Nvidia Jetson TX1 based autonomous drone for recognizing different bird and animal species in an area with bird call recognition and image classification.

BY- Arpit Chauhan

# Table of contents

## Introduction

## Experimental techniques and methods

## Discussion and results

## Summary /conclusion

## References

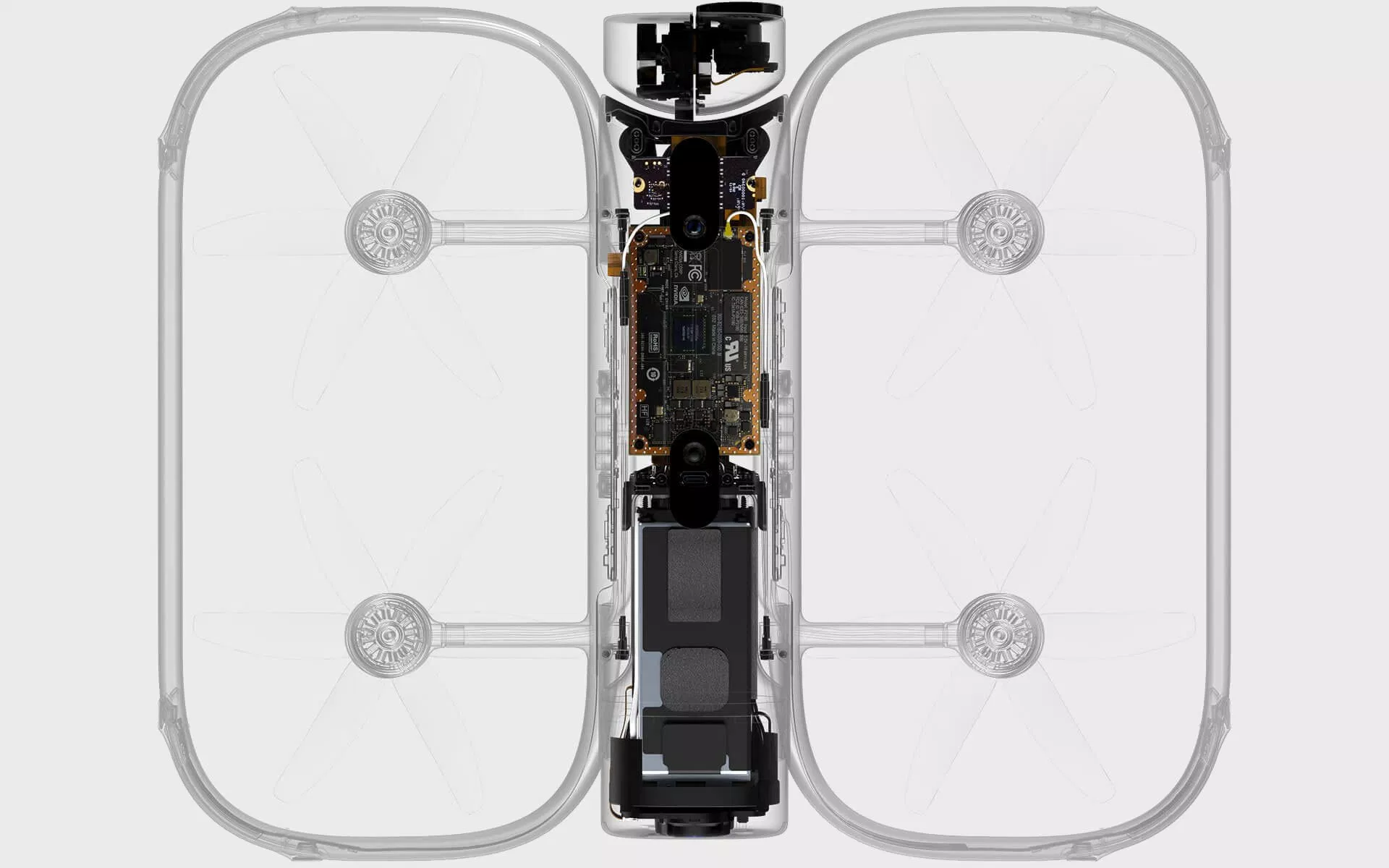
# Introduction

This technical report is on an Nvidia Jetson TX1 based project. It is supposed to be a hypothetical project which is also cost and time effective. In this project I have tried to use as much of the features and specifications of Jetson Platform as possible in this short time frame.

