Technical report
IT-CDT 2023-08

Detection of aircraft traces in nighttime all-sky camera images using deep learning

A new way to improve aerial safety in Satellite Laser Ranging operations.

Carlos Hernández Fernández
José Carlos Rodríguez
Universidad de Alcalá de Henares
Yebes Observatory / IGN Spain

7th November 2023

Unión Europea
Fondo Europeo de Desarrollo Regional
"Una manera de hacer Europa"
GOBIERNO DE ESPAÑA
MINISTERIO DE TRANSPORTES, MOVILIDAD Y AGENDA URBANA
INSTITUTO GEOGRÁFICO NACIONAL
YEBES

Acknowledgements

The image data employed for the work described in this report have been taken with an all-sky camera purchased within the YDALGO project, financed with Multi regional Operational Programme for Spain 2014-20 ERDF funds.

Note

The results obtained in this work have been presented at the International Laser Ranging Service (ILRS) Virtual International Workshop on Laser Ranging 2023, celebrated online on 16–20 October, 2023. The video of the talk given by Carlos Hernández can be found on the workshop website[1] (Session 2, New Technologies and Applications).

Contact

Instituto Geográfico Nacional
C/ General Ibáñez de Ibero, 3
28003, Madrid
Spain

email: jcrodriguez@mitma.es

[1] https://ilrs.gsfc.nasa.gov/lw23/index.html

Contents

Abstract

All-sky cameras imagery is employed in Satellite Laser Ranging (SLR) stations during operations to gather information about the observing conditions, typically as an sky awareness tool for operators. Due to the emission of laser light towards the sky in the SLR technique, it is a priority to ensure aerial safety prior to and during operations. For this purpose, a range of techniques are employed, such as active radars, cameras, airplane spotters operating kill-switches, and Automatic Dependent Surveillance-Broadcast (ADS-B) receivers. Thanks to the aviation regulatory environment in many parts of the world, which mandates the equipment of ADS-B for a broad set of aircraft, these devices are nowadays very popular in SLR stations. ADS-B is a cooperative surveillance technology wherein an aircraft determines its position through satellite navigation and transmits it periodically, enabling its precise tracking. Currently, the primary limitation with this technology is that not all aircraft are required to equip it, which calls for alternative detection systems. In this work we demonstrate the automatic detection of aircraft using neural networks to analyse nighttime images from an all-sky camera, and propose that such a system can provide complementary information to that obtained with ADS-B, therefore increasing the safety of the operations.

1 Introduction

Aerial safety stands as a pivotal concern in Satellite Laser Ranging. Accurate aircraft detection within the operational zone is paramount and ensures the required safety for these systems. Presently, the ADS-B (Automatic Dependent Surveillance-Broadcast) technology is employed as one of the primary methods for detecting aircraft in proximity to the operational zone.

The method proposed in this study aims to streamline the aircraft detection task by utilizing convolutional neural networks and binary semantic segmentation to analyze nocturnal images acquired through a wide-field-of-view camera. These tools will yield predictions for each given image, assigning a probability to each pixel of belonging to an aircraft trace. These techniques have been employed in similar approaches such as the detection of ship traces from satellite images, and have demonstrated their effectiveness compared to other methods such as using the Hough transform to detect lines in an image.

1.1 Automatic Dependent Surveillance-Broadcast

An ADS-B transponder, a radiofrequency device installed aboard aircraft, facilitates real-time knowledge of their position and heading. Some of the transmitted data include aircraft identification, altitude, elevation, and azimuth.

While traditional airborne radar tracks aircraft by reflecting transmitted radar waves, ADS-B actively broadcasts aircraft position data. In this process, participants independently determine their positions using navigation satellites and broadcast this information in L-band (1090 MHz).

This standard system is relatively young and has only been in existence since 2005. In Europe, ADS-B has been mandatory for large aircraft with a takeoff weight equal to or greater than 5.7 metric tons or a true airspeed greater than 250 KTAS (knots true airspeed) during flight, starting the 7th of December, 2020.

The primary issue this technology faces in the context of SLR operations is that not all aircraft are mandated to carry an ADS-B transponder on board. Typically, smaller aircraft such as gliders, ultralights, and certain military aircraft might be exempt from carrying an ADS-B transponder under specific circumstances. Regulations can vary based on the country and relevant civil aviation authority. Hence, the utilization of additional and complementary detection techniques is advisable to broaden the spectrum of detected aircraft. Thus, active radars have been in use in many SLR stations as the primary aircraft detection system. However, this technique can not be employed

in observatories that operate Very Large Baseline Interferometry (VLBI) for either geodesy or radioastronomy, due to radiofrequency interference. ADS-B data will be utilized in this study as the ground truth for aircraft positions, enabling comparisons with the examined detection system.

1.2 All-sky Camera

A whole sky camera is a specialized device used in meteorology and astronomy among other fields to capture an image of the entire sky. An example of an application is hemispherical photography, used to study plant canopy geometry and calculate near-surface solar radiation. At the Yebes Observatory, the all-sky camera is employed in SLR observations to obtain information about the atmospheric conditions, as a visual aid for system operators [1].

The images acquired by this camera will be employed in this study to train the network and conduct detections.

