www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
https://doi.org/10.5281/zenodo.14065813 Cosmos Impact Factor-5.86

SURVEILLANCE MONITORING DRONE WITH LIVE AND ALERTING SYSTEM

Dr T Sarada Devi[1], Sindhuja[2], R. Ankitha[3], P. Siri Priya[4]

[1] Assistant Professor, Dept. of EEE, Malla Reddy Engineering College for Women, Hyderabad,

[2 3 4]Research Student, Dept. of EEE, Malla Reddy Engineering College for Women, Hyderabad

---

Abstract -
This research paper investigates real-time target detection and video surveillance in UAV systems and focuses on computer vision models and deep learning algorithms. The first part examines computer vision algorithms such as Haar cascades, HOG, template matching, edge detection, and optical flow, while the second part focuses on deep learning algorithms, especially region-based detection and YOLO. Due to low power consumption, the use of deep learning in drones will cause problems. To solve this problem, this paper proposes a cloud computing system that can achieve the purpose of video detection and imaging. The results show that traditional computer vision algorithms are not fast enough for real-time tracking and deep learning algorithms are a suitable alternative system. These research papers contribute to the performance of unmanned aerial vehicles and instantaneous object searching through new ideas that can be used for various applications including vigilance, saving and protection, and agricultural inspection. The scheme can be extended to other applications that need to detect objects in resource-limited areas in real time.

Key Words: *Object-Detection, UAVs, Cloud Tracking, Drone, Region based detection, YOLO, SSD, Traditional computer vision algorithms, Deep learning.*

1. INTRODUCTION

Computer vision has improved significantly in recent years as a result of the advancement of deep learning algorithms [11], advances in hardware capabilities, and more data availability. Detecting items in a specific category such as people, cars, or animals within an image and reporting the location and extent of each object instance is one of the most commonly studied aspects of computer vision.Object detection, including object finding, scene assessment, crowd monitoring, segmentation, image captioning and activity recognition are key elements of a wide range of extremely complex computer vision tasks. Despite significant progress in developing broad object detection systems that can distinguish a wide range of items, there is still a need for precise and efficient object detection in the context of drone applications [14].

Drones are becoming more and more popular in a vast range of timely applications such as surveillance [26], delivery services [27], traffic tracking [28], agriculture [29], disaster management [30], and maritime security [31]. Amazon, for example, has been given federal authorisation to deploy drones as part of its delivery service and there are reports that drones may be an acceptable means of transporting medicinal products in rural areas. In the area of precision farming, drones are also expected to have a significant impact since they can assist farmers in tasks such as crop monitoring, analyses, and management, including selection of effective pesticides and optimisation of water supply. DJI, the world's leading drone maker, is developing drones that are equipped with sensors specific to protect agricultural crops from insects and weeds.

The history of drones dates back many years and it is possible to classify them on the basis of their flight speed, ability to stabilise position, hovering or loitering capability, environmental conditions as well as other characteristics. Various types of Unmanned Air Vehicles, each having its own.

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
https://doi.org/10.5281/zenodo.14065813 Cosmos Impact Factor-5.86

