How can Autonomous Drones improve Domestic Security?

Table of Contents

1. Introduction to Autonomous Drones in Domestic Security

Autonomous drones are increasingly being embedded into security systems, both in law enforcement and private security, as they can conduct **real-time surveillance**, **patrol,** and **monitoring tasks.** However, some autonomous drones are used in non-enforcement applications such as logistics and delivery. These drones are usually equipped with facial recognition systems, which Convolutional Neural Networks or CNNs often power to provide accurate face detection and identification in real-time, along with their ability to patrol pre-defined routes, allowing them to conduct automated patrols over residential and/or commercial areas. This is particularly useful in scenarios where **crime prevention**, **crowd surveillance,** and other matters may require a large number of human personnel to cover, making it limited or ineffective. Companies such as DJI or Autel Robotics are increasingly benign and embedded in and tested in both security and logistics services, which demonstrate their dual-use capabilities. (Vallee and Coste, 2019; Goodfellow, Bengio, and Courville, 2016).

Although these drones provide clear benefits, they raise concerns about privacy, cybersecurity (in regard to takeovers or interference), and most importantly, algorithmic bias. This is the apparent bias of an AI model against a certain ethnicity, gender, or race due to the data it is trained on, highlighting the need for proper regulations and ethical frameworks (Xu et al., 2024).

2. Core Technologies in Drone Surveillance

2.1 Convolutional Neural Networks (CNNs) for Face Detection

Convolutional Neural Networks which are explained in-depth in an MIT Press publication by Goodfellow and associate writers, are essential for computer vision implementations such as facial recognition and detection, CNNs are usually well-suited for drones as they can process images in real-time with relatively low computational power needed as they are typically optimized for **edge-computing devices**. Research has

shown that lightweight CNN architectures, such as MobileNet of FaceNet are able to achieve **low-latency performance** with minimal power consumption by using optimized interference models, this would be both beneficial to the drone's face detection efficiency and battery efficiency as it does not rely on extensive hardware, which is the case in a autonomous drone (Liu, Cao, and Duan, 2023). These CNNs can identify key facial features such as eyes and noses and match them within entries in a database, the CNN frameworks mentioned previously are able to work with low-resolution imagery and variable angles, which makes them effective outside of a controlled testing environment (Schroff, Kalenichenko, and Philbin, 2015).

CNNs use backpropagation which is a machine-learning technique that is essential in optimization, making it a better choice in both power-constricted and low-processing power devices. Additionally, CNNs take advantage of using gradient descent algorithms which make them highly effective when working with large-scale datasets and pattern recognition which means it can identify individuals in complex environments such as crowded public spaces with multiple individuals and disruptive backgrounds (Goodfellow, Bengio, and Courville, 2016).

2.2 Navigation Technologies and SLAM

Drones use a technology called **SLAM** (Sim
ultaneous Localization and Mapping) technology to create accurate real-time maps of the environment, which can ensure autonomous flight without GPS which is usually used when GPS signals become unreliable and can no longer pinpoint the active location of the drone. This can be achieved by using LIDAR or Ultrasonic sensors to scan the environment around them where the drone can compare its current data to previous data to calculate its local position, SLAM creates a map from the data received by the sensors which updates consistently as the drone moves (Xu et al., 2024). Additionally, ultrasonic sensors can be implemented in an object avoidance system where the processor within the drone adjusts its flight power and trajectory depending on the distance as ultrasonic sensors can find its relative position from other objects by emitting waves, in this case with a frequency of 40kHz, and receiving them, calculating the distance from the time taken to receive the emitted waves. This allows the drone to maintain a patrol path even in confined or crowded spaces, additionally, these sensors can be placed on the bottom side of the drone to adjust its altitude depending on its distance above ground (Vallee and Coste, 2019). However, SLAM does bring its own array of challenges, it requires high computational demand which requires significant processing power to analyze and execute instructions based on the real-time data from sensors, this requires a CPU on the system which is able to execute **edge-computing tasks** without drawing too much power, draining the power supply of the system, a good example of this is the ATMega2560 which has a clock speed of 16Mhz, while it may be small compared to processors on modern systems, it is extremely capable of processing large amounts of real-time data while drawing a relatively small voltage of 4-6 Volts.

2.3 Edge Computing & Data Privacy

Real-time processing is made possible by using **edge-computing** technology which allows PIC's in drones to function effectively. Edge computing plays a critical role in drones as it enables the usage of real-time data processing directly onto the device rather than leaving the processing to a cloud device which causes a significant delay in control and balance in flight. This ensures that low-latency operations are able to proceed, which allows them to process video feeds, detect data and make flight adjustments immediately, additionally, it reduces the dependence on network activity which is essential for drones operating in rural or areas where internet connectivity is limited. Moreover, it reduces the chance of unauthorised actions of the network between the drone and a processing server.

Why Edge Computing Matters:

- Sensitive biometric data such as faces stay on the device, minimizing the risk of data breaches due to the violation of interim communications between the drone and an external server.
- Optimized algorithms which are essential in systems which use edge-computing allow the drone to perform necessary tasks such as face identification using CNN-based face detection without draining

power, which would extend the overall flight time of the drone. Furthermore, optimized algorithms would allow complex corrections such as motor adjustments due to flight drift to be done 'on the fly' which would allow it to respond quickly to sudden changes in flight which may prevent crashes.

Security Measures

While edge-computing reduces privacy risks, both **encryption** and **data storage** are key in maintaining robustness in security systems which help prevent unauthorized access or actions taking place on the system. As drones collect a large amount of sensitive data, back-end developers have to develop a complex **cybersecurity framework** which may be developed in-house by the production company of the drone or developers from large security firms may be hired externally (Xu et al., 2024). A multitude of complex security implications must be considered and addressed to ensure safe and ethical deployment.

1. **Data Breaches**
    a. As drones are equipped with cameras and sensors that collect large amounts of sensitive information every day they can become targets for **cyberattacks**, unauthorized access to the system may lead to a breach in data which could be used in crimes such as **identity theft** (Finn and Wright, 2012). Although these types of crimes may not be limited to those with initial unauthorized access to the drone system, employees who track and maintain the drones over their deployment also have access to this information if unsecured. (Liu, Cao & Duan, 2023)
    b. Mitigation Strategies
        i. Data Encryption would be one of the most effective ways to defend against unauthorized access attacks as all the data stored and transmitted would be encrypted with protocols such as AES-256 which requires a unique decryption key which is impossible to decode. (Vallee and Coste, 2019)
        ii. Another strategy would be to limit the **access control** and enforce authentication by implementing role-based permissions such as preventing access to those without the proper clearance
2. **Secure Communication Channels**
    a. As drones can't fully operate independently as they will need to communicate with control stations or relay data to other systems, these communication links may be vulnerable to an array of attacks such as **MiM (Man-in-the-middle) Attacks** (Xu et al. 2024) which could result in altered data between the drone and the system it is communicating with
    b. Similar to the strategy used in securing the drone system from data breaches, **end-to-end** (Liu, Cao and Duan, 2023) encryption could be used which uses keys that are unique to each system, each data will be sent in an encrypted form and can only be accessed using the key available on the system, which prevents data being accessed in altered during communication.
3. **Compliance with Data Protection Regulations**
    a. Naturally, governments around the world have introduced strict regulations to monitor and dictate the use of surveillance technologies. For example, one of the strictest surveillance regulations, the **GDPR (General Data Protection Regulation)** mandated in the EU, requires organizations to minimize data collection, gain consent and protect all personal data collected by the organization, which leads to the accountability of the organization if a breach occurs. This may be an issue if used in a domestic area, as the argument could be conceived that it could be a crime to monitor those who have not pre-consented to their surveillance by the drone.

2.4 Power Management & Energy Efficiency

Power management is an important aspect of autonomous drone deployment, especially since they are programmed to execute energy-intensive tasks such as processing a variety of real-time data (Liu, Cao, and Duan, 2023), therefore optimizing energy efficiency is key in ensuring a longer flight time, better performance and an increase of operational reliability in various environments as drones may need to draw more power from its power supply in certain scenarios, such as in GPS-denied areas where SLAM takes over its navigation

functionalities over GPS. Additionally, drones will usually use a lithium-polymer (LiPo) battery which is limited in providing a large amount of power for an extended period, constraining flight duration (Vallee and Coste, 2019). Moreover, power consumption also depends on the firmware used to control and program the drone, if left unoptimised by the developer the system may draw more power than needed.

