DEEP LEARNING BASED AERIAL IMAGERY CLASSIFICATION FOR TREE SPECIES IDENTIFICATION

O. C. Bayrak [1] *, F. Erdem [2], M. Uzar [1]

[1] Dept. of Geomatics Engineering, Faculty of Civil Engineering, 34220 Esenler, Istanbul, Turkey - (onurcb, auzar) @yildiz.edu.tr
[2] Institude of Earth and Space Sciences, Eskisehir Technical University, 26555 Tepebasi, Eskisehir, Turkey – firaterdem@ogr.eskisehir.edu.tr

KEY WORDS: Deep Learning, Image Classification, Image Processing, Aerial Imagery, Tree Species

ABSTRACT: Forest monitoring and tree species categorization has a vital importance in terms of biodiversity conservation, ecosystem health assessment, climate change mitigation, and sustainable resource management. Due to large-scale coverage of forest areas, remote sensing technology plays a crucial role in the monitoring of forest areas by timely and regular data acquisition, multi-spectral and multi-temporal analysis, non-invasive data collection, accessibility and cost-effectiveness. High-resolution satellite and airborne remote sensing technologies have supplied image data with rich spatial, color, and texture information. Nowadays, deep learning models are commonly utilized in image classification, object recognition, and semantic segmentation applications in remote sensing and forest monitoring as well. We, in this study, selected a popular CNN and object detection algorithm YOLOv8 variants for tree species classification from aerial images of TreeSatAI benchmark. Our results showed that YOLOv8-l outperformed benchmark's initial release results, and other YOLOv8 variants with 71,55% and 72,70% for weighted and micro averaging scores, respectively.

1. INTRODUCTION

Tree species categorization and increasing stock volume calculation are not only critical forest management procedures, but they may also be viewed as central responsibilities of forest research in general. The categorization of tree species and the assessment of expanding stock volume are critical components in the development of forest resource surveys. They can also be utilized as crucial markers for assessing the potential of forests to sequester carbon. The acquisition of rising stock volume and the timely and correct identification of tree species can represent the total scale and level of forest resources in a country or area. They are also useful for estimating the amount of forest resources and assessing the regional ecological environment.

Remote sensing technology has long been seen to be a viable answer to the challenge of tree species categorization. Remote sensing data from various platforms and sensors are widely employed in this field of study (Wang et al., 2021). This technology have placed a significant role for the monitoring of forests (Holzwarth et al., 2020) and minimize fieldwork, which is time-consuming and expensive. The advent of high-resolution satellite and airborne remote sensing technologies have supplied image data with rich spatial, color, and texture information, resulting in enhanced fine categorization of tree species (Dian et al., 2015). Because of their different structures and morphologies, different tree species have varied spectral information (Deepak et al., 2019). Even under the same environmental conditions, trees at different development stages or health states might have differing spectral information (Brilli et al., 2013).

Deep learning refers to deep neural networks that were created by combining many layers with certain configurations in a specific order (LeCun et al., 2015). Convolutional Neural Networks (CNN), which automatically extract features using convolutional filters, are one of the most extensively used deep learning models. CNNs use convolutional filters of varying sizes to automatically extract differentiating properties. CNNs and other deep learning models are commonly employed in image classification, object recognition, and semantic segmentation applications in remote sensing (Zhu et al., 2017; Ma et al., 2019).

Deep neural networks-based algorithms are commonly used for classifying tree species from multispectral images Martins et al., 2021; Zhang et al., 2021). CNNs have been widely used for tree detection and classification tasks such as tree species discrimination (Fricker et al., 2019; Carpentier et al., 2018; Liu et al., 2019; Pelletier et al., 2019), forest damage detection (Hamdi et al., 2019; Safonova et al., 2019), and tree mortality mapping (Sylvain et al., 2019). In addition to CNN, deep Boltzmann machines (Guan et al., 2015) and deep belief networks (Zou et al., 2017) were used to recognize tree species. The tedious procedure of producing training sample labels is a key restriction on training a CNN architecture for image classification (Hamedianfar, et al., 2022).

