

DETECTION OF DEGRADED ACACIA TREE SPECIES USING DEEP NEURAL NETWORKS ON UAV DRONE IMAGERY

Anne Achieng Osio^{a,*} Hoàng-Ân Lê^b Samson Ayugi^a Fred Onyango^a Peter Odwe^a Sébastien Lefèvre^b

^a The Technical University of Kenya (TUK), Faculty of Engineering & Built Environment, Nairobi, Kenya

^b IRISA, Université Bretagne Sud (UBS), Vannes, France

Commission III, WG III/10

KEY WORDS: UAV, Object Detection, Deep Learning, Acacia degradation

ABSTRACT:

Deep-learning-based image classification and object detection has been applied successfully to tree monitoring. However, studies of tree crowns and fallen trees, especially on flood inundated areas, remain largely unexplored. Detection of degraded tree trunks on natural environments such as water, mudflats, and natural vegetated areas is challenging due to the mixed colour image backgrounds. In this paper, Unmanned Aerial Vehicles (UAVs), or drones, with embedded RGB cameras were used to capture the fallen Acacia Xanthophloea trees from six designated plots around Lake Nakuru, Kenya. Motivated by the need to detect fallen trees around the lake, two well-established deep neural networks, i.e. Faster Region-based Convolution Neural Network (Faster R-CNN) and Retina-Net were used for fallen tree detection. A total of 7,590 annotations of three classes on 256×256 image patches were used for this study. Experimental results show the relevance of deep learning in this context, with Retina-Net model achieving 38.9% precision and 57.9% recall.

1. INTRODUCTION

Forest detection has been embraced in many studies using different types of remotely sensed datasets, captured from different aerial and satellite sensor platforms. Assessment of fallen trees is an important step to characterize forest health. According to Torres et al. (2021), most of the studies have been applied mainly in America and Europe with the most frequently used data being multi-spectral imagery from Landsat sensors (She et al., 2015). The recently launched Sentinel-2 have shown less capability in detecting individual vegetation species (Nzimande et al., 2021). UAV remotely-sensed images are ideal for forest health assessment since they provide optical imagery with high geometric spatial resolution (10-40cm) (Naik et al., 2021). Conversely, coarse ground resolution satellite imagery from optical sensors does not allow to capture the geometric structure of fallen trees (Naik et al., 2021), and hence remains mainly used for classification at local and regional scales (Gorelick et al., 2017). Other studies have also shown that incorporating LiDAR data with multi-spectral imagery improved the prediction of tree height estimation and canopy detection models within natural forests (Manzanera et al., 2016).

Despite the reliability of LiDAR data, its means of acquisition remains more expensive than that of UAV based data (Amptzidis et al., 2019). Moreover, models based on LiDAR data (especially in relation to above ground biomass or fallen trees) highly depends on the type of forest under study and the level of tree degradation on site (Galidaki et al., 2017). Synthetic Aperture Radar (SAR) data with longer wavelengths and cross-polarization capabilities have also been used in fallen tree studies. Models created using SAR data produce uncertainty and variance in modelling accuracy of above ground biomass (Dalponte et al., 2018; Naik et al., 2021). On the contrary, Osio and Lefèvre (2021) confirmed that the use of SAR-C channels

captured in Single Look Complex mode in conjunction with machine learning models and object-oriented approach yielded the best results with an Overall Accuracy (OA) of 98.1% and a Kappa of 97.0%, hence improving the above ground biomass classification on the Acacia xanthophloea strands around Lake Nakuru, Kenya. Despite achieving results at local scale, such a model was not suitable to capture individual degraded Acacia xanthophloea target trees that are fallen around the lake. In recent times, the deep learning paradigm with models tailored at image classification and detection has become a standard methodology in remote sensing studies. The main advantage provided by deep learning over classical machine learning approaches is that models created using deep learning can learn and extract information directly from input data, not requiring a costly feature engineering step. These models can then be used to detect and predict similar features on the entire scene under investigation.

Deep learning has been used in many studies in recent times including automatic extraction of ice-wedge polygons using Mask R-CNN framework on both high-resolution imagery and UAV, producing F1 Scores of 72% and 70% respectively (Zhang et al., 2020). Santos et al. (2019) made a comparison of three different deep learning frameworks, namely YOLOv3, Faster Region Based Convolutional Neural Networks (Faster R-CNN), and RetinaNet to assess a time series of RGB images in the context of tree crown detection achieving an overall average precision (AP) of 92%. Other studies have used object-based image analysis approach to detect coarse wood debris (CWD) from unmanned aerial systems in conjunction with LiDAR point clouds (Thiel et al., 2020). The authors reported an overall average precision (mAP) of 85% and a recall of 69.2%. However, they pointed out that the results achieved using very high resolution imagery and their line detection algorithm over CWD areas could be improved using deep learning approaches (Jiang et al., 2019). Further concerns by the authors involved controversy about the application of deep learning frameworks on tree

* Corresponding author.