The project is basically a drone that utilises Nvidia Jetson platform to recognise different bird and animal species in a given area by utilising the subtle differences in their calls. It Autonomously hovers over a predefined area recording various sound samples and later analyse and find a match from a large map of logically arranged bird calls available to it. At the same time, it uses image classification in real time to classify different animals on a coarse level.

This will help zoologists and Ornithologists keep track of biodiversity of the area and keep a record of the endangered species in that area.

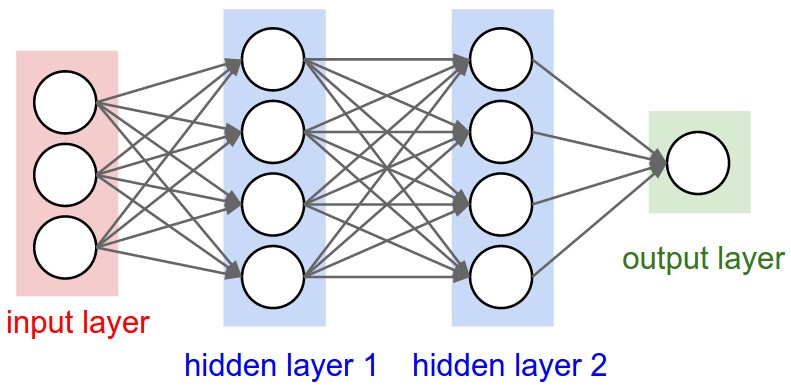
# Experimental techniques and methods used



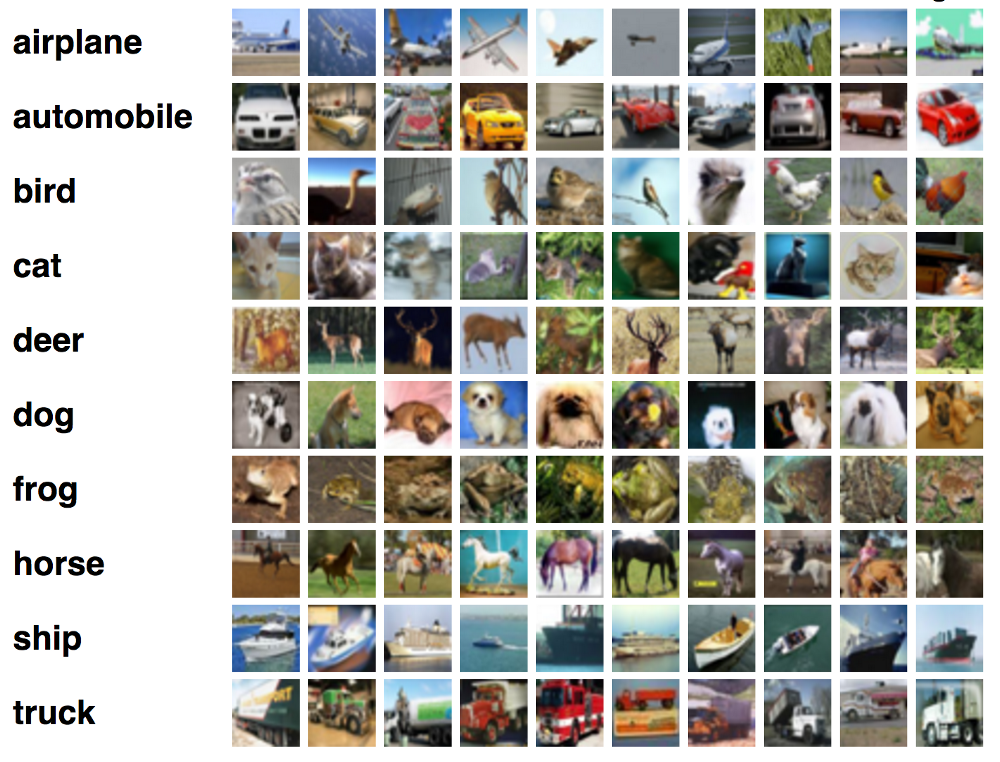
Skydio R1 cutaway showing Nvidia Jetson TX1 Module

