A REPORT

ON

(VISUAL DETECTION OF DRONES BY INTEGRATION OF FOG TIME IMAGES WITH AI)

BY

SUYASH SINGH	2020A3PS1540P
SHUBHI BHANDARI	2020AAPS1422G

AT

315 Army Base Workshop Pitoragarh

A Practice School – 1 Station

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

(June-July 2022)

A REPORT

ON

(VISUAL DETECTION OF DRONES BY INTEGRATION OF FOG TIME IMAGES WITH AI)

BY

Name(s) of the Student(s)	ID.No.(s)	Discipline(s)
SUYASH SINGH	2020A3PS1540P	B. E Electrical and Electronics
SHUBHI BHANDARI	2020AAPS1442G	B. E Electronics & Communications

Prepared in partial fulfillment of the Practice School-I Course Nos.
BITS C221/BITS C231/BITS C241

AT

315 Army Base Workshop Pitoragarh

A Practice School – 1 Station

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

(June-July 2022)

Acknowledgement

This project has been made possible by the direct and indirect cooperation of various people for whom I wish to offer my thanks. I would like to express my special thanks of gratitude to the Dean of Practice School, Prof. S. P. Regalla as well as the other members of the Practice School Division, BITS Pilani for providing me with the honor of being able to participate in this esteemed program. I would also like to thank Prof. Suparna Chakraborty for mentoring me throughout the Practice School Program. I would like to thank Lt. Col Antriksh Khanna, for providing me with such an opportunity and Nb Sub Naval Singh and Maj Ravinder Jaral and his team for guiding me throughout the project and giving me the exposure to the technical aspect of the project.

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI (RAJASTHAN)

Practice School Division

Station: 315 Army Base Workshop, Pitoragarh

Centre: EME Workshop

Duration: June-July 2022 (6 Weeks)

Date of Start: 30th May 2022

Date of Submission: 26th June 2022

ID No	Name(s)	Discipline(s)
2020A3PS1540P	Suyash Singh	B.E Electrical & Electronics
2020AAPS1442G	Shubhi Bhandari	B.E Electronics & Communication

Name(s) and Designation(s) of the expert(s):
Lt Col Antriksh Khanna, Nb Sub Naval Singh, Maj Ravinder Jaral

Name of the PS Faculty:
Ms. Suparna Chakraborty

Key Words:
AI, Object Detection, Drones

Project Areas:
AI Detection of Drones through fog time images

Abstract:
In this project, we have created an AI program which can detect and mark drones through fog time and daytime images. This program will be clubbed with other projects which involve AI detection of drones through smoke and HHTI images.

Suyash Singh

Shubhi Bhandari

Suparna Chakraborty

Signature(s) of Student(s)

Signature of PS Faculty

26/6/2022

26/6/2022

TABLE OF CONTENTS

INTRODUCTION

1) Scope

As drone attracts much interest, the drone industry has opened their market to ordinary people, making drones to be used in daily lives. However, as it got easier for drone to be used by more people, safety and security issues have raised as accidents are much more likely to happen colliding into people by losing control or invading secured properties. For safety purposes, it is essential for observers and drone to be aware of an approaching drone. My project is a step in this very same direction. The Indian Army should be aware of hostile drones approaching in regions which have low or reduced visibility during. Fog Time, Day Time, and Thermal AI Detection combined will provide well rounded drone detection system to the Indian Army.

2) Development of the AI Program

One of the backbone instruments of our project is YOLO. You only look once (YOLO) is a state-of-the-art, real-time object detection system. Prior detection systems repurpose classifiers or localizers to perform detection. They apply the model to an image at multiple locations and scales. High scoring regions of the image are considered detections. YOLO uses a totally different approach. It applies a single neural network to the full image. This network divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities.

Dataset was taken from the internet and fog was simulated on the images virtually. The entire training logic was written from scratch and implemented in a virtual environment using Google Colab.