Camera Information:

- Camera: OMEA-8C (Alcor System) [2]
- Internal: ZWO ASI294MC Pro
  - Back-illuminated CMOS, high QE
  - Large quantum well (for long exposures)
  - Cooled sensor (-35°C DeltaC)
  - Low noise and good sensitivity
  - Sky images at 2800x2800 (color)
- Automatic aperture
- Temperature and relative humidity sensors
- Anti-condensation system

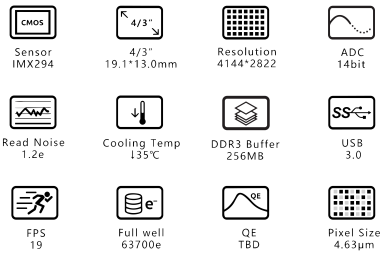

Figure 1: Characteristics of the sensor mounted in the all-sky camera.

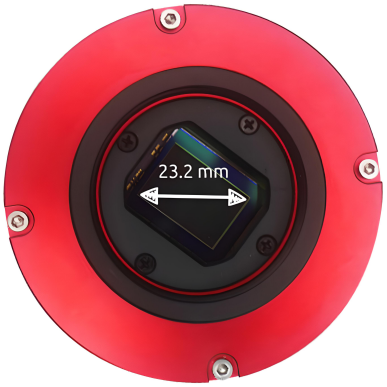

Figure 2: ZWO ASI294MC Pro.

1.3 AI, Machine Learning and Deep Learning

Artificial intelligence, commonly abbreviated as AI, entails the pursuit of endowing computers with capabilities that mimic human cognitive functions. These functions include tasks like reasoning, problem-solving, perception, language understanding, and learning from experiences. AI aims to create systems that can perform tasks that, when executed by humans, typically require human intelligence. It encompasses a spectrum of techniques and approaches, one of which is machine learning.

Machine learning is a subset of AI that revolves around the concept of enabling computers to improve their performance on a specific task by learning from data rather than being explicitly programmed.

Figure 3: All-sky camera in Yebes observatory

In traditional programming, a human programmer writes explicit instructions for a computer to execute. In contrast, machine learning algorithms adjust their internal parameters based on data examples, enabling them to make predictions or decisions without explicit programming for each scenario.

Deep learning represents a branch of machine learning characterized by the utilization of artificial neural networks. These networks are inspired by the interconnected structure of neurons in the human brain and consist of layers of interconnected nodes. Deep learning methods employ neural networks with numerous layers, hence the term "deep." These deep architectures enable the automatic extraction of hierarchical features from input data, allowing the system to discern intricate patterns that might be imperceptible through traditional approaches.

In essence, deep learning refines the concept of machine learning by employing complex neural networks capable of learning and representing data in progressively abstract ways. This ability to automatically uncover relevant features and representations has led to groundbreaking advancements in various fields, including image and speech recognition, natural language processing, autonomous vehicles and scientific research.

1.4 Classification, Object Detection, Semantic Segmentation and Instance Segmentation

In the realm of deep learning and computer vision, classification, object detection, semantic segmentation and instance segmentation are distinct tasks that deal with understanding and analyzing visual data in varying degrees of complexity.

Classification is one of the foundational tasks in machine learning and computer vision. It involves assigning a label or a category to an input image or object. In the context of images, this means determining what object or scene the image represents. For instance, given an image of a cat, a classification model would predict that the image contains a "cat."

Object detection takes classification a step further by not only identifying objects in an image but also localizing their positions. In object detection, the goal is to detect and locate multiple objects within an image, while also classifying them into different categories. Object detection algorithms output bounding boxes around the detected objects, along with their corresponding class labels.

Semantic segmentation is a more intricate task that involves pixel-level labeling of an image, where each pixel is assigned a class label.

Unlike object detection, which focuses on bounding boxes, semantic segmentation provides a detailed understanding of the spatial extent of different objects within an image. In other words, it divides the image into meaningful segments and assigns a label to each segment, effectively creating a pixel-wise map of object categories.

Instance segmentation combines elements of both object detection and semantic segmentation. This task goes beyond simply identifying object categories and their locations; it aims to provide a pixel-level understanding of objects within an image while distinguishing between individual instances of the same object class.

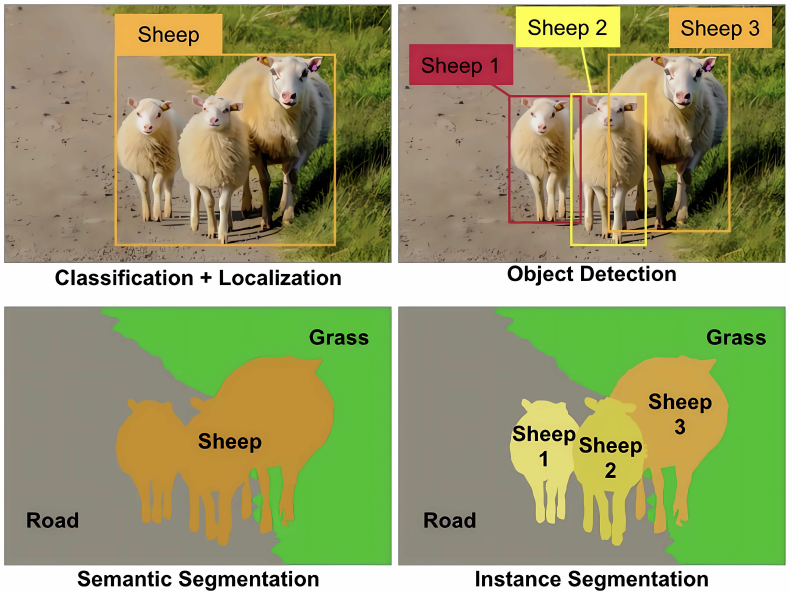

Figure 4: Differences between classification, object detection, semantic segmentation, and instance segmentation [3]

For this study, a model based on binary semantic segmentation will be trained. In this manner, a pixel-level class map will be obtained indicating the locations of aircraft traces.

