In: AIAA SciTech 2023 Forum, 23-27 January 2023, National Harbor, Maryland, USA
DOI: 10.2514/6.2023-2672

Self-Supervised Obstacle Detection During Autonomous UAS Taxi Operations

Muhammad Yousuf Shaikh[1] and Ivan Petrunin[2]
Cranfield University, Cranfield, Bedfordshire, MK43 0AL, United Kingdom

Argyrios Zolotas[3]
Cranfield University, Cranfield, Bedfordshire, MK43 0AL, United Kingdom

This research explores the application of self-supervised learning techniques for obstacle detection and collision avoidance during UAS auto-taxi. Autoencoders were used to detect obstacles as anomalies by comparison of reconstruction errors. RGB cameras and millimetre wave radars covering conflict free zones (CFZs) around the own-ship were chosen to provide inputs to autoencoders. Results demonstrated that autoencoders were able to detect obstacles as anomalies within the CFZs but with certain limitations at lay the foundations of further work and investigation within the research area.

I. Nomenclature

ACAS	=	Airborne Collision Avoidnce System
ASDE-X	=	Airport Surface Detection System - model X
ATM	=	Air Traffic Management
C2	=	Command & Control
CFZ	=	Conflict Free Zone
DAA	=	Detect and avoid
EGPWS	=	Enhanced Ground Proximity Warning System
JU	=	Joint Undertaking
RPA	=	Remotely Piloted Aircraft
RVR	=	Runway Visual Range
SESAR	=	Single European Sky ATM Research
TCAS	=	Traffic Collision Avoidance System
UAS	=	Unmanned Aircraft System

II. Introduction

One of the objectives of SESAR JU is the seamless integration of UAS along with measures that enable shared usage of airspace by civilian and military users. Current state of the art military UAS are primarily RPAs that are susceptible to C2 latency issues due to which taxi operations of remotely piloted UAS are less dynamic as slower taxi speeds allow for larger margin of errors catering for the link latencies. Addressing the requirement of civil and military vehicles using the same ATM infrastructure, the slower taxi speeds of remotely piloted UAS can hinder the flow of traffic on airport surfaces. With the above in mind, autonomous UAS taxi at airports independent of remote control can eliminate the issue of C2 link latencies. However completely autonomous taxi operations bring with it two major challenges. Firstly autonomous navigation across the airport and secondly autonomous decision-making with regards to collision avoidance manoeuvres. This research attempts to address the latter.

[1] Student, MSc Applied Artificial Intelligence, Cranfield University.
[2] Senior Lecturer, Centre for Autonomous and Cyber-Physical Systems, Cranfield University .
[3] Reader, Centre for Autonomous and Cyber-Physical Systems, Cranfield University .

Airport aprons and taxiways are parts of the ATM infrastructure where aircraft spacing is at its minimal Fig. 1. And separation is based on a principle of "seen and be seen" thus always requiring a human in the loop supported by advance surface detection and awareness systems such as ASDE-X [1]. In such an environment maintenance of separation in the longitudinal axis is essential for operations where the UAS is following another aircraft in a queue and the separation in the lateral axis of the UAS is important to ensure wingtip clearance with obstacles at the edge of the protected manoeuvring areas and from other aircraft that might be on a collision trajectory.

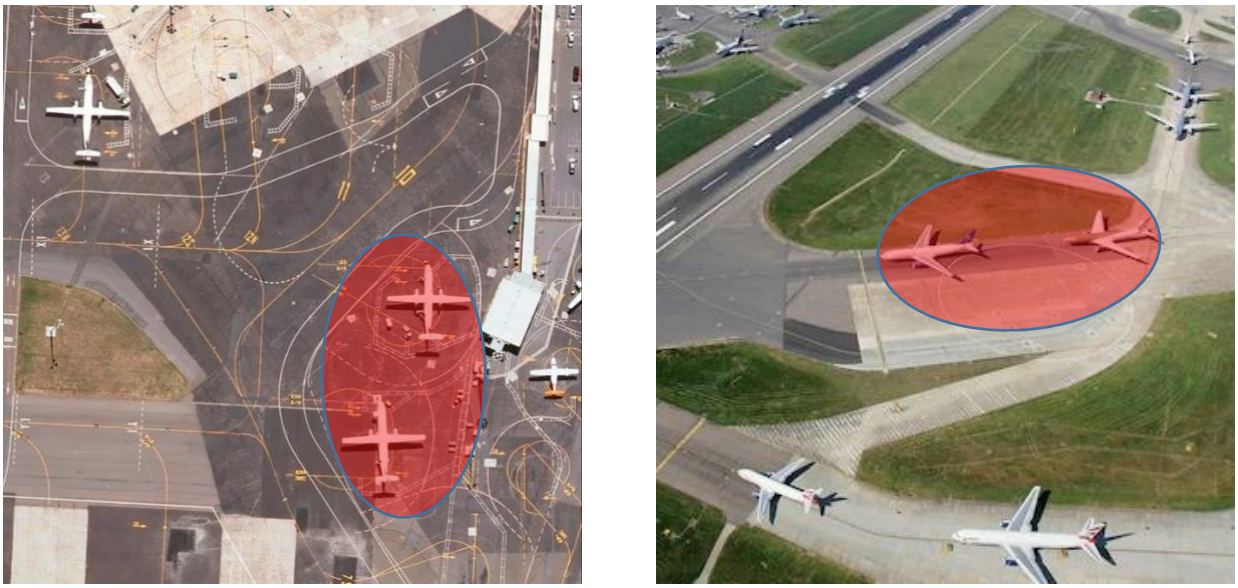


Fig. 1 Images of airport environments with high probability of trajectory conflicts and obstacles. Congested aprons (L) and busy taxiways (R). [4]