Email: osio@univ-ubs.fr, ansiyo22@gmail.com

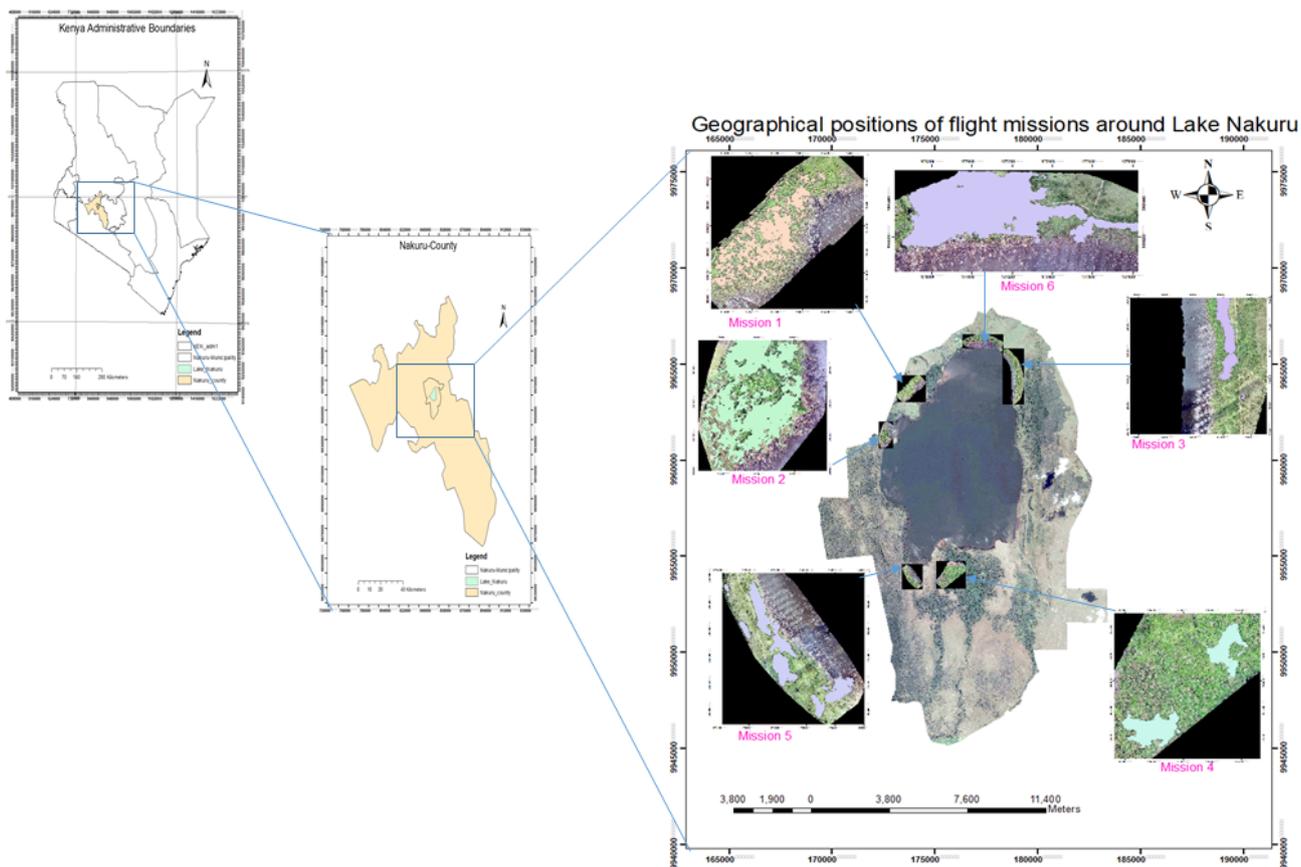


Figure 1. The geographic position of UAV flight missions around Lake Nakuru, projected at WGS84, UTM Zone 37S.

species, pointing out the challenge in model transferability onto similar scene.

In this paper, we evaluate off-the-shelf deep architectures in detecting fallen trees of the *Acacia Xanthophloea* trees around Lake Nakuru National Park, Kenya. The wetland was designated amongst wetlands of international importance (Odada et al., 2004) as Ramsar site number 476 on 5th June, 1990. Kenya, being a signatory to the Conference of Parties (COP) based on the Ramsar convention in 1971 (Davidson et al., 2019) is required to actively conserve and make wise use of Lake Nakuru wetlands in a sustainable manner.

It is important to quantify dead woody *Acacia xanthophloea* since, according to previous studies their presence improves biological diversity within live forests such as the introduction of mosses and lichens which attracts migratory birds in the National Park (Vareschi and Jacobs, 1985; Harmon et al., 1986; Nordén et al., 2008). The downed *Acacia xanthophloea* was not caused by any climatic factors but rather by the increased volume of water due to sedimentation (Iraddock et al., 2020) in Lake Nakuru which overflowed its bank (Osio et al., 2018), hence weakening the riparian trees from their roots, causing them to fall. Coarse fallen woody debris in water-bodies *i.e.* lakes and streams are known to increase channel complexity, which contributes to the improvement of habitat quality hence increasing nutrient retention inside stream systems (Cowden, 2002; Swanson and Franklin, 1992). Previous research by Bisson et al. (1992) reported reduction in habitat quality in streams that underwent traditional clearance of the deadwood in their stream systems.

