Deep Learning Based Long-Distance Optical UAV Detection: Color vs. Grayscale

Denis Ojdanić, Christopher Naverschnigg, Andreas Sinn, and Georg Schitter

Automation and Control Institute (ACIN), TU Wien, Gusshausstrasse 27-29, Vienna, Austria

ABSTRACT

This paper presents a comparison between grayscale and color based deep learning algorithms for long distance optical UAV detection using robotic telescope systems. Three deep learning object detection algorithms are trained with a custom dataset consisting of RGB images and the performance is evaluated against the same algorithms trained with the same dataset converted to grayscale. Network training from scratch and fine-tuning are evaluated. The results for all algorithms show that fine-tuning with RGB images maximizes the detection performance and scores about 5 % better in terms of mean average precision (mAP(0.5)) compared to fine-tuning on grayscale images.

Keywords: UAV detection, deep learning, color, grayscale

1. INTRODUCTION

Mini and micro drones, also known as unmanned aerial vehicles (UAVs), have gained massive popularity in recent years due to their versatility, ease of use and affordability. However, these characteristics paired with their small size, manoeuvrability and ability of carrying payload, make UAVs a modern safety hazard. UAVs have already been involved in several incidents including near collisions with air-planes,[1] unauthorized flyovers in the vicinity of nuclear power plants,[2] the illegal smuggling of goods over state boarders[3] and lately they play a significant role in modern warfare.[4] Given the diverse applicability of UAVs, which include potentially malicious utilization, and numerous accounts of incidents, research and development on proper detection and mitigation technologies is indispensable.

The detection of UAVs is an advanced research area with various approaches being investigated such as RADAR,[5] LiDAR,[6] radio frequency,[7] acoustics[8] and optics.[9] Most state of the art systems combine multiple of these technologies to benefit from the individual strengths forming a multispectral detection system.[10] Electro-optical systems are key components in all multispectral systems to ensure situational awareness, as captured images are easily interpreted. Therefore, systems equipped with cameras for object detection and classification are extensively researched. Optical systems are limited by their operational range and therefore, camera-based systems use narrow field of view (FoV) to extend the detection range. These devices are attached to mounts, which enable pan and tilt motion in order to observe a larger area and to track detected objects. Considering the selection of an appropriate camera, it is known that the achievable resolution, low light capabilities and quantum efficiency of monochrome sensors outperform color camera sensors due to the color filters used in the latter one.[11] Following this argumentation, the usage of monochrome cameras is advisable to improve detection distance.
However, state of the art object detection, which is usually facilitated via deep learning algorithms,[12] is mostly evolving around using color data as input, as the major publicly available datasets, e.g. COCO[13] or ImageNet,[14] which are often used for fine-tuning and transfer learning, suggest. Different studies show, that there is no consensus on whether color or grayscale images improve the object detection performance and analysis needs to be conducted for each specific field of application.[15,16]
The contribution of this paper is the evaluation of three state of the art deep learning based algorithms with respect to their UAV detection performance using either color or grayscale input images.

Further author information: Send correspondence to Denis Ojdanić
E-mail: ojdanic@acin.tuwien.ac.at, Telephone:+43 (0) 1 58801 376 520

Post-Print version of the article: Denis Ojdanić, Christopher Naverschnigg, Andreas Sinn, and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *In Pattern Recognition and Tracking XXXIV*, vol. 12527, pp. 80-84. SPIE, 2023. DOI: 10.1109/TAES.2023.3248560

2. UAV DETECTION

For UAV detection, three state of the art deep learning algorithms, FRCNN,[17] Retinanet[18] and SSD[19] are selected, which are predominantly used for efficient object detection. These three algorithms are trained once on color images and in a separate instance, on grayscale images, in order to evaluate, whether color or grayscale results in an improved deep learning based UAV detection performance. The main difference between color and grayscale images is the number of channels storing the data. For RGB images, an image contains three channels, e.g. red, green and blue with each channel storing the pixel values for the corresponding color. Grayscale images only contain one channel, which stores the grayscale information. Likewise, the filters of the input layers of deep learning algorithms contain three input channels for the color case. For grayscale images two network adaptions are suitable. First, the three input filter channels of the network can be reduced to one input channel to match the number of input image channels. Another possibility is to triplicate the grayscale input image to a three channel image, using the same input for each "color" channel. While the first approach is negligibly more computationally efficient, the second approach retains the same number of network weights, which allows a fair comparison between color and grayscale networks.