Recently, TreeSatAI benchmark dataset (Ahlswede et al., 2023) was published with aerial, Sentinel-1 and Sentinel-2 imagery for identification of dominant tree species in a specific region. By taking into account the constraints in satellite imagery such as temporal and geometric resolution issues, we considered only aerial image patches of TreeSatAI for fast response time in real-world applications. We, in this study, selected a popular CNN and object detection algorithm YOLOv8 variants for improving tree species classification performance. Our research questions examine which of (i) three different band combination is most suitable for the classification, and (ii) YOLOv8 architecture differs from others in terms of model size and accuracy.

2. MATERIAL AND METHODOLOGY

Prior to training YOLOv8 variants, different band combinations of TreeSatAI benchmark's aerial images were stacked to specify the most accurate combinations. The images were resized to 288 x 288 pixels, and the model performances were quantitatively evaluated using Precision, Recall, F1-Score, and mAP metrics (Figure 1).

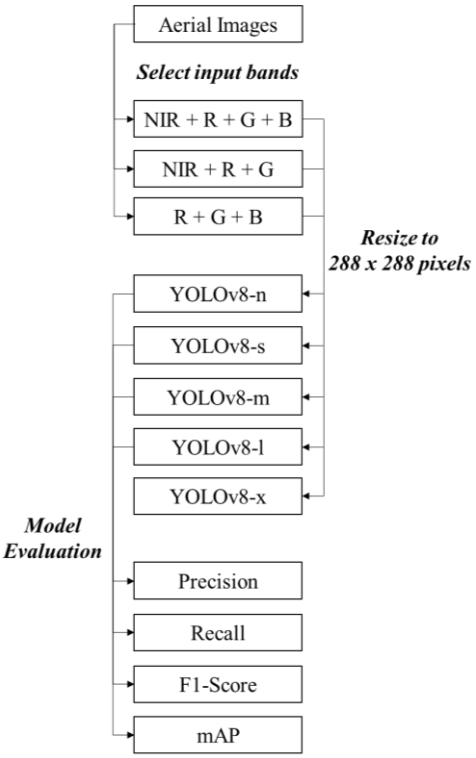

Figure 1. Flowchart of the study

2.1 TreeSatAI Dataset

TreeSatAI benchmark includes 12 age, 15 genus, and 20 tree species classes, with 50.381 (45337 train ,5044 test) synchronized image triplets of aerial, Sentinel-1, and Sentinel-2. Near-Infrared (NIR), Red (R), Green (G), and Blue (B) bands are available for aerial imagery with 0.2 m spatial resolution and these image patches were 304 x 304 pixels (Ahlswede et al. 2023). Figure A1 illustrates sample aerial images from the dataset.

Only aerial image patches were used and these patches resized to 288 x 288 pixels for training and testing. Three band combinations, R + G + B, NIR + R + G, and NIR + R + G + B (All) were included for train/test stage, to specify the outperforming band combinations. Similar to Ahlswede et al. (2023), we predicted genus classes for the following 15 types of trees: Abies, Acer, Alnus, Betula, Cleared, Fagus, Fraxinus, Larix, Picea, Pinus, Populus, Prunus, Pseudotsuga, Quercus and Tilia, respectively (Figure 3).

2.2 Image Classification with YOLOv8

YOLOv8 (Jocher et al., 2023), is an enhanced version of YOLO architectures. It builds upon the concepts of YOLOv5 (Jocher, 2020), such as the CSP idea from Wang et al. (2020), the feature fusion method (PAN-FPN) from Lin et al. (2017) and Liu et al. (2018), and the SPPF module. YOLOv8 also introduces models of different scales based on scaling coefficients, similar to YOLOv5, to address the requirements of various tasks. While retaining the core principles of YOLOv5, YOLOv8 incorporates the C2f module inspired by the ELAN structure in YOLOv7 (Wang et al., 2022). The majority of other components remain based on the original ideas of YOLOv5.