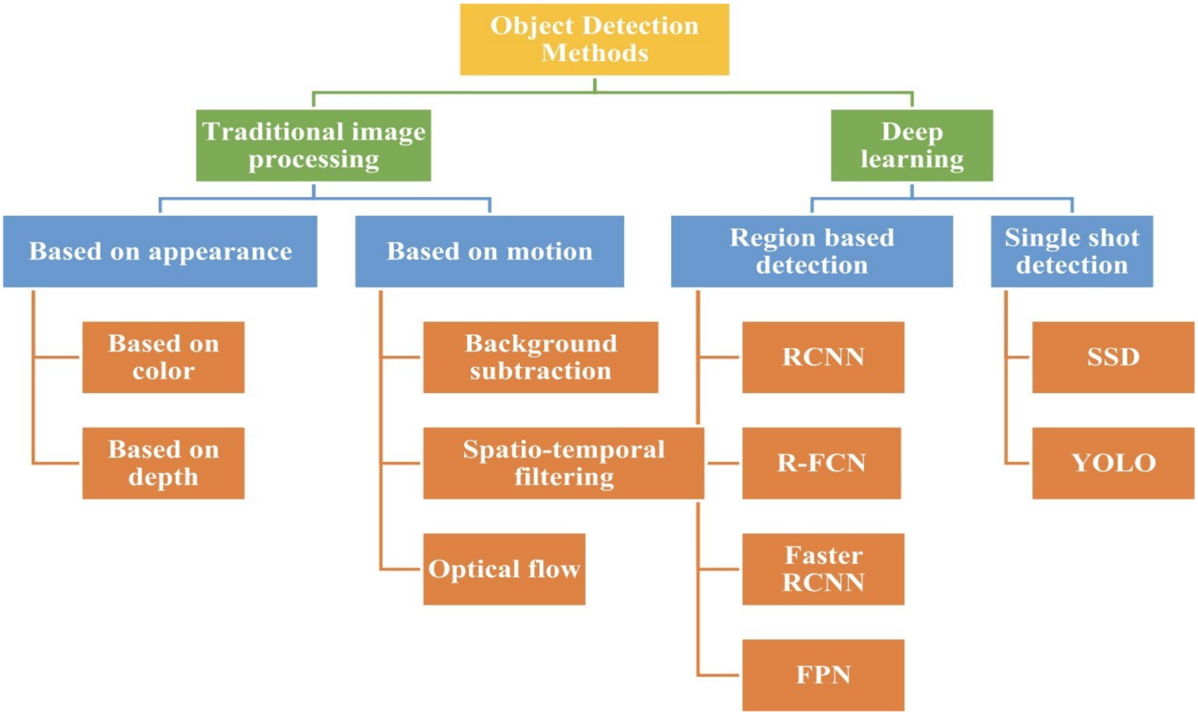


Fig1. Classification of object detection methods[11]

strengths and weaknesses, including but not limited to single UAVs [11], multiUAVs, fixed wing aircraft or hybrid UAVs. Autonomous vehicles are another area of research, with some drones able to execute flight plans without human interference, relying on Global Positioning System onboard and an ordered list of 3D points called waypoints[16].

Detecting objects from drones poses distinct challenges as opposed to conventional object detection. For instance, in the case of traffic monitoring, drones capture traffic activity from an aerial perspective, which provides more contextual information but makes object detection more challenging due to changes in viewpoint, scale, and aspect ratio. Bird's-eye-view object detection is further complicated by abrupt camera motion, motion blur, high object density, severe perspective distortion, and complex backgrounds. Additionally, aerial object detection studies often face the problem of biased datasets, as datasets need to be annotated. This means that object detection models trained on standard images may not be appropriate for detecting objects in aerial images.

Despite these challenges, researchers are making significant progress in developing accurate and efficient object detection algorithms for drone applications. One approach involves using deep learning models that can handle the complex, real-world conditions encountered in drone imagery. These models typically use convolutional neural networks (CNNs) to draw-out features from the images, followed by object proposal generation and classification of the proposed regions. Other approaches include adapting traditional object detection algorithms to the aerial domain, such as the Faster R-CNN[14] and YOLO algorithms[18].

In conclusion, object-detection in drone applications is a critical area of research with numerous real-world applications, from surveillance and delivery services to precision agriculture [29] and disaster management [30]. The use of specialized algorithms that can take into account the distinct characteristics of drones ' images is required to solve the problem of detecting objects in bird's eye view.

Further developments in the field of aerial object detection, which will lead to even more accurate and efficient detection systems in the future [11], will be expected as drone technology continues to advance.

2. LITERATURE REVIEW

Object detection is a computer vision task that involves many complex mathematical calculations and computation for identifying and localizing objects of interest within an image or video. This technique of detecting objects by a machine in a live video has got tremendous advancement in the field of robotics, autonomous vehicles, surveillance system, construction Industry and anymore.

