Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes

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

***Abstract*—Substantial technological development has made Unmanned Aerial Vehicles (UAVs) more versatile, cheaper and accessible to the public in recent years. Alongside many positive effects and use cases, safety concerns are increasing as a plethora of incidents demonstrate the destructive potential of UAVs. To counteract this development and thus protect people and critical infrastructure, UAV detection, tracking and defence gains more and more research attention. Whereas, different drone detection technologies like RADAR, radio frequency and acoustic detection are deployed within multi-spectral systems, optical detection and imaging of approaching objects provide key information to correctly assess the situation. As reaction time is a crucial parameter for successful UAV defence, the operating distance of the optical detection system needs to be improved further. This paper presents the analysis, development and evaluation of a telescope-based UAV detection system. The system consists of a high precision mount and a telescope equipped with a camera. UAVs are detected in the captured video frames by the deep learning algorithm YOLOv4 using a modified architecture. The proposed system, which uses a f/10 telescope with a focal length of f = 2540 mm and a camera equipped with a 7.3 mm x 4.1 mm sensor, allows a significant increase of the optical detection range to more than 3 km of UAVs down to 0.3 m in diameter under daylight conditions and sufficient contrast, extending the reaction time significantly for counter UAV systems.**

***Index Terms*—UAV detection, long distance detection, telescopes, deep learning**

I. Introduction

OVER the past decade advances in technology led to a massive grow in popularity of unmanned aerial vehicles both in professional and private sector [1]. Easy public accessibility is accompanied by an increased risk of potential misuse, ranging from relatively harmless violations to severe hazardous situations. Extensive research is conducted on possible future threat scenarios and an alarming number of real live incidents has already occurred [2], [3]. An analysis on the threat UAVs pose to nuclear facilities concludes that UAVs are resourceful tools to cause a variety of perturbations in the form of distraction, reconnaissance or kinetic attacks [3]. Likewise, incidents around airports show a development for potentially dangerous situations [4]. A famous example is the shutdown of the London Gatwick airport for more than a day due to a nearby UAV in 2018 [5]. Similar to the mentioned scenarios, numerous other threats exists that have already taken place like the illegal smuggling of goods over state borders [6] or incidents near governmental buildings [7]. Summarizing, the stated examples and studies emphasise that early detection and identification of uncooperative UAVs is essential in order to protect people and critical infrastructure.

Various technological solutions based on different sensors exist for the purpose of UAV detection and tracking. For example radio frequency (RF) signals, which are utilized for communication between the operator and drone, can be exploited to detect UAVs in a distance of up to 5 km [8], [9]. However, a major drawback of this technology is the inherent possibility for controlling the UAV without any communication for example using pre-programmed flight paths. In this scenario, the UAV remains completely invisible to the RF detection system.

RADAR does not suffer from the aforementioned limitation. It detects targets by emitting radio frequency waves and measuring the reflection of an obstructing object. The operational range can go up to several kilometres of distance depending on the size of the radar cross section of the object [10], [11]. Generally, a classification of the detected target is possible with the exploitation of the Doppler effect [12]. However, for objects like consumer drones, which have a small radar cross section, detection and identification becomes more challenging [11].

Acoustic sensors, in the form of microphone arrays, capture sound and detect UAVs by applying appropriate signal processing methods. Angular uncertainties for the target localization of about 4° can be achieved [13] and even identification is possible by using comprehensive databases [14]. The disadvantage of this technology is the small operational range of below 600 m [15], since it suffers greatly from ambient noise, which lies in a similar frequency band as the emitted UAV sound.

Finally, optical systems capture images with cameras using different spectral bands like the visible spectrum from 400 to 700 nm, but also the near infrared as well as thermal signatures. Paramount for the optical system performance in terms of detection distance is the selection of the field of view (FoV) of the camera system. A wider FoV allows the observation of a broader area with the expense of a shorter achievable distance as the resolution of camera is limited to a finite size. Therefore, most optical UAV detection systems are based on cameras with relatively narrow FoVs, mounted on pan-tilt devices [16]. These devices allow the observation of a large area through realignment of the camera orientation, while providing acceptable resolution. Additionally, these systems can support optical zoom to adapt the FoV to the given task [17]. Current state of the art optical systems are limited by the detection range and achieve about 1 to 2 km [18]

The authors are with the Automation and Control Institute (ACIN), TU Wien, 1040 Vienna. Further author information: (Send correspondence to Denis Ojdanić)
E-mail: ojdanic@acin.tuwien.ac.at, Tel.: +43 (0) 1 58801 376 515

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560

using for example a 4k image sensor with a focal length of 81.6 mm for UAV diagonal sizes of 350 mm. The short operational range allows little time to appropriately react to incoming threats. Whereas being susceptible to weather conditions like rain, snow and fog, the main advantage of optical systems over the previously mentioned methods is the inherent possibility to identify the incoming target and additional payloads over long distances. This is a crucial benefit, allowing a differentiated situational assessment in order to prepare for appropriate countermeasures. For the task of target detection within the captured images or videos, optical systems rely on computer vision algorithms. Since 2012 deep learning based approaches, which utilize pre-trained neural networks, offer the best performances for object detection [19]. This trend is co-evolving with the accessibility of high performance computers, especially GPUs, for fast and parallel computing. The application of deep learning models based on RCNN [20], SSD [21] and YOLO [22] for UAV detection have been studied extensively in recent years, with studies suggesting that YOLO gives a good trade off in terms of accuracy and speed [23] [24]. Likewise, the usage of deep learning as the main method for UAV detection in systems equipped with pan-tilt-zoom cameras is investigated by various studies [16] [17] [25]. Similarly, for the application of air to air visual detection, where UAVs detect other UAVs during flight, deep learning algorithms are investigated, emphasizing the significance of a diverse and generalized dataset. [26], [27]. Morphological filters can be used to detect small objects over long distances [28], and by combining these filters with neural networks the false positive rate can be reduced [29].

Each one of the presented sensor types offers individual advantages and disadvantages. In order to implement a functional UAV detection system the best advice is to combine several sensors to profit from the respective strengths. An exemplary use case is an unidentified object detected by a radar system in a significant distance. After this coarse localization, the optical component of the system is aligned towards the intruder to capture an image for identification purposes. Ctrl+sky [30] and Ihtar [31] are two multispectral systems that combine radar, RF, acoustics and cameras for UAV detection.