While there are many existing DAA solutions for the airborne phase of flight[2], [3], few if not none can be found that cater to the ground manoeuvring phase of UAS taxi. This research attempts to address that gap and explore the usage of self-supervised techniques as opposed to deterministic techniques used in the aerospace industry for collision avoidance. For this research the own-ship is considered to be completely autonomous fixed wing aircraft with dimensions comparable to a Boeing 737 with forward moving capability only. Non-cooperative sensors including millimetre wave radars and monocular cameras provide data for training and decision making with the use of autoencoders. The technique explored in this research is designed for a last resort DAA system that causes the aircraft to stop upon detection of an imminent collision. The research caters to only taxi operations on areas designated for taxi excluding the runway environment, moreover it is assumed that the maximum speeds are 30kts(15m/s) on straights and 10kts(5m/s) on turns. The weather conditions explored include a minimum visibility in terms of Runway Visual Range (RVR) of less than 400m but more than 75m.

Given the importance of airports in the ATM network and their potential to disrupt operations in case of accidents/incidents, the objectives of this research are to explore the anomaly detection capabilities of simple autoencoders for obstacle detection during completely autonomous taxi in conditions similar to airport operations.

III. Literature Review

Current methods of taxi include remote piloting and automatic guidance based on GPS coordinates[4]. However, both methods are limited by slower taxi speeds due to C2 latency and reliability/availability of satellites in the case of GPS. Autonomous UAS taxi could potentially solve such issues and make UAS ground operations more dynamic. Current collision avoidance systems in the aviation industry include air-air, air-ground solutions such as ACAS II and EGPWS for airborne phases of flight. For mitigation of incidents during ground manoeuvring, systems such as ASDE-

[4] Image credits Lauren O'Neil (L), Getty images (R).

X provide better situational awareness on the manoeuvring area of the airport. With some exceptions, all these systems incorporate humans in the loop as final decision makers. For this research the performance, limitations and implications of an independent agent are examined while it attempts to avoid obstacles by detecting anomalies outside of the normal operating environment.

While there is evidence of supervised and self-supervised techniques supported by computer vision for obstacle detection in self-driving cars [5]–[7] and while there is evidence of anomaly detection capabilities of autoencoders [8] with techniques including radar scans as inputs [9], obstacle detection for aircraft ground movement at aerodromes is still limited to supervised computer vision applications [10] and geometric camera angles [11]. This research therefore contributes to the state of the art by exploring the use of a self-supervised system using both computer vision and mmRadar data to detect anomalies by the use of simple auto encoders.

IV. Methodology

The research was conducted within a simulation environment supported by Carla UE4 which was customised to represent an FAA compliant airport with necessary obstacles and traffic simulated withing the environment. Training data was collected within this environment and then used to separately train autoencoders that were trained Google Colab GPUs. Testing scenarios were replicated within the same simulation environment while exploring variation in the airport environment and results were assessed based on the comparison of reconstruction errors in different obstacle scenarios.

Autoencoders used for this research are self-supervised networks that map data onto itself, thus learning how to de-construct (encode) and reconstruct (decode) the data. Once trained, the autoencoders are optimized to reconstruct the data they were trained on, thus if the input data differs significantly from the training data, the output is reconstructed with a larger reconstruction difference Fig. 2. This characteristic (reconstruction error) of the autoencoder is used to help detect obstacles as anomalies; by comparison of reconstruction errors between the input and the output data with a given threshold.

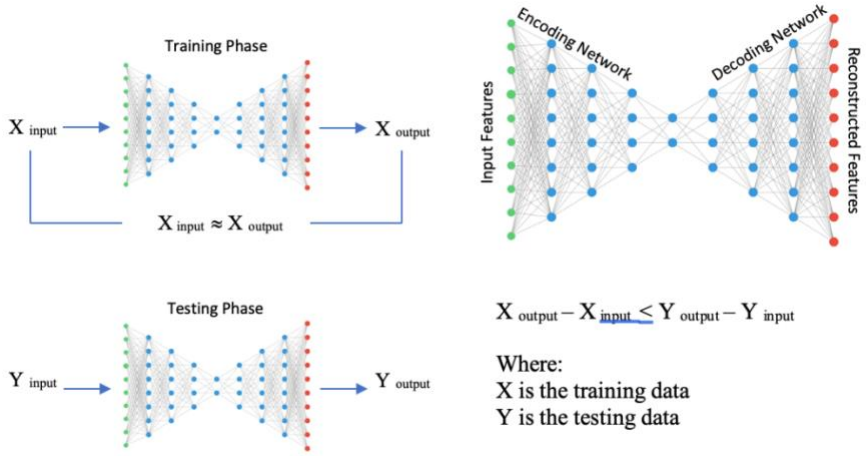


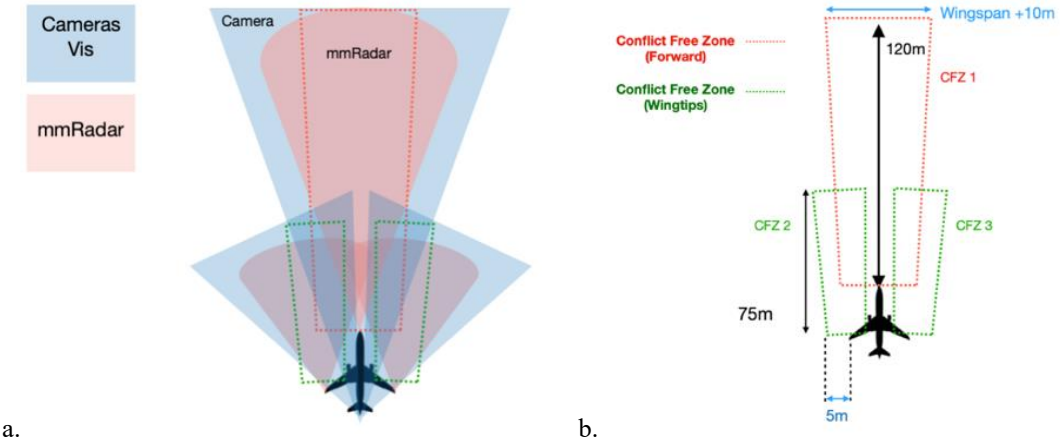
Fig. 2 Evaluating reconstruction errors with autoencoders