2 Datasets

Given that the all-sky camera stores images in FITS format, establishing the conversion parameters from FITS to PNG is necessary, as PNG is a widely used image format compatible with most image processing libraries and deep learning frameworks, in order to create different datasets. An astrometric calibration has been necessary to establish the relationship between positions in the image and celestial coordinates [4]. In total, no more than 2000 images have been employed for the training and testing of the network. The ADS-B data used was collected between the 10th of May, 2021, and the 14th of August, 2021. Certain adjustments have been made to the data, including the conversion of

CEST to UT format for the time values.

The preprocessing of images for their treatment in training and processing with neural networks is an important step for their success. Typically, images undergo certain preprocessing steps such as normalization, thresholding, removal or combination of different channels, smoothing, resizing, etc. In this study, two preprocessing methods have been experimented with, providing two different ways of visualizing the images. The first one follows the observatory's conventional approach, enabling images to be viewed naturally, resembling photographs.

For the second visualization method, a more intricate procedure has been undertaken. In the conversion from FITS to PNG one has take care of dealing with artifacts that may impact negatively the performance of the neural network, such as hot pixels, and the dynamic range of the pixel values is reduced to 8-bit per channel. A suitable pixel value scaling has to be employed, so that the relevant details of the image are suitably preserved. Here we have used a square root stretching with minimum and maximum threshold values determined empirically for our data set. For the second data set, prepared to enhance the signal to noise ratio of aircraft traces, and combat the effect of light pollution close to the horizon, we subtract the background from the image prior to its conversion to PNG. This was accomplished with the background estimation found in the software SExtractor [5]. Briefly, this consists in the creation of a background map interpolated from the local background in each mesh of a regular grid in the image, which are computed iteratively around the median pixel values. Details are found in the online documentation of SExtractor. With the background subtracted, the images are seemingly less affected by light pollution, and faint aircraft traces appear more prominent.

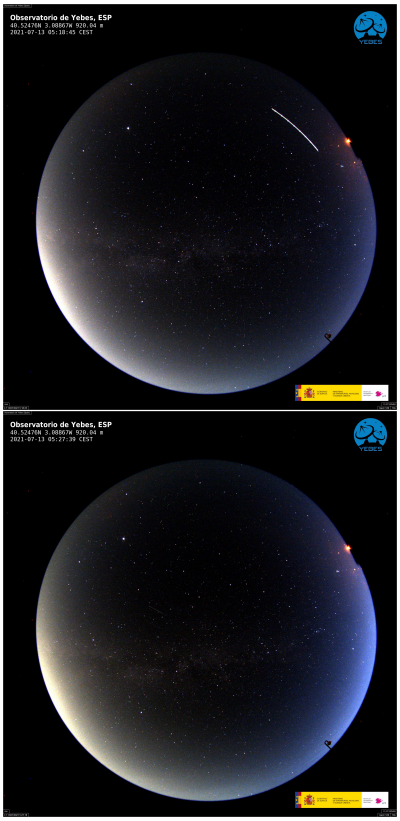

Figure 5: Images from the all-sky Camera

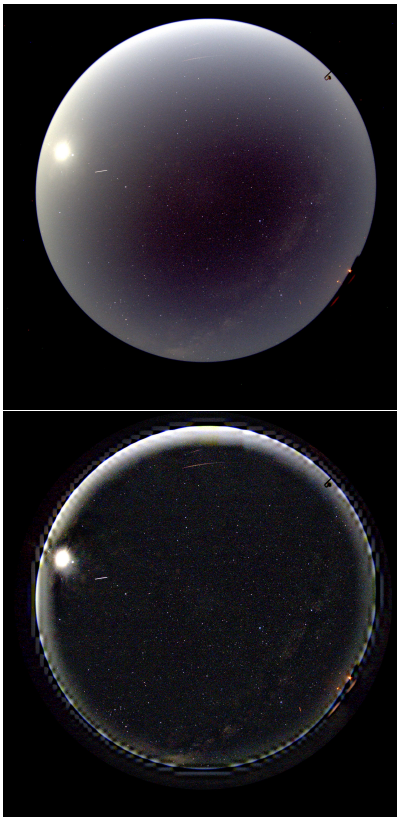

Figure 6: Classical (above) and enhanced (below) conversion methods

2.1 Data

The data utilized for this project comes in two forms:

- Nighttime all-sky imagery (FITS files) [6]
- ADS-B data for the same period (stored in Feather format after pre-processing)

For the training process, nocturnal images captured from the 28th of June, 2021, to the 30th of June, 2021, have been employed. Subsequently, for network testing, nocturnal images taken from the 01st of July, 2021, to the 21st of July, 2021, were utilized.

2.2 Training Dataset

For this project, only 10-second exposure images have been utilized, and a training data set has been constructed, consisting of 520 images along with their respective masks, using the "classical view" conversion method. To generate these images, a selection of frames displaying aircraft traces was made, and segments of these frames were extracted. VGG Image Annotator [7] tool was employed to generate binary masks for the images, assigning a value of 1 to aircraft pixels and 0 to the rest.

Using these images and their corresponding masks, the training dataset for the neural network is constructed. The ResNet50 network, a Residual Network with 50 layers, has emerged as one of the leading networks for semantic segmentation [8]. This work utilizes the network as the back-bone architecture to process the sky images. For the implementation of the network, the TensorFlow [9] and Keras libraries in Python were utilized.

One of the key steps to ensure good convergence and learning of the model is the selection of the network's hyperparameters. *Learning rate* is a hyperparameter that controls the magnitude of weight updates during the training process, influencing the convergence and learning speed of the model. *Batch size* represents the number of training examples used in each iteration of the optimization algorithm.