3) Tech Stack Used

For the realization of this project, familiarity with the given software/tools is required:

- Python
- Jupyter Notebook
- YOLOv4
- Google Colab
- Tensorboard

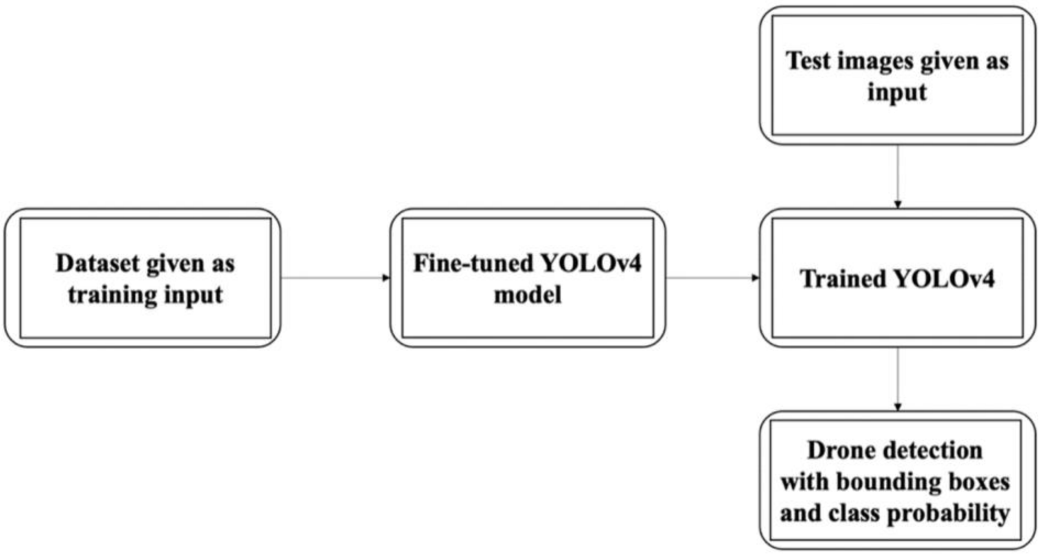
PROJECT DETAILS

Computer vision is extensively used to detect drones autonomously compared to other proposed solutions such as RADAR, acoustics, and RF signal analysis thanks to its robustness. Among these computer vision-based approaches, we see the preference of deep learning algorithms thanks to their effectiveness.

Since it is very difficult to obtain large number of fog time images of drones on the internet. We will use a normal data set of drone images and use the albumentation library to simulate fog on the normal day time images of drones. Albumentations is a fast and flexible image augmentation library. The library is widely used in industry, deep learning research, machine learning competitions, and open-source projects. Once we have the fog time data set, we were ready to train our AI program using YOLOv4 library which detects objects through a bounding box. A bounding box is an imaginary rectangle that serves as a point of reference for object detection and creates a collision box for that object. Data annotators draw these rectangles over images, outlining the object of interest within each image by defining its X and Y coordinates.

YOLO involves a single neural network trained end-to-end, which takes an image as input and directly predicts bounding boxes and class labels. The technique offers lower predictive accuracy (e.g., more localization errors) but operates at 45 to 155 frames per second (FPS), providing a speed-optimized version of the model. The YOLOv4 small used in this project has been trained from scratch for 10 epocs with a batch size of 32 and an image resolution of 640px by 640px. The initial dataset consisted of 1359 annotated drone images on which fog was simulated virtually. The trained YOLOv4 was evaluated using mAP, precision, recall, and F1-score all the graphs showcasing these results have been attached in the appendix for further analysis.

This project was limited to YOLO implementation only since various object detection algorithms require datasets to be labeled in certain formats, which is time consuming. In addition, speed was one of our considerations while choosing algorithms. In future work, a more diverse image dataset will be used to further improve the results. In addition, other object detection algorithms such as R-CNN, mobilenet, SSD, etc., will be trained and compared. YOLOv5 has already been released; thus, we will further use this version to see if the speed and accuracy improve. More objects will be added alongside the drone and bird images.



Weekly Work Plan

1) Week 1: Understanding Project Details and Technologies involved

Week 1 was consumed with the ideation of the project and how the project was going to be developed. We spent this time reading and familiarizing ourselves with the reading materials provided to us as well as deciding the workflow of the project. This had to be decided before starting the project so that the development can be done accordingly.

