Research Article
https://doi.org/10.33904/ejfe.1462335

Eur J Forest Eng 2024, 10(2):149-159

The Effect of Pre- and Post-Processing Techniques on Tree Detection in Young Forest Stands from Images of Snow Cover Using YOLO Neural Networks

Aleksey Portnov*, Andrey Shubin, Gulfina Frolova

Institute of Physico-Chemical and Biological Problems of Soil Science, Russian Academy of Sciences
142290 Pushchino, Russia.

Abstract

A neural network model for individual tree detection was developed based on the YOLOv4 architecture, which underwent additional preprocessing and postprocessing steps. The preprocessing step involved expanding the dataset by randomly cutting fragments from images, calculating anchor box sizes using the K-means clustering algorithm, and discarding anchor boxes that were too small a priori. The existing post-processing block of the YOLO architecture was modified by giving more weight to false positives in the error function and using the non-maximum suppression algorithm. Baseline neural networks from the YOLOv4 and YOLOv5 architectures, each in two versions (pre-trained and not pre-trained on the MS COCO dataset), were used for comparison without any additional modifications. In the overgrown experimental field, multi-season aerial copter surveys and ground counts were conducted on several sample plots to gather data. Comparison of multi-season aerial photographs with ground-count data showed that the best images in terms of the percentage of visually identifiable trees were those taken during the snowy season and when there was no foliage. Using these images and some additional images, we manually created a dataset on which we trained and tested neural network models. The model we developed showed significantly better results (2 to 10 times better) on the mAP 0.5 metric compared to the alternatives we considered.

Keywords: Individual tree detection, convolutional neural networks, YOLO, pre- and post-processing of data, aerial photography, young forest stands.

1. Introduction

Traditionally, direct ground-based measurement methods have been used for this purpose. However, these methods are associated with several logistical challenges (Li et al., 2023) and require significant labor (Condit, 1998; Gardner et al., 2008), as well as time costs (Luoma et al., 2017). Alternatively, remote sensing methods are actively used, such as satellite remote sensing (Bartalev et al., 2016; Aleksanin et al., 2019; Ershov et al., 2020; Schepaschenko et al., 2021; Zhang Y. et al., 2021) and unmanned aerial vehicle (UAV) remote sensing (Bennett et al., 2020; Neuville et al., 2021; Ivanova et al., 2021a; Ivanova et al., 2021b). Satellite data are often used for global and regional studies, as they are readily available. This saves scientists from having to create images independently. However, they have some limitations, such as a certain resolution and satellite revisit time (Al-Wassai and Kalyankar, 2013) and the presence of atmospheric interference (Kondratyev et al., 2013). The use of the PCA (radar aperture synthesis) method solves the problem of meteorological conditions, but it faces its own challenges. One of these challenges is the limited range-dependent sensitivity (Tanase et al., 2019). Another challenge is the decreasing sensitivity as we reach the "saturation" point (Joshi et al., 2017; Khati and Singh, 2022).

The use of unmanned aerial vehicles can significantly alleviate these limitations (Zhang et al., 2021). UAV data has a much higher resolution, up to 1 centimeter per pixel, which allows its use in the study of young forests on a local scale, including the development of methods for detecting individual trees.

The detection of individual trees within a forest provides opportunities for obtaining valuable data. Researchers from Japan (Fujimoto et al., 2019) are using single-tree detection technologies to develop a forestry decision-support system that accounts for forest carbon sequestration. Detection of individual crowns has been used in studies on biodiversity (Saarinen et al., 2018), forest health (Shendryk et al., 2016) and forest cover closure (Brūmelis et al., 2020). This method also finds its application in forest management (Koch et al., 2006; Berland et al., 2017), modeling ecosystem services

*Corresponding Author: Tel: +7 9674869437 E-mail: alekseyportnow@gmail.com
Received: 01 April 2024; Accepted: 06 June 2024

(Livesley et al., 2016), supply chain planning (Sparks et al., 2022) and reforestation projects (Pouliot et al., 2002).

Many studies that focus on tree detection currently use lidar data to generate dense point clouds (Wallace et al., 2016; Wu et al., 2016; Ayrey et al., 2017; Maschler et al., 2018). Other studies utilize high-resolution optical imagery to process images. Approaches based on the construction of dense point clouds using photogrammetry (Bohlin et al., 2012; Vastaranta et al., 2013; Rahlf et al., 2014; Holopainen et al., 2015; Puliti et al., 2017) as well as image processing algorithms (watershed algorithm, template matching) (Pollock, 1996a; Brandtberg and Walter, 1998; Pouliot et al., 2002) can be used. Algorithms based on machine learning (k-means algorithm, support vector machine, decision tree, random forest) (Malek et al., 2014; Chemura et al., 2015; Li et al., 2015), and deep learning algorithms (convolution neural network) (Secord and Zakhor, 2007; Cheang et al., 2017; Rizeei et al., 2018) have also been used. A number of studies have shown that the use of convolutional neural networks to solve the problem of tree detection yields the best results (Li et al., 2018; Sun et al., 2022).

One of the most popular solutions for object detection based on convolutional neural networks is the YOLO (You Only Look Once) model, which has been developed by Redmon et al. (Redmon et al., 2016) and has since been updated to various versions. YOLOv2 was released and outperformed all other neural networks in terms of accuracy and runtime efficiency (Redmon and Farhadi, 2017). YOLOv4 further improved upon YOLOv2 and outperformed other models such as SSD513, FPN+Faster R-CNN, and RetinaNet-101-800, both in terms of mAP-50 and execution speed (Bochkovsky et al., 2020), when tested on the MS COCO dataset (Lin et al., 2015). YOLOv5 maintains the high performance of YOLOv4 while also improving speed and memory efficiency (Nelson and Solawetz, 2024).

To improve the performance of convolutional neural networks, various image preprocessing and post-processing techniques are employed. One of the most used methods of data pre-processing is augmentation. This involves creating additional images from a given set of images. These new images can be created by taking fragments of existing images, rearranging parts of the images, rotating them, or modifying the color and brightness. Augmentation has been shown to significantly enhance the performance of neural networks by alleviating the limitations imposed by small and non-diverse datasets (Mumuni and Mumuni, 2022). Some of the post-processing techniques include non-maximum suppression (Song et al., 2019), morphological operations (Li et al., 2019), thresholding (Fang et al., 2023), and multi-scale feature extraction (Shen et al., 2022). Majority voting (Jhang, 2020) is also used in some cases. Various post-processing techniques can improve the performance of neural networks for detection, classification, and segmentation tasks (Salvi et al., 2021).