For classification in YOLOv8, the BCE loss is employed. The regression loss takes the form of CIOU loss + DFL, and VFL introduces an asymmetric weighting operation (Cao et al., 2020). YOLOv8 demonstrates greater extensibility compared to previous YOLO algorithms. It serves as a framework that supports previous YOLO versions and allows easy switching between them, facilitating performance comparisons.

The standout feature of YOLOv8 lies in its extensibility, enabling seamless compatibility with all YOLO versions and effortless comparison of their performance. This advantage greatly benefits researchers engaged in YOLO projects, making YOLOv8 the chosen baseline version. The backbone of YOLOv8 closely resembles that of YOLOv5, with the C3 module replaced by the C2f module based on the CSP concept (Figure 2). The C2f module draws inspiration from the ELAN idea in YOLOv7, combining C3 and ELAN to form an information-rich yet lightweight C2f module (Wang et al., 2022). At the end of the backbone, the popular SPPF module is retained, employing three sequential max-pooling of size 5 × 5 and concatenating the resulting layers. This approach ensures accurate object detection across various scales while maintaining a lightweight design. In the neck module, YOLOv8 continues to utilize the PAN-FPN feature fusion method, which enhances the fusion and utilization of feature layer information at different scales. The neck module of YOLOv8 comprises two upsampling operations, multiple C2f modules, and a final decoupled head structure as proposed by the authors.

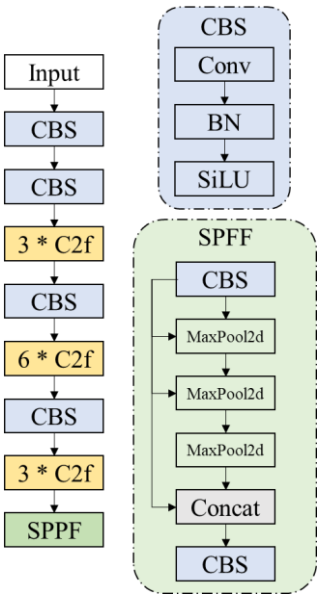

Figure 2. Backbone architecture of YOLOv8

YOLOv8-n, YOLOv8-s, YOLOv8-m, YOLOv8-l, and YOLOv8-x architectures were trained for a total of 20 epochs with Stochastic Gradient Descent (SGD) optimizer, 0.01 learning rate, batch size of 32, 0.937 momentum and 0.0005 weight decay on a RTX 2080 graphics card (Table 1).

Hyperparameter	Setting
Epochs	20
Batch Size	32
Optimizer	SGD
Learning Rate	0.01
Momentum	0.937
Weight Decay	0.0005
Input Size	288 x 288