2.1 Dataset

The dataset used for training consists of 5000 images depicting different UAVs during daytime and clear weather conditions. About 50 % of the images are captured through various optical systems ranging from standard cameras to telescopes[9] and the remaining images are generated by blending cropped images of UAVs over different backgrounds. Therefore, the dataset contains images with various degrees of atmospheric blurring and backgrounds. Furthermore, other flying objects, like birds, are visible in the images. All images are captured or simulated by using color cameras and images. To transform the color data to grayscale the following relation is used[20]

$$Gray = 0.299R + 0.587G + 0.114B, \tag{1}$$

with R, G and B representing the red, green and blue channels of the color image. Due to this color to grayscale transformation, it is assumed for the training and test dataset, that properties, like resolution, quantum efficiency and low light performance of a color and mono camera are equal. Fig. 1 depicts the distribution of the bounding box sizes of the training dataset, which indicates, that the majority the UAVs within the dataset are small relative to the image size with the median bounding box size being approximately 85 pixels x 85 pixels.

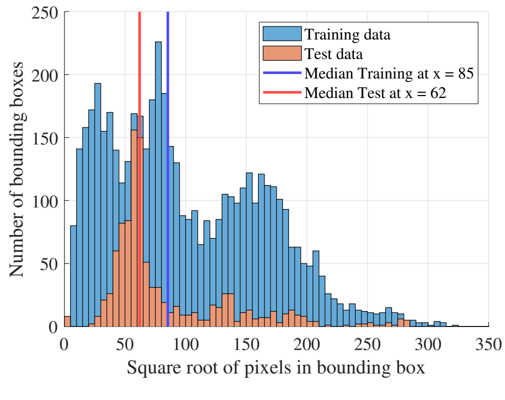


Figure 1. The histogram shows the distribution of the bounding box sizes within the training dataset with the median bounding box size being 85 x 85 pixels. Additionally, the distribution of the test dataset is depicted, which is used for evaluation in Section 3. The median bounding box size for the test dataset is 62 x 62 pixels.

Post-Print version of the article: Denis Ojdanić, Christopher Naverschnigg, Andreas Sinn, and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *In Pattern Recognition and Tracking XXXIV*, vol. 12527, pp. 80-84. SPIE, 2023. DOI: 10.1109/TAES.2023.3248560

2.2 Training procedure

To evaluate the algorithms two training approaches are applied. One method is training from scratch, where the algorithm weights are randomly initialized and optimized with the UAV dataset presented in Section 2.1. The second method is fine-tuning, whereas the algorithms are initialized with weights pre-trained on the COCO dataset. The COCO dataset[13] is selected, as it contains various classes, which are similar to the task of UAV detection, like "bird" or "plane". For both training procedures, the algorithms are trained for 30 epochs with a stepwise reduction of the learning rate. Table 1 shows the remaining parameters used for the training process. During the training process data augmentation is applied in the form of random flipping.

Table 1. The parameters used for training of each deep learning object detection algorithm.

Algorithm	Learning rate	Weight decay	Momentum
FRCNN	0.0004	0.0006	0.8
Retinanet	0.0009	0.0006	0.8
SSD	0.0007	0.0005	0.8

(a)

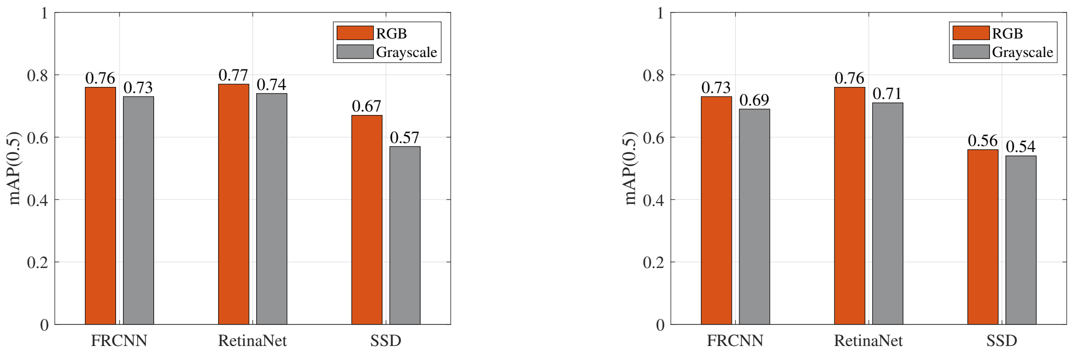
(b)

Figure 2. (a) An example color image showing a UAV. (b) The same image converted into grayscale according to Eq. 1.