It should be noted that, in most cases, tree detection is performed on summer aerial images. The formed crowns of trees have different colors, textures, and geometric shapes, depending on the species. Using these differences, algorithms have been developed to detect and classify different tree species, mainly for biodiversity studies (Pollock, 1996b; Brandtberg and Walter, 1998; Pouliot et al., 2002). However, several detection and classification studies have utilized winter images, where the contrast between forest vegetation and the ground surface is enhanced due to the presence of snow. For instance, a study by Kuzmin et al. (Kuzmin et al., 2016) identified pine, spruce, and birch trees in high-resolution winter aerial imagery using a linear discriminant analysis of segments created through object-oriented image processing. Winter satellite images were used in (Baumann et al., 2012) to differentiate between forested and non-forest areas. In quantifying bamboo undergrowth in a mixed forest (Wang et al., 2009) the use of winter imagery achieved an accuracy of 89% in detecting it. However, the possibility of automatically detecting individual trees in young deciduous forests with snow cover remains unexplored.

In this study, we tested two hypotheses: (1) aerial images of young forests during the snowy season can be used to detect trees of all size classes, including those that form subgrade layers; (2) the accuracy of detecting individual trees in aerial images can be enhanced by adding pre- and post-processing steps to the neural network model.

For this purpose, we tested the detection of individual trees in winter-spring aerial images of young forests on fallow land using five neural networks. These included the basic YOLOv4 and YOLOv5 models, each in two variations, with and without pre-training on the MS-COCO dataset (Lin et al., 2015), as well as a developed model based on YOLOv4 that was modified during the stages of data pre- and post-processing.

2. Materials and Methods

The studies were conducted on the territory of the experimental field station of the Institute of Basic Biological Problems of the Russian Academy of Sciences. The coordinates of the study site center are 54.832778 N, 37.570278 E (Pushchino, Moscow Region, Russia) (Figure 1).

Figure 1. The location of the overgrown experimental field where the measurements were taken (Google image source)

The object is located in an area of broadleaf forests (Suslova, 2019). The process of soil formation occurs on the cover of loess-like rocks, which are represented by medium and heavy loams and light clays, with a predominance of the dusty fraction (Alifanov, 1995). The microrelief of the area is characterized by alternating furrows and linear hillocks. The soil is gray postagrogenic (*Luvisol (Aric)*). The re-establishment of forest vegetation occurred on the site, which previously had been used for potato gardens that were abandoned between 2003 and 2004. The forest stand is mainly composed of birch (*Betula* spp.) and common aspen (*Populus tremula* L.). There are also single clumps of goat willow (*Salix caprea* L.) and some individual common pine trees (*Pinus sylvestris* L.). The average height of the trees is around 10 meters. The average undergrowth is negligible, consisting of the same species. The spatial distribution of trees on the plot is non-uniform. Ribbon-shaped zones extending from north to south and irregularly-shaped glades that are not covered by forest or only covered by single trees, as well as open stands, can be distinguished. These structures may have been formed due to various anthropogenic influences, such as ribbon-shaped zones without trees that could have arisen at the site of former roads dividing cultivated areas, and glades and open stands due to spring grass fires (Khanina et al., 2018). It should also be noted that there is a close proximity to settlements, which determines the anthropogenic influence in the study area.

Within the study area, eight temporary sample plots (TSPs) measuring 5 x 5 meters were laid out in the spring of 2022. The TSPs were oriented on the sides of the world, taking into account magnetic declination, and identifying signs were placed at their corners. At each TSP, a count of trees was conducted. This included species identification, measuring the trunk diameter at breast height, and assessing the vital status of the trees.

The same study area was captured in its entirety using UAVs. The images were taken in March, July, October 2022, and January and March 2023, using DJI Phantom IV Pro and DJI Mavic 2 quadcopters and Pix4Dcapture (Pix4D, Pix4Dcapture) and Ctrl+DJI (Pix4D, Ctrl+DJI) applications, over an area of 200×200 m, from a height of 40 m. The overlap between the images is 95%, and the direction of the shot is towards the nadir. The resolution of the images is 5742×3648 px. The diverse structure of plant growth allowed us to capture aerial images with various stand densities and tree arrangements.

A comparison of various seasonal aerial photographs of TPSs was carried out by overlapping identification marks. Then, for each TPS, the centers of the trunk bases were manually counted in winter-spring images. The centers of the trunk base and crown, which were preserved with foliage, were counted in autumn images, and the visually distinguishable crowns of the trees were counted in summer images.

The preparation of the dataset for training the neural network included manual processing of a photo array. After aerial photography, photos of young growths were supplemented with images of stands in winter taken from the internet. During processing, trees were distinguished in the photos by marking their bounding boxes, with the trunk and clearly visible branches forming the tree's crown. We selected and prepared 102 images, 40 of which were from our survey, and 62 were from the internet. The total number of trees selected was 7710. It should be noted that we took images of very sparse stands from the internet in order to improve the quality of delineation of individual trees in more closed stands. Thus, most of the identified trees belong to our study area.

We have used the YOLOv4 architecture as a basis for the neural network we used. YOLO (You Only Look Once) is a family of neural networks that integrate a neural network framework with feature extraction, bounding box prediction, non-maximum suppression algorithm, and contextual reasoning (Redmon et al., 2016). We optimized the neural network for images with a resolution of 608 x 608 pixels. We performed following data preprocessing:

- The set of images was expanded by cutting random fragments from the original images using

the RandomCrop method from the albumentations package (Buslaev et al., 2020) for the Python language.

- Using the K-means clustering algorithm, we calculated the sizes of anchor boxes used in the model.
- Anchor boxes smaller than a certain size were deemed unsuitable, as they corresponded to shrubs and not trees.

Changes were also made to the data post-processing block. Due to the complex fractal structure of tree crowns, the neural network selected multiple bounding rectangles for each tree. To minimize this effect, we redesigned the neural network's loss function and non-maximum suppression algorithm.

- The loss function has been rewritten to give higher values to false positives.
- The following changes have been made to the Non-Maximum Suppression (NMS) algorithm: Bounding boxes with an area less than 80% of the area of the smallest anchor box have been removed; Bounding boxes that overlap significantly have been merged; Bounding boxes within other bounding boxes have been removed if their area is n times smaller than the outermost bounding box (n is a configurable hyperparameter). A genetic algorithm has been used to select the NMS hyperparameters.

We used the standard metric, mAP 0.5, which is the mean AP (average precision) value for all images. This is calculated using an IoU (intersection over union) threshold of 0.5.

$$AP = \frac{TP}{TP+FP} \quad (1)$$

TP stands for the number of true positive detections, and FP stands for the number of false positives. A true positive is defined as a detection for which the $IoU \geqslant 0.5$, *i*.e., the area of intersection of the bounding rectangle selected by the human and drawn by the neural network is greater than or equal to half of the area of their union.