Detecting objects in videos or images can be accomplished using a variety of methods and techniques, such as deep learning-based approaches like SSD (Single Shot Detector) [15], YOLO (You Only Look Once) [15] and Faster R-CNN (Region-based Convolutional Neural Network) Or conventional computer vision methods like Haar cascades and HOG (Histogram of Oriented Gradients) [1].

Deep learning-based object detection methods have become increasingly popular in recent years due to their ability to achieve high accuracy in real-time applications. But we will first discuss traditional computer vision algorithm for object detection. Traditional computer vision techniques for object detection and video monitoring involve a wide range of methods and algorithms, many of which have been developed over several decades of research. Here are some of the most commonly used traditional techniques.

2.1. Traditional Computer vision Techniques

2.1.1. Haar cascades : One way to classify objects in images is by color. This method is often used in robotic soccer, where teams of robots compete against each other [1]. However, relying solely on color can be problematic.

The results of the international RoboCup contest have revealed that the lighting environment plays a significant role in determining the competition's outcome. Even minor variations in the surrounding illumination can significantly impact a team's ability to succeed in the event. Participants must recalibrate their systems multiple times due to minor changes in ambient light that occur throughout the day [3]. To detect objects in images, using only color is not very reliable.

A more advanced technique for identifying objects in images involves analyzing specific attributes or structures of the object. Viola and Jones created Haar-like features, which help overcome the challenge of performing computationally intensive feature calculations.A cascade classifier involves multiple stages, each containing weak learners, and scans an image with a sliding window. The classifier categorizes a specific area in each stage as either positive or negative. To function effectively, the classifier requires a low false negative rate in each stage, while it can tolerate a relatively high false positive rate.

For the cascade to function properly, it's necessary that each stage has a low false negative rate. This is because if an object is incorrectly classified as a non-object, then the classification for that branch stops, and unfortunately there's no other opportunity to correct the mistake later on. It is considered acceptable for individual stages in the object detection process to have a relatively high rate of falsely identifying non-objects as objects. If this occurs at a particular stage, the mistake can still be rectified in subsequent stages of the classifier, starting from the (n+1)th stage onwards [4].

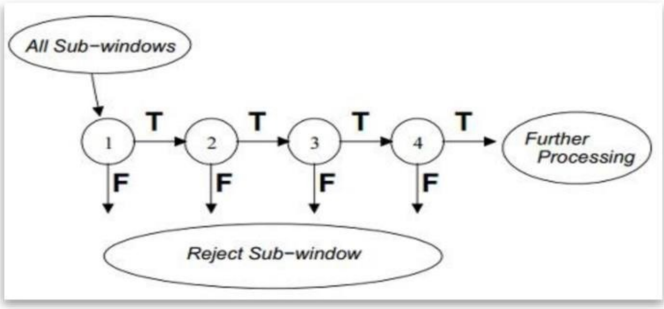


Fig2. Different Stages of the cascade classifier

2.1.2. HOG (Histogram of Oriented Gradients): To detect objects, Histogram of Oriented Gradients (HOG) is a technique used in computer vision that extracts features. The HOG approach involves calculating the gradient of pixel intensities in an image, which yields a set of gradient vectors. These vectors are then represented as histograms, and these histograms are utilized as features for object detection.[5].

The HOG feature extraction process involves several steps:

- Pre-processing: The input image is always pre-processed to enhance contrasting remove noise. This involves smoothing the image and applying a high-pass filter to extract edges.
- Gradient calculation: The gradient of the pixel intensities is calculated using a derivative filter, such as the Sobel operator. This produces two gradient components, one in the x-direction and one in the y-direction.
- Orientation binning: The gradient are binned into a set of orientation bins. The orientation bins divide the gradient angle varies into a set of discrete bins, such as 0-20 degrees, 20-40 degrees, and so on. The magnitude of each gradient vector is accumulated into the corresponding orientation bin.
- Block normalization: The histogram of gradient orientations is normalized over a local region of the image called a block. The block is typically rectangular and overlaps with neighbouring blocks. The normalization is performed to improve the robustness

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
https://doi.org/10.5281/zenodo.14065813 Cosmos Impact Factor-5.86

of the HOG features to variations in illumination and contrast.

