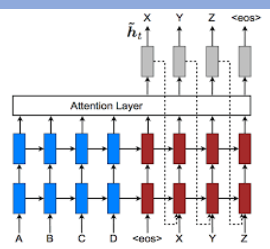
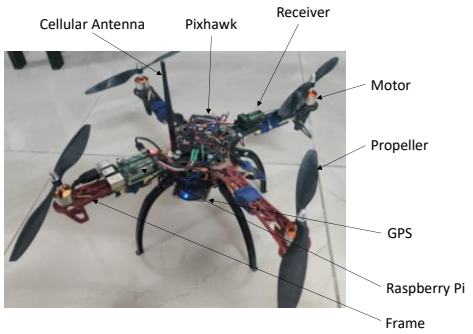
Voice command and Control of IOT Unmanned Aerial (Drones) and Ground Vehicles using Deep Learning, Cellular Communication and Dronekit

Neural Networks and Deep Learning (AttRNN)



Virtual Private Network (VPN) - ZeroTier

Drone



Drone Flight Test

Objective

The main objective of the project is to control a UAV/UGV with voice commands. This is carried out **in two parts:**

- **Development of DL model for Speech Recognition,**
- **Real time application by integrating this model into UAV/UGV device**

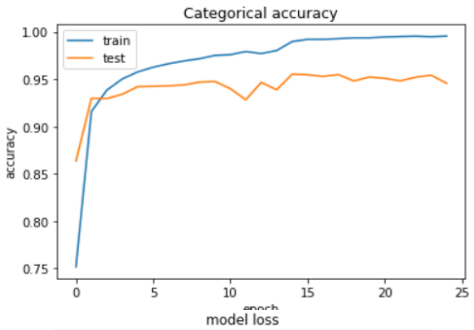
Attention RNN (AttRNN)

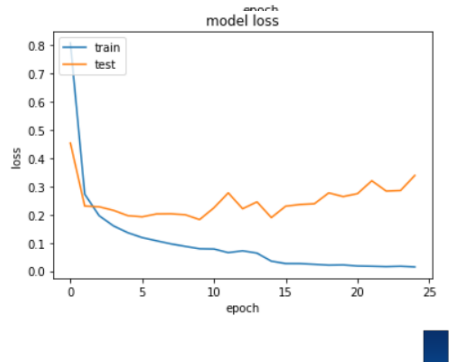
Recurrent Neural Networks (RNNs) is being used in successful tasks involving sequential data.
Attention is a technique combined with RNN that allows to focus on only certain parts of the input sequence when predicting a certain part of output sequence. It enables easy learning with high quality.

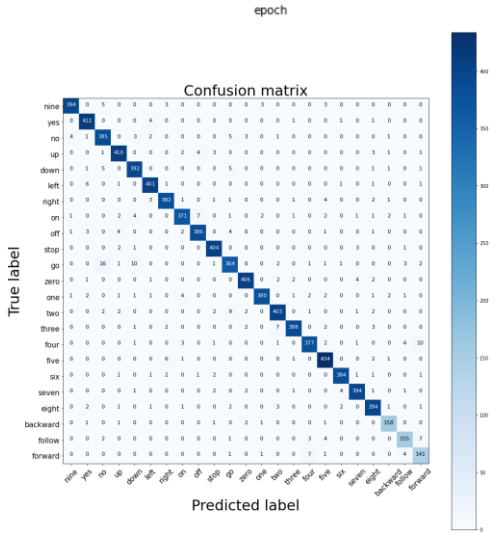
Dataset

The dataset used is Google speech Command Dataset V2 (sd_GSCmdV2) – Audio dataset of spoken words
A custom dataset of 23 commands: nine, yes, no, up, down, left , right, on, off, stop, go, zero, one, two, three, four, five, six, seven, eight, backward, follow, forward

Model







Command	TPR	FNR	TNR	FPR
Nine	0.9656	0.0344	0.9991	0.0009
Yes	0.9832	0.0168	0.9936	0.0064
No	0.9506	0.0494	0.9962	0.0038
Up	0.9647	0.0353	0.9980	0.0020
Down	0.9655	0.0345	0.9973	0.0027
Left	0.9733	0.0267	0.9984	0.0016
Right	0.9646	0.0354	0.9983	0.0017
On	0.9368	0.0632	0.9983	0.0017
Off	0.9601	0.0399	0.9985	0.0015
Stop	0.9829	0.0174	0.9984	0.0016
Go	0.9054	0.0946	0.9964	0.0036
Zero	0.9712	0.0288	0.9986	0.0014
One	0.9523	0.0477	0.9993	0.0007
Two	0.9504	0.0496	0.9980	0.0020
Three	0.9580	0.0420	0.9990	0.0009
Four	0.9425	0.0575	0.9983	0.0017
Five	0.9752	0.0248	0.9972	0.0028
Six	0.9746	0.0254	0.9989	0.0011
Seven	0.9704	0.0296	0.9986	0.0014
Eight	0.9656	0.0344	0.9973	0.0027
Backward	0.9575	0.0425	0.9988	0.0012
Follow	0.9011	0.0989	0.9987	0.0013
Forward	0.9096	0.0904	0.9971	0.0029

Applications

- Pick objects from one place and place them at another
- Can be used to reach small places in fire situation or highly toxic areas.
- In military, can be used to observe enemy camp.
- Product delivery
- Live streaming events

Advantages

- High accuracy speech recognition DL Model
- Can fly at high altitude
- Can fly at a larger range/distance
- More Flight Time

Limitations

- Can misinterpret the command
- Command can be treated as silence

Future Scope

- FPV (first-person view) camera
- Gesture controlled flying
- Live location tracking
- Return to user/controller can be used which makes the drone to return way back to the user/controller.
- Obstacle detection and avoidance
- Development of model without using GPS mode
- Can be implemented for Ground Vehicles