UAV flight missions around the lake revealed massive destruc-

tion of the trees, especially inside the water body and on the mudflats. Therefore, the purpose of this study was to provide a state-of-the-art model based on the detection and classification of fallen trees around the lake, hence enabling the wildlife and forest conservation managers to make informed decision on the fallen *Acacia xanthophloea* trees. More precisely, our main goal is to evaluate the performance of well-established deep neural networks over UAV-based fallen tree datasets. To the best of our knowledge, there are no known studies that have been carried out on the detection of fallen *Acacia xanthophloea* around Lake Nakuru, using UAV/RGB in conjunction with deep learning approaches.

2. MATERIALS AND METHODS

2.1 Study Site

Lake Nakuru National park is located in Nakuru County (see Fig. 1), approximately 170 km away from Nairobi. The Park is situated geographically at Latitudes $0^{\circ}18'S$ and $0^{\circ}27'S$ and Longitude $36^{\circ}1.5'E$ and $39^{\circ}9.25'E$ within the Kenyan Rift valley (Mubea and Menz, 2012). Before the recent flooding, its bottom was initially about 1,756 m above sea level while the surface of the water was at 1,758.5m above sea level. The altitude ranges from 1,760-2,080m above sea level (Iraddock et al., 2020). Mean annual rainfall ranges between 876mm and 1,050mm and has an inherent bi-modal pattern (Odada et al., 2004). The long rains start in March and end in June while the short rains occur between October and December. Mean daily minimum and maximum temperatures fluctuate between $8.2^{\circ}C$ and $25.6^{\circ}C$ (Ng'weno et al., 2010). Lake Nakuru has no outlets and hence evaporation is the only factor that accounts for

