão, L.E.O.C., 2019. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. Remote Sensing in Ecology and Conservation 5, 360–375. https://doi.org/10.1002/rse2.111

Weinstein, B.G., Marconi, S., Bohlman, S., Zare, A., White, E., 2019. Individual Tree-Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural Networks. Remote Sensing 11, 1309. https://doi.org/10.3390/rs11111309

Zhao, Yuan, Song, Ding, Lin, Liang, Zhang, 2019. Use of Unmanned Aerial Vehicle Imagery and Deep Learning UNet to Extract Rice Lodging. Sensors 19, 3859. https://doi.org/10.3390/s19183859