- Descriptor generation: The HOG features are generated by concatenating the normalized histograms from all the blocks in the image.

HOG has been used flourishingly in several computer vision applications, including face-detection, pedestrian-detection and object tracking. One of the advantages of HOG is that it is computationally efficient and can be used for real-time applications. However, HOG may not be as accurate as deep learning-based methods for object detection.

2.1.3. Template matching: It is a computer vision method that is often utilized to locate a sub-image in the targeted image that get matched a given template of image. It is a popular technique that finds applications in diverse fields such as robotics, medical imaging, manufacturing, and surveillance. Based on the method used for feature extraction, template matching approaches can further be categorized into two different groups: level histogram method and feature extraction method [6].

However, Performing template matching can be computationally intensive since it involves taking the template image and placing it in every possible position within a larger target image. This process requires the calculation of a numerical metric for each position.to determine the level of similarity. To address this, swarm intelligence algorithms have been considered as a solution in recent works. Swarm intelligence is a problem-solving approach inspired by the behaviour of social animals such as ants, birds, and fish. These animals display collective behaviour without a central control unit or any individual member knowing the overall goal of the group. Instead, they follow simple local rules that lead to emergent behaviour at the group level [7].

Swarm intelligence algorithms aim to replicate this behaviour in computational systems. They typically involve a population of agents (e.g., "ants" or "particles") that interact with each other and with their environment to collectively solve a problem. Each agent follows a set of simple rules that govern its behaviour, such as moving towards or away from certain stimuli or other agents, and updating its behaviour based on feedback from its environment.

The algorithms have been applied to a wide range of problems in optimisation, routing, classification, and other areas, and have been shown to be effective in many cases where traditional optimisation methods fail due to the complexity of the problem or the high dimensionality of the search space.

2.1.4. Edge Detection: Edge detection is a technique used to locate and identify sudden changes in the intensity of pixels within an image. These abrupt changes in intensity are known as discontinuities, and they often indicate the boundaries of objects within a scene.[8]. Classical edge detection methods use a 2-D filter, which is designed to highlight areas of the image with a large change in intensity. There are many different types of filters, each designed to work well for a certain type of edge. Noise in the image can make edge detection difficult, and attempt to reduce noise can result in a blurred or distorted edge. Some edges are not a sudden change in intensity, but instead a gradual change, which requires a different type of filter. There are two main types of edge detection: gradient-based and Laplacian-based, which use different mathematical techniques to find edges. The goal is to compare different edge detection methods to find the one that works best for different situations.

Edge detection can be performed using a variety of techniques, but these techniques can generally be of two categories.

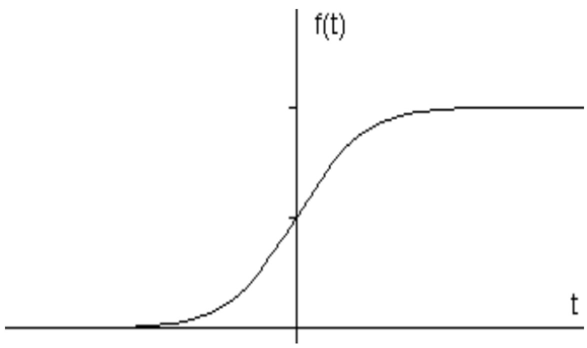
Gradient based Edge Detection:

The gradient-based method for detecting edges in an image involves identifying the highest and lowest points in the image's first derivative. This technique is used to locate the edges in an image.

Laplacian based Edge Detection:

The Laplacian approach detects edges in images by identifying the edges in an image can be identified by detecting zero crossings in the second derivative of the image. These edges take on a ramp-like shape and can be detected by calculating the derivative of the image. If there is a sudden change in intensity within the image, the derivative can be used to pinpoint the edge's location [8].