2) Week 2: Gathering Required Resources on the Internet

This week was spent reading and understanding research papers on various deep learning algorithms such as YOLO, FastRCNN etc. Ultimately, we decided to employ YOLOv4 algorithm due to its superior speed and precision compared to other recognition systems. We discussed as a group which labelled image dataset to use taking into consideration compatibility with YOLOv4 and found a library (albumentations.ai) that could perform image augmentation on our image dataset.

3) Week 3: Reading about the Tech Stack

Since many software tools are required for the project, week 3 was spent in learning about the details of the project and getting ourselves familiarized with the tools needed for implementation of our model including python, creating, and using a Jupyter notebook, Tensorboard and so on.

4) Week 4: Actual Implementation of the training algorithm

The YOLOv4 small was trained from scratch for 10 epocs with a batch size of 32 and an image resolution of 640px by 640px. The initial dataset consisted of 1359 annotated drone images further augmentations were done to the dataset to add fog to achieve a final program that would be able to detect to drones through fog time images efficiently. The entire training logic was written from scratch and implemented in a virtual environment using Google Colab.

5) Week 5: Implementation of Inference from the trained model (Upcoming)

The inference file was created for accessing the model results.
It allows users to put an image link from the internet and see the result.

6) Week 6: Bug Fixes and Finalization (Upcoming)

Testing of the product was done, and the bugs or glitches were fixed. The final product, along with the entire source code, was pushed to GitHub.

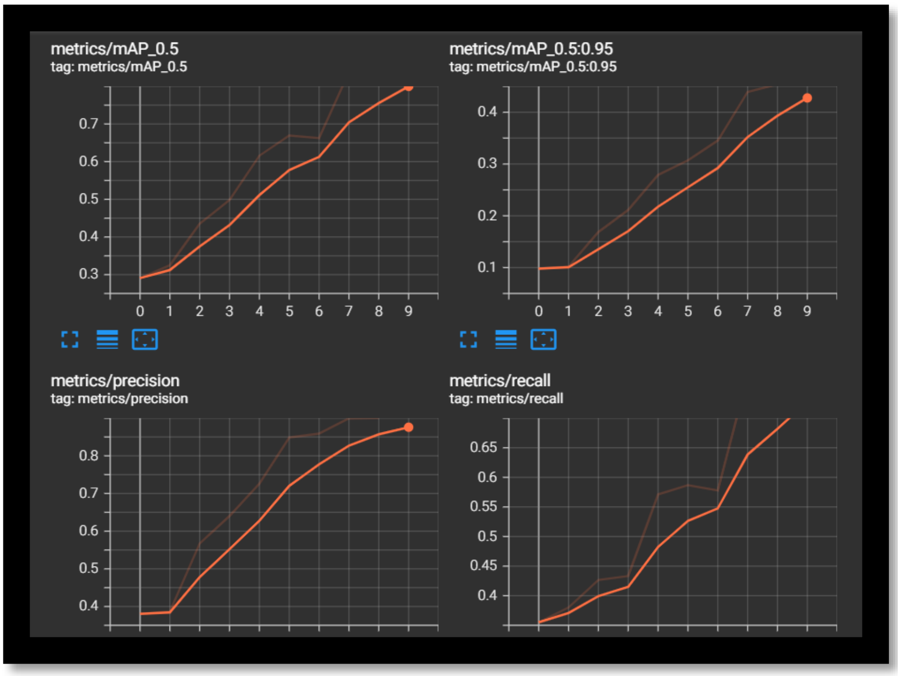
Conclusion

The detection, recognition, and identification of UAVs through projects like ours are crucial in discussing public safety and the threats posed by their existence. In plains and deserts, there is a huge issue of fog (and recently, smog) and thus any intelligence or seasoned adversary will try and use the bad weather to their own advantage. The project this year aims to provide a functional detection system to the Indian Army that would help them gain advantage in case of all weather conditions- be it daytime, night time, fog time combined with thermal detection.