After conducting the necessary experiments, it has been determined that the optimal combination of hyperparameters consists of a learning rate of 10^{-4}, a batch size of 16 with 50 epochs.

Another crucial parameter for the network is the optimizer, used to update the network's weights based on the computed gradients during backpropagation. The Adam optimizer [10] adapts the learning rate based on the distribution of parameters within the model.

Additionally, it is important to establish an appropriate loss function. The Jaccard loss function is a measure of dissimilarity between two sets of objects.

Lastly, during the training, the Intersection over Union (IoU) metric is calculated. IoU metric is a measure of similarity between two sets of pixels in an image. In the context of semantic segmentation networks, IoU is used to evaluate the quality of the segmentation.

2.3 Validation Dataset

The training data set is split into a training set and a test set using a 75-25 ratio. This aids in verifying the network's effective convergence across different combinations of hyperparameters.

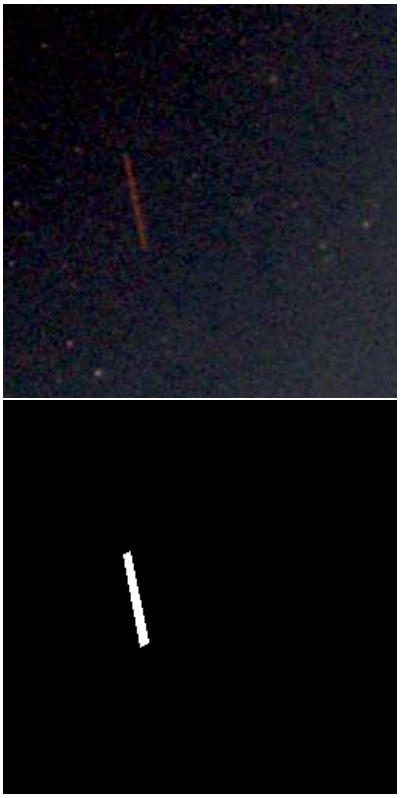

Figure 7: Example 1 of training image and mask

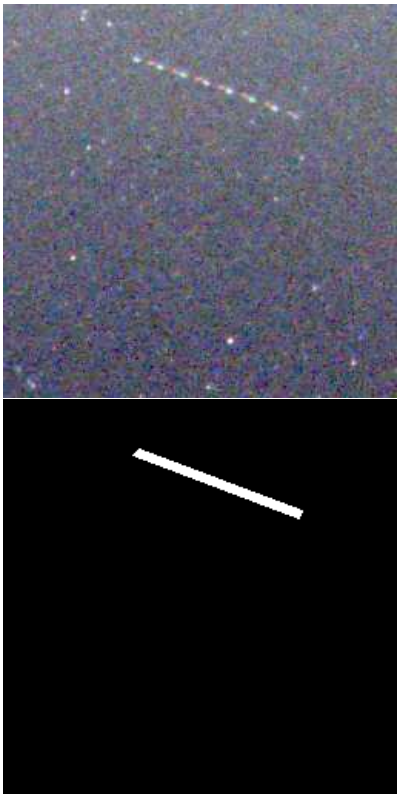

Figure 8: Example 2 of training image and mask

2.4 Test Dataset

Two training data sets have been constructed by selecting 750 random images taken between the 15th and 17th of July. One of the data sets was curated using "classical view", while the other utilized "enhanced view", enabling a comparative analysis of their outcomes.

From this data sets, metrics have been extracted by comparing predicted detections from the network with ADS-B data, including the comparison of the successful trace detection percentage corresponding to aircraft registered by ADS-B versus false positives.

These false positives can have different natures: they can be genuine false positives, i.e., detections made by the network where there is no actual trace, and predicted detections by the network where there is a trace that is not registered in ADS-B data.

2.5 Data augmentation

To enhance the network's performance, data augmentation [11] has been utilized. This process generates additional images during the training process by applying various transformations to the images within our data set, such as zooms, shears, flips, and shifts. This approach generates a new image for each training step. As our training data set consists of 520 images, and a batch size of 16 has been selected, each training epoch comprises 33 training steps. As we train the model for 50 epochs, the overall number of images generated due to data augmentation amounts to 1650. As a result, we effectively quadruple the size of our training data set.