1. **Energy-Efficient Onboard Processing**
    a. As the drone's face-detection system is going to be implementing the use of Convolutional Neural Networks (CNN) to detect and identify faces, it will have to use a lightweight model due to the raw processing power a CNN requires to process image or video data. Models such as **MobileNet** or **Tiny-YOLO** are designed for edge devices with limited computational resources, these architectures reduce the number of parameters and precision operations, lower the energy consumption and still manage to maintain accuracy. (Schroff, Kalenichenko, and Philbin, 2015). However, Liu et al. (2023) propose using a **multi-cycle** that processes data in smaller chunks by splitting raw data which could reduce the power required for real-time inference.
    b. As mentioned previously, algorithmic optimization is one of the most important aspects of maintaining an energy-efficient system. Techniques such as **quantization**, **model pruning,** and **weight sharing** are widely used to minimise the computational demands of CNNs. For example, quantization reduces the precision of CNNs from 32-bit floating-points to 8-bit integers, significantly decreasing power usage. Although floating-points require more computational power to process, converting them to an 8-bit integer still manages to retain their accuracy (Goodfellow, Bengio, and Courville, 2016).
2. **Standard Wiring Setup**
    a. Brushless DC Motors are usually the preferred form of thrust-power provided when it comes to lightweight flight vehicles due to their high efficiency and low maintenance requirements (Xu et al. 2024). However, each DC motor requires an Electronic Speed Controller (ESC) which regulates the speed and rotation of the drone, converting signals from a flight controller to the motor by controlling the voltage the motor receives which changes the RPM of the drone, reducing or increasing the lift power.
    b. Power Distribution Boards (PDB) are common when it comes to power management as it simplifies the power-handling process and prevents shorts or overvoltage to the DC motors such as the **Matek PDB-XT60**, reducing the chance of malfunction during flight and production. These PDBs have solder pads where the power lines of the ESC's are connected which are handled by the microcontrollers on the PDB to prevent any electrical faults. Furthermore, the specific model mentioned allows for an XT60 connector which allows for a quick connect-disconnect system with an XT60 plug which is usually connected to a LiPo battery. However, these batteries will have to be disconnected from their current connector, stripped and resoldered to work with an XT60 adapter. Moreover, PDBs allow a 12V and 5V output at a stable 5A current which allows for external processing boards such as a flight controller and a self-dependent camera, although this large current may cause an overcurrent in low-power components which may need to run through a pass-through such as a VIN input in an **Arduino Mega 2560** to reduce its current and/or voltage.
3. **Energy Management Systems (EMS)**
    a. In drone systems, it's effective to predict power demand in energy management systems which can estimate requirements for upcoming tasks based on sensor data which will ensure optimal battery usage, disabling or enabling components depending on its usage (Liu, Cao, and Duan, 2023)
    b. Furthermore, load balancing is where the circuitry distributes energy consumption across the components from the battery to prevent overloading components, this is especially needed during any climb stages where more power will have to be distributed to the motors of the drone for increased lift and thrust, although this does not mean all systems will be evenly distributed more power as it would destroy some of the smaller, low power components such as the ultrasonic sensor. Therefore, using the PDB mentioned before would automatically redirect and assign power to needed components in tandem with the Arduino which has an in-built power distribution system. (Goodfellow, Bengio, and Courville, 2016)
    c. While fully using the battery is great for extending its overall flight time, a portion of the battery should be reserved for critical functions such as emergency landings or balancing in the case of failure to improve safety. (Xu et al., 2024)

d. Various batteries could be used in a drone, separated into different application categories.
    i.

Power Solution	Advantages	Limitations	Applications
LiPo Batteries	Lightweight, widely available, does not require special infrastructure, supported by PDB's	Limited energy density and prone to explosions when agitated	Consumer-class Drones
Solid-State Batteries	Higher energy density, safer, robust	Expensive, less developed	Industrial Drones, which require higher durability and usage time
Solar Power Assisted	Renewable, unlimited energy source, long-range option	Weather-dependent, low efficiency, and may cause sudden faults due to improperly handled solar power	Long-range surveillance
Hydrogen Fuel Cells	Extremely long flight duration, low emissions	Bulky, requires special infrastructure, as it is highly explosive	Heavy-duty long-range delivery drones

    ii. As the rest of the battery options are much more expensive and seem to have a higher risk factor than using regular LiPo batteries, the implementation of the drone in the artefact will most likely use LiPo batteries as their wiring will have to be stripped and soldered onto a PDB for power control.

2.5 Specification List

To create the specification list of the drone, I will follow something similar to FAMEUSS framework, which covers all the essential components of an artefact and allows me a detailed framework to evaluate the final artefact against.

1. **Function**
    a. The **primary purpose** of the drone is to autonomously patrol a predefined area or act as a sentry that will scan and identify individuals by performing **real-time face recognition** to identify and verify individuals in both static and dynamic environments.
    b. Another core aspect of the drone is **navigation** using SLAM (Simultaneous Localization and Mapping) to generate maps of the environment and dynamically adjust routes or avoid obstacles based on its surroundings.
        i. To have an effective SLAM system, the incorporation of ultrasonic sensors and SONAR systems is essential in detecting and determining the distance between the drone and obstacles such as walls or moving objects
    c. Data collection and processing are required when it comes to scanning and identifying faces, this can be achieved by using lightweight CNN models such as MobileNet, or as an alternative drone's send data to a central processing server which will send recorded video or a series of photos which will be transmitted securly
    d. As an autonomous system is completely independent it will also have to be capable of mitigating emergencies by itself. This can be implemented by creating a fail-safe system which will force an emergency landing or shut down if a drone detects critical failures or catastrophic route paths such as extreme turning angles or strong g acceleration
    e. The drone system should also have a home-command system which would allow remote activation and deactivation of the drone's propellors along with an emergency killswitch. Furthermore, the module should be able to be accessed through any device with has the ability to access the drone's local network.
2. **Aesthetics**
    a. The drone should be compact with a quadcopter design for effective mobility and control, additionally, the drone's base should be carbon fiber for a high strength-to-weight ratio with have a sleek black finish.

b. Various status indicators should be placed on the drone to warn others of its operation mode. It should either use an LED system that uses green, yellow and amber to report its status or a vocal system that warns those around the drone when changing its operational status.
c. Furthermore, the camera which is required for face detection and analysis should be integrated into the body of the drone and should not be protruding from the base or overall chassis which will increase the risk of the camera being damaged.

3. **Materials**
   a. **Frame:** Constructed fully or mostly from carbon fiber for its reliable properties and high strength-to-weight ratio which ensure durability and lightweight performance, this will also keep the weight of the drone down in order to attach more components if required.
   b. **Propellers:** Propellors should be made out of nylon polymers as it would balance durability and flexibility, reducing the risk of damage during accidental impacts while testing and to be able to handle the heat generated from the DC motor.
   c. **Wiring and Electronics:** Use heat-resistant and weatherproof materials to ensure consistent performance in various environments, this will ensure that faults due to changing environments will be kept to a minimum.
4. **Ergonomics**
   a. A suitable user interface must be created in order to control the drone wirelessly if data needs to be viewed in real time or if specific aspects of the drone need to be controlled. Ideally, this would be a web app that would broadcast the drone network which would allow any device with an internet browser and the ability to connect to Wi-Fi to control the drone. However, this is a massive fault when it comes to network security as the SSID and password of any drone could be found out due to brute force attacks, this makes it much more applicable during the development process.
   b. A modular design allows for quick replacement of components and ease of maintenance, such as propellers and sensors, using either snap-in mechanisms or temporary attachments such as breadboards. This would make the development stage of the artefact simpler, as components can be swapped out, configured, or modified as needed.
      i. Ideally, each component should be able to communicate with each other without workarounds, as they may slow execution time during flight, which may cause issues concerning responsiveness. Realistically, the drone should be constructed with jumper wires, as it would allow for simple and quick wiring configurations and connections and be expendable. At this stage, a PCB would have to be manufactured every time a fault occurs, leading to a higher budget required.
      ii. Batteries should either be swappable or rechargeable as this would prevent having to re-strip and re-solder batteries to a PDB every time, not only leading the higher costs but it also pose a danger as if not done carefully, LiPo batteries can burst or wires can inter-cross and cause a spark and melt the battery together, causing an explosion.
5. **User**
   a. The drone should be as user-friendly as possible, as minimal training is required due to automated operation and autonomous decision-making, although a user should be able to learn the operations and control of the drone through the web app easily, in case of emergencies or changes which need to be made.
   b. This artefact is intended for a variety of audiences if it comes into production, but most notably it would be implemented into the security sector, whether private, domestic or public as the drone is capable of both SLAM and GPS navigation which would allow it to be used in a variety of scenarios which are either changing or static. The face-detection and recognition aspect of the drone could also help identify authorized individuals in the area it is patrolling in, leading to the increased robustness of security in the area, coupled with security personnel as the drone could cover larger areas or blind spots.
6. **Safety**
   a. As the drone will encounter unfamiliar spaces, especially those without a GPS signal, the drone will have to navigate through the area itself, a combination of ultrasonic sensors, cameras, and SLAM technology for real-time obstacle avoidance, additionally a redundant sensor system should be included in the artefact to ensure the safety of the drone and those

around it if a sensor fails. These sensors will communicate with the main board which will process and respond to the sensor data accordingly.

b. Additionally, the aspect of adding a fail-safe system is extremely important, especially if no operator is supervising the drone, in scenarios where the bank angle becomes too extreme, high-g deceleration or in the case of a low-battery error, the drone should have a killswtich or an emergency landing system should be implemented to prevent uncontrolled crashes by mitigating the potential damage the drone could face, guiding the drone down or at least reducing the overall damage.