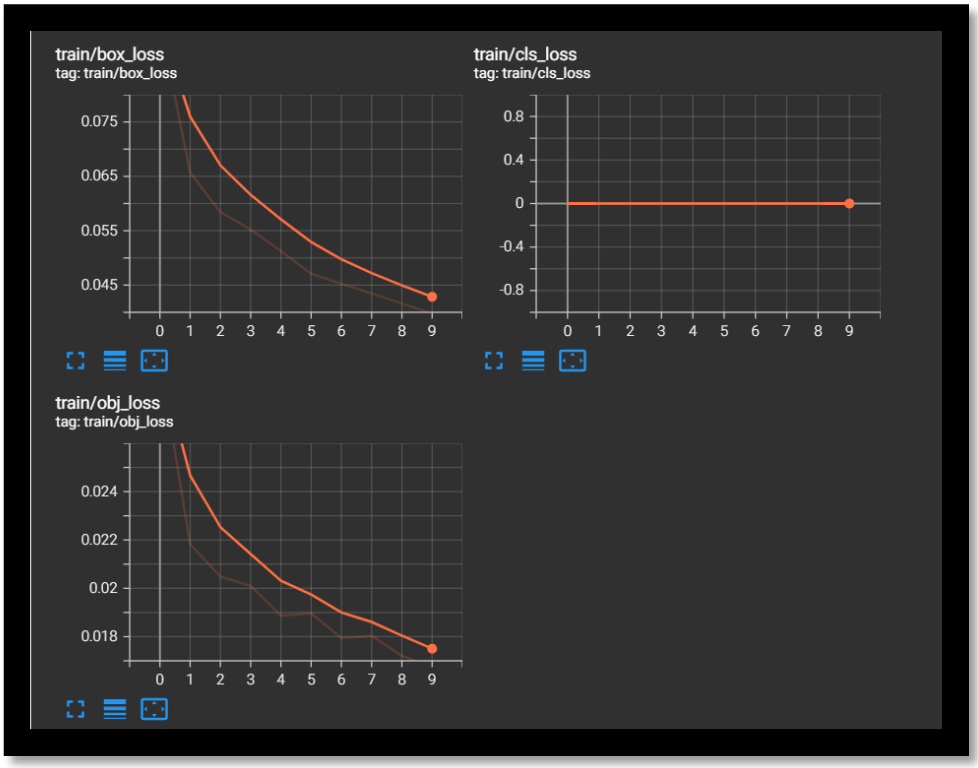
Through our project thus far, the user is able to successfully output the probability of an object being a drone even under fog time weather conditions. We have tested the project and all the bugs and glitches which we could find have been fixed. In the upcoming weeks, we plan on creating an application front-end that would allow the user to feed images from their own database and visually detect the presence of a drone. It is important to make the frontend design appealing, user-friendly and intuitive to the user while displaying the essential information from the database and providing the opportunity to alter the database from the frontend itself.