Data was collected via simulated mm-wave radars and RGB cameras placed on the own-ship (737-800). Data collected from normal obstacle-free operations on straight taxiways with imperfections and standard signboards was used to train Sensor-Autoencoder Pairs assigned to each monitoring area designated as a Conflict Free Zone (CFZ). Dimensions of the CFZs were dependent on vehicle stopping distance at max speed and ICAO wingtip separation requirements. In order to test the ability of Sensor Autoencoder Pairs, test scenarios were generated to compare the reconstruction errors of scenarios with obstacles against those without obstacles. With each scenario providing a reconstruction error, a statistically derived threshold was determined to classify a reconstruction error as a threat or not. Accuracy, missed obstacle rate and the false positive rates were calculated using a confusion matrix of results.

Key assumptions in this research were that the UAS simulated was in the aircraft design group III and the Taxiway design group 3. The UAS was considered to be completely autonomous operating in a straight line on paved taxiways without human input.

A. System Overview

The sensors distributed across the aircraft CFZs as shown inFig. 3.a and Fig. 3.b feed data into their respective trained autoencoders as shown in Fig. 3.c. Reconstruction errors are calculated using the trained autoencoders which are combined by assigning 60% weight to the radar autoencoder branch and 40% weight to the camera autoencoder branch in order to minimise the likelihood of false positives while ensuring adequate detection of obstacles. The combined reconstruction errors are then compared with a statistically determined threshold. If the reconstruction error was higher than the threshold, the inference is considered to contain an obstacle represented by an anomaly and a stop command is to be issued by the agent. If any CFZ infers a stop command the aircraft is to be stopped.



a. b.

Fig. 3 Distribution of sensors (a) and conflict free zone dimensions (b)

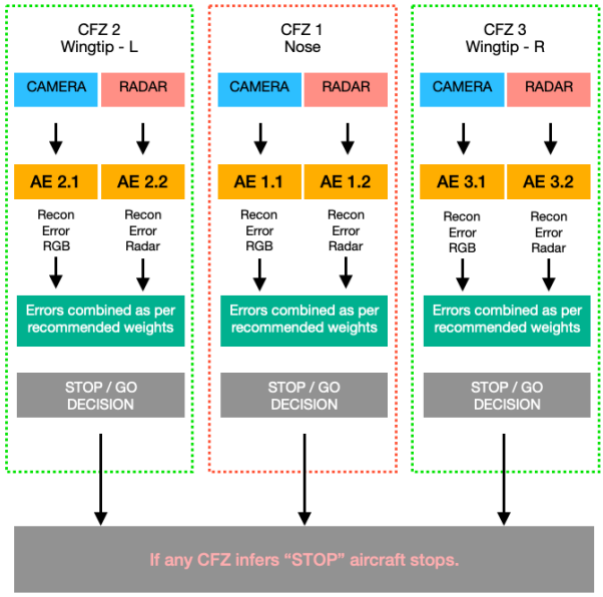


Fig. 4 An architecture of sensors autoencoders (yellow) and inferences distributed across the CFZs

B. Data

Data was generated using simulations of straight line taxi on FAA compliant airport environment built using Carla UE4. Training scenarios were a representation of normal operations including imperfections and obstacles that would

be present in normal circumstances such as minor cracks, skid marks, taxiway markings and signage. Testing scenarios included different conditions and obstacles relevant to the location of the CFZs.

Per training scenario there were a total of 12000 datapoints collected and preprocessed to remove null values and anomalies. This number included all 6 sensors placed on the own-ship with data collected at a rate of 12 data-points per meter over a 500m taxiway in the conditions simulated. Separate training environments were used for each CFZ. For CFZ 1 focus was on simulating taxiway markings and imperfections along with sign boards whereas for CFZ2/3 the focus was to have close by obstacles such as vehicles and parked airplanes to simulate a congested apron.

Because of computational limitations associated with the usage of Google Colab GPUs, the training of the model was limited to 3000 images for the camera autoencoder branches and 4500 radar data arrays (representing elevation information only) for the radar autoencoder branches. The training data and testing scenarios are escribed in table 1 below.

Table 1 Data used for training and testing scenarios

	CFZ 1	CFZ 2&3
TRAINING DATASET	Detail	Detail
Training Environment / Dataset used	• Clear weather • Straight Taxiways • Taxiway imperfections and signage included • 4500 Radar Arrays • 3000 Camera Images	• Clear weather • Straight Taxiways • Taxiway imperfections and signage included • Vehicles and airplanes parked close to taxiway • 4500 Radar Arrays • 3000 Camera Images
TESTING DATASET	Detail	Detail
No Obstacle / Different environment	Fog (No Obstacle) Visibility 110m	Fog (No Obstacle) Visibility 110m
No Obstacle / Different environment	Hold Markings (No Obstacle)	Hold Markings (No Obstacle)
No Obstacle / Different environment	Desert Landscape (No Obstacle)	Desert Landscape (No Obstacle)
With Obstacle / Environment similar to training environment	Animal on runway	Pole in front of wingtip
With Obstacle / Environment similar to training environment	Fire truck	Building in front of wingtip
With Obstacle / Environment similar to training environment	Building in front	Common Collision Scenario 1 Tail [10]
With Obstacle / Environment similar to training environment	Airplane Ahead within CFZ	Common Collision Scenario 2 Wingtip [10]
No Obstacle / Environment similar to training environment	19 Obstacle free scenarios similar to training dataset conditions.	17 Obstacle free normal scenarios similar to training dataset conditions.