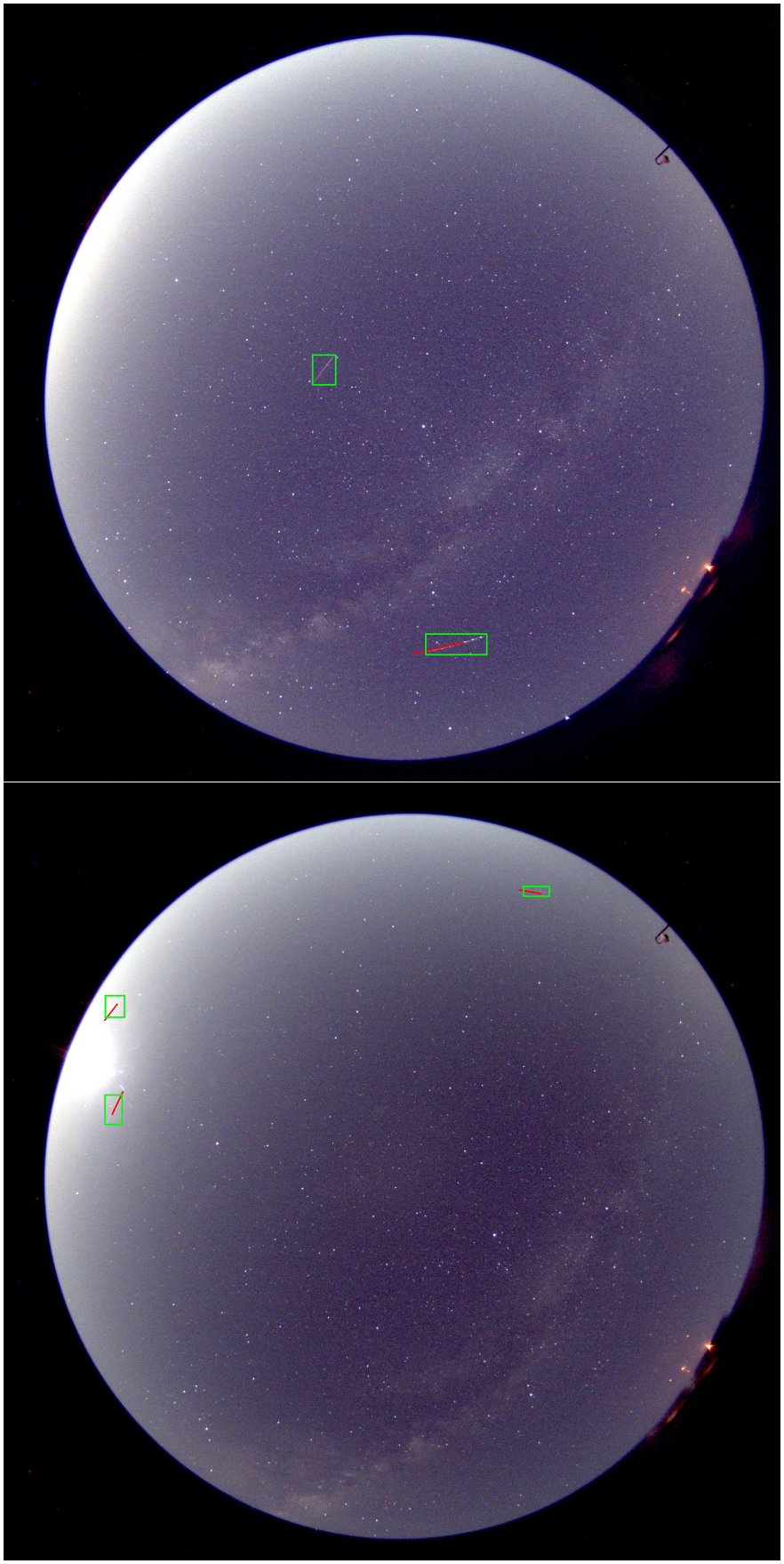

Figure 9: Output images showing ADS-B traces (red) and net detections (green)