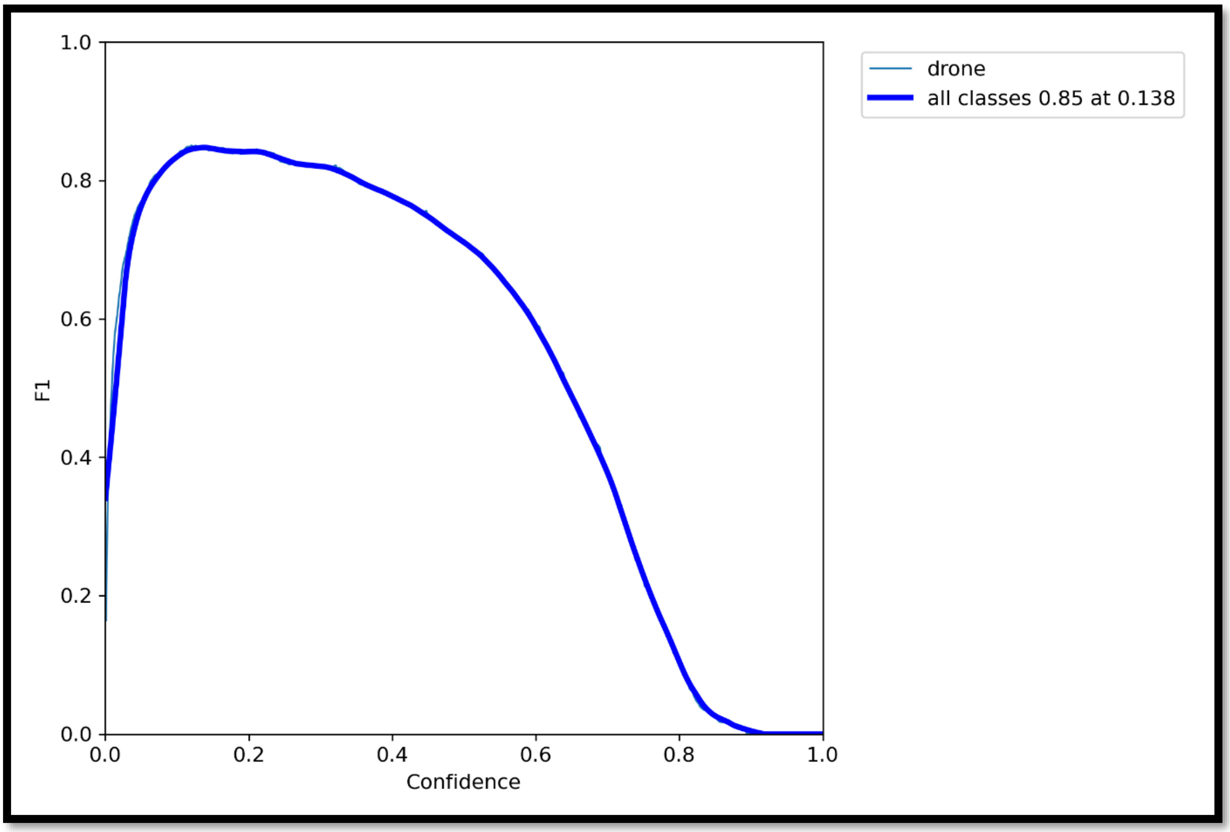
The future use of the application is thus very bright as it will be further developed to include a user-friendly front end as well as tie into the other AI projects this year. We are grateful to our mentors from the army base as well as our PS1 faculty Prof. Suparna Chakraborty for guiding us and making all the work we have been able to put in so far possible.