C. Comparison of reconstruction errors with thresholds for anomaly detection

An unseen portion of the training dataset containing normal scenarios was used to determine a threshold based on which an inference could be made whether there is an obstacle in the CFZs or not. The unseen obstacle free data was passed through trained models and the resulting distribution of reconstruction errors was used to determine thresholds used for anomaly detection. Typically, anomalies may be considered to occur beyond 3 standard deviation however there was evidence in literature [12], [13] where a different statistical measure was used based on the sensitivity of the model and the application. For this application mean plus 2 standard deviations of reconstruction errors from validation set were used as thresholds that would be used for comparison with reconstruction errors from the testing dataset in order to identify scenarios with obstacles. In the case of the camera autoencoder branch for CFZ2/3 it was observed that just the mean value of the reconstruction error distribution gave better results. Obstacles were detected by comparing differences in reconstructed images between obstacle free scenarios and those with obstacles. Threshold used to classify reconstruction errors detected for CFZs were calculated using the technique outlined below:

$$if:\ MSE_{Obstacle} > \left[\frac{1}{m}\sum_{i=1}^{m} MSE_i^{clear}\right] + n\left[\sqrt{\frac{\sum_{i=1}^{m}(MSE_i^{clear} - \overline{MSE_i^{clear}})^2}{m-1}}\right] \quad (1)$$

$$then: Obstacle = True$$

$$m: number\ of\ MSE\ values\ from\ obstacle\ free\ reconstruction$$

$$n = 2\ for\ all\ cases\ except\ CFZ2\&3\ Camera\ Autoencoder\ Branch$$

$$where\ n = 0$$

Fig. 5 below shows the absolute difference between original and reconstructed image pixel intensities for an obstacle free scenario. Since the autoencoder was trained on such scenarios, there are no prominent anomalies detected.

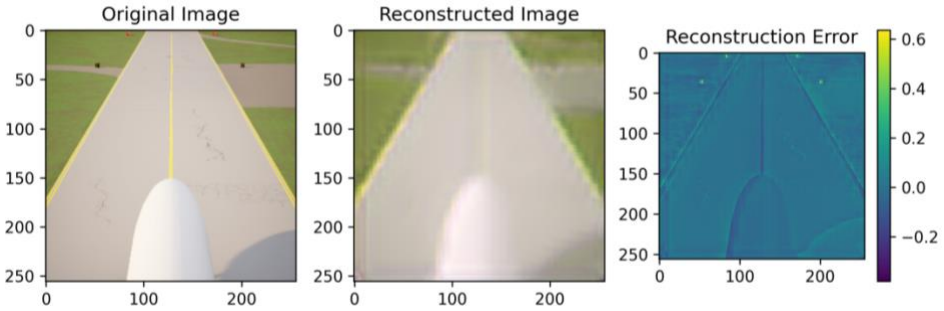


Fig. 5 Example of original and reconstructed images with 2D reconstruction error visualization for obstacle-free scenarios

Fig. 6 below shows the difference between reconstructions for a scenario with an aircraft taxying in front. It can be seen that the reconstruction error results in a hotspot within the image representing the reconstruction error.

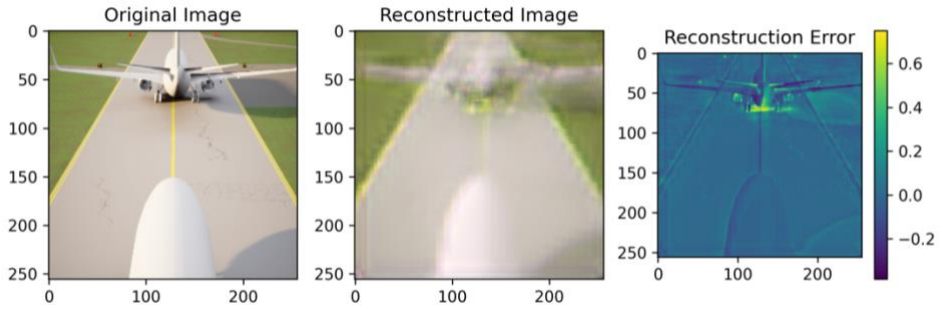


Fig. 6 Example of original and reconstructed images with 2D reconstruction error visualisation with an aircraft in CFZ1

Upon comparing both reconstructions in Fig. 7 below with the normal reconstruction providing a visual representation of a threshold/baseline for detection of the obstacle, the difference between the normal reconstruction and the reconstruction with the obstacle result in a positive detection of the aircraft.

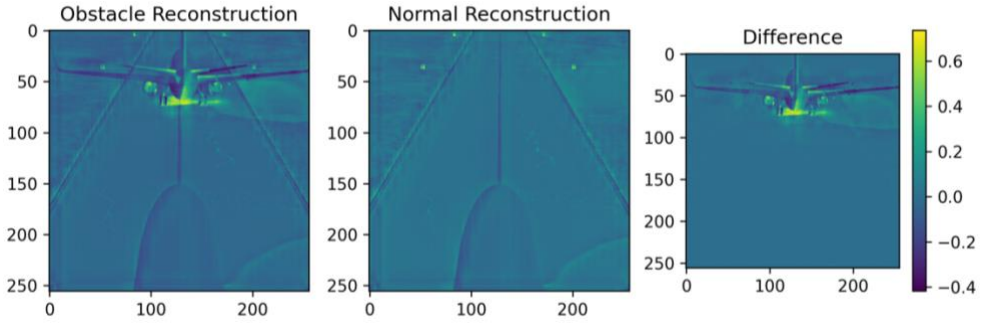


Fig. 7 Example of 2D visualisation of difference in reconstructions between normal scenario and aircraft as obstacle for CFZ 1