To select the hyperparameters for the non-maximum suppression algorithm, we used a metric that is based on the relation $\frac{\sum IoU}{ercoef}$ where $\sum IoU$ - is the sum of all intersection-over-union values in the image. The error rate, *ercoef*, is calculated as follows:

$$ercoef = \begin{pmatrix} \frac{h}{a}, a \geqslant h \\ \frac{a}{h}, h > a \end{pmatrix} \quad (2)$$

where, h is the number of trees that a human expert identifies in the image and a is the number that a neural network identifies.

For comparison, we also used neural networks based on the YOLOv4 and YOLOv5 architectures, both in two versions: pre-trained on the MS-COCO (Lin et al., 2015) dataset and without any additional training. In these cases, we did not make any additional changes to the neural network architecture or perform any additional pre- or post-processing on the data.

Neural networks based on the YOLO architecture utilize a parameter called confidence threshold, which is a predefined level of probability that a tree will be found within the predicted bounding box. The model will only include this information in the final output if the confidence threshold is reached. When training the unmodified YOLOv4 and YOLOv5 neural networks, a standard confidence threshold value of 0.25 was used. During post-processing, the confidence threshold for our model was increased to 0.6 in order to suppress false positive detections. When testing the neural networks, confidence thresholds were selected based on visualization convenience.

3. Results

The following results were obtained by counting trunk and crown bases in different seasonal images. In winter/spring images (when the snow cover had not melted, but small thawing had already occurred around many trees), it was possible to identify a proportion of trees (between 35 and 97% depending on the image) based on the number of trees identified in the site during the tree inventory process (avg. 62%) (Figure 2). Fall aerial photographs showed a proportion of between 13 and 65% (avg. 20%), while summer aerial photographs identified a lower number of trees than images from other seasons, between 9 and 39% (avg. 20%). The correlation coefficients between the number of trees identified in the images and the number of trees during the enumeration were 0.13 for summer images, 0.26 for fall images, and 0.48 for winter-spring images. However, all of these correlations were not statistically significant ($p > 0.1$). In contrast, the correlation coefficient between the number of trees detected in the images and the number of counted trees with a DBH of 3 cm or more (diameter at breast height, approximately 1.3 m) was 0.51, 0.61, and 0.88 for the summer, fall, and spring/winter images, respectively (see Table 1). However, only the correlation between the spring/winter images and trees with a DBH of 3 cm or more was statistically significant ($p = 0.003686$).

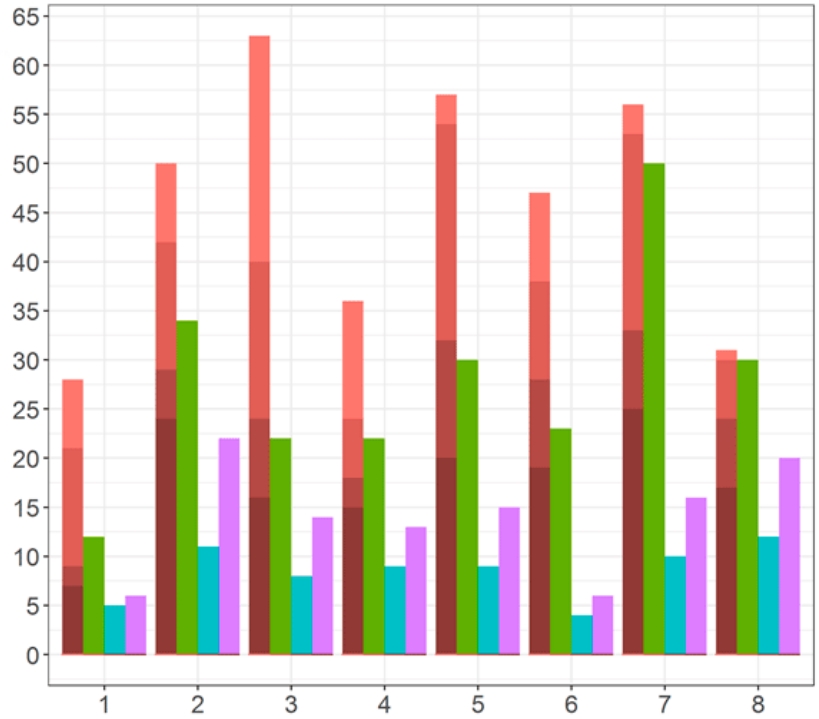

Figure 2. Comparison of the number of trees based on the enumeration results and those detected in different seasonal UAV images. On the X-axis are the temporary sample plots, and on the Y-axis is the number of detected trees. The red columns represents the enumeration data, the green columns represents data from winter-spring images, the blue columns represents data from summer images, and the purple columns represent data from fall images. The number of trees with different diameter at breast height (DBH) levels (0≤1cm, 1≤2cm, 2≤3cm, >3cm) based on the enumeration results are marked by different colors gradation

Table 1. Correlation coefficients between the number of trees detected in multi-seasonal images and the number of trees counted during ground surveys. A statistically significant correlation is highlighted in bold

	All accounted trees	Counted trees with DBH* > 3 cm
Fall	0.13	0.51
Summer	0.26	0.61
Winter/Spring	0.48	0.88

**DBH - diameter at breast height (~1.3 m)*

The fractions of trees identified from summer and autumn images are strongly correlated with each other (R = 0.96, p = 0.0001). However, the correlation with these fractions is smaller and not significant when using winter-spring images (R = 0.62, p = 0.10 in both cases). Therefore, we can conclude that the differences in the number of identified trees are primarily determined by season and secondarily by the stand structure at different locations. By training the neural networks until they achieved the best possible performance (with expert supervision to prevent overfitting), we obtained the mAP 0.5 values for shown in Table 2. Figures 3 and 4 show the results of processing one of the snapshots from the dataset using all five neural networks.