Table 1. Hyperparameter settings for training

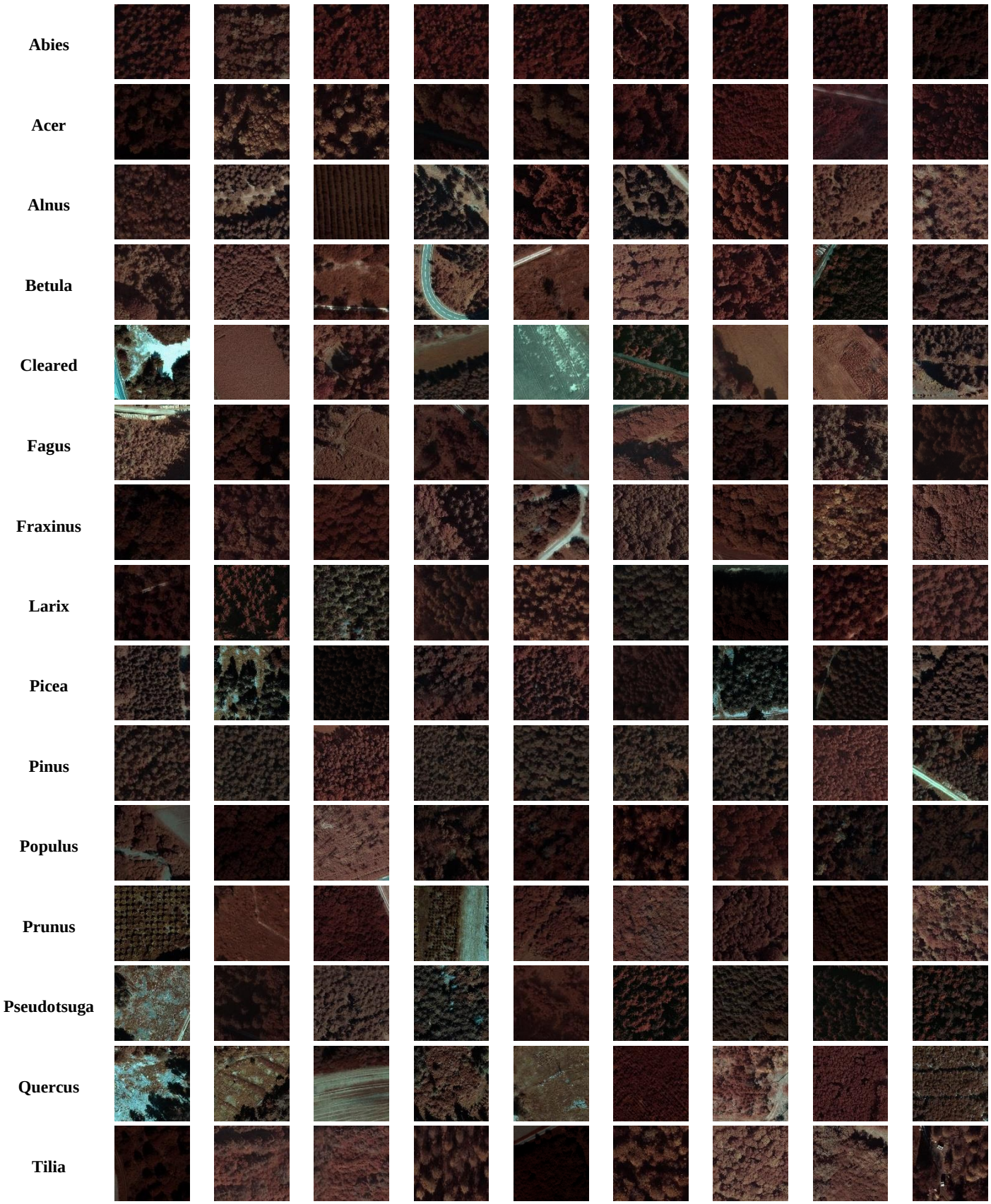

Figure 3. Sample aerial images of tree species from TreeSatAI benchmark

2.3 Accuracy Assesment

Model evaluation stage was performed with Precision, Recall, F1-Score (Goutte and Gaussier, 2005) and mAP (Mean Average Precision) (Everingham et al., 2010) metrics, respectively. Similar to Ahlswede et al. (2023), weighted and micro averaging were calculated due to model performance evaluation with class frequencies and global average, respectively.

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

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$F1-Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3)$$

$$mAP = \frac{1}{N}\sum_{k=1}^{k=N} AP_k \quad (4)$$

where TP, FP, TN, FN, N an AP represent the total numbers of true positives, false positives, true negatives, false negatives, number of classes, and average precision, respectively.