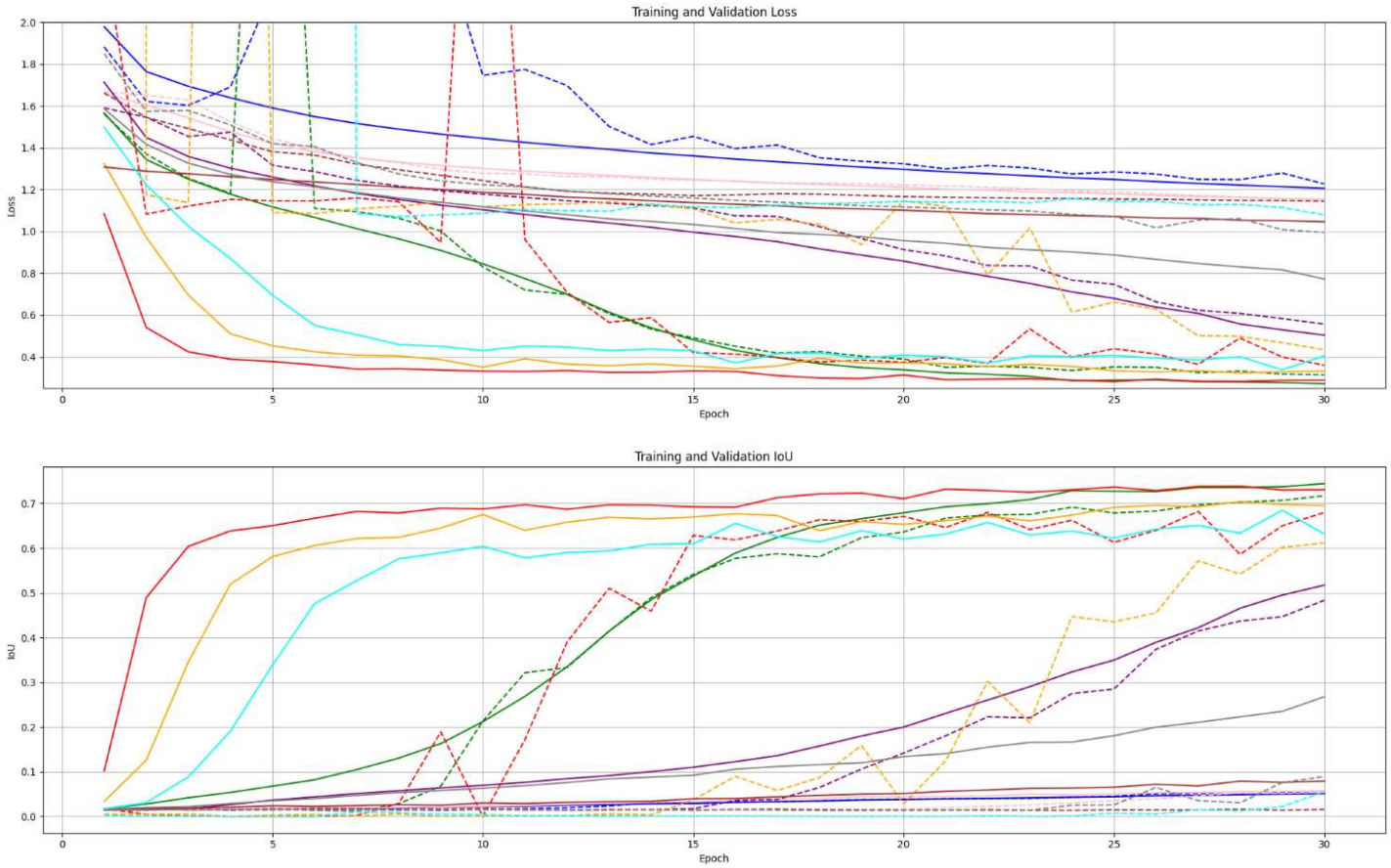

Figure 10: Training and Validation IoU comparison across different combinations of hyperparameters.

3 Methods

Once the chosen and trained neural network model is ready, a code can be constructed to enable the processing of new images. The code used in this research is an adaptation of the one published in [12]. Some adjustments have been made to this code, such as changing the BACKBONE from ResNet34 to ResNet50, and transitioning from working with TIF files to working with PNG files. At this juncture, the test data set will be employed to extract performance metrics for the model. The proposed code acquires images with dimensions of 2822×2744 pixels from the camera, removes black borders to achieve images of 2560×2560 pixels, and then segments each image into patches of 256×256 pixels, which is the expected input size for the network. Subsequently, each of these patches is processed by the network, resulting in a predicted mask. Once the 100 masks for an image are obtained, they are reassembled to generate the complete mask of the processed image. From this mask, contours are extracted for predicted traces that surpass a certain area threshold. These contours are then drawn onto the original image, enhancing clarity.

Similarly, by accessing the ADS-B data for the time period covered by the image, aircraft trajectories from that data are drawn on the image. After this process, nighttime sky images are obtained with ADS-B traces marked in one color and traces predicted by the network marked in another, as depicted in Figure 9.

Using these new images, deriving metrics such as the percentage of ADS-B traces detected by the network

$$\%\text{ of ADS-B traces detected} = \frac{\text{ADS-B traces} \cap \text{Predicted traces}}{\text{ADS-B traces}} \quad (1)$$

and the percentage of false positives becomes straightforward.

$$\%\text{ of false positives (based on ADS-B data)} = \frac{\text{Erroneous detections}}{\text{Total detections}} \quad (2)$$

Where erroneous detections encompass both network predictions in areas where there is no trace as well as network predictions where there is a trace, but it is not indicated by the ADS-B data.

4 Processing time

A key topic for the feasibility of an implementation of a detection system based on deep learning is the processing time required for the analysis. Although our efforts have focused on the question of detectability, we report here some timing tests, which are limited by our access to GPU units.

We have worked with images taken with a 10-second exposure time. This decision was made based on the availability of the acquired images. Further tests will be necessary to determine the minimum required exposure time for successful aircraft detection.

Table 1: Models used and their corresponding processing times.

Model	Type	Processing time (s)
Intel Xeon CPU @ 2.3 GHz (Colab)	CPU	25
Intel Core i5-8500 CPU @ 3 GHz	CPU	16.5
AMD Ryzen 5 7600X @ 4.7 GHz	CPU	10

Using GPU is crucial for achieving real-time implementation of the model. It's important to note that the processing time must be added to the 10-second exposure time of the image and the time it takes for the PC to receive the image from the camera. For this reason, it's essential to minimize the processing time of the network as much as possible if a real-time and reliable system is desired.

5 Results

After processing the 750 images twice (one for each visualization option), the network yields the following outcomes.

5.1 "Original view" data set

The results obtained using the "Original view" data set (Figure 11) exhibit an aircraft detection rate of 69% for those marked by ADS-B. Out of the total neural network detections, these ADS-B trace detections constitute 29.1%.

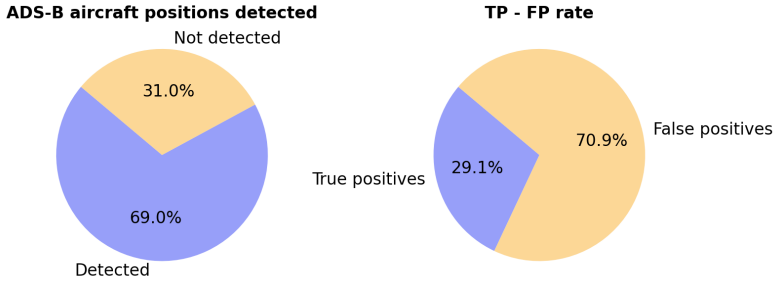

Figure 11: ADS-B detection rate (left) and TP-FP rate (right) with the "original view" test data set

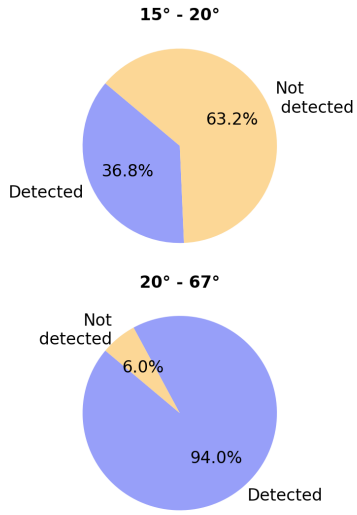

Figure 12: Plane detection based on elevation (Original view)

Another interesting metric that can be obtained by manually processing a smaller set of images (100 out of 750) is the percentage of detections made by the network that are not ADS-B traces but are visible traces of some kind. This set of traces may include aircraft without ADS-B transponders and satellites.