## Flying

For the drone itself I am using The Skydio R1 as a basis which is state of the art in its implementation of tracking and path planning algorithms. The hardware platform consists of a Jetson TX1 module, with 6 pairs of stereo cameras. Including the 4K video imager, this makes for 13 (!) cameras all together. Skydio R1 uses **Simultaneous Localization and Mapping (SLAM)** to construct and update a map of an environment while also keeping track of a person’s location within it. R1 interprets the context of each situation to figure out how to act appropriately. Through **deep learning** methods, it understands the difference between things it sees and factors that understanding into its behavior.

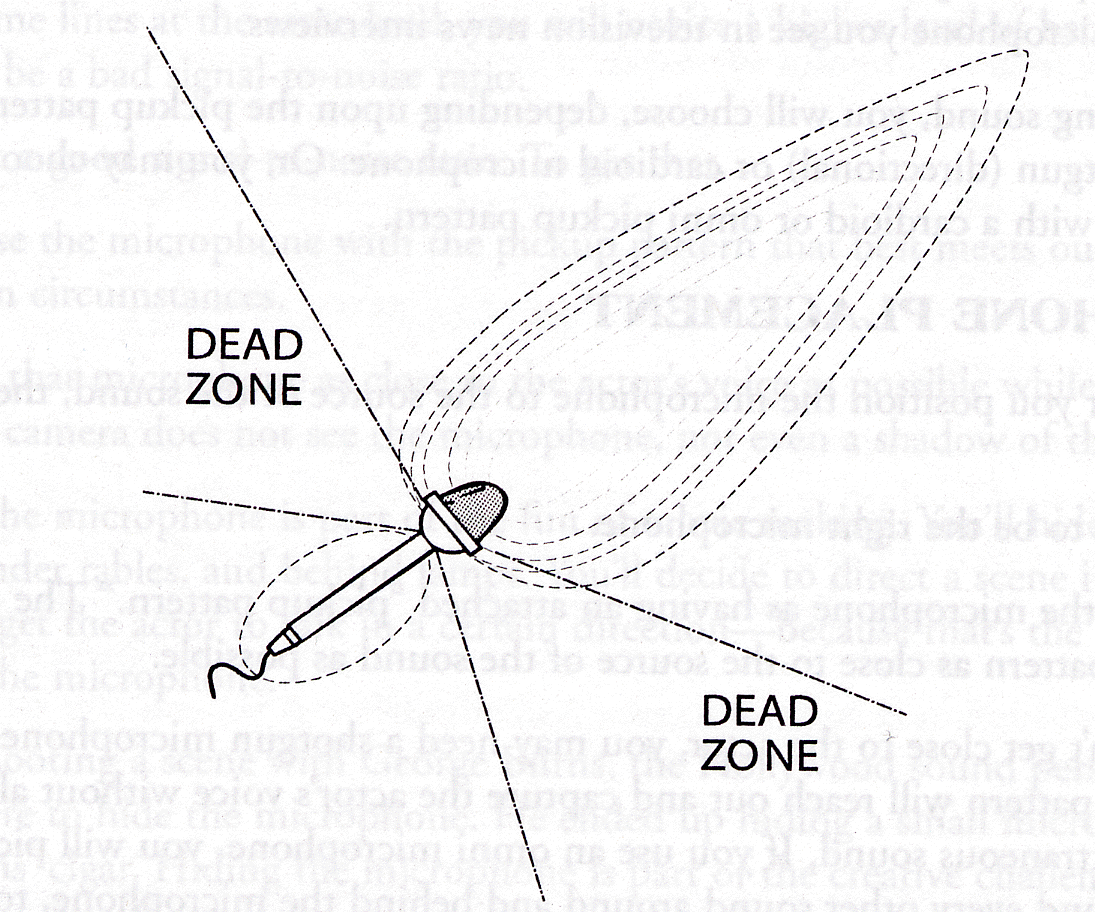


Neural network representation



## Visual

Machine learning also enables the autonomous drone to have eight different algorithms to track a subject in a scene. The stereo cameras help build a 360 degree view of the surrounding environment. There is a pair for each of the front, back, left, right, top and bottom of the aircraft. These camera views are used to train a **neural network** for identifying objects in the environment. This information allows the robot to not only classify subject in real time by exploiting deep learning image recognition systems like TensorFlow by Google, DeepFace by Facebook, Project Oxford by Microsoft or hosted APIs such as Google Cloud Vision, Clarifai, Imagga but also to avoid obstacles while flying.



Range of shotgun microphone

## Audio

The high fidelity audio is recorded and stored onboard using large range shotgun microphones to keep rotor sounds out of sound capturing area of microphones and additional nose cancelling microphones to further dampen the rotor sounds.

Also the ambient noise and rotor sounds can further be minimised in post processing. Then the sample data can be matched with the help of **acoustic fingerprinting.** It will be used to identify an audio sample or quickly locate similar items in an audio database. Bird sounds vary widely. Therefore machine learning is used to organize thousands of bird sounds. The computer is give only the audio. Using a technique called **t-SNE**, the computer creates a map, where similar sounds are placed closer together.



This is built by Kyle McDonald, Manny Tan, Yotam Mann, and friends at Google Creative Lab, Cornell Lab of Ornithology. The sounds are available in the Macaulay Library’s Essential Set for North America.