3. RESULTS AND DISCUSSIONS

This section covers the findings and interpretation of results. In this study, six different versions of YOLOv8 architecture were created for tree species classification from aerial imagery. As seen in Table 2 and 3, highest scores are highlighted bold. YOLOv8-l with the incorporation of NIR, R and G bands have the highest F1-Score (71,55% and 72,70% for weighted and micro averaging). Similarly, when all bands were included, the second highest F1-Score (71,11%) was obtained by YOLOv8-l. It should be noted that the results of this study are more accurate (71,54% F1-Score weighted averaging and 71,66% F1-Score weighted micro) than Ahlswede et al. (2023) where aerial imagery, S1 and S2 data were included.

Model	Average	Inputs		
	Weighted	R+G+B	NIR+R+G	All
YOLOv8-x	Precision	70,48	71,44	71,55
	Recall	71,98	72,26	**72,30**
	F1-Score	70,64	71,23	71,02
	mAP	78,16	**78,49**	**78,29**
YOLOv8-l	Precision	70,66	**71,49**	**71,77**
	Recall	71,71	**72,70**	72,20
	F1-Score	70,47	**71,55**	**71,11**
	mAP	77,79	78,33	78,26
YOLOv8-m	Precision	**71,65**	68,38	70,18
	Recall	**72,44**	69,34	71,27
	F1-Score	**71,37**	67,76	70,07
	mAP	**78,51**	75,33	77,53
YOLOv8-s	Precision	68,88	69,56	70,03
	Recall	70,18	70,71	70,97
	F1-Score	68,68	69,57	69,79
	mAP	76,55	76,8	77,27
YOLOv8-n	Precision	70,13	68,38	69,08
	Recall	71,19	69,34	69,98
	F1-Score	69,9	67,76	68,64
	mAP	77,43	75,33	75,7

Table 2. Different YOLOV8 models with weighted averaging results on different band combinations

Model	Average	Inputs		
	Micro	R+G+B	NIR+R+G	All
YOLOv8-x	Precision	71,98	72,26	**72,30**
	Recall	**71,98**	72,26	**72,30**
	F1-Score	71,98	72,26	**72,30**
	mAP	80,74	81,06	80,95
YOLOv8-l	Precision	71,70	**72,70**	72,20
	Recall	71,70	**72,70**	72,20
	F1-Score	71,70	**72,70**	72,20
	mAP	80,24	**81,13**	**81,05**
YOLOv8-m	Precision	**72,44**	69,34	71,27
	Recall	72,44	69,34	71,27
	F1-Score	**72,44**	69,34	71,27
	mAP	**81,03**	77,92	80,28
YOLOv8-s	Precision	70,18	70,71	70,97
	Recall	70,18	70,71	70,97
	F1-Score	70,18	70,71	70,97
	mAP	79,08	79,46	79,82
YOLOv8-n	Precision	71,19	69,34	69,98
	Recall	71,19	69,34	69,98
	F1-Score	71,19	69,34	69,98
	mAP	80,00	77,92	78,61

Table 3. Different YOLOV8 models with micro averaging results on different band combinations

In class-base performances (Table 4), Abies (84,52% F1-Score and 86,06 % mAP), Cleared (80,55% F1-Score and 88,97% mAP), Fagus (70,77% F1-Score and 79,62% mAP), Larix (81,13% F1-Score and 89,66% mAP), Picea (82,73% F1-Score and 88,61% mAP), Pinus (88,5% F1-Score and 95,79% mAP), Pseudotsuga (82,99% F1-Score and 89,71% mAP), and Quercus (74,34% F1-Score and 83,62% mAP) are higher than Ahlswede et al. (2023).