D. Combining results

For all CFZs, Radar Autoencoder branches were able to perform better in conditions where environmental factors such as weather and runway markings led to the detection of false positives by the camera-autoencoder branches. In terms of inferences made during normal operations with no weather disruptions and runway markings, the camera-autoencoder branches demonstrated consistent true negatives in terms of obstacle detection whereas the radar-autoencoder branches demonstrated high variation in data thus resulting in spurious detections. In order to maintain radar's ability to be unaffected by extreme weather and large surface markings, while taking advantage of camera-autoencoder pairs' consistency, inferences were combined together by assigning 60% weight to radar-autoencoder inferences and 40% weight to camera-autoencoder inferences.

V. Results and Discussion

Three different categories were explored as part of the test cases. Firstly, obstacle free scenarios with high environmental variation were tested to identify gaps in the model for conditions that excessively deviate from the environmental conditions of the training set (plots in magenta color Fig. 8 and Fig. 9). Secondly common obstacle scenarios in environmental conditions similar as those of the training set (plots in yellow color Fig. 8 and Fig. 9); to determine the model's obstacle detection capability. Thirdly, obstacle free cases with environmental conditions similar to those of the training set (plots in blue color Fig. 8 and Fig. 9); to check for false positives in ideal environments. The types of obstacles tested were different for CFZ1 and CFZs 2&3, since CFZ 1 considered the forward looking aspect whereas 2&3 were for wingtip protection.

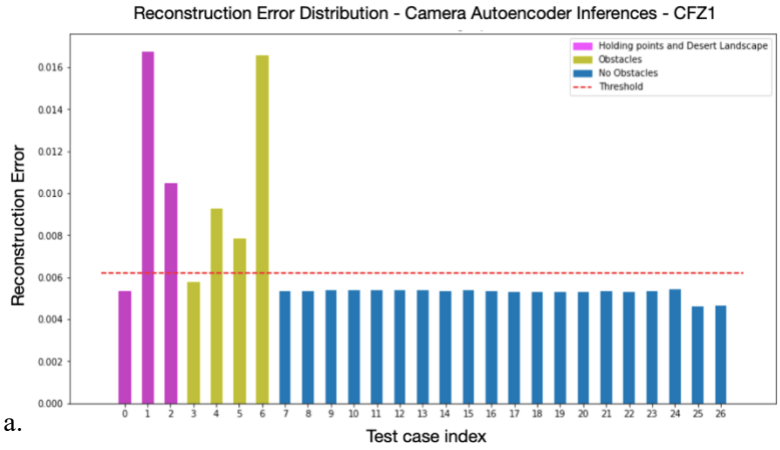
A. Results for CFZ 1

With reference to Fig. 8, obstacles were identified based on their reconstruction error being higher than a threshold that was set to the mean reconstruction error of obstacle free scenarios from the validation dataset plus two standard deviations of the same. Given this threshold, any reconstruction error higher than this value was considered to be a representation of an obstacle. Among obstacle free cases 0,1,2 with high environmental variation, cases 1 and 2 representing large taxiway markings and a change in landscape resulted in false positives for auto encoders working with camera data. Regarding cases 0,1 and 2, there was no evidence of false positives in inferences made by autoencoders working with radar data as the radar data was not affected by changes in environment or large visual markings aimed at grabbing the attention of the observer. Performance in terms of detecting obstacles in cases 3,4,5 and 6, was similar for both autoencoders working with radar and camera data. It was observed that small obstacles such as animals were difficult to detect as compared with obstacles that had more presence in images and/or had a larger radar cross section such as intruding aircraft and ground vehicles. For obstacle free cases 7 to 26 in clear weather conditions, autoencoders working with camera data demonstrated no false positives whereas the autoencoders working with radar data resulted in higher potential for false positives which was attributed to the high inherent variance within

the radar data extracted from the simulation. Looking at the individual performance of autoencoders working on radar and camera inputs, it was decided to combine reconstruction errors from the two along with the threshold values giving 40% weight to camera reconstructions and 60% weight to radar reconstructions. This ratio provided better detection of obstacles, and lower likelihood of false positives in obstacle free cases where the testing environment is the same is the training environment. As a trade-off, the ability of radar to accurately ignore extreme changes in environment was compromised considering that this short-coming could be overcome by including different environmental conditions in the training set for autoencoders working with camera data.