As observed in Figure 13, out of the total network detections, 35.2% correspond to ADS-B aircraft detections. 48.4% of the detections correspond to traces not identified by ADS-B, which could be aircraft without onboard ADS-B transponders or satellites. Lastly, only 16.5% of the detections are false positives. We note that many false positives found in these data set were found in images where the Moon was present, creating spurious reflections. This circumstance is not problematic as it can be easily filtered. In a similar way, other situation where fleeting false positives can be found are the presence of a particularly bright planet in the images.

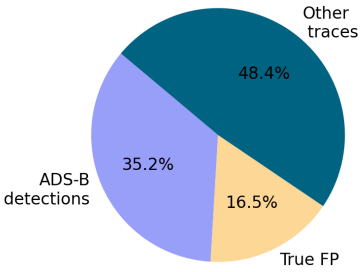

Figure 13: Other traces detected by the net

In Figure 14, we can observe that the network's main challenge lies in low-elevation aircraft. This is attributed to the level of light pollution present along the horizon in the all-sky camera images. Also, the maximum distance for an aircraft at the zenith is its altitude, approximately up to 12 km. However, in the case of an aircraft at low elevation, the distance between the aircraft and the camera increases. Above 20 degrees of elevation, the system has been capable of detecting 94% of the aircraft whose positions are known through ADS-B.

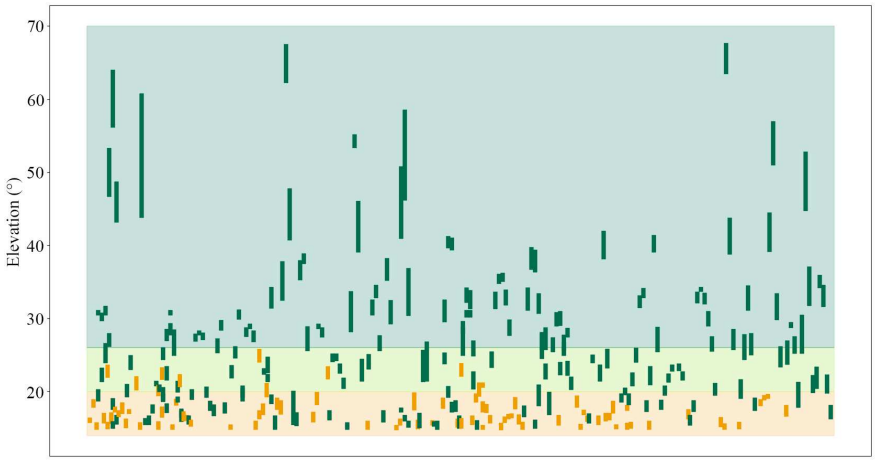

Figure 14: Plane detection based on elevation

5.2 "Enhanced view" data set

When processing the images using the "Enhanced view," we observe similar results. 72.1% of the ADS-B aircraft are detected. Out of the total neural network detections, these ADS-B trace detections constitute 6.6%. As evident in this case, with the "Enhanced view," the detection of ADS-B traces improves slightly, while the number of false positives increases significantly.

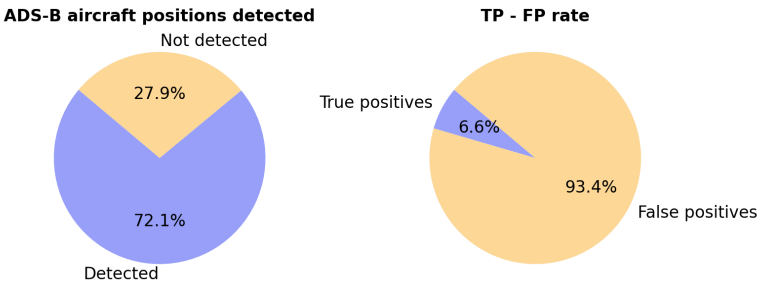

Figure 15: ADS-B detection rate (left) and TP-FP rate (right) with the "enhanced view" test dataset

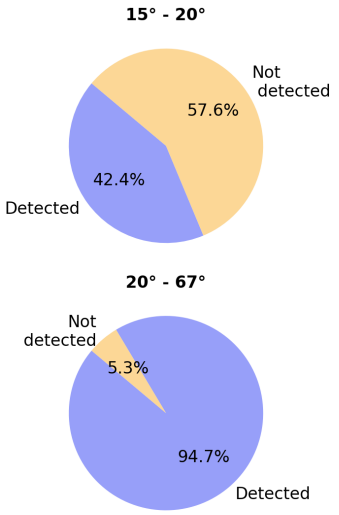

Figure 16: Plane detection based on eleva-

The increase in the number of false positives using the "enhanced view" data set may be attributed to several factors. Firstly, the training data set still consists of images with the "original view," which could potentially confuse the neural network. Ideally, for this case, it would be optimal to construct a sufficiently large and diverse data set with enhanced images. Another reason for this increase could be that when enhancing the aircraft traces, certain image features such as edges or sets of aligned stars are also enhanced, which may create the impression of being a trace.