Species	Precision	Recall	F1-Score	mAP
Abies	**84,52**	**76,34**	**80,23**	**86,06**
Acer	53,74	28,01	36,83	**42,8**
Alnus	61,67	**56,68**	59,07	**61,37**
Betula	56,79	**49,11**	52,67	**56,83**
Cleared	77,46	**83,89**	**80,55**	**88,97**
Fagus	67,28	**74,65**	**70,77**	**79,62**
Fraxinus	43,65	43,14	43,39	**43,04**
Larix	**85,07**	**77,53**	**81,13**	**89,66**
Picea	77,98	**88,09**	**82,73**	**88,61**
Pinus	86,34	**90,78**	**88,5**	**95,79**
Populus	53,13	20,99	30,09	36,28
Prunus	38,46	17,24	23,81	20,99
Pseudotsuga	**86,38**	**79,86**	**82,99**	**89,71**
Quercus	69,25	**80,23**	**74,34**	**83,62**
Tilia	0	0	0	0,0

Table 4. Class-wise YOLOv8-l results with NIR + R + G band combination.

However, prediction results for imbalanced classes such as Prunus, Populus and Tillia were poor. As stated in Ahlswede et al. (2023) the incorporation of both deep neural networks and multi-layer perceptron produced more accurate results than using only deep neural networks. Implementing such approaches into YOLOv8 architectures could improve the results. Another major finding was absence of correlation between model depth and model accuracy, since for R+G+B bands, YOLOv8-m obtained the best performance whereas YOLOv8-l was the most succesful for NIR + R + G and All except micro averaging scores in Table 3.

As illustrated in Figure 4, number of true positives for Abies, Alnus, Betula, Cleare, Fagus, Fraxinus, Larix, Picea, Pinus, Pseudotsuga and Quercus classes are relatively higher than underrepresented classes Populus, Prunus, and Tilia. It is thought that the small number of training and test data causes this situation. Especially for Tilia class, no correct predictions could be performed. In order to solve this problem, it is foreseen to use machine learning classifiers as previous study carried out by Ahlswede et al. (2023).

		Predicted														
	Class	Abies	Acer	Alnus	Betula	Cleared	Fagus	Fraxinus	Larix	Picea	Pinus	Populus	Prunus	Pseudotsuga	Quercus	Tilia
Reference	Abies	71	0	1	1	0	5	0	2	2	5	0	0	6	0	0
	Acer	1	79	8	8	11	59	43	1	11	4	1	0	0	56	0
	Alnus	1	3	140	18	7	14	14	4	15	4	4	2	0	21	0
	Betula	0	3	19	138	15	20	6	6	15	20	4	0	1	34	0
	Cleared	0	2	8	10	354	14	2	0	9	4	0	2	0	17	0
	Fagus	0	17	8	18	10	477	21	2	20	9	0	3	2	52	0
	Fraxinus	0	22	9	2	2	43	110	2	3	4	4	0	0	54	0
	Larix	4	0	3	2	3	6	3	245	25	17	0	0	6	2	0
	Picea	2	4	5	6	13	16	0	5	503	6	0	0	7	4	0
	Pinus	2	2	2	17	3	6	2	9	9	689	0	0	9	9	0
	Populus	0	2	11	5	3	7	19	1	0	1	17	1	0	14	0
	Prunus	0	1	2	0	3	6	7	1	1	0	1	5	0	2	0
	Pseudotsuga	1	0	0	1	4	4	0	8	22	13	0	0	222	3	0
	Quercus	2	11	5	16	29	29	22	2	10	20	1	0	4	617	1
	Tilia	0	1	6	1	0	3	3	0	0	2	0	0	0	6	0

Figure 4. Confusion matrix of YOLOv8-l prediction results (Green: True Positives, Red: False Positives and False Negatives)

4. CONCLUSION

In this study, we utilized a popular object detection algorithm YOLOv8, for tree species classification from aerial imagery. Our results reveal the advantages of object detection algorithms against commonly used networks for image classification tasks. Although prediction results were improved in general with YOLOv8, predictions on imbalanced classes were performed poorly. According to our results, no correlation was found between model depth and accuracy. YOLOv8-l, which is the second deepest model, achieved a higher classification score than other variants. This work can be extended using state-of-art and advanced networks such as Transformer, etc. In addition, it is thought that the results can be improved by including different machine learning classifiers into the classification part of utilized networks. It should be addressed that by including satellite images as well as aerial imagery, texture information can be enriched and more accurate results can be obtained.