The limitation in the current state of the art is the operational range offered by optical systems as small UAVs like consumer drones are not detectable over long distances with conventional optical systems. As stated above, optical sensors offer the most conclusive information for classification, as visual material can be classified by neural networks or even cross-checked and understood by a human operator. Therefore, it is paramount to increase the optical detection distance in order to gain more time to decide and prepare for appropriate countermeasures.

The contribution of this paper is the analysis, implementation and evaluation of a telescope-based optical detection system which is capable to detect UAVs over significantly larger distances and thus enabling early threat assessment. Following an analysis of the optical components, suitable hardware is selected to meet the requirements. Deep learning algorithms process the captured video frames and are used for UAV localization. Based on the extracted information, commands are provided to a mount controller to precisely actuate the telescope to keep the UAV within the FoV.
Section II presents an analysis on the most relevant design parameters to achieve long distance UAV tracking. In Section III the system is described together with details about the deep learning algorithms and control strategies in use. The performance of the proposed solution is evaluated in Section IV. Finally, a conclusion is given in Section V.

II. System Analysis

For a conclusive optical system analysis and design, first the expected UAV size and operation range are defined. The UAV classes examined within this work are commercially available drones which fall into the categories $C0$ to $C3$ according of European regulations [32]. Expressed in numbers, the smallest UAV sizes to be detected are about 0.3 m in diameter and the goal is to significantly extend the state of the art optical detection distances, which are currently around 1 to 2 km for daylight conditions [18]. In the following, an analysis of the most relevant design parameters and their impact on the system performance is given.

A. Resolution

One crucial parameter to distinguish objects in an image is the theoretically achievable resolution of the optical system. In optimal environmental conditions this value is limited by the diffraction limit which can be calculated using the Rayleigh criterion [33] given in Equation 1

$$\theta = 1.22\frac{\lambda}{D}, \tag{1}$$

where λ is the wavelength of light, D is the aperture size of the optical device and θ the smallest resolvable angle between two objects using the approximation $sin(\theta) \approx \theta$. A wavelength of 530 nm is selected in accordance to the peak response of typical color cameras for green light. Following this equation an increase of the telescope aperture is necessary to achieve ambitious requirements. However, in practice this is accompanied by two drawbacks. First, larger apertures need bigger telescope structures as seen in Fig. 1, where the weight of typical telescopes is depicted. A larger system mass increases the inertia, which inherently limits the achievable bandwidth for pan and tilt motions. Second, besides the fundamental limit described in Equation 1, the atmosphere further decreases the achievable resolution depending on the Fried parameter r_0 [34]. Local variations due to temperature, pressure and humidity differences affect the refracting index of air and the cumulation of these disturbances along the optical path results in beam spread, wander and intensity fluctuations [35]. Using Equations 2 and 3 to describe the variance of the beam tilt α^2_{jit} and spread σ^2_{comp} [35]:

$$\alpha^2_{jit} = 0.182\left(\frac{D}{r_0}\right)^{\frac{5}{3}}\left(\frac{\lambda}{D}\right)^2 \tag{2}$$

$$\sigma^2_{comp} = 0.134\left(\frac{D}{r_0}\right)^{\frac{5}{3}}, \tag{3}$$

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560

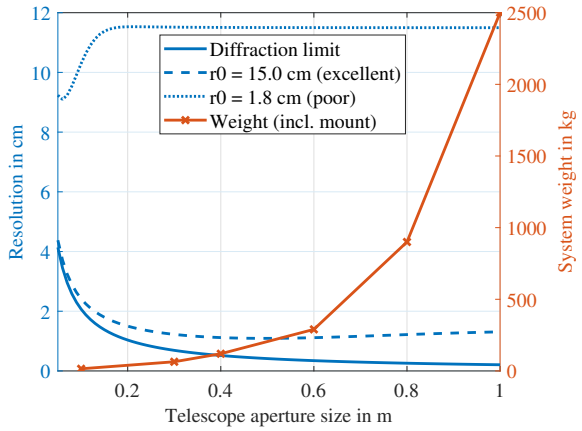


Figure 1: Theoretically achievable resolution of a telescope depending on the aperture size in a distance of 3.2 km. The resolution limited by diffraction (solid blue line) and atmosphere (dashed blue line) under poor (r_0 of 1.8 cm) and excellent (r_0 of 15 cm) conditions is depicted.

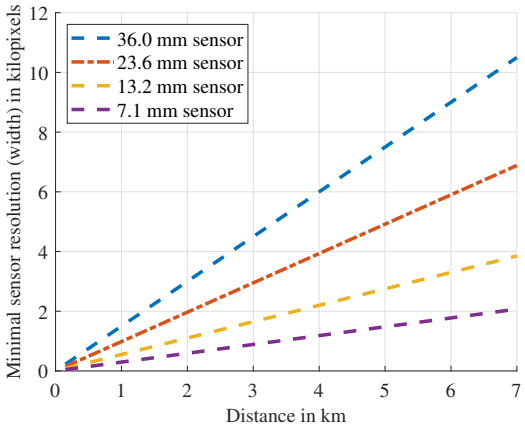


Figure 2: Minimum necessary resolution to obtain 15x15 pixels for an object of 0.3 m size at a given distance using a telescope with a focal length of 1200 mm.

the long exposure Strehl ratio can be calculated [35]:

$$S = \frac{e^{\sigma^2_{comp}}}{1+\left(\frac{2.22\alpha_{jit}D}{\lambda}\right)^2}\frac{1-e^{\sigma^2_{comp}}}{1+\left(\frac{D}{r_0}\right)^2}. \tag{4}$$

Incorporating the atmospheric disturbances to the diffraction limit from Equation 1 results in Equation 5 [35],

$$\theta_{atm} = \theta\frac{Q}{S} \tag{5}$$

$$Q = \sqrt{\frac{e^{2\sigma^2_{comp}}}{1+\left(\frac{2.22\alpha_{jit}D}{\lambda}\right)^2}\cdot\frac{1-e^{\sigma^2_{comp}}}{1+\left(\frac{D}{r_0}\right)^2}}, \tag{6}$$

with Q taken from [35]. Plotting the calculations for different aperture sizes and expected values for r_0 during daylight, Fig. 1 demonstrates that larger telescope diameters not necessarily increase the achievable resolution [36], [37]. Depending on environmental conditions smaller telescopes may perform similar as larger ones.