Table 2. The results of applying different neural networks to the same set of images

Model	mAP 0.5
YOLOv4	0.073
YOLOv4 (MS-COCO)	0.064
YOLOv5	0.251
YOLOv5 (MS-COCO)	0.335
Our model	0.606

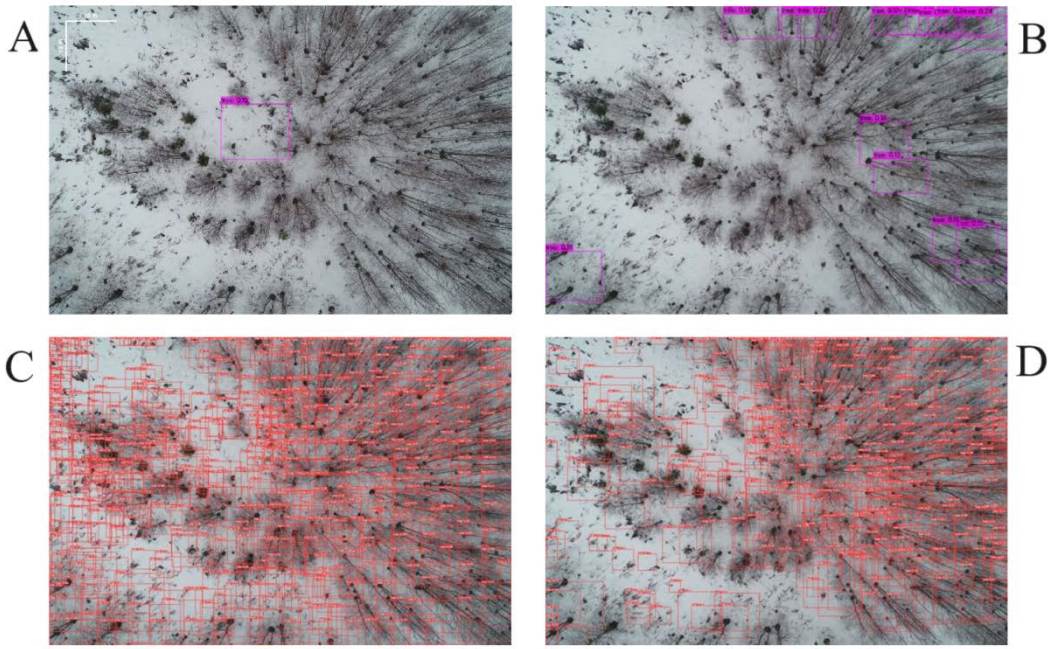

Figure 3. Tree detection results on the same image using different neural networks with a confidence threshold of 0.15 and pre-trained on our own dataset. (A) is the baseline YOLOv4 model, (B) is YOLOv4 pre-trained on the MS-COCO dataset, and (C) is the baseline YOLOv5 model. (D) is YOLOv5 pre-trained on the MS-COCO dataset.

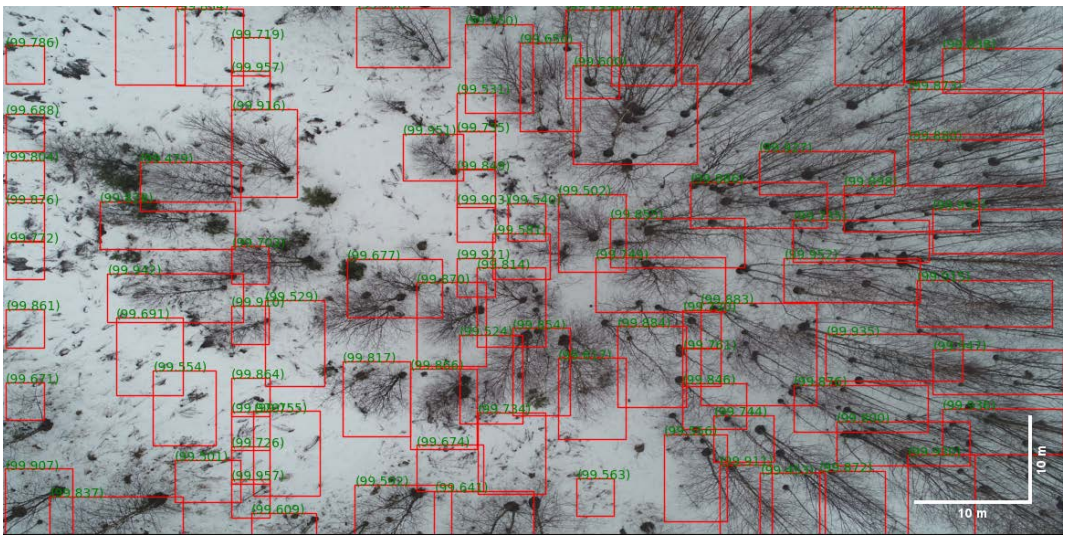

Figure 4. Tree detection results on the same image using our YOLOv4-based model with a confidence threshold of 0.9945

Based on the results above, we can see that YOLOv4, when used as-is and trained on our data set, tends to produce false negatives (missed existing trees), and YOLOv5, on the other hand, tends to generate false positives (detection of trees that do not exist). Both models also have very low confidence in their predictions (no trees were identified with a confidence level of 0.3 or higher). Our model, however, has managed to somewhat reduce the severity of these issues.

The data preprocessing block we added to YOLOv4 resulted in a significant reduction in the number of false negatives, but at the same time the number of false positives increased, bringing the results closer to those of YOLOv5. In turn, changes in the post-processing block reduced the severity of this problem.

Among its own issues, there is a slight decrease in accuracy on the densest images and a tendency for the bounding boxes to be over-fitted. We hypothesize that these issues can be corrected by a more careful selection of the hyperparameters.

During the research, we started to have doubts about whether mAP was the best metric for determining the efficiency of neural networks in detecting trees. We found that even the overfitted neural network on our data did not achieve mAP values greater than 0.69. We hypothesize that the problem lies in the fractal nature of deciduous tree structures in images during the leaf-less season. Due to this fractal property, one of the main sources of false positives is the neural network's identification of large tree branches as separate trees, leading to underestimation of accuracy. It has been shown in the (Padilla et al., 2021) that the performance evaluation of a convolutional neural network can depend significantly on the metric used and that some metrics

are better suited for certain classes of objects. Thus, for the task of tree detection (and possibly all objects with a fractal structure), it may be more efficient to use other metrics that more accurately reflect the progress achieved by the neural network.

Nevertheless, even with the mAP values, we can conclude that the YOLOv5 baseline neural network performs better than YOLOv4 for tree object detection. However, YOLOv4 with the modifications made to the pre- and post-processing data significantly outperforms the baseline YOLOv5, even. In earlier studies it was noted that YOLOv5 is not unambiguously better than YOLOv4, it depends on the type of task. In the (Kutyrev et al., 2023) it was found that YOLOv4 performed 10.2% better on the task of recognizing apple tree fruit on the tree crown.

It should be noted that pre-training the neural network on the MS-COCO dataset shows inconsistent results. This is consistent with the results of a study (Zoph et al., 2020) that found that pre-training has limited utility for object detection, which is even more severely reduced for highly augmented data.

Even with these considerations, the metric values achieved by the baseline neural networks seem to be extremely low, which can be explained by the size of the used dataset. Therefore, we can conclude that our proposed pre- and post-processing methods can help improve the performance of neural networks trained on smaller datasets.

