Elsa – Autonomous Assistance Drone

# Technology Stack

Advanced Deep learning AI Algorithms consolidated with latest Jetson Xavier NX processor with PX4 firmware in building Elsa (Autonomous Assistance Drone).

## Hardware components used

Disintegrated Parts of Elsa
- Nvidia Jetson Xavier NX
- M.2 NVMe drive
- Raspberry pi3
- Pixhawk Flight controller with PX4 software
- Ublox NEO M8N compass
- FlySky FS-i6 2.4G 6CH PPM RC Transmitter With FS-iA6B Receiver
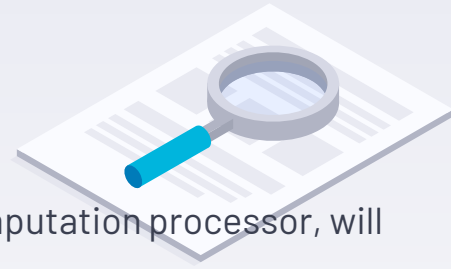- FPV Cameras for stereo vision

## Software technologies involved

- ❑ Voice based Authentication and control with **Deep learning AutoEncoder Semi-Supervised models.**

- ❑ Speech Recognition with **Quartz net** and **BERT base model** for Language understanding.

- ❑ OpenCV image processing for Stereo vision

- ❑ **Drone-kit** library to connect and control drone.

- ❑ Extended Kalman filter algorithms in localizing and stabilizing the drone to precise points.

- ❑ **MobileNet-YOLOv3** to object detection and customizing the objects in frame
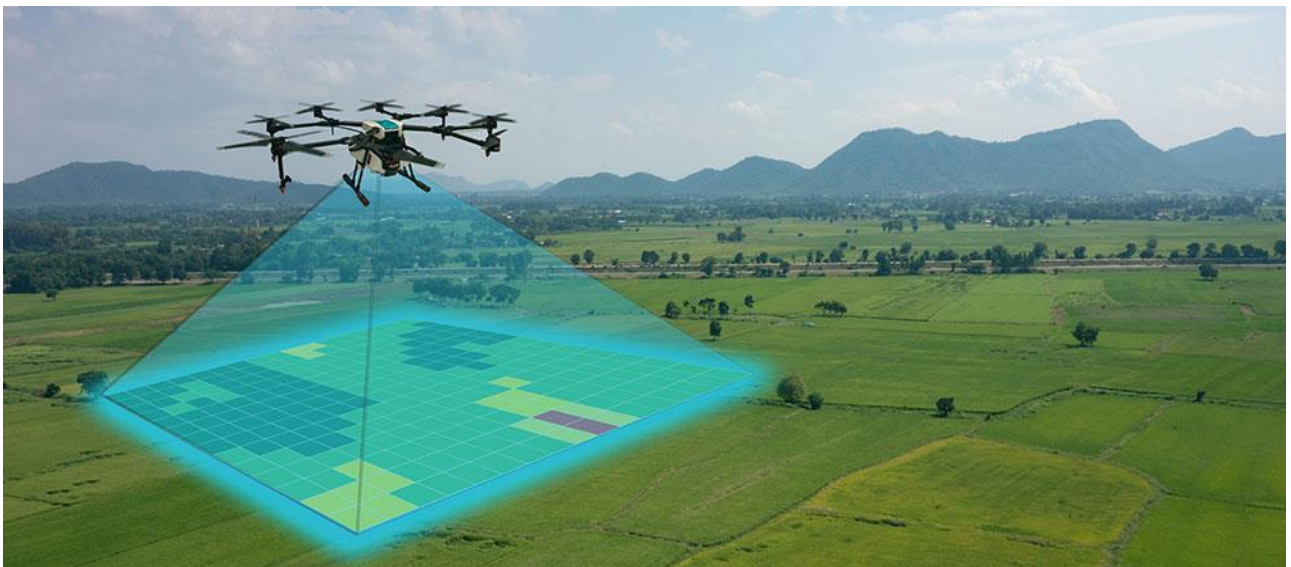
# Technology Stack

Features

▸ Complete autonomous behavior from inbound computation processor, will **not allow any network or link to hack the drone**.

▸ **Automatic charging** from charge stations

▸ Water and temperature invariant

▸ Advanced continuous updating **object detection algorithms** to detect most of usual objects and to track

▸ Voice based authentication, operating, summary extraction, triggering and scheduling tasks.

## Disintegration of Elsa operation

Elsa works with inbound processor of Nvidia Jetson NX by giving controlling signals to PX4 flight controller. Manual controls can also be received to Elsa from outbound platform connected with either Raspberry processor which is not connected directly to Elsa, App.

All the manual controls, Summary extraction and path changing signals can be possible from the outbound processor(Raspberry), Inbound processor (NX) or Application from any other device only once after the authentication is done.

# Technology Stack

❑ Voice based Authentication and control with **Deep learning AutoEncoder Semi-Supervised models.**

Deep Learning Auto Encoder Transformer based models have used in building and understanding speech patterns of large volume of speech samples. Speaker Identification model has trained rigorously on the LibriSpeech dataset, can understand the voice patterns in Authentication. Which enables one shot learning to understand and apply speech patterns to newly added user voice information and to classify it more precisely.

❑ Speech Recognition with **Quartz net** and **BERT base model** for Language understanding.

Quartznet is an end-to-end neural acoustic model for automatic speech recognition. The model is composed of multiple blocks with residual connections between them. Each block consists of one or more modules with 1D time-channel separable convolutional layers, batch normalization, and ReLU layers. It is trained with CTC loss. The proposed network achieves near state-of-the-art accuracy on LibriSpeech and Wall Street Journal, while having fewer parameters than all competing models.

❑ OpenCV image processing for Stereo vision

In **stereo vision**, eyes capture two different views of a three-dimensional object. Retinal images are fused in the brain in a way that their disparities (or parallaxes) are transformed into depth perception, yielding a three-dimensional representation of the object in the observer's mind.
Here OpenCV based Image processing have used in disintegrating the object points and to find the angle, position and distance associated with objects.

# Technology Stack

❑ **Drone-kit** library to connect and control drone.

DroneKit-Python allows developers to create apps that run on an onboard <u>companion computer</u> and communicate with the <u>ArduPilot</u> flight controller using a low-latency link. Onboard apps can significantly enhance the autopilot, adding greater intelligence to vehicle behaviour, and performing tasks that are computationally intensive or time-sensitive (for example, computer vision, path planning, or 3D modelling).

## Deploying DL models with NVIDIA – Docker

Elsa uses 3 main Deep learning models.
- MobileNet YOLO Object Detection.
- Speaker Identification AutoEncoder model.
- Language Understanding BERT Base model.

To enable efficient use of system resources, faster delivery, portability here NVIDIA- Docker containers have used. Nvidia Docker allows CUDA programming models to use Graphics effectively.