ACKNOWLEDGEMENTS

The authors would like to thank The National Aeronautics and Space Administration (NASA), for covering the costs of the symposium under the "Early Career Graduate" programme.

REFERENCES

Ahlswede, S., Schulz, C., Gava, C., Helber, P., Bischke, B., Förster, M., Arias, F., Hees, J., Demir, B., Kleinschmit, B., 2023: TreeSatAI Benchmark Archive: A multi-sensor, multi-label dataset for tree species classification in remote sensing. Earth System Science Data, 15(2), pp.681-695. doi.org/10.5194/essd-15-681-2023.

Brilli, L., Chiesi, M., Maselli, F., Moriondo, M., Gioli, B., Toscano, P., Zaldei, A., Bindi, M., 2013. Simulation of olive grove gross primary production by the combination of ground and multi-sensor satellite data. International journal of applied earth observation and geoinformation, 23, pp.29-36. doi.org/10.1016/j.jag.2012.11.006.

Cao, Y., Chen, K., Loy, C.C., Lin, D., 2020. Prime sample attention in object detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11583-11591).

Carpentier, M., Giguere, P., Gaudreault, J., 2018, October. Tree species identification from bark images using convolutional neural networks. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 1075-1081). IEEE. doi.org10.1109/IROS.2018.8593514.

Deepak, M., Keski-Saari, S., Fauch, L., Granlund, L., Oksanen, E., Keinänen, M., 2019. Leaf canopy layers affect spectral reflectance in silver birch. Remote Sensing, 11(24), p.2884. doi.org/10.3390/rs11242884.

Dian, Y., Li, Z., Pang, Y., 2015. Spectral and texture features combined for forest tree species classification with airborne hyperspectral imagery. Journal of the Indian Society of Remote Sensing, 43, pp.101-107. doi.org/10.1007/s12524-014-0392-6.

Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A., 2010: The pascal visual object classes (voc) challenge. International journal of computer vision, 88, pp.303-338. doi.org/10.1007/s11263-009-0275-4.

Fricker, G.A., Ventura, J.D., Wolf, J.A., North, M.P., Davis, F.W., Franklin, J., 2019. A convolutional neural network classifier identifies tree species in mixed-conifer forest from hyperspectral imagery. Remote Sensing, 11(19), p.2326. doi.org/10.3390/rs11192326.

Goutte, C., Gaussier, E., 2005: A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In Advances in Information Retrieval: 27th European Conference on IR Research, ECIR 2005, Santiago de Compostela, Spain, March 21-23, 2005. Proceedings 27 (pp. 345-359). Springer Berlin Heidelberg. doi.org/10.1007/978-3-540-31865-1_25

Guan, H., Yu, Y., Ji, Z., Li, J., Zhang, Q., 2015. Deep learning-based tree classification using mobile LiDAR data. Remote Sensing Letters, 6(11), pp.864-873. doi.org/10.1080/2150704X.2015.1088668.

Hamdi, Z.M., Brandmeier, M., Straub, C., 2019. Forest damage assessment using deep learning on high resolution remote sensing data. Remote Sensing, 11(17), p.1976. doi.org/10.3390/rs11171976.

Hamedianfar, A., Mohamedou, C., Kangas, A., Vauhkonen, J., 2022. Deep learning for forest inventory and planning: a critical

review on the remote sensing approaches so far and prospects for further applications. Forestry, 95(4), pp.451-465. doi.org/10.1093/forestry/cpac002.

Holzwarth, S., Thonfeld, F., Abdullahi, S., Asam, S., Da Ponte Canova, E., Gessner, U., Huth, J., Kraus, T., Leutner, B., Kuenzer, C., 2020: Earth observation based monitoring of forests in Germany: A review. Remote Sensing, 12(21), p.3570. doi.org/10.3390/rs12213570.