We plan to further develop this study by estimating the biomass of young stands through the segmentation of detected tree trunks and the determination of their heights. Another potential direction for the study's development could involve investigating the applicability of additional metrics for tree detection, including in the context of young stands and images taken in winter.

4. Conclusion

Thus, we tested five different neural networks on a set of images from overgrown fields: the basic YOLOv4 and YOLOv5 networks, each with two variants (pre-trained on the MS-COCO dataset and not pre-trained), and our own modified YOLOv4 model, which underwent pre- and post-processing changes. These modifications partially solved the problems caused by the small size and homogeneity of the dataset and improved the performance of the neural network. Our model exhibited significantly fewer false negatives than unmodified YOLOv4 and fewer false positives than unmodified YOLOv5. In addition, the confidence of the model estimates was significantly improved. However, we believe that using the mAP 0.5 metric for tree detection may not be optimal. The use of images collected during the snow cover season opens new opportunities for forest ecology research. Specifically, automated measurement of key allometric tree characteristics detected by neural networks can be used to estimate the carbon storage capacity of the trees.

Acknowledgements

The authors would like to thank Elena Babenko, Lukyan Mirny, Aleksey Nikonov and Svetlana Urabova for their assistance in preparing the data set. We would also like to extend our gratitude to Anastasia Iovcheva, who provided valuable advice on soil classification, and Vladimir Shanin, Candidate of Science, for his expertise on the structure and design of this article.

This work was conducted within the framework of the state assignment of the Federal Research Centre for Scientific Biology of the Russian Academy of Sciences, registration number 122111000095-8.

Ethics Committee Approval: N/A.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept: A.P.; Design: A.S., A.P.; Supervision: A.S.; Resources: A.S., G.F.; Data Collection: G.F.; Analysis: A.P., G.F.; Literature Search: A.P.; Writing Manuscript: A.P.; Critical Review: G.F., A.S.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors declared that this study has received no financial support.

Cite this paper as: Portnov, A., Shubin, A., Frolova, G. 2024. The Effect of Pre- and Post-Processing Techniques on Tree Detection in Young Forest Stands from Images of Snow Cover Using YOLO Neural Networks, *European Journal of Forest Engineering*, 10(2):149-159.

References

Al-Wassai, F.A., Kalyankar, N.V. 2013. Major limitations of satellite images. *Journal of Global Research in Computer Science*, 4(5):51–59.

Aleksanin, A.I., Kim, V., Morozov, M.A., Fomin, E.V. 2019. Obnaruzhenie rubok otdel'nyh derev'ev po tenyam na osnove snimkov pribora "Geoton" sputnika "Resurs-P" [Detection of individual tree felling by shadows on the basis of images of Geoton instrument of Resurs-P satellite]. *Sovremennye problemy distancionnogo zondirovaniya Zemli iz kosmosa*, 16(5):174–182. [In Russian]. DOI:10.21046/2070-7401-2019-16-5-174-182.

Alifanov, V.M. 1995. Klyuchevoj uchastok "Pushchino". In: Paleokriogenez i sovremennoe pochvoobrazovanie [Pushchino key site. In: Paleocryogenesis and modern soil formation]. Izdatel'stvo ONTI Pushchinskogo nauchnogo centra Rossijskoj akademii nauk, Pushchino. pp. 95–156. [In Russian].

Ayrey, E., Fraver, S., Kershaw, Jr. J.A., Kenefic, L.S., Hayes, D., Weiskittel, A.R., Roth, B.E. 2017. Layer stacking: A novel algorithm for individual forest tree segmentation from LiDAR point clouds. *Canadian Journal of Remote Sensing*, 43(1):16–27. DOI:10.1080/07038992. 2017.1252907.

Bartalev, S., Egorov, V., Zharko, V., Loupian, E., Plotnikov, D., Khvostikov, S., Shabanov, N. 2016. Sputnikovoe kartografirovanie rastitel'nogo pokrova Rossii [Land cover mapping over Russia using Earth observation data]. IKI RAN, Moscow. 208 p. [in Russian].

Baumann, M., Ozdogan, M., Kuemmerle, T., Wendland, K.J., Esipova, E., Radeloff, V.C. 2012. Using the Landsat record to detect forest-cover changes during and after the collapse of the Soviet Union in the temperate zone of European Russia. *Remote Sensing of Environment*, 124:174–184. DOI:10.1016/j.rse. 2012.05.001.

Bennett, G., Hardy, A., Bunting, P., Morgan, P., Fricker, A. 2020. A transferable and effective method for monitoring continuous cover forestry at the individual tree level using UAVs. *Remote sensing*, 12(13):2115. DOI:10.3390/rs12132115.

Berland, A., Shiflett, S.A., Shuster, W.D., Garmestani, A.S., Goddard, H.C., Herrmann, D.L., Hopton, M.E. 2017. The role of trees in urban stormwater management. *Landscape and Urban Planning*, 162:167–177. DOI:10.1016/j.landurbplan.2017.02.017.

Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M. 2020. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934. DOI:10.48550/arXiv.2004.10934.

Bohlin, J., Wallerman, J., Fransson, J.E.S. 2012. Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM. *Scandinavian Journal of Forest Research*, 27(7):692–699. DOI:10.1080/02827581. 2012.686625.

Brandtberg, T., Walter, F. 1998. Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. *Machine Vision and Applications*, 11:64–73. DOI:10.1007/s001380050091.

Brūmelis, G., Dauškane, I., Elferts, D., Strode, L., Krama, T., Krams, I. 2020. Estimates of tree canopy closure and basal area as proxies for tree crown volume at a stand scale. *Forests*, 11(11):1180. DOI:10.3390/f11111180.

Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A. 2020. Albumentations: fast and flexible image augmentations. *Information*, 11(2):125. DOI: 10.3390/info11020125.

Cheang, E.K., Cheang, T.K., Tay, Y.H. 2017. Using convolutional neural networks to count palm trees in satellite images. arXiv preprint arXiv:1701.06462. DOI:10.48550/arXiv.1701.06462.

Chemura, A., van Duren, I., van Leeuwen, L.M. 2015. Determination of the age of oil palm from crown projection area detected from WorldView-2 multispectral remote sensing data: The case of Ejisu-Juaben district, Ghana. *ISPRS Journal of Photogrammetry and Remote Sensing*, 100:118–127. DOI:10.1016/j.isprsjprs.2014.07.013.

Condit, R. 1998. Total Cost for a Large Plot. In: Tropical Forest Census Plots: Methods and Results from Barro Colorado Island, Panama and a Comparison with Other Plots. Springer, Berlin. pp. 97–99.