B. Field of view

Another paramount system property is the FoV. A large FoV allows simultaneous coverage of more area during the search for UAVs and an increased margin for tracking errors. However, increasing the observed area requires camera sensors with a higher number of pixels to ensure sufficient resolution. The width of the FoV is determined according to

$$FoV_{rad} = 2\arctan(\frac{s_w}{2f}), \tag{7}$$

with s_w being the camera sensor width and f the focal length of the telescope. A similar formula is applied for calculating the height. The achievable FoV in meter is given in Equation 8

$$FoV_m = 2d\cdot\tan(\frac{FoV_{rad}}{2}), \tag{8}$$

and is used to determine the necessary number of pixels needed to detect objects of different sizes for a given distance d. An exemplary calculation for a system with a typical focal length of 1200 mm and different sensors sizes demonstrates the requirements on the hardware of varying FoVs. An increased sensor size results in a larger FoV as per Equation 7. To detect and identify an object with a diameter of 0.3 m reliably, a minimum number of 15x15 pixels is needed, as discussed in Section IV. The red dash-dotted line in Fig. 2 shows that for the presented system a 23 mm wide APS-C sensor would require a 4k image resolution to detect a UAV reliably in 4 km of distance.

Ultimately the choice of a suitable FoV depends on the speed of the object, the frame rate and the expected detection distances. Assuming a camera frame rate of 60 FPS, Fig. 3 depicts the number of frames captured of a typical consumer UAV flying at a velocity of 31 m/s horizontally through the FoV, if the optical system is not moving. Referring back to the previous example system with a FoV of 1.1°, it is possible to capture 74 images of the UAV in a distance of 2 km with sufficient resolution to accurately detect the object as seen in Fig. 2. Additionally, the 74 captured frames provide enough margin for the object detection algorithm to detect the UAV in one of the frames. Reducing the distance between telescope and UAV leads to a reduction of the number of captured images in such a scenario.
Therefore, to implement a system capable of observing long and short distances, a combination of multiple optical systems with different FoVs is necessary. Pairing a large FoV, which observes a broader area over shorter distances, with a narrow FoV, which offers high resolving capabilities for detection and

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560

identification over long distances, establishes a system with a wide and flexible operational range.

III. System Implementation

The main architecture of the system is depicted in Fig. 4. The system consists of a telescope and camera as the imaging devices and a commercially available mount to pan and tilt the optical components. The extraction of the UAV position from the video is performed on a PC running a deep learning algorithm. The UAV position is sent to an FPGA, which implements suitable control strategies for feedback control of the mount to keep the UAV within the FoV. In the following an overview of the most relevant components of the system is given.

A. Optical design

For the optical system a Meade Schmidt Cassegrain telescope (LX200-ACF, Meade Acquisition Corp., Watsonville, USA) is selected, with a focal length of 2540 mm and an aperture size of 254 mm. The choice for the telescope is justified by the calculated resolution using Equation 5 as seen in Fig. 1. A larger aperture, leading to an increased system weight, does not significantly improve the resolution. As an imaging device an ASI 385 MC-Cool (ZWO Company, Suzhou, China) camera is used, with a sensor width of 7.3 mm, a sensor height of 4.1 mm, and it provides up to 67 frames per second (FPS) at a resolution of 1936x1096 pixels. A FoV of 0.16° on the horizontal axis and 0.09° on the vertical axis is reached using the telescope and camera. As stated in Section II-B, to extend the optical detection distance a narrow FoV is necessary. A DJI Mini 2 with a size of 0.29 m x 0.05 m in a distance of 3 km for example covers roughly 65x11 pixels in an image, which is enough to perform a detection using current state of the art algorithms [16].

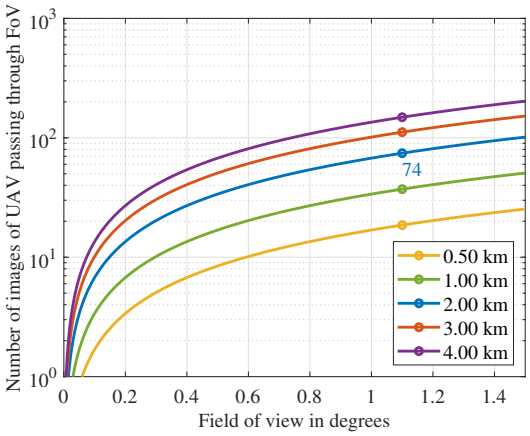


Figure 3: The number of images obtained of a UAV passing the camera FoV horizontally with a speed of 31 m/s at various distances. The camera is running with 60 FPS. For a FoV of 1.1°, 74 images are captured of a UAV passing at a distance of 2 km.

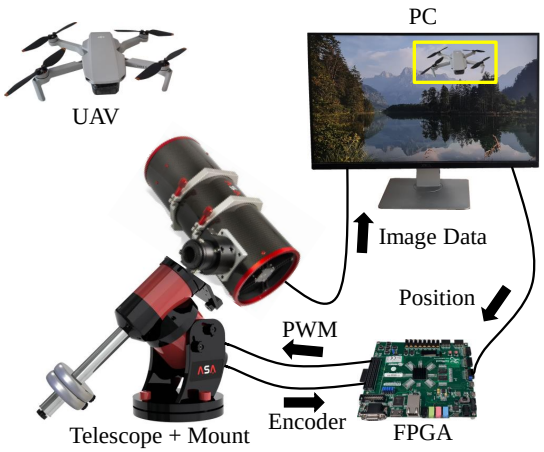


Figure 4: Overview of the system architecture consisting of the telescope and camera as the imaging system, a PC running computer vision algorithms and an FPGA implementing feedback control to actuate the telescope mount to keep the UAV within the FoV.

B. Object detection

After capturing a video stream with the camera, appropriate algorithmic methods for UAV localization within each video frame are needed. For this task a deep learning framework is selected and modified. To achieve acceptable results within short processing times, the algorithm runs on a PC equipped with a RTX 3080 GPU (Nvidia Corporation, Santa Clara, California, USA) with 10 GB of GPU RAM, an AMD Ryzen 3900 CPU (Advanced Micro Devices, Inc., Santa Clara, California, USA) and 32 GB of RAM.