3. EXPERIMENTS AND RESULTS

For the experimental analysis a test dataset is created, which consists of 1065 labeled images, which are captured during daytime and clear weather conditions. The bounding box size distribution of the test dataset is visible in Fig. 1. An example test image of a UAV is depicted in Fig. 2(a) and using Eq. 1, this image is converted to grayscale as shown in Fig. 2(b). As stated in Section 2.1, the color and test dataset appear to be captured by cameras with equal properties in terms of resolution, quantum efficiency and low light performance, as a consequence of using Eq. 1.
The results of applying the test dataset to the trained models are displayed in Fig. 3. As an evaluation metric the mean average precision (mAP) is used with an overlap threshold of 0.5 Comparing the two training methods, the fine-tuned models outperform the models, which are trained from scratch by an average of about 4 % in terms of mAP(0.5). It appears that the fine-tuned models learned to generalize better, due to being exposed to more data during the training process. Evaluating the results of comparing color to grayscale, the color models outperform the grayscale models by 5 % mAP(0.5) when fine-tuned and by 3 % when trained from scratch. For the case of fine-tuning, the difference between color and grayscale is larger, which appears to be a consequence of the COCO-dataset consisting of color images, which is advantageous for the later color fine-tuning phase.
Summarizing the results, fine-tuning as a training strategy achieves on average 4 % higher mAP(0.5) than training from scratch. Furthermore, the fine-tuned color models achieve the overall best performance, outperforming the fine-tuned grayscale models and the color models trained from scratch by 5 % respectively in terms of mAP.

Post-Print version of the article: Denis Ojdanić, Christopher Naverschnigg, Andreas Sinn, and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *In Pattern Recognition and Tracking XXXIV*, vol. 12527, pp. 80-84. SPIE, 2023. DOI: 10.1109/TAES.2023.3248560



(a) Fine-tuned models. (b) Models trained from scratch.

Figure 3. (a) shows the mAP(0.5) for three object detection algorithms fine-tuned on RGB and on grayscale images. (b) depicts the results of the algorithms trained from scratch. The fine-tuned color models show the overall best performance and are therefore, the preferred training strategy.

4. CONCLUSION

Three state of the art deep learning object detection algorithms have been trained for the task of UAV detection and an evaluation has been performed whether color or grayscale images offer an advantage in terms of detection performance. The evaluation is performed using data captured by color cameras during daytime and clear weather conditions. The grayscale transformation implies that color and mono cameras both perform equally in terms of resolution, quantum efficiency and low light performance, when capturing the test data. Based on the evaluation it can be concluded, that models trained with color images outperform models trained with grayscale images for both chosen training strategies. Best overall results are achieved, when fine-tuning the models on color images, as it outperforms the fine-tuned grayscale models and the color models trained from scratch by about 5 % according to the mAP(0.5) respectively.
Future work will focus on a holistic system evaluation, which extends the presented algorithm analysis by an evaluation of different cameras and daytime conditions.

ACKNOWLEDGMENTS

This publication is funded by the Austrian defence research programme FORTE of the Federal Ministry of Finance (BMF).
The authors gratefully acknowledge the cooperation with ASA Astrosysteme GmbH and thank for their support and valuable expertise.

REFERENCES

[1] Serna, J., "Lufthansa jet and drone nearly collide near LAX." Los Angeles Times (2016). Accessed March 2023.

[2] Phillips, C. and Gaffey, C., "Most French Nuclear Plants 'Should Be Shut Down' Over Drone Threat." Newsweek Magazine (Feb. 2015). Accessed Feb 2022.

[3] "Charges over drone drug smuggling into prisons." BBC News (Mar. 2018). Accessed Feb 2022.