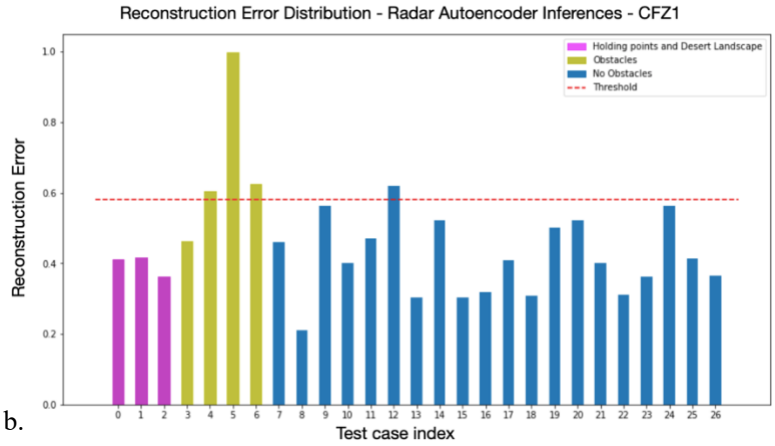
B. Results for CFZ 2/3

With reference to Fig. 9, obstacles were identified based on the same technique mentioned in part A of this section. Any reconstruction error higher than the thresholds were considered to be a representation of an obstacle. Among obstacle free cases 0,1,2 with high environmental variation, cases 0, 1 and 2 representing fog, large taxiway markings and a change in landscape respectively resulted in false positives for auto encoders working with RGB data. There was no evidence of false positives in inferences made by autoencoders working with radar data as the radar data was not affected by major visual changes in the environment. For cases 3,4,5 and 6, autoencoders working with cameras were able to perform best at detecting wingtips whereas autoencoders working with radar data struggles to detect wingtips (case 6) as obstacles. Small obstacles such as light poles (case 3) were difficult to detect as compared with obstacles that had more presence in images and/or had a larger radar cross section. For obstacle free cases 7 to 24 in clear weather conditions, autoencoders working with camera data demonstrated no false positives whereas the autoencoders working with radar data resulted in higher potential for false positives which was attributed to the high inherent variance within the radar data extracted from the simulation. Considering individual performance of autoencoders working on radar and camera inputs, reconstruction errors and thresholds were combined giving 40% weight to camera reconstructions and 60% weight to radar reconstructions.



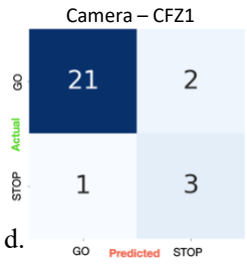
a.

Legend	
Case	Condition
0	Fog
1	Holding point markings
2	Desert landscape
3	Small animal (fox)
4	Fire truck
5	Building
6	Airplane taxying in front
7-26	Obstacle free scenarios in environment similar to training set

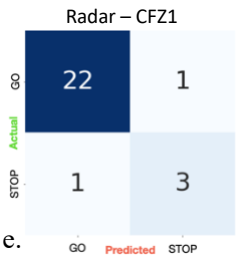


b.

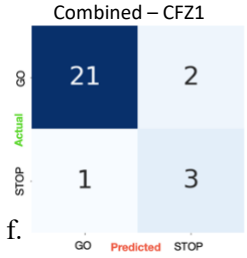
Confusion Matrices



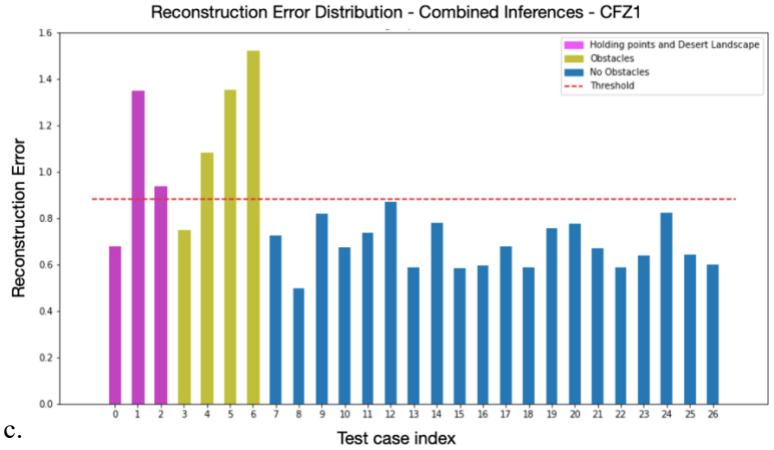
d.



e.



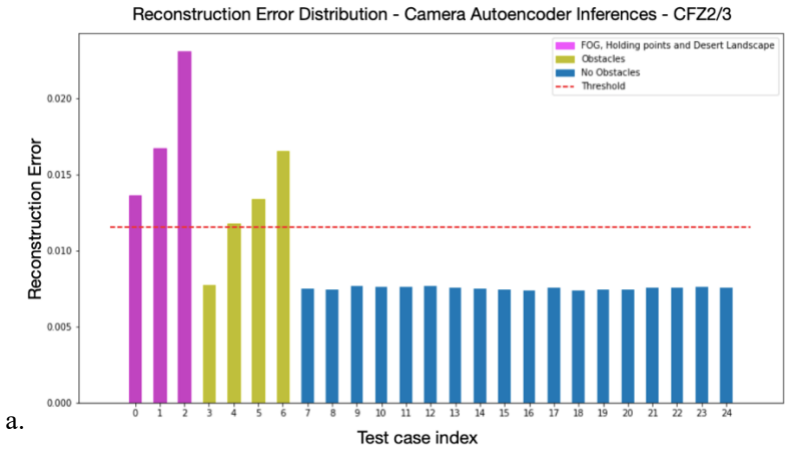
f.



c.

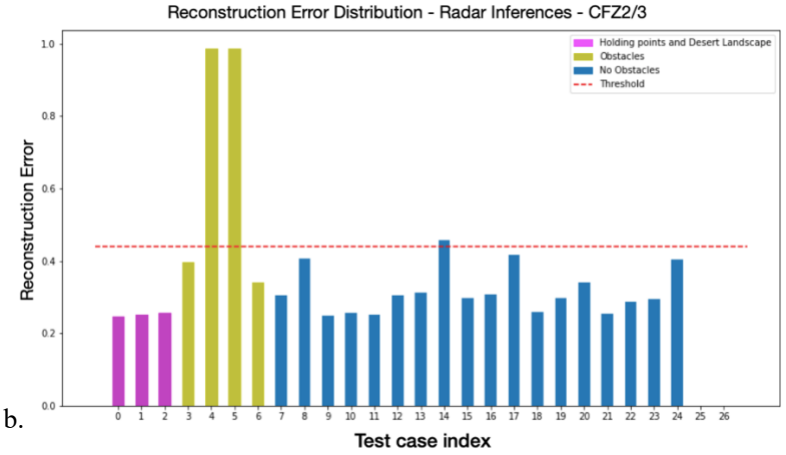
Results		
Technique	Accuracy	Misclassification Rate
RGB Auroencoder	0.889	0.111
Radar Autoencoder	0.925	0.074
Combined Inferences	0.889	0.111