A crucial perquisite to work with deep learning algorithms is an extensive and diverse dataset. A custom dataset is created consisting of six UAV multi-rotor models ranging in a diagonal size from 0.3 m to 2 m. Video material of these UAV models is gathered during flight on different days, locations and backgrounds. The video material is gathered using an ASI 385 MC-Cool and with four different telescopes to capture several field of views and distances as shown in Table I. From these videos, approximately 5.000 images are extracted every few seconds and manually labeled. Additional data is generated synthetically by blending images of drones onto background images. Using this method, data can easily be generated with the corresponding label. For this purpose several images of drones are manually cropped along their edges and in a variety of poses. These images are then blended pixel wise onto a random position within different background images. Furthermore, Gaussian blur is applied at a varying degree to these images. Using synthetic data generations, another 5.000 images are generated to create a dataset consisting of 10.000 labeled images. Fig. 6 shows the distribution of the data with regards to the bounding box size in pixels. Roughly 21 % of the images contain UAV bounding box sizes below 15x15 pixels, 25 % are between 15x15 pixels and 50x50 pixels, and the

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560

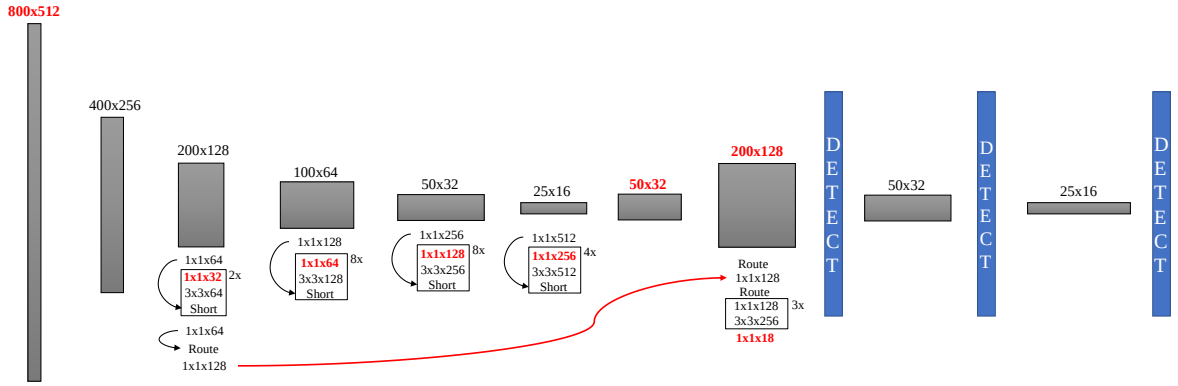


Figure 5: Proposed YOLO architecture. Marked in red are the changes made to the default architecture. Summarizing the adaptions, the network input resolution size is set to 800x512 and the number of filters is reduced in the residual blocks. Furthermore, routing from an earlier stage in the network is established as depicted by the red arrow, paired with appropriate upsampling to allow correct concatenation.

remaining 54 % are larger than 50x50 pixels. Finally, during the training process data augmentation is facilitated via the implemented methods supported by YOLOv4 [22]. Images are randomly cropped, resized and the saturation, exposure and hue are changed randomly.

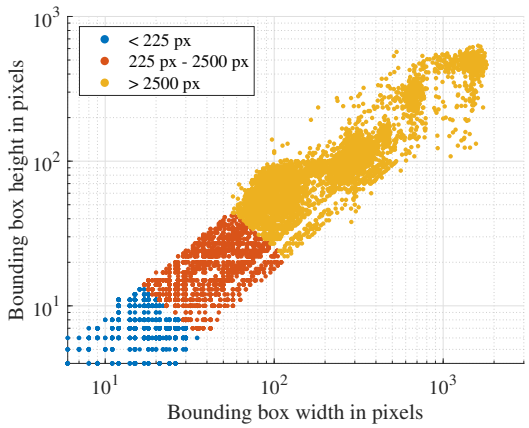


Figure 6: Distribution of the bounding box sizes of the used dataset on a logarithmic scale. The blue dots show bounding boxes smaller than 15x15 or 225 pixels, the red dots are between 225 pixels and 2500 pixels and the yellow dots represent boxes larger than 50x50 or 2500 pixels.

Table I: Telescopes used to capture training data.

Telescope	Focal length f	Aperture D
Celestron	1.25 m	0.125 m
ASA	1.14 m	0.3 m
Meade	2.54 m	0.254 m
Meade Adventure	0.36 m	0.06 m

YOLOv4 is used as deep learning algorithm for the task of object detection as it provides accurate results while operating at a high frame rate [22], [38]. In order to meet the requirements of long operational distances for the proposed setup, the capability to detect small objects comprising of only a few pixels, is desirable. Therefore, a modified YOLO architecture is proposed based on the standard YOLO approach and the alterations are summarized in Fig. 5.
YOLOv4 uses CSPDarknet53 as a backbone, which is responsible for feature extraction and deriving from these features, detection is performed with the standard YOLO head [22]. In order to improve the detection performance for small objects, the backbone is modified. In the first step, the camera frames are resized to fit the network input image size. If the network resolution is low, an object covering only a few pixels in the camera frame will almost diminish after resizing, making detection more challenging. Hence, a simple possibility for improvement is the increase of the network input resolution, as the frame provided by the camera is resized to the network input size. Additionally, a rectangular input size of 800x512 is chosen to better fit the aspect ratio of the frames provided by the ASI camera. The detection distance can be further increased, by cropping the camera frame to the network input resolution, however, this reduces the field of view of the setup, as a smaller portion of the sensor is utilized. Increasing the resolution comes with a drawback of more memory consumption and lower achievable frame rates. To compensate for this, one might reduce the number of filters, however, a network with a smaller amount of filters learns less features and therefore suffers from poor detection and classification accuracies. A known strategy to improve computation performance is the introduction of a 1x1xN block, before a 3x3xN convolution within the residual layers in order to reduce the number of filters while keeping as much information as possible [39]. In comparison to the default YOLOv4 configuration the value N, for residual layers is bisected to improve the frame rate and

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560

allow for higher input resolutions.
Similar to the residual layers, route layers are utilized within YOLOv4 to facilitate signal propagation through the deep network and to connect finer details observable in earlier layers with layers further back. This technique allows to preserve information from the earlier stages by avoiding the in-between processing [39]. For small object detection it is desirable to route from layers with a high resolution, thus the network is configured to route information from layer number 23 back to layer number 129, which brings information with the size of 208x152 to the later stages of the network [22]. To ensure mathematically correct concatenation the upsampling layers are set to 4 beforehand [22]. The modified route for small object detection is depicted in Fig. 5 by the red arrow.
Finally, the classical YOLO head consisting of three detection layers concludes the network architecture [22]. Each detection layer is responsible to detect objects of varying sizes based on anchor boxes learned from the training dataset. For the application to the custom UAV dataset, these anchor boxes are recalculated using the standard k-means clustering approach [22]. For the training process of both architectures a learning rate of 0.0007, a momentum of 0.949 and a decay of 0.0005 is set. Both models are trained for 20.000 iterations, whereas a stepwise reduction of the learning rate after 16.000 and 18.000 steps is facilitated via multiplication by a factor of 0.1. For both models the training ended with an average loss of below 0.4 % and the last model weights are chosen for the evaluation.