Ershov, D.V., Gavrilyuk, E.A., Tikhonova, E.V., Braslavskaya, T.Y., Koroleva, N.V., Bavshin, I.M., Grummo, D.G., Sudnik, A.V., Titovets, A.V. 2020. Veroyatnostnaya ocenka raspredeleniya klyuchevyh biotopov v lesah nacional'nogo parka “Smolenskoe Poozer'e” na osnove sputnikovyh i topograficheskih dannyh [Stochastic Evaluation of Key Biotopes Allocation in the Forests of “Smolenskoe Poozer’ye” National Park Based on Satellite and Topographic Data]. *Lesovedenie* 1:17–34. [In Russian]. DOI:10.31857/S0024114820010052.

Fang, L., Zhao, J., Pan, Z., Li, Y. 2023. TPP: Deep learning based threshold post-processing multi-focus image fusion method. *Computers and Electrical Engineering*, 110:108736. DOI:10.1016/j.compeleceng.2023.108736

Fujimoto, A., Haga, C., Matsui, T., Machimura, T., Hayashi, K., Sugita, S., Takagi, H. 2019. An end to end process development for UAV-SfM based forest monitoring: individual tree detection, species classification and carbon dynamics simulation. *Forests*, 10(8):680. DOI:10.3390/ f10080680.

Gardner, T.A., Barlow, J., Araujo, I.S., Ávila-Pires, T.C., Bonaldo, A.B., Costa, J.E., Esposito, M.C., Ferreira, L.V., Hawes, J., Hernandez, M.I.M, Hoogmoed, M.S., Leite, R.N., Lo-Man-Hung, N.F., Malcolm, J.R., Martins, M.B., Mestre, L.A.M., Miranda-Santos, R., Overal, W.L., Parry, L., Peters, S.L., Ribeiro-Junior, M.A., Da Silva, M.N.F., Motta, C.D.S., Peres, C.A. 2008. The cost-effectiveness of biodiversity surveys in tropical forests. *Ecology Letters*, 11(2):139–150. DOI:10.1111/j.1461-0248.2007.01133.x

Heinrich, V.H.A., Dalagnol, R., Cassol, H.L., Rosan, T.M., de Almeida, C.T., Silva Jr., C.H., Campanharo, W.A., House, J.I., Sitch, S., Hales, T.C., Adami, M., Anderson, L.O., Aragão, L.E. 2021. Large carbon sink potential of secondary forests in the Brazilian Amazon to mitigate climate change. *Nature Communications*, 12(1):1785. DOI:10.1038/s41467-021-22050-1.

Holopainen, M., Vastaranta, M., Karjalainen, M., Karila, K., Kaasalainen, S., Honkavaara, E., Hyyppä, J. 2015. Forest inventory attribute estimation using airborne laser scanning, aerial stereo imagery, radargrammetry

and interferometry-finnish experiences of the 3d techniques. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2:63–69.DOI:10.5194/isprsannals-II-3-W4-63-2015.

Ivanova, N.V., Shashkov, M.P., Shanin, V.N. 2021a. Obtaining tree stand attributes from unmanned aerial vehicle (UAV) data: the case of mixed forests. *Vestnik Tomskogo gosudarstvennogo universiteta. Biologiya = Tomsk State University Journal of Biology,* 54:158–175. [In Russian, English Summary]. DOI:10.17223/19988591/54/8.

Ivanova, N.V., Shashkov, M.P., Shanin, V.N. 2021b. Study of pine forest stand structure in the Priosko-Terrasny State Nature Biosphere Reserve (Russia) based on aerial photography by quadrocopter. *Nature Conservation Research*, 6(4):1–14. [In Russian, English Summary]. DOI:10.24189/ncr.2021.042.

Jhang, K. 2020. Voting and ensemble schemes based on CNN models for photo-based gender prediction. *Journal of Information Processing Systems*, 16(4):809–819. DOI:10.3745/ JIPS.02.0137.

Joshi, N., Mitchard, E.T., Brolly, M., Schumacher, J., Fernández-Landa, A., Johannsen, V.K., Marchamalo, M., Fensholt, R. 2017. Understanding 'saturation' of radar signals over forests. *Scientific Reports*, 7(1):3505. DOI:10.1038/s41598-017-03469-3.

Khanina, L.G., Smirnov, V.E., Romanov, M.S., Bobrovsky, M.V. 2018. Effect of spring grass fires on vegetation patterns and soil quality in abandoned agricultural lands at local and landscape scales in Central European Russia. *Ecological Processes*, 7(1):1–19. DOI:10.1186/s13717-018-0150-8.

Khati, U., Singh, G. 2022. Combining L-band Synthetic Aperture Radar backscatter and TanDEM-X canopy height for forest aboveground biomass estimation. *Frontiers in Forests and Global Change* 5:918408. DOI:10.3389/ffgc.2022.918408.

Koch, B., Heyder, U., Weinacker, H. 2006. Detection of individual tree crowns in airborne lidar data. *Photogrammetric Engineering & Remote Sensing*, 72(4):357–363. DOI:10.14358/PERS.72.4.357.

Kondratyev, K.Y., Kozoderov, V.V., Smokty, O.I. 2013. Remote sensing of the Earth from space: atmospheric correction. Springer Science & Business Media.

Kutyrev, A.I., Smirnov, I.G., Andriyanov, N.A. 2023. Sravnitel'nyj analiz modelej nejronnyh setej dlya raspoznavaniya plodov yabloni na krone dereva [Neural network models of apple fruit identification in tree crowns: comparative analysis]. *Sadovodstvo i vinogradarstvo*, 5:56–63. [In Russian]. DOI:10.31676/0235-2591-2023-5-56-63.

Kuzmin, A., Korhonen, L., Manninen, T., Maltamo, M. 2016. Automatic segment-level tree species recognition using high resolution aerial winter imagery. *European Journal of Remote Sensing*, 49(1):239–259. DOI:10.5721/EuJRS20164914.

Larson, A.J., Lutz, J.A., Donato, D.C., Freund, J.A., Swanson, M.E., HilleRisLambers, J., Sprugel, D.G., Franklin, J.F. 2015. Spatial aspects of tree mortality differ strongly between young and old-growth forests. *Ecology*, 96(11):2855–2861. DOI:10.1890/15-0628.1.

Li, C.T., Chung, P.C., Tsai, H.W., Chow, N.H., Cheng, K.S. 2019. Inflammatory cells detection in H&E staining histology images using deep convolutional neural network with distance transformation. In: New Trends in Computer Technologies and Applications: 23rd International Computer Symposium, ICS 2018, Yunlin, Taiwan, December 20-22, 2018, Revised Selected Papers 23. Springer, Singapore. pp. 665–672. DOI:10.1007/978-981-13-9190-3_73.