7. **Size**
    a. To follow aviation regulations for small unmanned aircraft, the weight of the drone in total has to be under 2kg, which allows me to fly it in both public and private spaces for testing, additionally, a payload capacity of 0.5kg should act as a buffer in case more components need to be added such as additional sensors.
    b. Furthermore, the drone should be easily storable as it should span across 50cm including its propellors, and a maximum height of around 20cm ensuring portability and compatibility with tight spaces. In addition, a smaller drone could help reduce redundancies which may add unnecessary weight with minimal additional function.
8. **Addendum**
    a. The drone should have an easy assembly as it should be able to be taken apart and fully assembled again within 30 minutes for easy accessibility for long-term maintenance. Moreover, the drone should be able to boot and calibrate its systems **in less than two minutes**, allowing for rapid deployment when required. It should last around 30 minutes per battery with the battery's rechargeable capability. Regarding ease of maintenance, propellors and motors should be able to be disassembled without taking apart the whole of the system as these parts need to be cleaned and maintained regularly.
    b. Although drone’s are inherently fragile, the artifact is required be durable to withstand testing conditions. First of all, the drone should be able to be shock resistant and should have a reinforced frame or use a strong material such as carbon fiber to withstand minor impacts during unprecedented crashes. Likewise, the propellors and motors itself will have to last at least 500 flight hours without noticeable performance degradation, in the case of failure the propellers and the motors should be completely modular.

3. Applications and Case Studies

3.1 Law Enforcement and Surveillance

Autonomous drones have emerged as one of the most powerful tools in law enforcement and public surveillance, serving unparalleled efficiency and coverage in urban and rural areas. Their ability to perform real-time facial recognition, monitor large crows and the ability to tracking individuals of interest has made them indisposable for modern policing and the safety of the general public. This area explores the various applications of law enforcement, supplemented by case studies and analysis of their benefits and challenges.

1. **Crowd Monitoring And Patrol**
    a. Drones which are equipped with facial recognition systems whether using a thin or thick client models can be frequently deployed to monitor large gatherings sucg as public events, concerts or protests, these drones can monitor these developing situations by hovering above the crowd and monitoring each individual.
        i. Persons of interest are able to be detected using facial recognition algorithms which can cross-reference captured faces with law enforcement databases or other databases to detect individuals who have a BOLO, a VIP or any other watchlist (Schroff, Kalenichenko, and Philbin, 2015).
        ii. Furthermore, AI models can be trained to analyze patterns that lead to crime or suspicious behavior, this can be implemented in large crows by analyzing both individual and group movements to alert and flag potential threats.

iii. **Usage in Tokyo Olympics (2021)**
    1. During the Tokyo Olympics, authorities equipped drones with cameras powered with AI-driven analytics were used to monitor crowd density and ensure compliance with COVID-19 restrictions at the time, as drones would headcount pre-defined areas and alert authorities to disperse crowds. Moreover, the implementation of facial recognition systems allowed authorities to identify individuals who did not adhere to mask mandates as they were trained on unmasked face data and cross-referenced it to capture images of individuals (Vallee and Coste, 2019).

iv. **Law Enforcement**
    1. As of late, drones are starting to be implemented in border patrol operations to detect and prevent illegal crossings into neighboring countries. Their mobility and high-altitude capability make them ideal for covering vast and complex areas, as well as covertly monitoring individuals.
    2. In hard-to-reach or accessible areas drones which are equipped with thermal cameras and face recognition are being deployed on the **US-Mexico border** which have helped identify and track individuals who have attempted unauthorized entry into the United States. These individuals are flagged and alerted to border control authorities, allowing them to track and detain them easily. Furthermore, they serve as an early-warning system for ground patrol officers as they can detect and predict areas of attempted intrusions, reducing response times (Finn and Wright, 2012).
    3. Moreover, police units in Dubai have started to implement high-speed drones in routine operations such as monitoring traffic, enforcing laws by alerting nearby units of criminal activities and a short-term solution to responding to emergencies. Furthermore, these drones use **real-time image processing** and **facial recognition to identify vehicles or individual**s who are involved in criminal activities or to monitor individuals.

b. As public safety sectors are formalizing drone units there are a few **key issues** that need to be addressed which come with these advancements.
    i. An issue persists in algorithmic bias which lies within the training data the drone is provided, leading to misidentifications that disproportionately affect certain demographics. For example, this issue persists in face-detection and recognition systems in the United States which often misidentifies people of African-American descent which leads to false arrests or false information being provided to law enforcement units (Schroff, Kalenichenko, and Philbin, 2015).
    ii. As drones function they collect information and sensitive personal data which is either processed on the drone or sent to process on servers, this data needs to be properly encrypted when sent or processed to prevent interceptions by unauthorized users. Additionally, there is a general public distrust or apprehension about surveillance drones are widespread (Finn and Wright, 2012) which is due to the potential violations of personal privacy as these units are able to cover wide and tight areas autonomously, this can be prevented by assigning no fly-zones in high-density residential areas which can be assigned within the drone's software as a software-based lock which disallows the drone to continue operations in a restricted area once it has passed through a specific GPS threshold.

c. However, there are also significant benefits to the integration of autonomous systems in the public/domestic sector.
    i. As these systems are much less reliant on human-based patrols which each individual has multiple factors to be accounted for, such as fatigue or focus, autonomous systems will usually require a central operator and maintenance or recovery teams. Unlike human-based systems these autonomous systems will always be performing at maximum efficiency as long as the system itself is well maintained, this is potentially useful when it comes to patrol as drones can cover more ground in less time which allows the re-arrangement of other resources and more effectively (Schroff, Kalenichenko and Philnin, 2015).

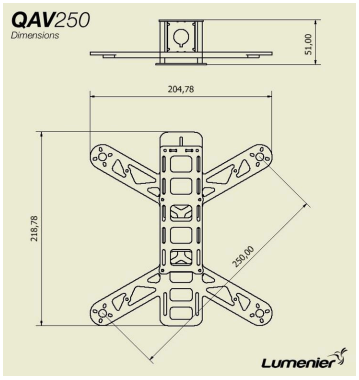
ii. Human and autonomous systems do not have to be mutually exclusive as these systems can be symbiotic with one another. For example, by providing real-time situational awareness and forward reconnaissance, it can minimize the risk to security officers during potentially dangerous encounters as these risks will be monitored and flagged by the operator (Goodfellow, Bengio, and Courville, 2016).

iii. While the initial investment in drone technology is significant, drones reduce long-term costs by minimizing manpower requirements and wages that need to be paid. Furthermore, these drones can be upgraded and repaired over time as they are maintained, which will allow them to last for a significant period (Liu, Cao, and Duan, 2023)

4. Development & Experimentation

4.1 Initial Development:

Before compiling a list of components to create the drone artefact, a chassis or a frame must be chosen that is both lightweight and has sufficient space for all components and electronic speed controllers for the motor. Due to these conditions, I have decided to use the QAV250 quadcopter frame by Luminier as the frame comes with a mixture of aluminum and carbon fiber parts which offers a good strength-to-weight ratio.



(fig. 1, Dimensions of QAV250)

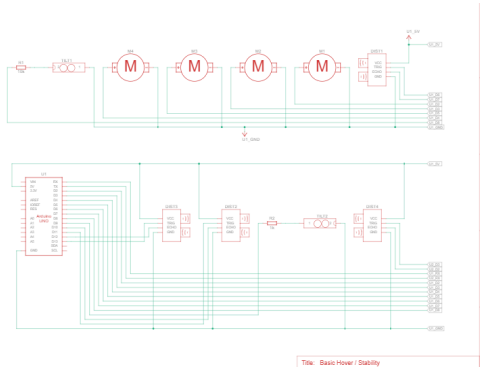
As seen in (fig. 1) the drone frame has a length of 220mm and a width of 205mm with a diagonal length of 250mm which is ample space to fit the Arduino and the DC motors which are 53mm x 100mm and 45mm x 24 x 9mm respectively. After the frame has been chosen and finalized a component list can be compiled which would comply with the sizing of the frame.