C. Control of the system

The control of the telescope mount is based on the system developed by Riel et al. [40]. For the current implementation a larger mount model, a DDM100 (ASA Astrosysteme GmbH, Neumarkt, Austria), is used. Additionally, the outer control loop is modified to a cascaded velocity (PI) and positional (PID) feedback loop to allow the system to react to the unpredictable motion of an approaching UAV. The velocity loop determines the current which serves as a reference input for an underlying field-oriented-control of the PMSMs motor of the DDM100 mount. The control loops are implemented on an FPGA (Avnet Inc., Phoenix, USA), and the communication between PC, which is connected to the camera and runs the deep learning algorithms for UAV detection, and FPGA is established via an UDP interface.

IV. Experiments and Results

For the experimental analysis and the evaluation of the performance in terms of detection distance, field tests are performed with the implemented setup depicted in Fig. 7. Additionally, the proposed YOLOv4 architecture is tested and compared to the default YOLOv4 configuration on a separate dataset consisting of labeled videos and images.

A. Resolution evaluation

To get an estimate of the resolving power of the presented system, experiments involving a modified version of the United States Air Force (USAF) resolution test chart are performed. Targets consisting of two white and three black beams are placed in a distance of 3.2 km to validate the imaging system as seen in Fig. 8. The width of the beams is reduced for each target to find the smallest resolvable pattern.

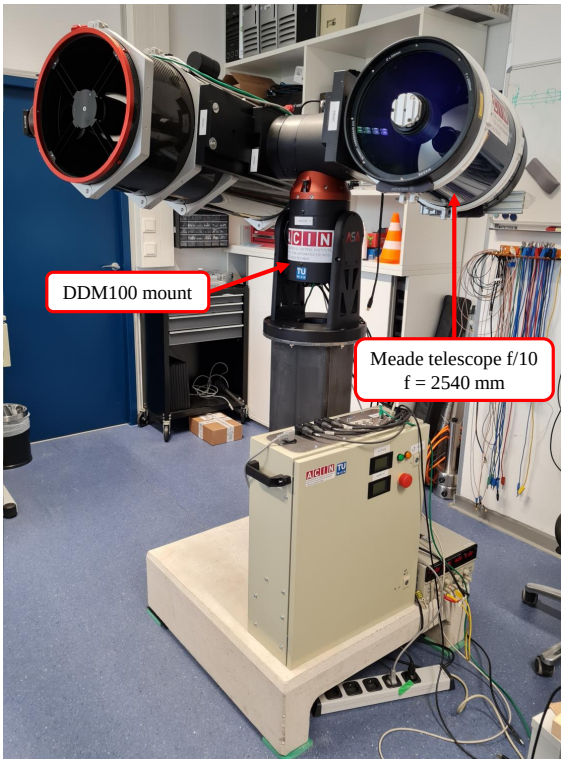


Figure 7: Implemented system with ASA DDM100 mount and two telescopes. For the evaluation of long distance UAV detection, only the Meade Schmit Cassegrain telescope is used.

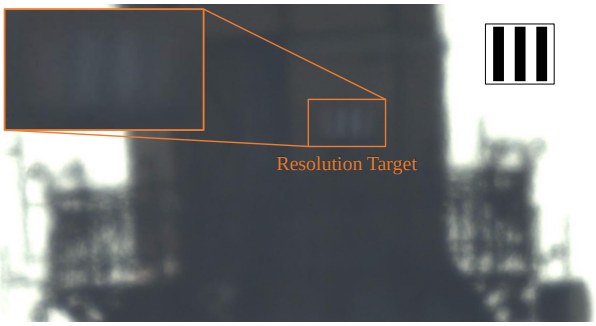


Figure 8: Example of a resolvable target with a beam width of 72 mm in a distance of 3.2 km under challenging conditions. In the top left the target, consisting of alternating black and white beams, is magnified and in the top right, the used target is displayed. The distribution of the pixel value intensities is shown in Fig. 9.

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560

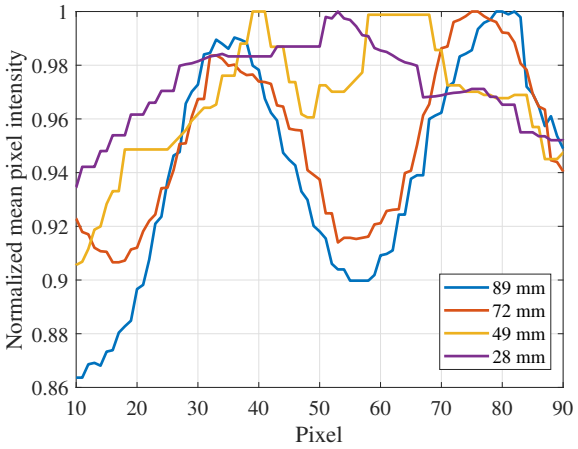


Figure 9: Mean pixel values obtained from resolution targets with various beam widths in a distance of 3.2 km. Higher pixel intensity values correspond to the white coloured beam and lower values to the black beam. Targets down to a beam width of 72 mm are distinguishable.