Li, H., Chen, L., Yao, Z., Li, N., Long, L., Zhang, X. 2023. Intelligent Identification of Pine Wilt Disease Infected Individual Trees Using UAV-Based Hyperspectral Imagery. *Remote Sensing*, 15(13):3295. DOI:10.3390/rs15133295.

Li, L., Dong, J., Tenku, S.N., Xiao, X. 2015. Mapping oil palm plantations in Cameroon using PALSAR 50-m orthorectified mosaic images. *Remote Sensing*, 7(2):1206–1224. DOI:10.3390/rs70201206.

Li, W., Dong, R., Fu, H., Yu, L. 2018. Large-scale oil palm tree detection from high-resolution satellite images using two-stage convolutional neural networks. *Remote Sensing*, 11(1):11. DOI:10.3390/rs11010011.

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C.L., Dollár, P. 2015. Microsoft COCO: Common Objects in Context. arXiv preprint arXiv:1405.0312v3.DOI:10.48550/arXiv.1405.0312.

Livesley, S.J., McPherson, E.G., Calfapietra, C. 2016. The urban forest and ecosystem services: Impacts on urban water, heat, and pollution cycles at the tree, street, and city scale. *Journal of Environmental Quality*, 45(1):119-124. DOI:10.2134/jeq2015.11.0567.

Luoma, V., Saarinen, N., Wulder, M.A., White, J.C., Vastaranta, M., Holopainen, M., Hyyppä, J. 2017. Assessing precision in conventional field measurements of individual tree attributes. *Forests*, 8(2):38. DOI:10.3390/f8020038.

Malek, S., Bazi, Y., Alajlan, N., AlHichri, H., Melgani, F. 2014. Efficient framework for palm tree detection in UAV images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(12):4692–4703. DOI:10.1109/JSTARS.2014.2331425.

Maschler, J., Atzberger, C., Immitzer, M. 2018. Individual tree crown segmentation and classification of 13 tree species using airborne hyperspectral data. *Remote Sensing*, 10(8):1218. DOI:10.3390/ rs10081218.

Maslov, A., Gulbe, A., Gulbe, Y., Medvedeva, N., Sirin, A. 2016. Otsenka situatsii s zarastaniem selskohozyaystvennyh zemel lesnoy rastitelnostyu na primere Uglichskogo rayona Yaroslavskoy oblasti

[Assessment of the overgrowing of agricultural land by forest vegetation on the example of the Uglich district of the Yaroslavl region]. *Ustoychivoe losepolzovanie* 48:6–14 [in Russian].

Mumuni, A. and Mumuni, F. 2022. Data augmentation: A comprehensive survey of modern approaches. *Array,*16:100258. DOI:10.1016/j.array.2022.100258.

Nelson, J. and Solawetz, J. 2024. YOLOv5 is Here: State-of-the-Art Object Detection in 140 FPS. https://blog.roboflow.com/yolov5-is-here/(Accessed: 26 March 2024).

Neuville, R., Bates, J.S., Jonard, F. 2021. Estimating forest structure from UAV-mounted LiDAR point cloud using machine learning. *Remote Sensing,* 13(3):352. DOI:10.3390/rs13030352.

Padilla, R., Passos, W.L., Dias, T.L.B., Netto, S.L., da Silva, E.A.B. 2021. A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit. *Electronics,* 10(3):279. DOI:10.3390/electronics10030279.

Pix4D. Pix4Dcapture. URL: https://www.pix4d.com /product/pix4dcapture (Accessed: 26 March 2024).

Pix4D. Ctrl+DJI. URL: https://ctrl-dji.en. aptoide.com/ app (Accessed: 26 March 2024).

Pollock, R. 1996a. The Recognition Task. In: The automatic recognition of individual trees in aerial images of forests based on a synthetic tree crown image model. Diss. University of British Columbia. pp. 17–29. DOI:10.14288/1.0051597.

Pollock, R. 1996b. Background. In: The automatic recognition of individual trees in aerial images of forests based on a synthetic tree crown image model. pp. 5-17. Diss. University of British Columbia. DOI:10.14288/1.0051597.

Pouliot, D.A., King, D.J., Bell, F.W., Pitt, D.G. 2002. Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. *Remote Sensing of Environment,* 82(2-3):322–334. DOI:10.1016/S0034-4257(02)00050-0.

Pugh, T.A., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., Calle, L. 2019. Role of forest regrowth in global carbon sink dynamics. *Proceedings of the National Academy of Sciences* 116(10):4382–4387. DOI:10.1073/pnas.1810512116.

Puliti, S., Gobakken, T., Ørka, H.O., Næsset, E. 2017. Assessing 3D point clouds from aerial photographs for species-specific forest inventories. Scandin*avian Journal of Forest Research,* 32(1):68–79. DOI:10.1080/02827581.2016.1186727.

Rahlf, J., Breidenbach, J., Solberg, S., Næsset, E., Astrup, R. 2014. Comparison of four types of 3D data for timber volume estimation. *Remote Sensing of Environment,* 155:325–333. DOI:10.1016/j.rse.2014.08.036.

Rebane, S., Jõgiste, K., Kiviste, A., Stanturf, J.A., Metslaid, M. 2020. Patterns of carbon sequestration in a young forest ecosystem after clear-cutting. *Forests,* 11(2):126. DOI:10.3390/f11020126.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A. 2016. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). pp. 779–788. DOI:10.1109/CVPR.2016.91.

Redmon J. and Farhadi A. 2017. YOLO9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). pp. 7263–7271. DOI:10.1109/CVPR.2017.690.

Rizeei, H.M., Shafri, H.Z., Mohamoud, M.A., Pradhan, B., Kalantar, B. 2018. Oil palm counting and age estimation from WorldView-3 imagery and LiDAR data using an integrated OBIA height model and regression analysis. *Journal of Sensors,* 2018:1–13. DOI:10.1155/2018/2536327.

Saarinen, N., Vastaranta, M., Näsi, R., Rosnell, T., Hakala, T., Honkavaara, E., Wulder, M.A., Luoma, V., Tommaselli, A.M.G, Imai, N.N., Ribeiro, E.A.W., Guimarães, R.B., Holopainen, M., Hyyppä, J. 2018. Assessing biodiversity in boreal forests with UAV-based photogrammetric point clouds and hyperspectral imaging. *Remote Sensing,* 10(2):338. DOI:10.3390/rs10020338.

Salvi, M., Acharya, U.R., Molinari, F., Meiburger, K.M. 2021. The impact of pre-and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine,* 128:104129. DOI:10.1016/j.compbiomed.2020.104129.