The parts list will include the following:

- Arduino Mega 2560
- X4 A2212 1000KV DC Brushless Motors
- MPU6050 Gyrometer/Accelerometer IMU
- ESP32-DEVKITV1
- ESP32-CAM
- Matek Power Distribution Board
- x4 40A ESC
- X5 HC-SR04 Ultrasonic Sensor

To be able to synthesize sensor fusion, I would need a powerful microcontroller board with an abundance of digital/analog pins such as the Arduino Mega 2560 as this board uses a 16Mhz processor which provides ample processing power for the data that the board will handle. Furthermore, the microcontroller has 256Kb of memory which is required when the microcontroller will need to potentially store some data and the large instruction set that will be programmed into the microcontroller. To assist the Arduino Mega I have also

included the ESP32-DEVKITV1 which has a high processing speed of 240Mhz to handle extensive serial communications in tandem with web communication from controller devices which can adjust the values of the drone which will be forwarded to the main microcontroller board (Arduino Mega) to apply these changes to the ESC's themselves. On paper, the ESP32 microcontrollers that I have chosen for auxiliary functions outperform the Arduino Mega in both storage and processing power. However, these ESP32 boards lack the abundance of digital, PWM, and analog pins that are extremely important when it comes to adding components in the future and synchronizing the data between components. Moreover, all boards will communicate at a 57600 baud rate compared to their default 9600 baud rate to be able to send, process, and receive more data at the sacrifice of using more power as baud rates are the rate of how many signal changes can be executed per second.



(fig. 2, Circuitry plan for a basic balancing system)

As these components are fairly non-standard, the initial circuit plan for the drone *(fig. 2)* will replace some components with more basic components such as the MPU6050 being represented by two tilt sensors which effectively have similar functionality. The schematics also lack an ESC for the motors as in the initial code schematics controlled the drone directly, which will not be the case in the implementation of the final artifact. Additionally, the HC-SR04 will work in tandem with the MPU-6050 as part of the SLAM system initially covered in the research phase of the project, these HC-SR04 sensors will constantly send data to the Arduino to essentially create a theoretical digital map of its surroundings as it sends an ultrasonic signal which reflects off of surfaces and back to the sensor, the constant speed and the time it takes for the signal to return to its sensor dictates the distance between points in its surroundings and the drone. Along with the MPU6050 which tracks gyroscope and accelerometer data it sends raw gyro/acceleration data to the Arduino Mega to prevent drift and sudden tilts which could cause the drone to crash. Moreover, the Arduino Mega will conclude if a sensor reports a point that is too close to the drone which would allow to set a temporary 'goal' for the gyroscope to achieve which would cause a bias in a side of the motors which would allow the system to evade and avoid any obstacles which are near or incoming.

4.1.1 SLAM

1. A Proportional-Integral-Derivative or simply PID controller is widely used in autonomous drone flight stabilization and obstacle avoidance systems, with the combination of the HC-SR04 ultrasonic sensor which would serve as obstacle detection sensors and the MPU6050 IMU these sensors can provide real-time feedback which would assist in adjusting the drone's flight dynamically. Additionally, implementing **SLAM technology and techniques** will allow the drone to map its environment while navigating
    a. *Roll, pitch, and yaw control:* MPU6050 gyroscope and accelerometer
    b. *SLAM-based Navigation:* Optimizing drone position relative to the environment from sensor data
    c. *Altitude stabilization:* HC-SR04 Ultrasonic Data
2. **PID Control Equation** (Goodfellow, Bengio, and Courville, 2016)
    a. $$U(t) = K_p e(t) + K_i \int e(t)dt + K_d \frac{de(t)}{dt}$$
    (fig. 3, PID control equation)
        i. U(t) denotes the control signal sent to the motors

ii. e(t) denotes the difference between the desired and actual state
iii. Kp denotes the proportional gain which determines the reaction to the current error
iv. Ki denotes the integral gain which eliminates steady-state error by considering past errors
v. Kd denotes the derivative gain which predicts future errors and prevents overshoot

b. In the context of altitude control where PIDs are extremely effective the following factors are calculated as follows (Vallee and Coste, 2019):
   i. Error calculations which were stated earlier denoted as *e(t)* is the difference between the desired altitude and the current altitude which can be measured by an HC-SR04 sensor.
   ii. These error calculations play a factor in the overall control output which will adjust the motor speed by sending signals to the ESC based on the PID output which will be determined by the formula stated earlier, this would smoothly adjust the drone's altitude to maintain stability.

c. Another context the PID would be essential is the balance of the drone during a hovering altitude, the PID would be responsible in determining motor speeds by monitoring the roll, pitch, and yaw which are synthesized by a combination of gyroscope and accelerometer data.
   i. The sensor responsible for monitoring and synthesizing the data is the MPU6050 which is a 6-DOF IMU that also contains fusion algorithms such as Kalman Filters to integrate data for accurate orientation (Thrun, Burgard, and Fox, 2005).
   ii. The error calculation remains similar to the altitude control calculations as *e(θ)* represents the difference in each axis of the IMU (roll, pitch, yaw) which will be integrated into the PID calculation mentioned above.

1. $$\theta = \arctan\left(\frac{\text{Acc}_Y}{\sqrt{\text{Acc}_X^2 + \text{Acc}_Z^2}}\right)$$
$$\phi = \arctan\left(\frac{-\text{Acc}_X}{\sqrt{\text{Acc}_Y^2 + \text{Acc}_Z^2}}\right)$$

(fig. 4, Pitch and Roll calculations)

2. These formulas in the *(fig. 4)* are responsible for calculating the pitch (θ) and roll (ϕ) angles as Acc(n) is the accelerometer readings along the different X, Y, and Z axes respectively as it measures the gravitational forces along these axes.

3. $$\psi_t = \psi_{t-1} + \omega_Z \cdot \Delta t$$

(fig. 5, Yaw calculation)

4. Unlike pitch and roll, the yaw is determined using gyroscope readings as the gyroscope provides angular velocity in degrees per second, leading to an approximate estimation using numerical integration where:
   a. ψt denotes the current yaw angle
   b. ψt−1 denotes the previous yaw angle
   c. ωZ denotes the gyroscope reading along the Z-axis
   d. Δt denotes the time step between measurements

d. When calculating the pitch and roll values, sometimes it may be necessary to integrate gyroscope data into accelerometer-based data to improve accuracy by implementing a complementary filter that combines both data through a coefficient usually set between 0.90 and 0.98.

$$\theta = \alpha(\theta_{\text{previous}} + \omega_Y \cdot \Delta t) + (1 - \alpha)\theta_{\text{acc}}$$

e. $$\phi = \alpha(\phi_{\text{previous}} + \omega_X \cdot \Delta t) + (1 - \alpha)\phi_{\text{acc}}$$

(fig. 5, Pitch and Roll calculation with sensor fusion)

3. Combining these two different sensors allows *IMU-based SLAM* which will primarily use the MPU6050 combined with sensor fusion which will allow the drone to maintain stability in GPS-denied environments and independence.

4.1.2 CNN-based Face Recognition

1. Due to the removal of the ESP-32S1 and S2 chips series support in ESP-WHO which is a CNN-based face detection module, the drone will rely on server-side processing of images that are forwarded from the ESP32 which will act as a webcam which will continuously send image data to the processing server, this server will generate a simple web-page showing the preview of the video stream and if an individual is recognized or not. The integration of server-side CNN processing for face detection allows for real-tine recognition while offloading a heavy computational workload which may be too overbearing for an ESP32 to handle as CNN's are a relatively heavy artificial neural network. To implement this workflow, the server is implemented using Python, Flask, OpenCV and face_recognition Python modules where the workflow uses a client-server architecture where the ESP32 streams video as the serve processes individual frames and performs CNN-based face encoding and matches detected faces against a pre-enroled dataset, such as a folder of images or a simple database.
    a. The overall architecture of the face recognition system is split into two parts, the ESP-32 CAM client which streams the live video over HTTP with an FFMPEG encoding and acts as a wireless feed provided with minimal on-board processor in order to prevent overheating, and the Python flask server which captures the ESP32 video stream using OpenCV and downscales the frames to make the process more efficient. Once downscaled, the frame is analyzed as the CNN model extracts facial features for encoding which is matches the detected face against a preloaded image dataset which the CNN is trained on.
2. Unlike using the trivial architecture of Histogram of Oriented Gradients (HOG)-based methods which would be much more computationally efficient than CNN models but struggle with complex lighting and angles which may cause an issue during face-detection with an indoor environment. Meanwhile, a CNN-based face detection model leverages the usage of deep learning models which are trained on large datasets that allows for improved accuracy (Goodfellow, Bengio, and Courville, 2016).
    a. The system attached to the drone will be split into four different processing sections, creating an overall pipeline.
        i. Video frames are streamed from the ESP32 WebSerber program to the Python flask server by allowing the server to access thewebaddress of the WebServer's stream. To execute this process, both devices will need to be connected to the same WiFi SSID, which in this case will be a shared wireless network instead of a self-hosted network.
        ii. As the frames streamed are considerably heavy to process due to their size, the frames are resized and downscaled to reduce the computational load on the server and converted to an RGB format as the CNN model which the system will be using is trained in color spaces. While the system could use a lower frame resolution during the streaming process, the server also hosts a webpage where individuals can observe the frames which are being streamed and the individuals detected which needs to be in a higher quality frame than the processed frames.
        iii. In order to process the faces using a CNN model, the system will be using the face_recogintion module in Python which applies a pre-trained deep CNN model to locate faces (Schroff, Kalenichenko, and Philbin, 2015). This CNN will also create a dataset consisting of known faces which will help it recognize known faces during its runtime.
        iv. After extracting the stream it is compared against the stored reference encodings by using Euclidean distance metrics, as these values re compared a threshold value can be set in the system, starting from 0.0 and ideally ending at 1.0 where the higher the value equates to the system being more lenient overall.