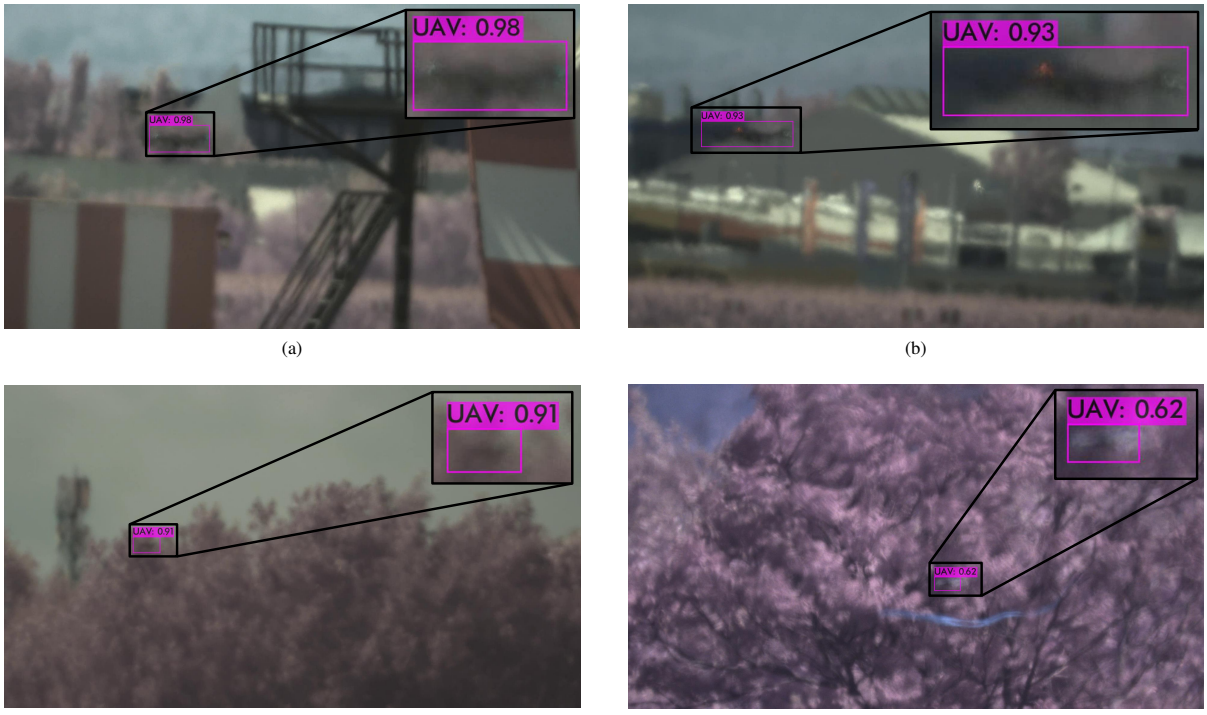
To obtain a clear image, the focus is adjusted manually without dedicated auto focus algorithm. The experiment is performed under challenging conditions, as the sun is located behind the target due to the daytime. Considering the application of UAV detection, this scenario corresponds to the detection of a UAV, which is not directly illuminated by the sun as, for example, it is flying in the shadow of a building. Fig. 9 shows the distribution of pixel intensities over the resolution target. The values are obtained by cropping the image around the target and calculating the mean of pixel intensities in the direction of the beams. Finally, the values are normalized to the corresponding maximal value for better comparison. The two peaks representing the white beams are clearly visible in Fig. 9 for target beam widths of 89 and 72 mm. As the width of the black and white beams is reduced, the two peaks become smaller and not clearly distinguishable. The decision criterion to classify a target as resolvable is linked to the number of frames captured in the video that produce distinguishable results as in Fig. 9 only a single observation can be shown. The 49 mm target from Fig. 9 represents the best image in a video sequence of roughly 2 minutes and produces a subjectively barely visible target, whereas the 72 mm bar width is constantly distinguishable throughout the whole video. Therefore, the experimentally obtained resolution is given with 72 mm in a distance of 3.2 km corresponding to a Fried parameter of 2.9 cm, which represents a typical value for daylight observations [36], [37]. This analysis shows, that detection of small UAVs, like a DJI Phantom 4 with a diagonal size of 350 mm is possible, as the resolution suffices to discern enough features. For the analysis it is assumed that the image provides enough contrast to allow a human to discriminate an object like a drone from the background.
The experiments are repeated for the vertical axis of the camera sensor and similar results are obtained as the height and width ratio of the sensor size have a similar ratio to the horizontal and vertical number of pixels.

B. Evaluation of YOLOv4 architecture

For the test phase an additional dataset is needed along with corresponding labels, which have not been used for the training phase. Therefore, the dataset described in Section III is separated into two parts before the training process, namely a training and test set. The test set is further divided into three subsets according to the ground truth bounding box sizes marking the UAVs. The separation into less than 15x15 pixels, 15x15 to 50x50 pixels and larger than 50x50 pixels is necessary to evaluate the algorithm accuracy on different object sizes. The test set consists of images showing UAVs over various backgrounds like sky, trees, buildings, clouds etc. Fig. 10 shows some example images of UAVs in front of different backgrounds. It is important to note, that the test images are obtained from different video sequences than the training image set to guarantee that the evaluation is performed with previously unseen and completely new data.

The results of applying the standard and proposed YOLOv4 configurations to the test datasets are shown in Fig. 11. The proposed YOLOv4 architecture outperforms the standard configuration around the operating frame rate of the used ASI camera considerably for object sizes below 15x15 pixels. At 65 FPS the proposed architecture achieves a mean average precision (mAP) of 0.61 compared to 0.42 of the default architecture. Likewise, for object sizes between 15x15 to 50x50 pixels the proposed model achieves a mAP of 0.86 compared to 0.81 of the default configuration. This indicates that a larger input resolution of the network, paired with a reduced number of filters, to save memory and increase execution speed, yields an improved performance. Introducing early routing within the network architecture combined with upsampling layers additionally improves the detection performance for small objects. For large objects both approaches leverage similar results. The measurements for different frame rates are obtained by altering the network input resolution showing the overall trend that higher resolutions produce better prediction capabilities. The false positive rate for both YOLOv4 architectures is neglectable when no UAV is located within the image. Table II shows the application of the models on the whole dataset with various intersection over union (IOU) thresholds for the average precision. As mentioned before, the improved precision is attributed to the success rate of the proposed model on small objects. Examining the average precision (AP) for different thresholds, the AP for an IOU threshold of 0.75 of the default architecture is significantly reduced compared to the proposed architecture. Considering the application of reliable long distance UAV detection, Table III shows the recall and precision values, when applying both YOLOv4 configurations to the three test datasets. For precision values close to 1, the respective recall indicates the amount of correctly detected UAVs, as each prediction is a correct prediction and the recall is calculated by the number of correct predictions divided by all possible ground truth objects. Therefore, it can be

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560



(a) (b) (c) (d)

Figure 10: Typical examples of detected UAVs with the corresponding bounding box in front of various backgrounds. On the top right of each image a magnification of the detection is shown.

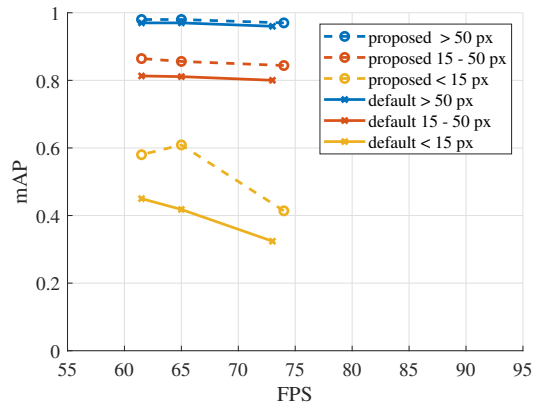


Figure 11: Mean average precision (mAP) for the three test datasets obtained with the default and the proposed YOLOv4 architecture. For small object sizes, the proposed architecture significantly outperforms the default architecture.

concluded, that the probability to detect a UAV with a size of 15x15 pixels using the proposed YOLOv4 architecture is 58% and for UAVs larger than 15x15 pixels the probability increases to 82%.