Jiang, S., Yao, W., Heurich, M., 2019. Dead wood detection based on semantic segmentation of VHR aerial CIR imagery using optimized FCN-Densenet. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, pp.127-133. doi.org/10.5194/isprs-archives-XLII-2-W16-127-2019.

Jocher, G. 2020. YOLOv5 by Ultralytics (Version 7.0). doi.org/10.5281/zenodo.3908559.

Jocher, G., Chaurasia, A., Qiu, J. (2023). YOLO by Ultralytics (Version 8.0.0). doi.org/10.5281/zenodo.3908559.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. nature, 521(7553), pp.436-444. doi.org/10.1038/nature14539

Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017. Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117-2125).

Liu, J., Wang, X., Wang, T., 2019. Classification of tree species and stock volume estimation in ground forest images using Deep Learning. Computers and Electronics in Agriculture, 166, p.105012. doi.org/10.1016/j.compag.2019.105012.

Liu, S., Qi, L., Qin, H., Shi, J., Jia, J., 2018. Path aggregation network for instance segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 8759-8768).

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS journal of photogrammetry and remote sensing, 152, pp.166-177. doi.org/10.1016/j.isprsjprs.2019.04.015

Martins, G.B., La Rosa, L.E.C., Happ, P.N., Coelho Filho, L.C.T., Santos, C.J.F., Feitosa, R.Q., Ferreira, M.P., 2021: Deep learning-based tree species mapping in a highly diverse tropical urban setting. Urban Forestry & Urban Greening, 64, p.127241. doi.org/10.1016/j.ufug.2021.127241.

Pelletier, C., Webb, G.I., Petitjean, F., 2019. Temporal convolutional neural network for the classification of satellite image time series. Remote Sensing, 11(5), p.523. doi.org/10.3390/rs11050523.

Safonova, A., Tabik, S., Alcaraz-Segura, D., Rubtsov, A., Maglinets, Y., Herrera, F., 2019. Detection of fir trees (Abies sibirica) damaged by the bark beetle in unmanned aerial vehicle images with deep learning. Remote sensing, 11(6), p.643. doi.org/10.3390/rs11060643.

Sylvain, J.D., Drolet, G., Brown, N., 2019. Mapping dead forest cover using a deep convolutional neural network and digital aerial photography. ISPRS Journal of Photogrammetry and Remote Sensing, 156, pp.14-26. doi.org/10.1016/j.isprsjprs.2019.07.010.

Wang, C.Y., Bochkovskiy, A., Liao, H.Y.M., 2023. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 7464-7475).

Wang, C.Y., Liao, H.Y.M., Wu, Y.H., Chen, P.Y., Hsieh, J.W., Yeh, I.H., 2020. CSPNet: A new backbone that can enhance learning capability of CNN. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (pp. 390-391).

Wang, Y., Wang, J., Chang, S., Sun, L., An, L., Chen, Y., Xu, J., 2021. Classification of street tree species using UAV tilt photogrammetry. Remote Sensing, 13(2), p.216. doi.org/10.3390/rs13020216

Zhang, C., Xia, K., Feng, H., Yang, Y., Du, X., 2021: Tree species classification using deep learning and RGB optical images obtained by an unmanned aerial vehicle. Journal of Forestry Research, 32(5), pp.1879-1888. doi.org/10.1007/s11676-020-01245-0.

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), pp.8-36. doi.org/ 10.1109/MGRS.2017.2762307.

Zou, X., Cheng, M., Wang, C., Xia, Y., Li, J., 2017. Tree classification in complex forest point clouds based on deep learning. IEEE Geoscience and Remote Sensing Letters, 14(12), pp.2360-2364. doi.org/10.1109/LGRS.2017.2764938.