3. Using a CNN-based recognition gives superior accuracy than traditional HOG methods (Schroff, Klaenichenko, and Philbin, 2015) along with the scalability of server-side implementation which could allow the integration of multiple ESP32-CAM. As I am using a pre-trained CNN model it uses its own dataset which has trained it to recognize individuals, the model may need some fine-tuning with a custom dataset as the model may have demographic biases (Goodfellow, Bengio, and Courville, 2016).

4.1.3 Command-Over-WiFi

1. As the drone requires fine-tuning adjustments during its testing period and constantly uploading new code to change small parameters would take a considerable amount of time. Additionally, the drone should be able to be switched off and on and switch modes remotely as mentioned by the specification. To execute this effectively I attached an ESP32DEVKIT-V1 to the frame of the drone and wired it to the Arduino with the following:
    a. VCC to shared 5V High-Current Power rail
    b. GND to bridged Arduino GND
    c. RX(0) to TX(1)
    d. TX(0) to RX(1)
2. The ESP32 used in this system has an inbuilt WiFi module allowing it to host its network and display webpages flashed as part of the software on the microcontroller. The system will be able to facilitate command transmission from a webpage to the Arduino Mega over UART which is a form of communication through the serial pins of these devices, the Mega will be able to receive these commands by parsing its serial input from the ESP32 and execute commands when needed, such as adjusting PID parameters or an emergency kill switch. This architecture allows for low-latency wireless communication which allows users to effectively control the drone either for safety or adjustments without re-flashing the firmware.
3. The frontend UI will be an ESP32 web server which will be hosted on its own local network which the ESP32 creates, this network can be accessed by any device that is authorized to authenticate itself into the network. The HTML-based control panel will be accessible via a browser and the user must submit each new input, this sends commands over an HTTP request to the ESP32. Once received it will process the request and convert it into a command understood by the Arduino Mega over serial communication. In return, the Arduino will return any debug outputs to the web-bage allowing the user to see both the internal serial output of the Arduino and the communication exchange between the ESP32 and the Arduino. Both devices will be connected at the same baud rate of 57600 which gives a good balance between speed and stability, which is essential especially for sending killswitch commands as a fast response time is needed for operator safety.
4. Finally, the Mega receives the serial communication from the ESP32 and parses the incoming commands, if any of the commands sent from the ESP32 matches a set value of ‘known commands’ the Mega will adjust the drones parameters or execute predefined actions and send a status response back to the ESP32, confirming or informing the user of the Mega’s actions.

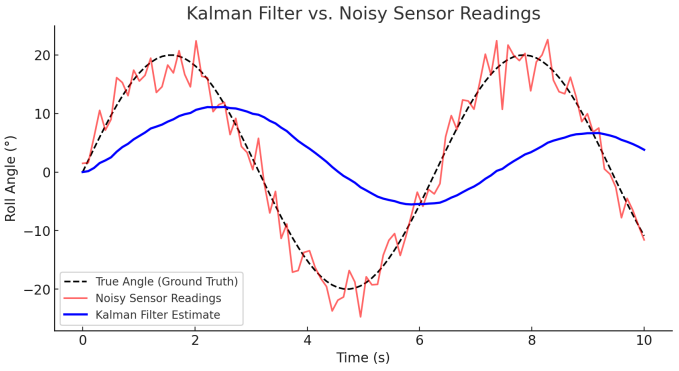
4.2 Testing and Improvements

A Proportional-Integral-Derivative (PID) controller is an essential component in flight control and balancing systems, it needs to be manually fine-tuned to accommodate for its weight distribution and any flaws the IMU module may have. Accurate PID values allow for precise adjustments in roll, pitch, yaw and altitude stabilization by correcting the natural or manufactured deviations from a setpoint. However, this process is extremely manual with hours of testing and adjustments to ensure the Kp, Ki, and Kd values responsible for PID adjustments mentioned above are as accurate as possible giving the drone the mid-flight stability and autonomous foundations it requires.

Furthermore, the CNN-based face recognition and detection system will require adjustments to optimize the processing of streamed images, as multiple faces in the stream will substantially increase the processing time required to process and scan the faces, which would not be ideal in a real-world environment as faces need to be scanned and processed almost instantaneously.

4.2.1 Manual PID Fine-Tuning for Stability and Control

1. As mentioned in *fig. 3* previously, understanding the PID equation is essential to understanding the role of each parameter. Each term affects the drone's response differently, with Kp directly affecting if the drone overshoots or undershoots the corrections the system makes. Moreover, an excessive Ki may lead to excessive correction which would cause an uncontrollable chain reaction of corrections attempting to override each other, causing mass instability. Finally, the Kd value of the PID controller affects the damping value and the response delay of the PID controller, with an excessive value leading to heavy damping and a slow response, which may allow the drone to bank at a certain angle for too long which would disallow the PID to regain stability of the system.
2. Additionally, the MPU6050's accelerometer and gyroscope readings, which are essential for estimating roll, pitch, and yaw, are noisy, which may lead to spikes in readings, causing an overreaction in the system unless proper handling is implemented. To mitigate this, the sensor values will go through a Kalman filter which is a technique that combines multiple sensor readings, reducing noise and improving stability (Thurn, Burgard, and Fox, 2005).
    a. The Kalman filter is an optimized recursive algorithm that can estimate 'hidden' system states such as the actual roll, pitch, and yaw from noisy sensor measurements. It provides the system with a predictive and corrective approach, where it continuously refines and estimates by considering past values, current measurements, and uncertainties. The filter uses the previous state and motion model to estimate the next state and adjusts the estimate based on new sensor readings, this is contrary to simple moving average or low-pass filters as the Kalman filter can adapt dynamically to measurement noise and changing system conditions, making it robust for autonomous systems. However, the Kalman filter requires matrix calculations which can make it computationally expensive for microcontrollers. Below is a visual representation from simulated noise and simulated true values and how a Kalman filter creates a stable product of noisy readings.



(fig. 6, Kalman filter graphical representation.)

3. Since the drone system has a Command-Over-WiFi system, when adjusting the PID values during testing, the system does not have to flash new firmware after each attempt. Instead, new K parameters can be loaded into the drone's firmware through the drone's personal control network, allowing for dynamic adjustments. Every drone system will have different 'K parameters' as a 1.0 value on every 'K parameter' may be viable for one drone system but it most likely would not work for another system if transferred due to differences in motor power, weight balancing, and other unseen factors.

-

4.2.2 Optimizations of CNN-Based Face Recognition

1. Given that deep learning models, such as the one used in this face recognition are computationally expensive, necessary optimizations are necessary to balance processing speed without sacrificing recognition accuracy. There are common bottlenecks in these systems:
    a. High latency may occur due to processing each frame with a CNN model which can cause unnecessary delays especially if adjacent frames are extremely similar. This is because each frame extracts facial features of the individual frames, comparing and matching the extracted features with the stored references to identify individuals. This could be improved by skipping every nth frame as the camera streams up to 60 frames per second, with sub-30 frames per second under load, a portion of the streamed frames can be skipped with minimal issues in the context of image processing and extraction.
    b. As the *face_recognition* python module is a relatively heavy deep learning face-detection model, which will lead to an extensive overhead in processing time, the processing server must be able to process multiple frames simultaneously. While using a more optimized module would decrease the processing time and overhead, it may lead to a less accurate recognition of multiple faces. Therefore, integrating GPU acceleration inside the Python flask instead of running the CNN on a single-core process would give the server superior performance as CNN models involve millions or parameters that must be processed efficiently. GPU's are inherently designed for parallel computation, making them significantly faster for matrix operations and convolutional layers. Simultaneous frame processing with a CPU is inefficient as CPUs process operations sequentially inherently, whereas GPUs process thousands of operations in parallel. RTX GPU's, in particular, contain AI Tensor cores which are dedicated to deep learning workloads, and CUDA cores for optimized general-purpose parallel processing with the usage of Nvidia's CuDNN tool to accelerate the usage of their CUDA cores for artificial neural network usage or development, using both of these architectures the recognition server can offload complex CNN computations to specialized hardware, reducing processing time from hundreds of milliseconds.