Table II: Average precision over different IOU thresholds (0.25 to 0.75) after application to whole test dataset. For strict thresholds of 0.75, the AP of the default architecture significantly drops compared to the proposed architecture.

Configuration	$AP_{0.25}$	$AP_{0.5}$	$AP_{0.75}$	FPS
Default YOLOv4	0.89	0.87	0.59	65
Proposed YOLOv4	0.93	0.91	0.70	65

Table III: Comparison of precision and recall results using an IOU threshold of 0.5. For precisions close to 1 recall values resemble the percentage of correctly detected UAVs in the absence of false positives.

	$<$ 15 px		15 - 50 px		$>$ 50 px	
Configuration	Prec.	Recall	Prec.	Recall	Prec.	Recall
Default YOLOv4	0.83	0.38	0.99	0.78	1.0	0.91
Proposed YOLOv4	1.0	0.58	0.99	0.82	1.0	0.98

In summary it is shown that the proposed system design paired with a modified deep learning algorithm extends the optical detection range significantly compared to conventional approaches, successfully demonstrating that UAVs are detected with a mAP of 86% in a distance of 3 km.

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560

V. Conclusion

A UAV detection system is developed based on a telescope and camera for image acquisition, a precise mount for alignment of the optical components with the UAV, and a deep learning algorithm for UAV localization. Following an analysis of parameters influencing the system performance, suitable hardware components are selected to create a system that increases the optical detection range of small UAVs compared to the state of the art. Proper training and modification of the deep learning algorithm show that UAVs, covering between 15x15 to 50x50 pixels, are detected with a mAP of 86%. Paired with the proposed optical system using a f/10 telescope with a focal length f of $2540\,\mathrm{mm}$ and a camera equipped with a $7.3\,\mathrm{mm}$ x $4.1\,\mathrm{mm}$ sensor, the detection range for small unidentified UAVs down to a diagonal size of $0.3\,\mathrm{m}$ is extended to more than $3\,\mathrm{km}$ under daylight conditions and sufficient contrast.

Future work will include the integration of a second telescope to add various FoVs for better situational awareness. Additionally, the impact of different light conditions on the achievable detection distances will be investigated.

Acknowledgments

This publication is funded by the Austrian defense research programme FORTE of the Federal Ministry of Agriculture, Regions and Tourism (BMLRT).

The authors gratefully acknowledge the cooperation with ASA Astrosysteme GmbH and thank for their support and valuable expertise.

Additionally, the support of the Municipal Department 31 of Vienna, Austria (Vienna Water, MA31) for providing access to the Favoriten Water Tower during measurements is gratefully acknowledged.

References

[1] C. F. Liew, D. DeLatte, N. Takeishi, and T. Yairi, "Recent developments in aerial robotics: A survey and prototypes overview," *ArXiv*, vol. abs/1711.10085, 2017.

[2] J. Pyrgies, "The UAVs threat to airport security: risk analysis and mitigation," *Journal of Airline and Airport Management*, vol. 9, no. 2, pp. 63–96, Dec. 2019.

[3] A. Solodov, A. Williams, S. A. Hanaei, and B. Goddard, "Analyzing the threat of unmanned aerial vehicles (UAV) to nuclear facilities," *Security Journal*, vol. 31, no. 1, pp. 305–324, Apr. 2017.

[4] "Air transportation safety investigation report a17q0162 - in-flight collision with drone," Sky Jet M.G. INC., Québec/Jean-Lesage International Airport, Quebec, Tech. Rep., Oct. 2017.

[5] BBC News, "Gatwick drone attack possible inside job, say police," 2010, accessed Sept 2021. [Online]. Available: https://www.bbc.com/news/uk-47919680

[6] B. Shields, "Air traffic control: How mexican cartels are utilizing drones to traffic narcotics into the united states," *Penn State J. Law Int. Aff.*, vol. 5, no. 1, p. 207, 2017.

[7] J. Bart, "Drone crash at white house reveals security risks," 2015, accessed Dez 2021. [Online]. Available: https://eu.usatoday.com/story/news/2015/01/26/drone-crash-secret-service-faa/22352857

[8] S. Yang, H. Qin, X. Liang, and T. Gulliver, "An improved unauthorized unmanned aerial vehicle detection algorithm using radiofrequency-based statistical fingerprint analysis," *Sensors*, vol. 19, no. 2, p. 274, Jan. 2019.

[9] Dedrone, "DedroneSensor RF-360," accessed Jan 2022. [Online]. Available: https://www.dedrone.com/products/hardware/rf-sensors/rf-360

[10] A. Duque De Quevedo, F. Ibanez Urzaiz, J. Gismero Menoyo, and A. Asensio Lopez, "Drone detection and RCS measurements with ubiquitous radar," in *2018 International Conference on Radar (RADAR)*. IEEE, Aug. 2018.

[11] G. Lykou, D. Moustakas, and D. Gritzalis, "Defending airports from UAS: A survey on cyber-attacks and counter-drone sensing technologies," *Sensors*, vol. 20, no. 12, p. 3537, Jun. 2020.

[12] K. Kang, J. Choi, B. Cho, J. Lee, and K. Kim, "Analysis of micro-doppler signatures of small UAVs based on doppler spectrum," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 3252–3267, Oct. 2021.

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

[14] A. Sedunov, D. Haddad, H. Salloum, A. Sutin, N. Sedunov, and A. Yakubovskiy, "Stevens drone detection acoustic system and experiments in acoustics UAV tracking," in *2019 IEEE International Symposium on Technologies for Homeland Security (HST)*. IEEE, Nov. 2019.

[15] M. Benyamin and G. H. Goldman, "Acoustic detection and tracking of a class i UAS with a small tetrahedral microphone array," Tech. Rep., Sep. 2014.

[16] H. U. Unlu, P. S. Niehaus, D. Chirita, N. Evangeliou, and A. Tzes, "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.

[17] A. G. Haddad, M. A. Humais, N. Werghi, and A. Shoufan, "Long-range visual UAV detection and tracking system with threat level assessment," in *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, Oct. 2020.