Fig. 8 Results summary for CFZ 1. Reconstruction error distributions: Camera (a), Radar (b), Combined (c). Confusion matrices: Camera (d), Radar (e), Combined (f)

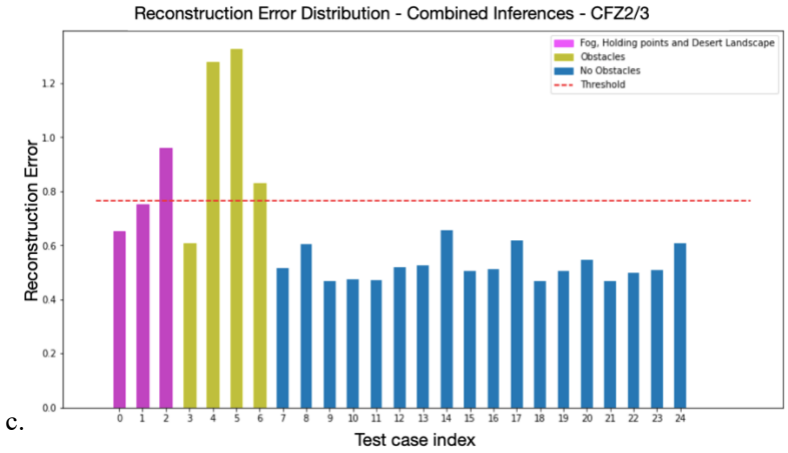


a.

Legend	
Case	Condition
0	Fog
1	Holding point markings
2	Desert landscape
3	Small obstacle (light pole)
4	Building
5	Common Collision Scenario 1 - Approaching a tail section
6	Common Collision Scenario 1 - Approaching a wingtip
7-24	Obstacle free scenarios in environment similar to training set



b.



c.

d. e. f.

Results		
Technique	Accuracy	Misclassification Rate
RGB Auroencoder	0.889	0.111
Radar Autoencoder	0.925	0.074
Combined Inferences	0.889	0.111

Fig. 9 Results summary for CFZ 2/3. Reconstruction error distributions: Camera (a), Radar (b), Combined (c). Confusion matrices: Camera (d), Radar (e), Combined (f)

C. Discussion

Holding point markings on taxiways indicate the beginning of a runway environment and thus are surface markings on the taxiway that are deliberately prominent by design Fig. 10. The autoencoder using camera data thus detected these markings as anomalies and gave false positives while detecting them as obstacles.

While global reconstruction errors were considered for the results shown previously, it was observed that the system can benefit from a more localised approach towards identifying reconstruction errors while minimising false positives caused by environmental changes and large markings on the taxiway. Fig. 10 shows inferences from both the image and radar data reconstruction. The image reconstructions detect reconstruction errors (red box) whereas the radar profile doesn't reflect the same in the elevation profile. Fig. 11 demonstrates that both image reconstructions and the radar elevation profile reconstruction demonstrate anomalies. Thus an anomaly/obstacle confirmed by inferences from both sensors in the case of Fig. 11. Moreover, it was observed that the range sensitive nature of the radar allowed for detection of objects as anomalies in scans beyond a distance of 60m thus providing more reliable results with reconstruction errors in the areas specified by the green bounding boxes in both Fig. 10 and Fig. 11.

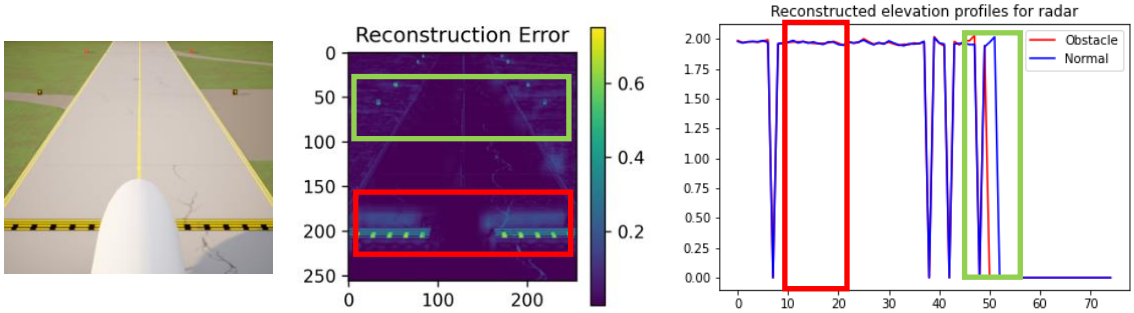


Fig. 10 Inference of holding point markings (L) as image anomalies (C) and radar scan anomalies (R)

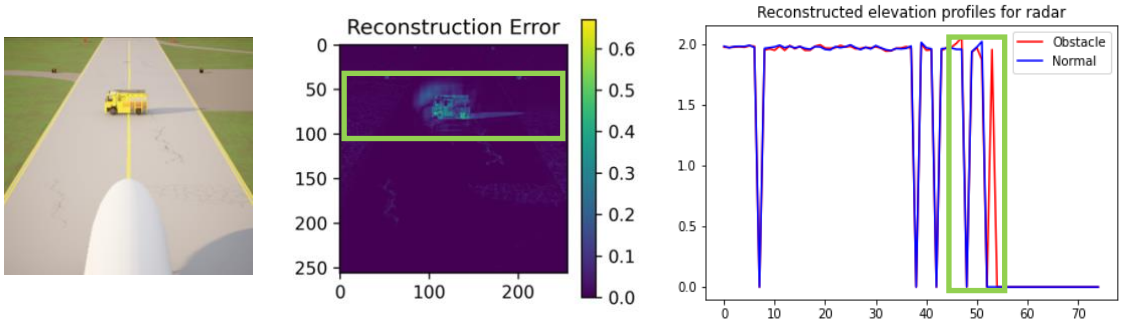


Fig. 11 Inference of a fire truck (L) as image anomalies (C) and radar scan anomalies (R)