Schepaschenko, D., Moltchanova, E., Fedorov, S., Karminov, V., Ontikov, P., Santoro, M., See, L., Kositsyn, V., Shvidenko, A., Romanovskaya, A., Korotkov, V., Lesiv, M., Bartalev, S., Fritz, S., Shchepashchenko, M., Kraxner, F. 2021. Russian forest sequesters substantially more carbon than previously reported. *Scientific Reports* 11(1):12825. DOI:10.1038/s41598-021-92152-9.

Secord, J., Zakhor, A. 2007. Tree detection in urban regions using aerial lidar and image data. IEEE *Geoscience and Remote Sensing Letters,* 4(2):196–200. DOI:10.1109/LGRS.2006.888107.

Shen, X., Weng, L., Xia, M., Lin, H. 2022. Multi-scale feature aggregation network for semantic segmentation of land cover. *Remote Sensing* 14(23):6156. DOI:10.3390/rs14236156.

Shendryk, I., Broich, M., Tulbure, M.G., McGrath, A., Keith, D., Alexandrov, S.V. 2016. Mapping individual tree health using full-waveform airborne laser scans and imaging spectroscopy: A case study for a floodplain eucalypt forest. *Remote Sensing of Environment,* 187:202–217. DOI:10.1016/j.rse.2016.10.014.

Soloviev, S.V., Miller, G.F., Bezborodova, A.N., Filimonova, D.A. 2018. Sukcessiya na molodyh i

srednevozrastnyh zalezhah lesostepnoj zony Zapadnoj Sibiri v predelah Novosibirskoj oblasti [Succession on young and medium age fallows of the forest-steppe zone of Western Siberia in the Novosibirsk region]. *Mezhdunarodnyj zhurnal prikladnyh i fundamental'nyh issledovanij*, 10:116120. [In Russian]. DOI:10.17513/mjpfi.12427.

Song, Y., Pan, Q.K., Gao, L., Zhang, B. 2019. Improved non-maximum suppression for object detection using harmony search algorithm. *Applied Soft Computing*, 81:105478. DOI:10.1016/j.asoc.2019.05.005.

Sparks, A.M., Corrao, M.V., Smith, A.M.S. 2022. Cross-Comparison of Individual Tree Detection Methods Using Low and High Pulse Density Airborne Laser Scanning Data. *Remote Sensing*, 14(14):3480. DOI:10.3390/rs14143480.

Sun, C., Huang, C., Zhang, H., Chen, B., An, F., Wang, L., Yun, T. 2022. Individual tree crown segmentation and crown width extraction from a heightmap derived from aerial laser scanning data using a deep learning framework. *Frontiers in Plant Science,* 13:914974. DOI:10.3389/fpls.2022.914974.

Suslova, E.G. 2019. Lesa Moskovskoj oblasti [Forests of the Moscow region]. Ekosistemy: ekologiya i dinamika 3(1):119-190. [In Russian] https://ecosystemsdynamic.ru/lesa-moskovskoj-oblasti/ (Accessed 26 March 2024).

Tanase, M.A., Villard, L., Pitar, D., Apostol, B., Petrila, M., Chivulescu, S., Leca, S., Borlaf-Mena, I., Pascu, I.-S., Dobre, A.-C., Pitar, D., Guiman, G., Lorent, A., Anghelus, C., Ciceu, A., Nedea, G., Stanculeanu, R., Popescu, F., Aponte, C., Badea, O. 2019. Synthetic aperture radar sensitivity to forest changes: A simulations-based study for the Romanian forests. *Science of the Total Environment,* 689:1104–1114. DOI:10.1016/j.scitotenv.2019.06.494.

United Nations. 1992. UN Framework Convention on Climate Change. https://unfccc.int/resource/docs/convkp/conveng.pdf (Accessed: 26 March 2024).

United Nations. 1998. Kyoto Protocol to the United Nations Framework Convention on Climate Change. https://unfccc.int/resource/docs/convkp/kpeng.pdf (Accessed: 26 March 2024).

Vastaranta, M., Wulder, M.A., White, J.C., Pekkarinen, A., Tuominen, S., Ginzler, C., Kankare, V., Holopainen, M., Hyyppä, J., Hyyppä, H. 2013. Airborne laser scanning and digital stereo imagery measures of forest structure: Comparative results and implications for forest mapping and inventory update. *Canadian Journal of Remote Sensing,* 39(5):382–395. DOI:10.5589/m13-046.

Wallace, L., Lucieer, A., Malenovský, Z., Turner, D., Vopěnka, P. 2016. Assessment of forest structure using two UAV techniques: A comparison of airborne laser scanning and structure from motion (SfM) point clouds. *Forests*, 7(3):62. DOI:10.3390/f7030062.

Wang, T., Skidmore, A.K., Toxopeus, A.G., Liu, X. 2009. Understory bamboo discrimination using a winter image. *Photogrammetric Engineering & Remote Sensing,* 75(1):37–47. DOI: 10.14358/PERS.75.1.37.

Wu, B., Yu, B., Wu, Q., Huang, Y., Chen, Z., Wu, J. 2016. Individual tree crown delineation using localized contour tree method and airborne LiDAR data in coniferous forests. *International Journal of Applied Earth Observation and Geoinformation,* 52:82–94. DOI:10.1016/j.jag.2016.06.003.

Zamolodchikov, D. 2011. Sistemy ocenki i prognoza zapasov ugleroda v lesnyh ekosistemah [Systems of assessment and forecasting of carbon stocks in forest ecosystems]. *Ustojchivoe lesopolzovanie*, 29:15–22. [In Russian]. https://www.elibrary.ru/download/elibrary_ 22517667_72914254.pdf (Accessed: 26 March 2024).

Zhang, X., Tan, X., Chen, G., Zhu, K., Liao, P., Wang, T. 2021. Object-based classification framework of remote sensing images with graph convolutional networks. *IEEE Geoscience and Remote Sensing Letters* 19:1–5. DOI:10.1109/LGRS.2021.3072627.

Zhang, Y., Dian, Y., Zhou, J., Peng, S., Hu, Y., Hu, L., Han, Z., Fang, X., Cui, H. 2021. Characterizing spatial patterns of pine wood nematode outbreaks in subtropical zone in China. *Remote Sensing*, 13(22):4682. DOI:10.3390/rs13224682.

Zoph, B., Ghiasi, G., Lin, T. Y., Cui, Y., Liu, H., Cubuk, E. D., Le, Q. 2020. Rethinking pre-training and self-training. *Advances in Neural Information Processing Systems,* 33:3833–3845. https://proceedings.neurips.cc/paper/2020/hash/27e9661e033a73a6ad8cefcde965c54dAbstract. html (Accessed: 25 June 2024).