The code is freely available from their repository on Github.

# Hardware Specifications

SIZE

13”x16”x2”

WEIGHT

2.2 lbs

BATTERY

16 minutes each

MICROPHONE

Neumann KMR82i (Shotgun Microphone)

TOP SPEED

25 mph

 CAPTURE

* 4k, 30fps
* 1080p, 30, 60fps
* Sony IMX377 -- 1/2.3" sensor
* Vibration isolated
* 150° FOV
* 64GB onboard storage
* 1.5 hrs @ 4K, 4.5 hrs @ 1080p30
* Stabilized  
  (2 axis mechanical gimbal,  
  1 axis flight control)
* Pitch range: +50° to -105°
* Roll range: +35° to -35°

 COMPUTER

* 256-core Nvidia Jetson GPU
* Quad-core 64-bit ARM CPU
* 4GB RAM

 LIVE VIDEO STREAM

* 720p, 15fps
* Wi-Fi range 300 ft (Depending on phone)

 SENSORS

* 12x navigation cameras for omnidirectional vision
* 1x user video camera
* 4x IMU
* 1x GPS
* 1x barometer

 SHARING

* Instant clip creation after every flight via the Skydio app
* Full flight recording can be downloaded, edited over USB

# Results/ Summary

The drone will be able to autonomously fly over an area avoiding obstacles in its path, calculate the trajectory of objects 4 seconds ahead into the future, plan out paths and change them as required in real time.

Capture audio samples and store them onboard for sample matching later with the database available in order to recognize the different species of birds available in that area(only birds because the data is available for birds only as of now).

Further it will be able to recognize basic differences between animals and classify them accordingly(as dog, cat, lion zebra,etc. Links to sample videos given in references.).

Also it will be able to stream the video in real time to the base station over wifi with a range of over 300 ft.

The training of neural network will require a lot of sample data and a lot of time but as the drone is used it will also analyse the data it records and get better with time.

# References

<https://www.skydio.com/technology/>

<https://www.jetsonhacks.com/2018/02/14/skydio-r1-jetson-tx1-based-self-flying-camera/>

<https://experiments.withgoogle.com/bird-sounds>

<http://macaulaylibrary.org/guide/audio/essential-set-for-north-america>

<https://github.com/googlecreativelab/aiexperiments-bird-sounds>

<https://en.wikipedia.org/wiki/Acoustic_fingerprint>

<http://www.dvinfo.net/forum/all-things-audio/239274-looking-suggestions-shotgun-mic-long-distances.html>