If we now take another look at the elevations of ADS-B aircraft, we can observe that the results are practically identical to those obtained

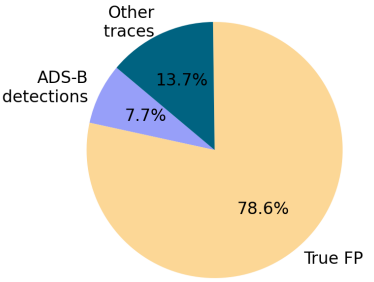

Figure 17: Other traces detected by the net

with the "Original view." The main issue with the network still lies in detecting aircraft at low elevations. In this case, starting from 20 degrees of elevation, the network is capable of detecting 94.7% of the aircraft. Examining the low-elevation detections (<20°) and comparing them to the previous results, we can observe an improvement. 42.4% of the low-elevation ADS-B traces are detected, compared to 36.8% in the previous case (using the "original view" data set).

As observed in Figure 17, out of the total network detections, 7.7% correspond to ADS-B aircraft detections. 13.7% of the detections correspond to traces not identified by ADS-B, which could be aircraft without onboard ADS-B transponders or satellites. Lastly, 78.7% of the detections are false positives.

As mentioned before, in this new image representation, a higher percentage of ADS-B trace detections is achieved. On the other hand, the percentage of false positives increases significantly as a trade-off.

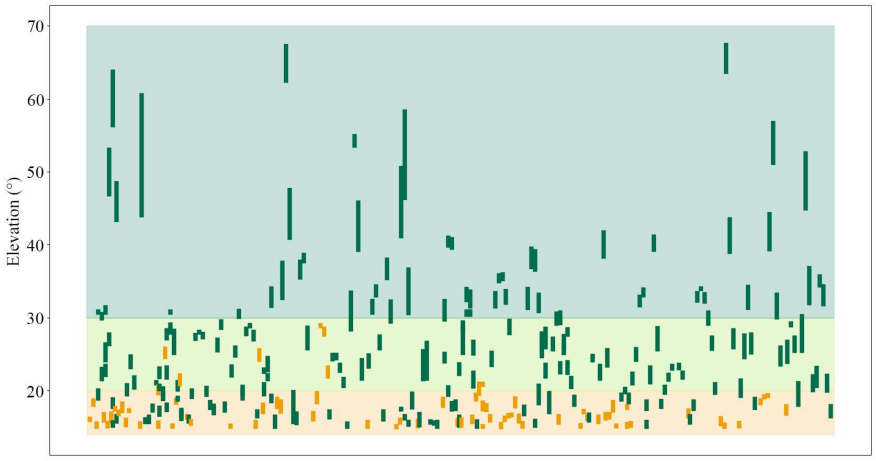

Figure 18: Plane detection based on elevation

Figure 18 illustrates how the distribution of ADS-B aircraft detection concerning elevation follows a similar pattern to the previous case. The network encounters challenges in identifying all traces below 20°. Between 20° and 30°, there are some detection losses, but in this range, the vast majority are detected. Above 30°, 100% of the ADS-B traces are detected.

6 Conclusions and future work

The presented work demonstrates the potential for enhancing aircraft detection using neural networks. This system would enable the identification of aircraft that currently lack an onboard ADS-B transponder. Given a system with sufficient GPU capabilities, real-time implementation would be feasible. The primary challenge of the network is the

detection of aircraft at elevations lower than 20 degrees due to the high levels of light pollution present at the horizon. Future work in the same direction could involve exploring a multi-class semantic segmentation network, allowing for the detection of not only aircraft traces but also clouds, satellites, meteors, and more. Additionally, the creation of a more comprehensive training data set, incorporating a greater number of images and more precisely generated masks, has the potential to enhance the model's performance. Finally, for the implementation, a user interface could be developed to consolidate both ADS-B data and network predictions into a single platform, streamlining the workflow for technicians and operators.

References

[1] A. García Marín et al. *YLARA sensors installation and outdoor tests*. Tech. rep. Yebes Observatory, 2022. URL: https://icts-yebes.oan.es/reports/doc/IT-CDT-2022-6.pdf.

[2] *Alcor System - AllSky Camera*. Consulted on 15-Sept-2023. URL: https://www.alcor-system.com/new/AllSky/Omea_camera.html.

[3] *Image Classification vs Semantic Segmentation vs Instance Segmentation*. Consulted on 15-Sept-2023. URL: https://nirmalamurali.medium.com/image-classification-vs-semantic-segmentation-vs-instance-segmentation-625c33a08d50.

[4] J.C. Rodríguez et al. *Astrometric calibration of Yebes all-sky camera*. Tech. rep. Yebes Observatory, 2023. URL: https://icts-yebes.oan.es/reports/doc/IT-CDT-2023-4.pdf.

[5] E. Bertin and S. Arnouts. 'SExtractor: software for source extraction'. In: *Astronomy and Astrophysics Supplement* (317 1996).

[6] *Yebes Observatory All Sky Images*. Consulted on 15-Sept-2023. URL: https://icts-yebes.oan.es/allsky/.

[7] *VGG Image Annotator*. Consulted on 15-Sept-2023. URL: https://annotate.officialstatistics.org/.

[8] Kaiming He et al. 'Deep Residual Learning for Image Recognition'. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[9] *TensorFlow ResNet50*. Consulted on 15-Sept-2023. URL: https://www.tensorflow.org/api_docs/python/tf/keras/applications/resnet50.

[10] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2017. arXiv: 1412.6980 [cs.LG].

[11] Luke Taylor and Geoff Nitschke. 'Improving Deep Learning with Generic Data Augmentation'. In: *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*. 2018, pp. 1542–1547. DOI: 10.1109/SSCI.2018.8628742.

[12] Bhattiprolu Sreenivas. *Python for microscopists*. 2021. URL: https://github.com/bnsreenu/python_for_microscopists/blob/master/216_mito_unet__xferlearn_12_training_images.py.