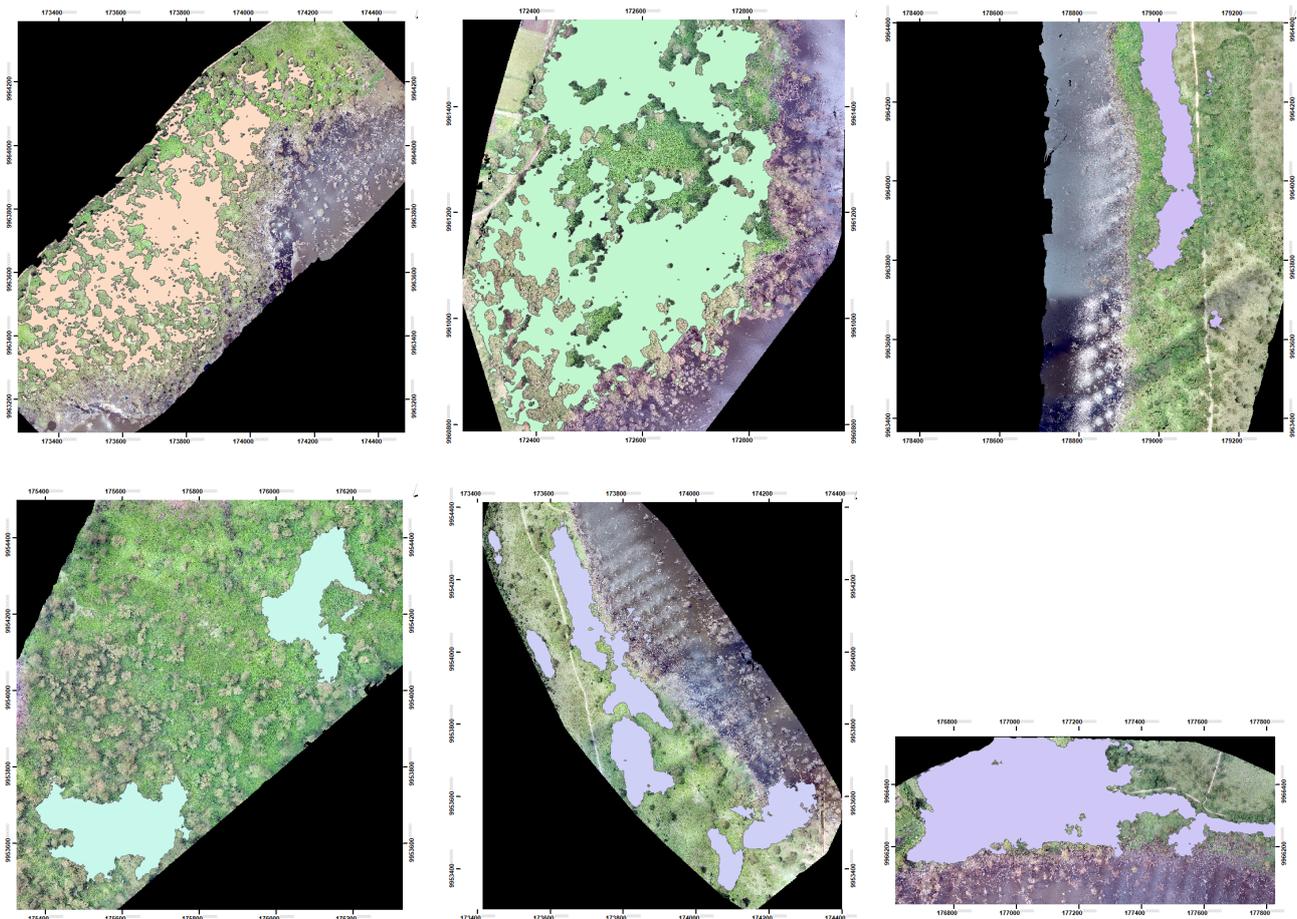


Figure 2. The six flight missions captured from different sites around the Lake using DJI Phantom 4, SDK Drone. The overlay vector polygons were derived from QGIS-based forest detection plugin, Mapflow.ai representing the detected trees on each imagery.

water loss. Four seasonal rivers feed the lake, *i.e.* Lamurdiak, Makalia, Enderit and Enjoro. *Acacia xanthophloea* tree patches have been in existence on the Northern, Southern, Eastern and Western side of the lake for decades. The soils on the shores of this lake are volcanic and shallow in nature. Underneath the *Acacia* savanna were the open grasslands thriving on soils and ashes that were well-drained, friable to sandy clay loams (see Fig. 2). Recent studies have shown that the health of *Acacia xanthophloea* trees have been degrading since the year 2010 due to the persistent flooding around the lake (Osio et al., 2020; Osio and Lefèvre, 2021).

2.2 Data Capture

Unmanned Aerial Vehicles (UAV) drone imageries were captured across six sites around the lake in the early September 2021 for 5 consecutive days, see Fig. 2. The UAV drone used for the ground surveys was a DJI Phantom 4 RTK SDK, all its specification are reported in Tab. 1. Images with a mean Ground Sampling distance (GSD) of 4.84cm and covering approximately 6.1 square kilometers were captured across six designated sites around the Lake Nakuru National Park. The surveyed areas were mainly near the lake shorelines where a large number of degradation were observed during the flight planning stage. A total of 9,056 training image patches were generated from the six images (see Figure 3) with 7,590 tree samples annotated from the training sets. The fallen trees are classified into 3 classes according to their background, namely Water (W), Land (L), and Mudflat (M) as shown in Fig. 3.

2.3 General Workflow

The study consisted of 4 steps. On the first step, images are acquired from a UAV/RGB drone platform. Then, patches of non-overlapping 256-pixel squares are extracted from the acquired images and randomly sampled to training and test set. The fallen *Acacia xanthophloea* trees from the image patches are annotated with bounding boxes using the LabelImg¹ tool. Finally, the training image patches with their annotations are used to train a deep convolutional neural network (CNN) while the mutually exclusive test set is used to report the performance of the network on unseen data.

Deep learning for object detection Object detection is made of 2 subtasks, (1) localizing an object of interest in an image with a bounding box and (2) categorizing the box with the correct class. The tasks could be performed in a 2-step process, with the bounding boxes first proposed then classified, or all at once, leading to the so-called 1-stage detectors. The exemplar model for each type of architecture include Faster Region based Convolutional Neural Networks (Faster RCNN) (Ren et al., 2015) and RetinaNet (Lin et al., 2017), respectively. In this study, we evaluate these two models for detecting fallen trees around the Lake Nakuru.

Faster RCNN The Region-based CNN (RCNN) laid the groundwork for deep-learning-based object detection (Girshick et al., 2014) with bounding boxes proposed by Selective

¹ <https://github.com/tzutalin/labelImg>



Figure 3. The three sample classes and numbers of annotations in the dataset: from left to right, dead trees on land (2,514 boxes), dead trees on mudflat (2,077 boxes), and dead trees on water (2,999 boxes).

Search (Uijlings et al., 2013). Fast R-CNN (Gkioxari et al., 2015) speeds up the process by introducing the region of interest (ROI), pooling layer and inputting the full image instead of just the proposals to the deep network. Significant acceleration is achieved by Faster RCNN when a sub-network, called Region Proposal Network (RPN), is used to generate the proposal boxes in place of Selective Search and both 2 stages can be trained end-to-end.

The vanilla Faster RCNN architecture extracts low-level image features by passing an input image through several convolutional blocks, called a backbone (sub-)network. The features are then shared between both the region proposal network (RPN) and region of interest (ROI) head. The Region Proposal Network (RPN) generates a number of proposal boxes, 2000 by default, from which the corresponding features are obtained and classified by the ROI head (sub-)network. The classifier is a multi-layer perceptron (MLP).

RetinaNet RetinaNet (Lin et al., 2017) is a single-stage architecture for object detection. It consists of (i) a classification sub-network which predicts the probability of an object occurrence at each spatial location for each annotated box and object class, (ii) a regression sub-network that regresses the offset for the bounding boxes from the annotated boxes for each ground-truth object, (iii) a bottom-up pathway which consists of the backbone network (ResNet) whose role is to calculate the feature maps at different scales, and (iv) a top-down pathway that up-samples the spatially coarser feature maps from higher pyramid levels. Lateral connections are included to merge top-down layers and bottom-up layers with the same spatial size. Specifically, the Feature Pyramid Network (FPN) is proposed for feature extraction through upsampling or downsampling approach and relies on the focal loss objective function.