[18] J. Farlik, M. Kratky, J. Casar, and V. Stary, "Multispectral detection of commercial unmanned aerial vehicles," *Sensors*, vol. 19, no. 7, p. 1517, Mar. 2019.

[19] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep learning for generic object detection: A survey," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 261–318, Oct. 2019.

[20] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Jun. 2014.

[21] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," in *Computer Vision – ECCV 2016*, 2016, pp. 21–37.

[22] A. Bochkovskiy, C. Wang, and H. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," *CoRR*, vol. abs/2004.10934, 2020.

[23] B. K. S. Isaac-Medina, M. Poyser, D. Organisciak, C. G. Willcocks, T. P. Breckon, and H. P. H. Shum, "Unmanned aerial vehicle visual detection and tracking using deep neural networks: A performance benchmark," in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. IEEE, Oct. 2021.

[24] H. Liu, K. Fan, Q. Ouyang and N. Li, "Real-time small drones detection based on pruned YOLOv4," *Sensors*, vol. 21, no. 10, 2021.

[25] J. Park, D. H. Kim, Y. S. Shin, and S. Lee, "A comparison of convolutional object detectors for real-time drone tracking using a PTZ camera," in *2017 17th International Conference on Control, Automation and Systems (ICCAS)*. IEEE, Oct. 2017.

[26] Y. Zheng, Z. Chen, D. Lv, Z. Li, Z. Lan, and S. Zhao, "Air-to-Air Visual Detection of Micro-UAVs: An Experimental Evaluation of Deep Learning," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1020–1027, apr 2021.

[27] F. Gonzalez, R. Caballero, F. J. Perez-Grau, and A. Viguria, "Vision-based UAV Detection for Air-to-Air Neutralization," in *2021 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. IEEE, oct 2021.

[28] R. Carnie, R. Walker, and P. Corke, "Image processing algorithms for UAV "sense and avoid"," in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.* IEEE, 2006.

[29] J. James, J. J. Ford, and T. L. Molloy, "Learning to Detect Aircraft for Long-Range Vision-Based Sense-and-Avoid Systems," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4383–4390, oct 2018.

[30] "ctrl+sky drone detection and neutralization system - datasheet," Advanced Protection Systems LLC, New Yersey, USA, Sydney, Australia, 2018, accessed Dez 2021. [Online].

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560

Available: https://www.southerncrossdrones.com/download/aps-counter-drone-technology-sxd-.pdf
[31] Aselsan, "IHTAR anti-drone system - datasheet," ASELSAN A.S., Ankara, Türkiye, 2018.
[32] *Easy Access Rules for Unmanned Aircraft Systems (Regulation (EU) 2019/947 and Regulation (EU) 2019/945)*, European Union Std., 2019.
[33] Rayleigh, "XXXI. Investigations in optics, with special reference to the spectroscope," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 8, no. 49, pp. 261–274, Oct. 1879.
[34] D. L. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," *Journal of the Optical Society of America*, vol. 56, no. 10, p. 1372, Oct. 1966.
[35] R. Tyson, *Principles of Adaptive Optics*. CRC Press, Sep. 2010.
[36] A. V. Sergeyev and M. C. Roggemann, "Monitoring the statistics of turbulence: Fried parameter estimation from the wavefront sensor measurements," *Applied Optics*, vol. 50, no. 20, pp. 3519–2528, Jul. 2011.
[37] T. Özışık and T. Ak, "First day-time seeing observations at the TÜBiTAK national observatory in turkey," *Astronomy & Astrophysics*, vol. 422, no. 3, pp. 1129–1133, Jul. 2004.
[38] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Scaled-YOLOv4: Scaling cross stage partial network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2021, pp. 13 029–13 038.
[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385
[40] T. Riel, A. Galffy, G. Janisch, D. Wertjanz, A. Sinn, C. Schwaer, and G. Schitter, "High performance motion control for optical satellite tracking systems," *Advances in Space Research*, vol. 65, no. 5, pp. 1333–1343, Mar. 2020.

Denis Ojdanić is a doctoral researcher at the Automation and Control Institute (ACIN), group for Advanced Mechatronic Systems (AMS) at TU Wien, Vienna, Austria. He received his MSc. degree in electrical engineering from TU Wien in 2019.
His primary research interests are object detection, tracking and identification using high-performance telescope systems.

Andreas Sinn is a doctoral researcher at the Automation and Control Institute (ACIN), group for Advanced Mechatronic Systems (AMS) at TU Wien, Vienna, Austria. He received his MSc. degree in electrical engineering from TU Wien in 2016. His primary research interests are adaptive optics and system integration for optical ground stations, as well as high-performance telescope systems for object tracking and identification.

Christopher Naverschnigg is a doctoral researcher at the Automation and Control Institute (ACIN), group for Advanced Mechatronic Systems (AMS) at TU Wien, Vienna, Austria. He received his MSc. degree in electrical engineering from TU Wien in 2021. His primary research interests are modelling and contol of mechatronic systems, especially high-performance telescope systems.

Georg Schitter (Senior Member, *IEEE*) received the B.Sc. degree in electrical engineering from TU Graz, Graz, Austria, in 2000, and the M.Sc. degree in information technology, and Ph.D. degrees in control engineering and nanotechnology from ETH Zurich, Zurich, Switzerland, in 2004.
He is Professor for Advanced Mechatronic Systems with the Automation and Control Institute (ACIN), TU Wien. His primary research interests are on high-performance mechatronic systems, particularly for applications in the high-tech industry, scientific instrumentation, and mechatronic imaging systems, such as AFM, scanning laser and LIDAR systems, telescope systems, adaptive optics, and lithography systems for semiconductor industry. Prof. Schitter was the recipient of the Journal Best Paper Award of *IEEE/ASME Transactions on Mechatronics* (2018), of the *IFAC Mechatronics* (2008–2010), of the *Asian Journal of Control* (2004–2005), and the 2013 *IFAC Mechatronics Young Researcher Award*. He was an Associate Editor for *IFAC Mechatronics, Control Engineering Practice*, and for the *IEEE TRANSACTIONS ON MECHATRONICS*.

Post-Print version of the article: Denis Ojdanić, Andreas Sinn, Christopher Naverschnigg and Georg Schitter, "Feasibility Analysis of Optical UAV Detection Over Long Distances Using Robotic Telescopes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 5, pp. 5148-5157, 2023. DOI: 10.1109/TAES.2023.3248560