4.3 Results and Experimentation

It is critical to assess the efficacy of autonomous-based systems in comparison to human performance to gauge the vast difference between efficacy and performance between these systems. AI-driven facial recognition systems have been increasingly adopted to combat the issues which lie in human visual identification systems which may be susceptible to cognitive bias and slower recognition speeds.

By evaluating both human and AI/CNN performance under controlled experimental conditions, it is possible to conclude the strengths and weaknesses of each approach and determine which system performs better overall intending to provide empirical evidence for the superiority of CNN-based facial recognition systems in security applications

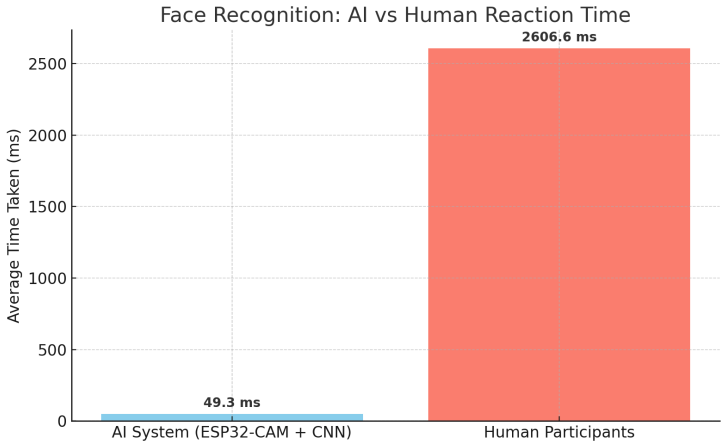
4.3.1 Methodology for Experimental Evaluation

1. To systematically assess the performance of humans in face identification, participants will be subjected to a single recognition task. The task will be conducted under timed conditions to measure both the accuracy and reaction times of participants.
2. A dataset of fifteen faces will be presented to the participant, there will be three stages of the task with each stage including a single reference image within the dataset. Once the participant is ready, the participant will observe the reference image and will be handed the dataset instantaneously along with a timer being started, the participant will not have constant access to the dataset ensuring the participant cannot study the dataset beforehand. Once the participant selects the correct face in the dataset according to the reference, the timer will be stopped, any incorrect identifications will be recorded as wrong attempts. On the other hand, the AI-based recognition system will be modified to code a timer within the system, to prevent the most accurate timing results possible. To achieve this, the program runs a timer on a separate thread of the CPU, making its processing separate from the facial recognition which does not add overhead. Once a face is detected, the timer will start and then stop once the correct face is detected.

a. The selection of faces in the dataset will have to be visually similar and must not have any unique facial features or artifacts such as glasses or braces. The traditional face recognition problem assumes moderate variation among individuals, allowing for relatively high accuracy in matching tasks. However, in some real-world scenarios or a worst-case scenario, it often requires distinguishing between highly similar-looking individuals. To test both the AI and human recognition capabilities under these constraints, a dataset will have to be chosen without bias, which can be curated from a synthetic GAN-based face generator that produces high-quality and randomized facial images.
b. The specific GAN-based generator used to curate the faces in this data set is the website ThisPersonDoesNotExist.com which employs a generative adversarial network that is trained on real-world face datasets to synthesize unique images that do not correspond to real individuals. This tool allows a controlled similarity selection as the curator can filter through generated images for structural resemblance, race similarity, and other factors. This will allow the task to present an increased difficulty to both the participant and the AI. Furthermore, as all the faces are synthetic, no ethical issues will arise from using personal biometric data which can be a breach of privacy.

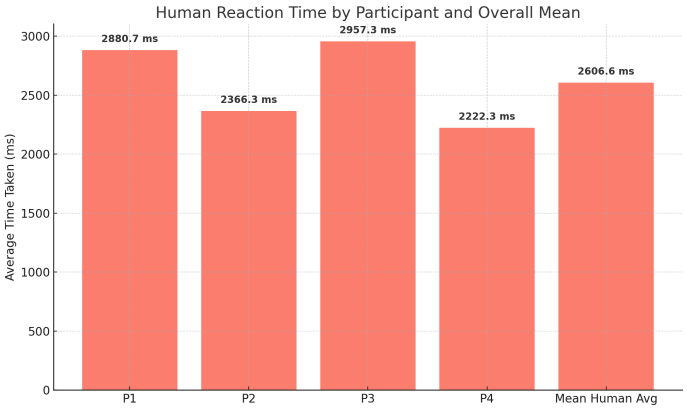
3. It is hypothesized that the CNN-based facial recognition system will significantly outperform human participants in terms of speed and accuracy due to two key factors.
    a. The CNN-based facial recognition system allows for a form of processing called batch processing which allows multiple faces to be processed simultaneously by utilizing the parallel processing method, unlike humans which can only analyze one face at a time. This allows the AI to compare a reference image to a dataset of a large number of faces in milliseconds whereas human participants must manually scan each face sequentially, introducing delays (Philips and O'Toole, 2014).

4.3.2 Experimentation Results



(fig. 7, mean results of the AI system against human participants)

1. The experiment was conducted with four human participants, with two participants with active experience in the educational sector as teachers. All participants were able to accurately identify faces on the first attempt, meaning both the AI system and human participants achieved an accuracy rate of 100%, suggesting that they are as accurate as eachother. This value may change when conditions are non-ideal unlike the experimental setup where light, angles or quality of image could affect decision making by both the AI system and human participants.



(fig. 8, human participants individual reaction time and mean)

2. Participants *P2* and *P4* are individuals employed as teachers while the other two participants are individuals from the general population, with no formal training of professionale xperience in tasks involving facial recognition or identification. The hypothesis initially predicted the AI would outperform *all human participants* in the context of identification speed which was confirmed. However, human participants performance varied depending on previous or active experience which revealed distinctions that allowed for further analysis

Participant	Role	Average Reaction Time (ms)
P1	General Population	2880.7
P2	Teacher	2366.3
P3	General Population	2957.3
P4	Teacher	2222.3
Human Mean	-	2606.7
AI System	-	49.3

(fig. 9, quantitive summary of human participants and AI identification time)

3. While both P2 and P4 participants had slower recognition times compared to the AI system, they were consistently faster than general participants which mah be attributed to the professional cognitive conditioning inherent in teaching roles.
    a. Daily exposure of a large group of individuals may develop sharper short-term visual memory and pattern recognition skills, which may lead to a heightened ability to detect subtle facial differences, especially in situations where they must observe behavioural nuances in students (Murre, 2000). Furthermore, the observational skillset required in classrooms which are required by teachers could improve face familiarity heuristics, which could be better applied in more ambiguous situations.
    b. It should be noted that participants *P3 and P4* are male participants and out of the three samples used to identify to mean reaction time, the third sample of male gender was consistently identified faster by male particpants, with the same situation occuring with the first and second samples which were female that were identified faster by *P1 and P2,* female participants.

    i. Own-gender bias is a documented phenomenon which occurs in psychology and face recognition research, which individuals that matched the gender they were identifying tend to perform better (Rehnmann and Herlitz, 2007). Which is believed to be developed from differential social exposure during childhood and early adulthood in tandem with attentional encoding, as people tend to spend more cognitive effort processing faces they can more readily relate to.
    ii. While AI-systems can also be biased towards certain genders, it is not due to own-gender bias and instead due to sampling data bias where the ratio of male-to-female samples may be uneven, leading to a bais in identifying a certain gender. The same philosophy can be applied when it comes to race and facial deformities.
4. Despite the professional advantage and training, teachers were behind in performance compared to the CNN-based system in terms of raw recognition time due to the batch and parallel processing executed by the CNN system as it could detect multiple faces and compare a face multiple times in a few milliseconds which can maintain consistent recognition time without fatigue, distraction or variation across repeated trials.
5. While these results clearly show that the AI system is far superior in identification to the human participants, it is important to note that human-based systems may still hold contextual advantages. Humans may exhibit a greater accuracy when compared to observing expressional behaviour in faces which may not be interpreted accurately by AI systems due to limited sampling data. Furthermore, AI systems are always limited by the quality of training data which depends on the development of the system itself. Additionally, with lighting/angle constraints, human perception may be able to adapt to non-standard visual cues (Phillips and O'Toole, 2014).