[4] Bi, Z., Chen, H., Hu, J., et. al., "Analysis of UAV Typical War Cases and Combat Assessment Research," in [*2022 IEEE International Conference on Unmanned Systems (ICUS)*], 1449–1453 (2022).

[5] A. D. De Quevedo, F. I. Urzaiz, J. G. Menoyo, and A. A. Lopez, "Drone Detection and RCS Measurements with Ubiquitous Radar," in [*2018 International Conference on Radar (RADAR)*], 1–6 (2018).

[6] Ojdanić, D., Gräf, B., Sinn, A., Yoo, H. W., and Schitter, G., "Camera-guided real-time laser ranging for multi-UAV distance measurement," *Appl. Opt.* **61**, 9233–9240 (Nov 2022).

Post-Print version of the article: Denis Ojdanić, Christopher Naverschnigg, Andreas Sinn, and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *In Pattern Recognition and Tracking XXXIV*, vol. 12527, pp. 80-84. SPIE, 2023. DOI: 10.1109/TAES.2023.3248560

[7] Yang, S., Qin, H., Liang, X., and Gulliver, T., "An Improved Unauthorized Unmanned Aerial Vehicle Detection Algorithm Using Radiofrequency-Based Statistical Fingerprint Analysis," *Sensors* **19**, 274 (Jan 2019).

[8] Baron, V., Bouley, S., Muschinowski, M., Mars, J., and Nicolas, B., "Drone localization and identification using an acoustic array and supervised learning," in [*Artificial Intelligence and Machine Learning in Defense Applications*], 11169F SPIE (Sep 2019).

[9] Ojdanić, D., Sinn, A., Naverschnigg, C., and Schitter, G., "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems* , 1–10 (2023).

[10] Farlik, J., Kratky, M., Casar, J., and Stary, V., "Multispectral Detection of Commercial Unmanned Aerial Vehicles," *Sensors* **19**, 1517 (Mar. 2019).

[11] Chakrabarti, A., Freeman, W. T., and Zickler, T., "Rethinking color cameras," in [*2014 IEEE International Conference on Computational Photography (ICCP)*], 1–8 (2014).

[12] Unlu, H. U., Niehaus, P. S., Chirita, D., Evangeliou, N., and Tzes, A., "Deep learning-based visual tracking of UAVs using a PTZ camera system," in [*IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*], IEEE (Oct. 2019).

[13] Lin, T., Maire, M., Belongie S., et. al., "Microsoft COCO: Common Objects in Context," in [*Computer Vision–ECCV 2014, Proceedings, Part V 13*], 740–755, Springer (2014).

[14] Deng, J., Dong, W., and Socher, R., et. al, "Imagenet: A large-scale hierarchical image database," in [*2009 IEEE Conference on Computer Vision and Pattern Recognition*], 248–255 (2009).

[15] Funt, B. and Zhu, L., "Does colour really matter? Evaluation via object classification," in [*Color and Imaging Conference*], **26**, 268–271, Society for Imaging Science and Technology (2018).

[16] Bui, H. M., Lech, M., Cheng, E., et. al., "Using grayscale images for object recognition with convolutional-recursive neural network," in [*2016 IEEE Sixth International Conference on Communications and Electronics (ICCE)*], 321–325 (2016).

[17] Ren, S., He, K., Girshick, R., and Sun, J., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in [*Advances in Neural Information Processing Systems*], **28**, Curran Associates, Inc. (2015).

[18] Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P., "Focal loss for dense object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **42**(2), 318–327 (2020).

[19] Wei, L., Anguelov, D., Erhan, D., et. al., "SSD: Single Shot MultiBox Detector," in [*Computer Vision – ECCV 2016*], 21–37, Springer International Publishing, Cham (2016).

[20] Kanan, C. and Cottrell, G. W., "Color-to-grayscale: does the method matter in image recognition?," *PloS ONE* **7**(1), e29740 (2012).

Post-Print version of the article: Denis Ojdanić, Christopher Naverschnigg, Andreas Sinn, and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *In Pattern Recognition and Tracking XXXIV*, vol. 12527, pp. 80-84. SPIE, 2023. DOI: 10.1109/TAES.2023.3248560