2.4 Experimental setup

In this study, we use the implementation of Faster RCNN and RetinaNet provided in the Detectron2 library as previously implemented by Wu et al. (2019).

In relation to this experiment, all the annotations from the three classes, namely dead tree on Land (L), Water (W) and Mudflat (M) were sampled into two parts consisting of training/validation and testing set. From each given class of the dataset, 80% of the samples across all the missions were used for training/validation while 20% remained for testing. This

particular design was adopted to cater for imbalances within the datasets across the 6 missions (Fig. 2). Missions 3 and 4 have particularly fewer annotations compared to the rest of the missions.

Metrics We follow the Common Objects in Context (COCO) challenge for quantitative assessment, as demonstrated by Lin et al. (2014). The challenge employs the standard definition of precision (P) and recall (R) based on the notion of true positive (TP), false positive (FP), and false negative (FN):

$$P = \frac{TP}{TP + FP} \quad R = \frac{TP}{TP + FN} \quad F1 = 2 \frac{P \times R}{P + R} \quad (1)$$

The COCO metric differs in the definition of a *positive* box prediction, for which the intersection over union (IoU, or Jaccard index) of it with ground truth boxes are computed. IoU measures the ratio between the overlapping area of the two boxes divided by the area covered by their union. As such, the number of positive boxes changes according to the IoU level: higher IoU threshold (max of 1 or 100%) results in fewer positive predictions and thus, lower measurement.

At an IoU threshold, the boxes with at least (or higher than) the given level are considered positive, true or false depending on the predicted class, and are used to compute precision and recall. In this paper, the precision are computed for each of pre-defined recall values (101 values from 0 to 1 with step of 0.01), which are used to plot the precision-recall (PR) curve. We also report the average precision (AP), average recall (AR), and subsequently F1-score using Eq. 1 for 2 IoU levels, 0.50 and 0.75. The average precision is taken across all recall levels and equal to the area under the PR-curve:

$$\sum_{j=1}^N p(k) \Delta r(k), \quad (2)$$

where N is the total number of images in the collection, $p(k)$ is the precision at a cutoff of k images, and $\Delta r(k)$ is the change in recall that happened between cutoff $k - 1$ and cutoff k .

3. RESULTS

We provide in Fig. 6 and Fig. 7 the Precision-Recall curves for the two deep neural networks considered in our study. When



Figure 4. Sample results from different scenes using RetinaNet, with average precision (AP) averaging over all IoU levels in [0.5, 0.95], step size 0.05 and (50 being precision at IoU \geq 0.5) on three classes: Dead Tree in Water (W), on Land (L), and on Mudflat (M). Ground truths are shown on the left and predicted bounding boxes on the right. The confidence levels (%) are shown on the bounding boxes.



Figure 5. Sample results from different scenes using Faster-RCNN, with average precision (AP) averaging over all IoU levels in [0.5, 0.95], step size 0.05 and (50 being precision at IoU \geq 0.5) on three classes: Dead Tree in Water (W), on Land (L), and on Mudflat (M). Ground truths are shown on the left and predicted bounding boxes on the right. The confidence levels (%) are shown on the bounding boxes.

comparing RetinaNet and Faster RCNN, one can observe that the former delivers higher average AP results (38.85%) than the latter (37.21%) considering an IoU threshold of 0.5. Nevertheless, the difference remains within a small margin. In terms of Average Precision per class, RetinaNet performs best for both classes “Land” and “Water” (with improvement of 2-4%) but worse for class “Mudflat” (degradation of 2%).

Considering a higher IoU threshold of 0.75, the methods reach lower APs. Again, RetinaNet performs better than Faster RCNN (also with a small margin). A classwise analysis leads to somehow different conclusions, with RetinaNet achieving better results for classes “Mudflat” and “Water”, Faster RCNN for “Land”. The margin being very small, one should take these results with caution.

We then show in Fig. 4 and Fig. 5 some visualization results achieved by RetinaNet and Faster RCNN per each given class. All classes were predicted with a high confidence level as shown at the edge of the bounding boxes. Illustrations also include the average Precision (AP) reported per class and mean Average Precision (mAP) on all the bounding boxes of a specific class (shown in red values).

4. DISCUSSION

The model that achieves the best performances (Precision and Recall) was the RetinaNet with a setting $IoU \geq 0.5$, as shown in Tab. 2. These results corroborates with recent studies carried out by Santos et al. (2019) where RetinaNet outperformed two other variants, namely YoloV3 and Faster-RCNN. Similar studies by Alon et al. (2019) involving tree crown detection with UAV orthophotos and LiDAR point clouds had promising results with RetinaNet as well.

