Key aspects from 3 SUAS 2024 competition

1. ODLC (Object Detection, Localization, and Classification) Pipeline:

These are the approaches that have caught the attention out of the numerous teams who are building the ODLC system using a variety of approaches and strategies.

- First Method:

classification for ODLC the requirements include protection of targets classification based on shape color and alpha numeric characters

To test ODLC performance images were captured and various backgrounds after AI inference a popup in the UI shows Target attributes the operator can log the target and activate the Drone for autonomous draw and mission for our AI contingency plan, A GUI is created with direct drone control operator can halt or move the Drone by selecting objects in the GUI using real time pixel coordinates accuracy of the AI models are as follows these are the results with weing

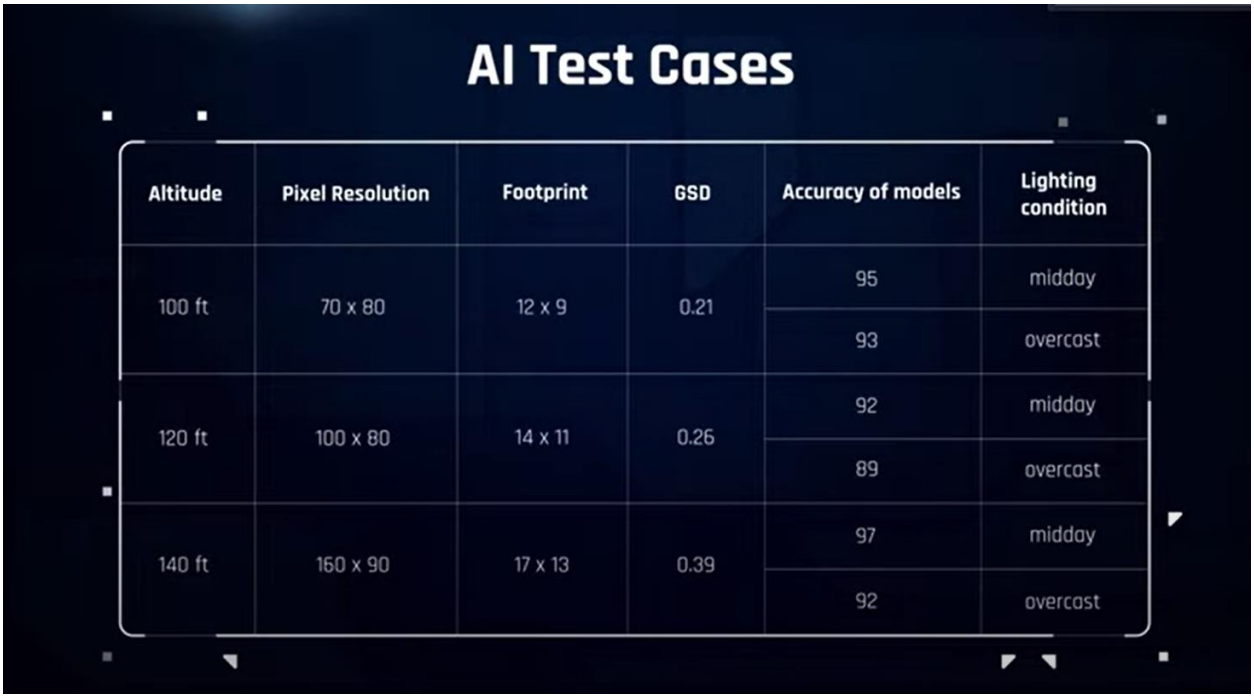

Figure 1.1

- Localization part

A CNN is used to detect and localize the region of Interest based deep learning model from a given 4K image after the targets are detected and localized the detected crops with an average size of 160 cross 90 pixels are further sent for shape segmentation that is performed to extract the shaped mask an image head classification model is trained for wide range of RGB values to identify color

For letter identification pre-processing is used with shape color based adaptive thresholding and image rotation passing images through an OCR model for car recognition via a GPU accelerated pipeline

Then an inference time of 300 MCS is achieved per image cropped image metadata holds results matching GPS to image timestamp gives desired Target coordinates for calculating intended location for localization

- Second Method:

A two-stage processing pipeline can be developed to detect and classify Target objects first live images can be captured with the camera, these images are then processed using custom YOLO model to identify objects of Interest both emergent and static split them into a list of targets and enhance them for the second stage after using a second YOLO model and the color masking algorithm to classify the target's
character shape and two colors the onboard computer checks its bottle queue
which is predefined Target

- Localization part

The classified Target is uploaded during setup if one matches a drop is performed after the exact drop location is determined by a Geor referencing algorithm from 400
simulated images and 50 real life images of targets from our flights at 80 ft (The numbers are just an example)

Then, 386 of the simulated images and 40 of the real-life images are classified successfully.

- Third Method

Three machine learning models are used

A model to detect and classify standard and emergent objects to detect standard objects

A YOLO V8 object detection model was fine-tuned, and by using a synthetic data set to reduce the need for data collection and modeling, a data set was generated by using open CV to place different shapes and characters on a pavement background and assigned bounding boxes accordingly

The object detection model was used to detect the shape and color of the object then feed it into the par seek text recognition model to read the character from a detected shape to detect emergent objects

The YOLO V8 model was also used which is pre-trained on the Coco data set which includes humans as its class when using the model for prediction in the competition

High resolution images were taken of 5472 by 3648 pixels that cover the whole Runway area where the target will be placed, we will slice the image into smaller chunks before feeding it into

- Localization

The YOLO V8 model since it is trained on small images roughly 640x 640 pixels and if the full image is feeded the individual object will be very small then the detection result is combined and make the airdrop decision the camera that is used is the phx2 200s made by Lucid Vision Labs the sensor has a pixel size of 2.4x 2.4 microns and the lens has a focal length of 12 mm this gives out a ground sample distance of 46 cm per pixel an 8.5x 8.5 Target at a height of 75 ft should have a resolution of around 47x 47 pixels