Appendix

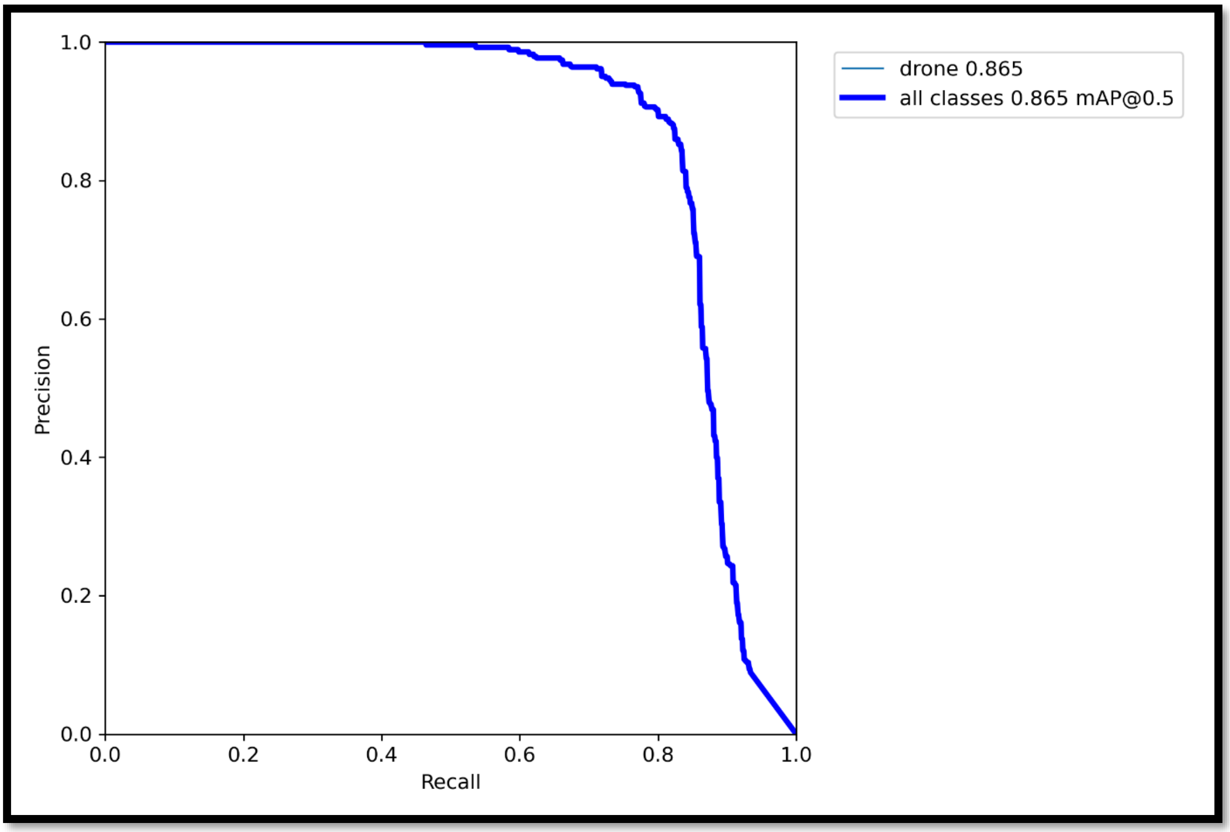


Training Curves (Tensorboard)





(F1 Curve)



(PR Curve)

Output (Bounded Box Detection)

References

- https://ieeexplore.ieee.org/document/8539442

- https://www.researchgate.net/publication/329490192_Drone_Detection_and_Identification_System_using_Artificial_Intelligence

- https://www.youtube.com/watch?v=b59xfUZZqJE&ab_channel=edureka%21

- https://www.mdpi.com/2504-446X/5/3/95/htm

- Deep Convolutional Neural Networks", Computer Vision and Pattern Recognition, https://arxiv.org/ftp/arxiv/papers/1901/1901.06032.pdf

Project Repository: -

https://github.com/suyash-03/PS-ArmyBaseWorkshop

Training Model
https://colab.research.google.com/drive/1q5173ZN_qVQ5J-5O0ISEZLnut618cTJu?usp=sharing

Inference File
https://colab.research.google.com/drive/1tYUGpJcGy9-i7dl5qpw8SSAyV3v4Xycb?authuser=1#scrollTo=GFeLD0KDCyZ3

Glossary

1. **YOLOv5**- YOLO is an acronym that stands for You Only Look Once. We are employing Version 5, which was launched by Ultralytics in June 2020 and is now the most advanced object identification algorithm available. It is a novel convolutional neural network (CNN) that detects objects in real-time with great accuracy. This approach uses a single neural network to process the entire picture, then separates it into parts and predicts bounding boxes and probabilities for each component. These bounding boxes are weighted by the expected probability. The method “just looks once” at the image in the sense that it makes predictions after only one forward propagation run through the neural network.

2. **Image Augmentation**- Image augmentation is a technique of altering the existing data to create some more data for the model training process. In other words, it is the process of artificially expanding the available dataset for training a deep learning model.

3. **Computer Vision**- Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs — and take actions or make recommendations based on that information. If AI enables computers to think, computer vision enables them to see, observe and understand.

4. **Deep Learning**- Deep learning is a subset of machine learning, which is essentially a neural network with three or more layers. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—allowing it to “learn” from large amounts of data. While a neural network with a single layer can still make approximate predictions, additional hidden layers can help to optimize and refine for accuracy.

5. **Neural Network**- Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another. Artificial neural networks (ANNs) are comprised of a node layers, containing an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, connects to another and has an associated weight and threshold. If the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next layer of the network.

6. **TensorBoard-** TensorBoard provides the visualization and tooling needed for machine learning experimentation such as tracking and visualizing metrics such as loss and accuracy, visualizing the model graph (ops and layers), viewing histograms of weights, biases, or other tensors as they change over time, projecting embeddings to a lower dimensional space, displaying images, text, and audio data, profiling TensorFlow programs and much more.

7. **UAV-** An unmanned aerial vehicle (UAV), commonly known as a drone, is an aircraft without any human pilot, crew, or passengers on board. UAVs are a component of an unmanned aircraft system(UAS), which includes adding a ground-based controller and a system of communications with the UAV.