The overall Average Precision (AP) reported with RetinaNet was 43.85%. Usually AP or Area Under the Curve are normally used on imbalanced datasets. Datasets with low true positive rates (TPR) and high false positive (FP) and false negative (FN) rates usually produce low Precision and Recall. In the case of our UAV dataset, annotations were carried out on the fallen trees which exhibited different backgrounds, *i.e.* mudflat, water and land. Class “water” had the highest detection rates with an $AP \doteq 43.85\%$ due to uniformity of their background, while class “mudflat” shows the opposite behavior due to the background non-uniformity.

It is well-known that a deep network’s performance is impacted by the number of annotations and image patches derived from the UAV dataset, as already shown in recent studies by Hägele et al. (2020). Although a large number of annotations are employed in this study (9,056 patches with 7,590 annotations), the performance is incomparable with tree crown detection using RetinaNet (Santos et al., 2019) which reportedly achieves AP of 92.64% on 392 image patches. We observed that two different areas might come with very different visual features and landscapes, hence problems with varying difficulty to tackle. More precisely, the tree crown detection problem is much simpler (thus leading to higher detection rates) due to the tree crown uniformity (shape factor) than our use case, where we had to deal with coarse wood debris (tree biomass) which come in different shapes and backgrounds, hence the creation of noise within networks.

We observed that the complexity of the proposed approaches which are based on both networks were trained on a cluster

UAV Features	DJI Phantom 4 RTK
Frequencies used	GPS:L1/L2; GLONASS:L1/L2
Positional Accuracy	H: 1.5cm; V:1cm; Both +1ppm (RMS)
Image Sensor	CMOS 1”
Max resolution	4864×3648 (4:3); 5472×3648 (3:2)
Field of view	84°
Mechanical Shutter	8-1/2000s
Data format	Photo (JPEG), Video (MOV)

Table 1. Da-Jiang Innovations Science & Technology Co. Ltd (DJI) Real Time Kinematic (RTK) of Phantom 4 Specifications according to (Phantom, 2018).

IoU	RetinaNet			Faster RCNN		
	Precision	Recall	F1 Score	Precision	Recall	F1 Score
0.50	38.9%	57.9%	46.5%	37.2%	53.2%	43.8%
0.75	14.1%	29.0%	19.0%	14.0%	27.7%	18.8%

Table 2. Quantitative results obtained with the two deep neural networks and two IoU thresholds.

node with 2 CPU x 20 E5-2687W v3 @ 3.10GHz, 396G RAM, and 1 NVIDIA GeForce GTX Titan-X of 12.2GB VRAM, which are shared among cluster users. The Faster RCNN with ResNet50 backbone and FPN contains 17,260,319 parameters while the RetinaNet model with the same backbone contains around 14,460,660 parameters.

5. CONCLUSION

In this study, we dealt with the detection of fallen Acacia xanthophloea trees, on which we evaluated two competitive CNN models, *i.e.* Faster-RCNN and RetinaNet using images captured by UAV with RGB Cameras on board. The networks were trained and assessed using a dataset made of 9,056 image patches and 7,590 annotations on bounding boxes. RetinaNet achieved overall Precision of 38.9% and Recall of 57.9%.

These results indicate that RGB Cameras embedded on UAV in conjunction with deep neural networks could possibly lead to the development of operational tools for the detection of fallen Acacia xanthophloea trees on different environmental backgrounds. This could also help in carrying out demographic surveys on fallen Acacia xanthophloea trees around the Park and in similar environments. Fallen Acacia xanthophloea tree demography could help ecologists and conservationists in quantifying the magnitude of Acacia tree degradation around Lake Nakuru. This model can be applied on other areas with similar characteristics such as fallen trees along the riparian reserve of other Rift valley lakes. Future studies should look into combining UAV-based images and their point clouds for tree detection and classification using deep networks.

6. ACKNOWLEDGEMENTS

The authors acknowledge: Kenya National Research Fund (K-NRF) and Campus France through Pamoja PHC, Kenya National Council for Science Technology & Innovation (K-NACOSTI) and Kenya Wildlife Services for providing permit to enable Drone Surveys in Lake Nakuru.

REFERENCES

Alon, A. S., Festijo, E. D., Juanico, D. E. O., 2019. Tree detection using genus-specific retinanet from or-

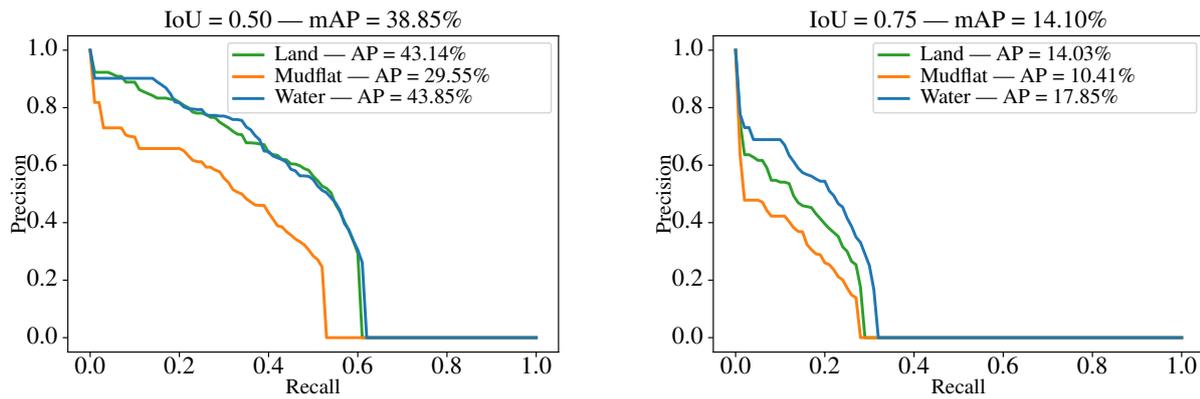


Figure 6. Precision-Recall curves for RetinaNet and two IoU thresholds.