4.3.3 Experimentation Evaluation

1. The experiment provided a critical insight into the capabilities and the limitations of both human-based recognition and CNN-driven autonomous identification systems. However, the experiment did not include the factors of diversity, realism and complexity of the dataset, as well as the environmental context that these samples were displayed in which must be examined in-depth.
2. A notable limitation of the current setup was the usage of synthetic images from a GAN-based face image generator. While these images did provide statistically randomized and prevented repetition bias, they are constructed through a *Generative Adversarial Network* which may not fully represent real-world facial variations such as asymmetry, makeup, changing facial hair or lighting inconsistencies as well as other ethnographic characteristics which the autonomous system my fail to recognise which is highlighted by Phillips et al. as facial recognition systems often perform less accurately on racially diverse populations due to the biases embedded in the training data these systems are developed in, which is unique for each CNN system (Phillips et al., 2011).
3. Furthermore, the current batch of reference and test images lacked movement or any real-time variation such as the face recognition system receiving data from an individual which turns their faces or changes environment as in real world applications such as surveillance and law enforcement, subjects are rarely stationary. A study by Kortli et al. (2020) emphasizes that CNN's can struggle with face tracking under variable illumination and motion blur which limits their deployment unless trained robustly against such conditions.
4. Ultimately, the initial experimentation strongly indicates the advantages of a CNN-based system in accuracy and reaction time over human capability. However, its evaluation shows that only by robustly testing the system against broader ethnic, environmental and dynamic diversity in datasets can the system by fully validated as a fit system for deployment in a real-world domestic security scenario, especially with the current state of the drone which is limited to slow and controlled movements which may not be suitable in scenarios where the subject of interest may be moving at fast speeds or in the event that the subject may not be able to provide stable information to the CNN-based system for proper and reliable identification.

5. Evaluation & Reflection

5.1 Summary of Key Findings

This project set out to investigate whether an autonomous drone working in-tandem with a CNN-based face detection and recognition system could outperform human operators within a domestic security and recognition context. Through research, development and experimentation, several critical insights have been made clear:

1. Within a controlled trial, the CNN-driven system consistently identified and matched faces in under 50 milliseconds consistently, which was far faster than the average human reaction time which was made by 'regular' individuals and individuals who would be more accustomed to recognizing faces which yielded around 2.6 seconds (mean across participants) under identical conditions. This advantage was most pronounced when processing large batches of test images which was the concept of the facial recognition experiment, illustrating the benefit and superiority of parallel batch processing on GPU hardware based LLM systems (Goodfellow, Bengio, and Courville 2016).
2. Under a non-controlled environment with changing lighting conditions and angle conditions, the CNN system could potentially struggle (by a unknown certain factor until further testing) depending on the dataset it is trained with, as it may contribute to providing false identifications or simply not detecting a face at all. Therefore, highlighting the need of a potential multi-modal sensor fusion such as LiDAR and Infrared instead of simply depending on an optical input. Whilst tests which were recorded during development proccess did prove that the face recognition system was able to still recognize faces under changing illumination and facial angles, further tests must be performed under strict experimentation guidelines to test the true capabilities of the systems upper limit of functionality.
3. While the performance gains of using an autonomous system is clear, there is still a significant concern around data security, consent, and algorithmic fairness (Finn and Wright 2022). Which is a byproduct of those who develop the systems themselves, as fairness may depend on the data given to the system as samples, as insufficient data will increase the risk of false positives and in some extreme cases lead to discrimination based on a persons ethnicity or gender depending on the robustness of the intitial dataset given to the drone. Moreover, data security is an excruciatingly important detail which cannot be overlooked, as any data transmission between the drone and processing servers could be intercepted unless data transmission is handled properly and correctly such as using a form of encryption, unless mitigated by having all processing and storage on the device itself which could significantly reduce the vulnerability of the data being processed, although this remains highly unrealistic as datasets would have to be extremely large and would cost a significant amount of processing power to handle CNN-based face detection on-the-fly.
4. The main factor that could make this project feasible to be implemented in the private or public sector of domestic security is the core component of a small mobile object that would hold the face recognition system, although this has raised a few issues with initially calibrating the balancing and gyroscopic systems which are crucial to its flight. Developing a PID from scratch leads to a lot of testing and tweaking especially in a light object which requires a lot of variables to consider, this results in a lot of the project's time being spent on fine-tuning the balancing system itself instead of further developing further capabilities such as GPS navigation.

5.2 Reflection and Improvements

1. While at the end of the project a proof of concept drone and autonomous security system was created, it did not achieve a lot of the additional features which were envisioned at the start of the project which can mostly be placed on the duration it took to complete the gyroscoping calibration and PID systems due to issues with computing power and issues with the quality of the components used in the PID system. This can be improved in the future by having a more specialized time slot during the development of the drone to adjust and fully develop the PID system or potentially purchasing a flight controller which could be integrated within the drones main system which would allow the development of additional components and features. Although, due to the current project, a working PID controller framework has been developed to work between Arduino and MPU systems, allowing the usage of this same framework in a future iteration in a already working capacity.
2. During the development of this artefact, a plethora of new skills have been developed, such as implementing flask-based servers for offloading processing, PID control development and increased knowledge of designing and creating embedded systems. An area which was extremely surprising

which was able to be integrated into the artefact was CUDA-accelerated CNN's which allowed for parallel processing which could not be initially done at the start of the artefact due to the constriction of running the CNN on the chip itself which is now offloaded to a separate processing server.

3. While I researched the feasibility of implementing a drone as a large part of a domestic security system, I discovered the intricacies behind the ethical aspect of autonomous security system which I did not consider at the start of the project and therefore was not integrated into the final artefact. A future iteration of this artefact could have a larger focus on data security in the scope of data transmission and storage as the face data in the current Flask server is publicly available and can be accessed by a user who is knowledgeable about server hosting.
4. Although the experimentation was a success in supporting the hypothesis that an autonomous face recognition system is superior to human recognition and detection, the environment of the tests were very clinical and static. A future iteration of this project should explore the possibility of testing the system in an uncontrolled environment with variable lighting and different angles of perception, which can then be compared to the results of an 'ideal environment' to determine its fitness in a deployment setting.
5. To ensure my observations and evaluation remain unbiased as possible, an expert in the field of Software Engineering with a Masters of Engineering and a Bachelor of Science along with experience of working in Microsoft evaluated the report and artefact, which is also further analyzed.
    a. Feedback can be found here.
    b. The evaluator has taken into the consideration the complexity of the project, labelling it as ambitious and exploring different fields, demonstrating technical depth and societal awareness. Furthermore, the expert has stated the proficiency of embedded-software and circuitry skills to create a working proof of concept, with additional consideration in adapting the plan and management of the project when requirement with transparency.
    c. However, the expert has stated tha the would like to see a more robust testing framework, such as accuracy benchmarks and more enhanced stress testing for the CNN-based face recognition system to systematically evaluate recognition accuracy and system resilience under varied conditions, simulating real-life scenarios during deployment. Furthermore, it has also been advised to strengthen the remote-control interface by adding encryption, authentication, and threat-model analysis to mitigate potential attack vectors to the system, which is currently in threat of MiTM (Man-in-The-Middle) attacks from individuals on the same network as the drone, which can lead to a hostile takeover.

6. Bibliography

Books

- Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning*. Cambridge, MA: MIT Press, 2016. https://www.deeplearningbook.org/.
- Thrun, Sebastian, Wolfram Burgard, and Dieter Fox. *Probabilistic Robotics*. Cambridge, MA: MIT Press, 2005.

Journal Articles

- Finn, Rachel L., and David Wright. "Unmanned Aircraft Systems: Surveillance, Ethics and Privacy in Civil Applications." *Computer Law & Security Review* 28, no. 2 (2012): https://doi.org/10.1016/j.clsr.2012.01.005.