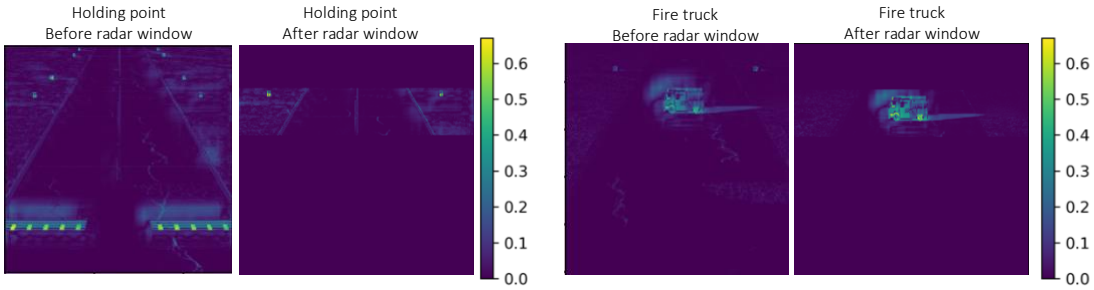


Fig. 12 Inferences after application of radar window thresholds for holding point from Figure 8 (L) and fire truck from Figure 9 (R)

Areas represented by the green bounding boxes therefore ensure adequate provision of reliability regarding autoencoders working on both sensors. Thus, the green bounding boxes can be used to define regions of interest to detect local reconstruction errors by use of a "radar window" defining the region of interest as shown in Fig. 12.

VI. Conclusion

This research aimed at exploring the performance and limitations of using simple autoencoders to detect obstacles as anomalies by comparing reconstruction errors of obstacle free and with obstacle scenarios. While supervised learning methods and deterministic techniques exist, this approach demonstrates 88.9% accuracy in terms of making correct inferences. However due to the unbalanced testing dataset the performance is better examined by looking at individual obstacle scenarios in more detail, the examination of which uncovered the possibility of false positives in obstacle detection with cameras due to severe changes in weather, terrain and large taxiway markings.

References

[1] "Airport Surface Detection Equipment, Model X (ASDE-X)."

[2] "tcas ii v7.1 intro booklet".

[3] Institute of Electrical and Electronics Engineers, IEEE Aerospace and Electronic Systems Society, American Institute of Aeronautics and Astronautics, and American Institute of Aeronautics and Astronautics. Digital Avionics Technical Committee, *38th DASC, Digital Avionics Systems Conference : 2019 conference proceedings : San Diego, California, USA, September 8-12, 2019*.

[4] "Taxi Assist System Concept and Requirements Definition," 2022.

[5] R. Okuda, Y. Kajiwara, and K. Terashima, "A survey of technical trend of ADAS and autonomous driving," in *Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test, VLSI-DAT 2014*, 2014. doi: 10.1109/VLSI-DAT.2014.6834940.

[6] H. H. Jebamikyous and R. Kashef, "Autonomous Vehicles Perception (AVP) Using Deep Learning: Modeling, Assessment, and Challenges," *IEEE Access*, vol. 10, pp. 10523–10535, 2022, doi: 10.1109/ACCESS.2022.3144407.

[7] T. Ohgushi, K. Horiguchi, and M. Yamanaka, "Road Obstacle Detection Method Based on an Autoencoder with Semantic Segmentation."

[8] L. Sabetti and R. Heijmans, "Shallow or deep? Training an autoencoder to detect anomalous flows in a retail payment system," *Latin American Journal of Central Banking*, vol. 2, no. 2, p. 100031, Jun. 2021, doi: 10.1016/j.latcb.2021.100031.

[9] F. Picetti, G. Testa, F. Lombardi, P. Bestagini, M. Lualdi, and S. Tubaro, "Convolutional Autoencoder for Landmine Detection on GPR Scans," in *2018 41st International Conference on Telecommunications and Signal Processing, TSP 2018*, Aug. 2018. doi: 10.1109/TSP.2018.8441206.

[10] J. Gauci and D. Zammit-Mangion, "Obstacle detection around aircraft on ramps and taxiways through the use of computer vision," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2009. doi: 10.2514/6.2009-5866.

[11] "Wing-Tip-flyer".

[12] J. Kim, "Using median as a threshold in determining anomaly in back-end authentication," in *Proceedings - 3rd International Conference on Computational Intelligence and Applications, ICCIA 2018*, Jul. 2018, pp. 249–254. doi: 10.1109/ICCIA.2018.00055.

[13] Y. N. Nguimbous, R. Ksantini, and A. Bouhoula, "Anomaly-based Intrusion Detection Using Auto-encoder; Anomaly-based Intrusion Detection Using Auto-encoder," 2019.

Self-supervised obstacle detection during autonomous UAS taxi operations

Shaikh, Yousuf

2023-01-19

Shaikh MY, Petrunin I, Zolotas A. (2023) Self-supervised obstacle detection during autonomous UAS taxi operations. In: AIAA SciTech 2023 Forum, 23-27 January 2023, National Harbor, Maryland, USA. Paper number AIAA 2023-2672

https://doi.org/10.2514/6.2023-2672