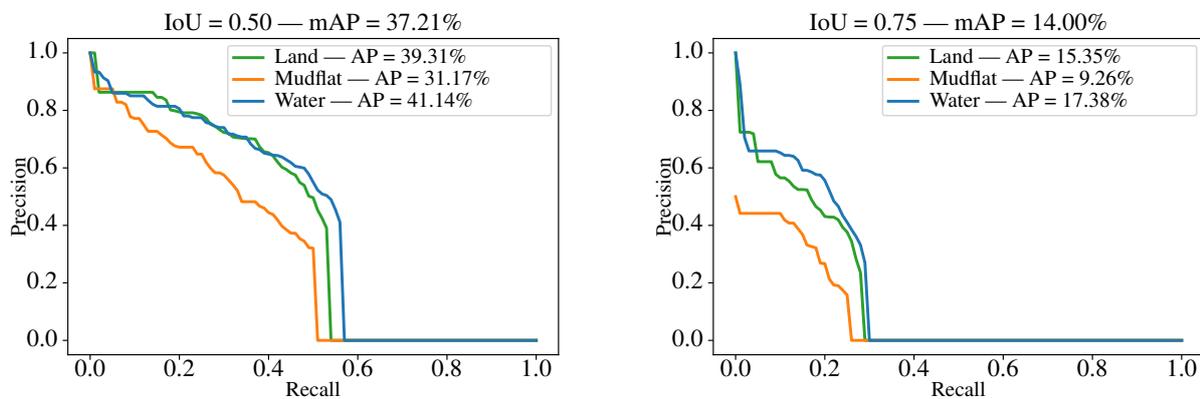


Figure 7. Precision-Recall curves for Faster RCNN and two IoU thresholds.

thophoto for segmentation access of airborne lidar data. *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, IEEE, 1–6.

Ampatzidis, Y., Partel, V., Meyering, B., Albrecht, U., 2019. Citrus rootstock evaluation utilizing UAV-based remote sensing and artificial intelligence. *Computers and Electronics in Agriculture*, 164, 104900.

Bisson, P. A., Quinn, T. P., Reeves, G. H., Gregory, S. V., 1992. Best management practices, cumulative effects, and long-term trends in fish abundance in pacific northwest river systems. *Watershed management*, Springer, 189–232.

Cowden, M. M., 2002. A study of the current range and habitat of fuzzy sandozi conks (*Bridgeoporus nobilissimus*) throughout Pacific Northwest forests.

Dalponte, M., Ene, L. T., Gobakken, T., Næsset, E., Gianelle, D., 2018. Predicting selected forest stand characteristics with multispectral ALS data. *Remote Sensing*, 10(4), 586.

Davidson, N., Dinesen, L., Fennessy, S., Finlayson, C., Grillas, P., Grobicki, A., McInnes, R., Stroud, D., 2019. A review of the adequacy of reporting to the Ramsar Convention on change in the ecological character of wetlands. *Marine and Freshwater Research*, 71(1), 117–126.

Galidaki, G., Zianis, D., Gitas, I., Radoglou, K., Karathanassi, V., Tsakiri-Strati, M., Woodhouse, I., Mallinis, G., 2017. Vegetation biomass estimation with remote sensing: focus on forest and other wooded land over the Mediterranean ecosystem. *International Journal of Remote Sensing*, 38(7), 1940–1966.

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 580–587.

Gkioxari, G., Girshick, R., Malik, J., 2015. Contextual action recognition with r* cnn. *Proceedings of the IEEE International Conference on Computer Vision*, 1080–1088.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202, 18–27.

Hägele, M., Seegerer, P., Lapuschkin, S., Bockmayr, M., Samek, W., Klauschen, F., Müller, K.-R., Binder, A., 2020. Resolving challenges in deep learning-based analyses of histopathological images using explanation methods. *Scientific reports*, 10(1), 1–12.