- Liu, Xinyu, Chenhong Cao, and Shengyu Duan. “A Low-Power Hardware Architecture for Real-Time CNN Computing.” *Sensors* 23, no. 4 (2023): 2045. https://doi.org/10.3390/s23042045.
- Vallee, Nicolas, and Marc Coste. "Drone-Based Autonomous Surveillance Systems: Design and Implementation." *Robotics and Autonomous Systems* 126 (2019): 35–47. https://doi.org/10.1016/j.robot.2019.103451.
- Wang, Lijie, et al. "Ultrasonic Obstacle Avoidance and Full-Speed-Range Hybrid Control for Intelligent Garages." *Sensors* 24, no. 6 (2024): 129–143. https://doi.org/10.3390/s24061029.
- Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "FaceNet: A Unified Embedding for Face Recognition and Clustering." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 815–823, 2015. https://doi.org/10.1109/CVPR.2015.7298682.
- Hancock, Peter J.B., Vicki Bruce, and A. Mike Burton. *"Recognition of Unfamiliar Faces." Trends in Cognitive Sciences* 4, no. 9 (2000): 330–37. https://doi.org/10.1016/S1364-6613(00)01519-9.
- Murre, Jaap M. J. *"Learning and Categorization in Modular Neural Networks."* Hillsdale, NJ: Lawrence Erlbaum Associates, 2000.
- Kortli, Yassine, Mohamed Jridi, A. Al Falou, and M. Atri. “Face Recognition Systems: A Survey.” *Sensors* 20, no. 2 (2020): 342. https://doi.org/10.3390/s20020342.

Web Articles

- "Drone Sensors: Types and Uses." *Mach34 Aerospace Blogs*. Accessed October 12, 2024. https://mach34aerospace.com/sensors-drone-types.
- “ATMega2560 Data Sheet” Atmel Microchip Technologies, Accessed November 21, 2024. ATmega640/1280/1281/2560/2561 datasheet
- Phillips, P. Jonathon, and Alice J. O'Toole. "Comparison of Human and Computer Performance Across Face Recognition Experiments." *Image and Vision Computing* 32, no. 1 (2014): 74–85. https://www.sciencedirect.com/science/article/abs/pii/S0262885613001741, Accessed February 17th, 2025

Literature Review

Sources	CRAAP Test
• Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning*. Cambridge, MA: MIT Press, 2016. https://www.deeplearningbook.org/.	✓

• Thrun, Sebastian, Wolfram Burgard, and Dieter Fox. *Probabilistic Robotics*. Cambridge, MA: MIT Press, 2005.	✓
• Finn, Rachel L., and David Wright. "Unmanned Aircraft Systems: Surveillance, Ethics and Privacy in Civil Applications." *Computer Law & Security Review* 28, no. 2 (2012): https://doi.org/10.1016/j.clsr.2012.01.005.	✓
• Liu, Xinyu, Chenhong Cao, and Shengyu Duan. "A Low-Power Hardware Architecture for Real-Time CNN Computing." *Sensors* 23, no. 4 (2023): 2045. https://doi.org/10.3390/s23042045.	✓
• Vallee, Nicolas, and Marc Coste. "Drone-Based Autonomous Surveillance Systems: Design and Implementation." *Robotics and Autonomous Systems* 126 (2019): 35–47. https://doi.org/10.1016/j.robot.2019.103451.	✓
• Wang, Lijie, et al. "Ultrasonic Obstacle Avoidance and Full-Speed-Range Hybrid Control for Intelligent Garages." *Sensors* 24, no. 6 (2024): 129–143. https://doi.org/10.3390/s24061029.	✓
• Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "FaceNet: A Unified Embedding for Face Recognition and Clustering." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 815–823, 2015. https://doi.org/10.1109/CVPR.2015.7298682.	✓
• "Drone Sensors: Types and Uses." *Mach34 Aerospace Blogs*. Accessed October 12, 2024. https://mach34aerospace.com/sensors-drone-types.	✓
• "ATMega2560 Data Sheet" Atmel Microchip Technologies, Accessed November 21, 2024. ATmega640/1280/1281/2560/2561 datasheet	✓

• Hancock, Peter J.B., Vicki Bruce, and A. Mike Burton. *"Recognition of Unfamiliar Faces." Trends in Cognitive Sciences* 4, no. 9 (2000): 330–37. https://doi.org/10.1016/S1364-6613(00)01519-9.	✓
• Murre, Jaap M. J. *"Learning and Categorization in Modular Neural Networks."* Hillsdale, NJ: Lawrence Erlbaum Associates, 2000.	✓
• Kortli, Yassine, Mohamed Jridi, A. Al Falou, and M. Atri. "Face Recognition Systems: A Survey." *Sensors* 20, no. 2 (2020): 342. https://doi.org/10.3390/s20020342.	✓

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning*. Cambridge, MA: MIT Press, 2016:

Currency:

- The book was published in 2016 which makes the text relatively current within the scope of the fast-evolving field of deep learning and neural networks. Although newer research has emerged and advancements have been made, the foundational theories and techniques that are discovered and discussed remain highly relevant.

Relevance:

- The source provides a comprehensive overview of deep learning, covering theoretical concepts and how they are implemented in practical situations in a broad scope, this makes it a valuable resource for both academic research concerning the technology which will be implemented into the autonomous security system itself and the practical application during the development of the artifact.

Authority:

- The publication is published by MIT Press which reinforces the source's academic credibility, it does not guarantee the reliability of the source itself. Although, the authors are highly respected and knowledgeable individuals in the field. Ian Goodfellow is renowned for his discovery on generative adversarial networks (GANs) while Yoshua Bengio is a pioneer in the field of deep learning.

Accuracy:

- This source is cited in a large number of academic journals and the publication itself is rigorously peer-reviewed within MIT Press as it offers detailed mathematical explanations and empirical evidence to support their findings and conclusions. The precision and depth it provides ensure that the information can be well understood along with being reliable and accurate.

Purpose:

- Designed as a textbook and as reference work, the source aims to educate readers on deep learning fundamentals as well as more in-depth and advanced topics such as types of neural networks and their structure. I evaluate that this source is intended for a scholarly audience as it provides clear and in-depth insight into the field

Thrun, Sebastian, Wolfram Burgard, and Dieter Fox. Probabilistic Robotics. Cambridge, MA: MIT Press, 2005.

Currency:

- While this source is published 2005 making the source relatively old it contains many core concepts and methodologies which are implemented in modern technology in probabilistic robotics which continue to be influential and form the basis for research in this field.

Relevance:

- This book is central to understanding the methods of how probabilistic programming is implemented in systems, this is particularly relevant during the creation of my artefact when I get to develop the PID of my system as the source covers the application of uncertainty handling, sensor fusion, and general autonomous systems.

Authority:

- Authored by experts and well-established individuals in robotics which include individuals such as Sebastian Thrun who is a prominent figure in autonomous vehicle research along with Wolfram Burgard and Dieter Fox which leads to this book carrying significant authority and reliability. Furthermore, this source is also published by MIT Press which further reinforces its academic reliability.

Accuracy:

- This source is known for its rigorous and well-documented approach as its work contains detailed methodologies along with valid experiments which support their conclusions and statements. Moreover, they extensively cite studies when it comes to secondary research and conclusions they come to; numerous studies and subsequent research in the field of robotics have advocated their scientific rigor.

Purpose:

- This source aims to provide a comprehensive breakdown of probabilistic robotics as it serves to be both a foundational text and a practical guide into the field. It targets an academic audience that seeks to understand the theoretical functions of complex algorithms and their application in managing uncertainty in robotic systems.

Finn, Rachel L., and David Wright. "Unmanned Aircraft Systems: Surveillance, Ethics and Privacy in Civil Applications." ***Computer Law & Security Review*** **28, no. 2 (2012):**

Currency:

- Even though this source was published in 2012 which makes it relatively old concerning UAV technology as autonomous flight vehicle technology has majorly improved since then, it still retains the fundemental principles of ethics and legality.

Relevance:

- The source directly addresses surveillance and the usage of drones and associated privacy/ethics concerns, with a focus on domestic security and autonomous based surveillance. The source provides a detailed legal analysis and case studies which cann be useful for framing ethical frameworks along with regulatory considerations

Authority:

- Rachel L. Finn is a lecturer in information law and David Wright is a senior consultant in privacy and data protection who are both recognized as experts in technological ethics and law. Furthermore, the publisher, *Computer Law & Security Review* is a peer-reviewed journal which is a reputable academic publisher in the field of law and technology

Accuracy:

- The source maintains and objective legal-analytic tone, though it advocates for strong privacy safeguards and considers the possibility of mitigating privacy breaches within an autonomous surveillance systems. The arguments presented are supported by references to statutory law, regulatory guidance and have analyzed and documented case studies. Due to the publisher being a peer-review journal, the source has underwent rigorous academic peer review, ensuring methodological soundness and factual reliability.

Purpose:

- Designed to inform policymakers and technological developers about the ethical and privacy implications of a civil drone deployment for the purpose of surveillance and security with an aim of a professional and scholarly audience in law, security or technology. The source does have a bias on privacy protection, although the journal article still presents the benefits and risks of UAV surveillance equally.