Harmon, M. E., Franklin, J. F., Swanson, F. J., Sollins, P., Gregory, S., Lattin, J., Anderson, N., Cline, S., Aumen, N., Sedell, J. et al., 1986. Ecology of coarse woody

- debris in temperate ecosystems. *Advances in ecological research*, 15, 133–302.
- Iradukunda, P., Sang, J. K., Nyadawa, M. O., Maina, C. W., 2020. Sedimentation effect on the storage capacity in lake Nakuru, Kenya. *Journal of Sustainable Research in Engineering*, 5(3), 149–158.
- Jiang, S., Yao, W., Heurich, M. et al., 2019. Dead wood detection based on semantic segmentation of vhr aerial cir imagery using optimized fcn-densenet. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 127–133.
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollar, P., 2017. Focal loss for dense object detection. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C. L., 2014. Microsoft COCO: Common objects in context. *European Conference on Computer Vision*, Springer, 740–755.
- Manzanera, J. A., García-Abril, A., Pascual, C., Tejera, R., Martín-Fernández, S., Tokola, T., Valbuena, R., 2016. Fusion of airborne LiDAR and multispectral sensors reveals synergic capabilities in forest structure characterization. *GIScience & Remote Sensing*, 53(6), 723–738.
- Mubea, K., Menz, G., 2012. Monitoring land-use change in Nakuru (Kenya) using multi-sensor satellite data.
- Naik, P., Dalponte, M., Bruzzone, L., 2021. Prediction of Forest Aboveground Biomass Using Multitemporal Multispectral Remote Sensing Data. *Remote Sensing*, 13(7), 1282.
- Ng'weno, C. C., Mwasi, S. M., Kairu, J. K., 2010. Distribution, density and impact of invasive plants in Lake Nakuru National Park, Kenya. *African Journal of Ecology*, 48(4), 905–913.
- Nordén, B., Götmark, F., Ryberg, M., Paltto, H., Allmer, J., 2008. Partial cutting reduces species richness of fungi on woody debris in oak-rich forests. *Canadian Journal of Forest Research*, 38(7), 1807–1816.
- Nzimande, N., Mutanga, O., Kiala, Z., Sibanda, M., 2021. Mapping the spatial distribution of the yellowwood tree (*Podocarpus henkelii*) in the Weza-Ngele forest using the newly launched Sentinel-2 multispectral imager data. *South African Geographical Journal*, 103(2), 204–222.
- Odada, E., Raini, J., Ndeti, R., 2004. Experiences and lessons learned brief, Lake Nakuru. Lake Basin Management Initiative: Main Report.
- Osio, A., Lefèvre, S., 2021. Object-Based Change Detection on Acacia Xanthophloea Species Degradation Along Lake Nakuru Riparian Reserve. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 347–352.
- Osio, A., Lefèvre, S., Ogao, P., Ayugi, S., 2018. Obiased monitoring of riparian vegetation applied to the identification of degraded acacia xanthophloea along lake nakuru, kenya. *GEOBIA 2018-From pixels to ecosystems and global sustainability*, 18–22.
- Osio, A., Pham, M., Lefèvre, S., 2020. Spatial Processing of Sentinel Imagery for Monitoring of Acacia Forest Degradation in Lake Nakuru Riparian Reserve. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3, 525–532.
- Phantom, D., 2018. RTK User Manual v1. 4. *DJI: Shenzhen, China*.
- Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28, 91–99.
- Santos, A. A. d., Marcato Junior, J., Araújo, M. S., Di Martini, D. R., Tetila, E. C., Siqueira, H. L., Aoki, C., Eltner, A., Matsubara, E. T., Pistori, H. et al., 2019. Assessment of CNN-based methods for individual tree detection on images captured by RGB cameras attached to UAVs. *Sensors*, 19(16), 3595.
- She, X., Zhang, L., Cen, Y., Wu, T., Huang, C., Baig, M. H. A., 2015. Comparison of the continuity of vegetation indices derived from Landsat 8 OLI and Landsat 7 ETM+ data among different vegetation types. *Remote Sensing*, 7(10), 13485–13506.
- Swanson, F., Franklin, J., 1992. Promoting the Science of Ecology. *Ecological Applications*, 2(3), 262–274.
- Thiel, C., Mueller, M. M., Epple, L., Thau, C., Hese, S., Voltersen, M., Henkel, A., 2020. UAS Imagery-Based Mapping of Coarse Wood Debris in a Natural Deciduous Forest in Central Germany (Hainich National Park). *Remote Sensing*, 12(20), 3293.
- Torres, P., Rodes-Blanco, M., Viana-Soto, A., Nieto, H., García, M., 2021. The Role of Remote Sensing for the Assessment and Monitoring of Forest Health: A Systematic Evidence Synthesis. *Forests*, 12(8), 1134.
- Uijlings, J., van de Sande, K., Gevers, T., A.W.M., 2013. Selective Search for Object Recognition. *International Journal of Computer Vision*, 104, 154–171.
- Vareschi, E., Jacobs, J., 1985. The ecology of Lake Nakuru. *Oecologia*, 65(3), 412–424.
- Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., Girshick, R., 2019. Detectron2. <https://github.com/facebookresearch/detectron2>.
- Zhang, W., Liljedahl, A. K., Kanevskiy, M., Epstein, H. E., Jones, B. M., Jorgenson, M. T., Kent, K., 2020. Transferability of the deep learning mask R-CNN model for automated mapping of ice-wedge polygons in high-resolution satellite and UAV images. *Remote Sensing*, 12(7), 1085.