When we compute the gradient of this signal, which involves taking the initial derivative w.r.t time in one dimension, the resulting signal is as follows.[8]:



www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
https://doi.org/10.5281/zenodo.14065813 Cosmos Impact Factor-5.86

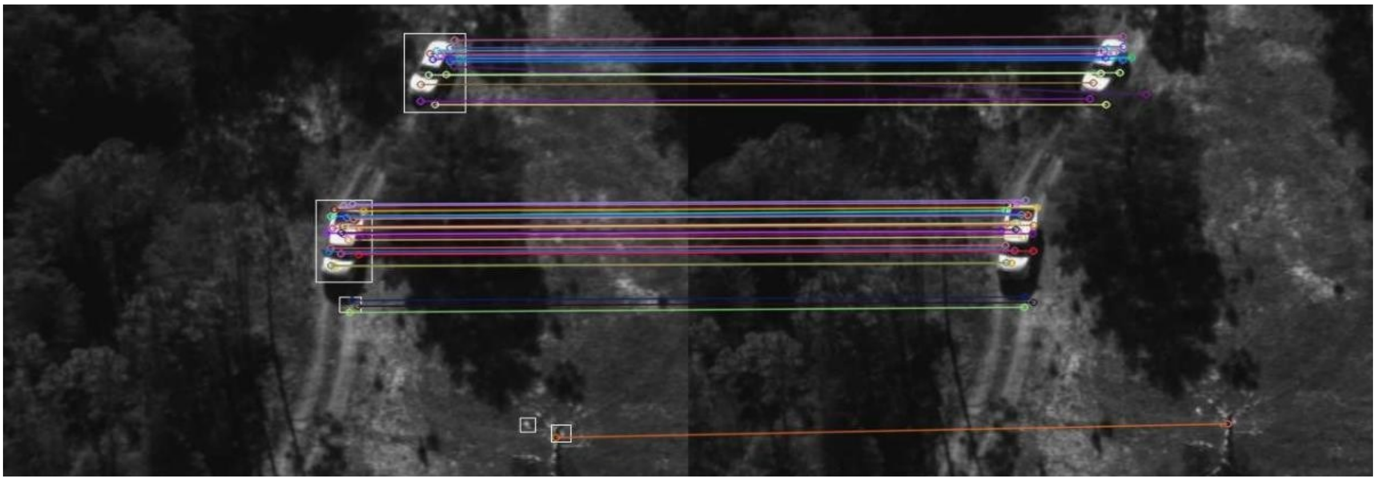


Fig3. *Test result for feature detection on EgTest05[10]*

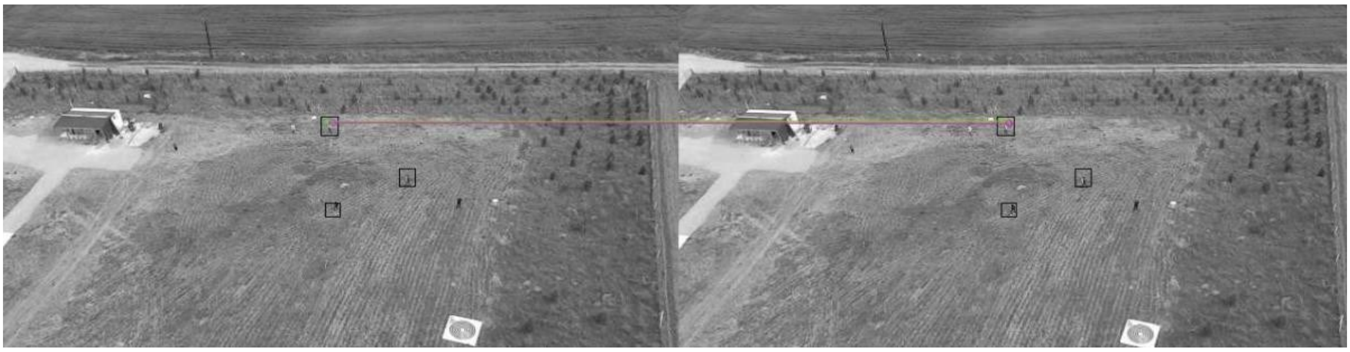
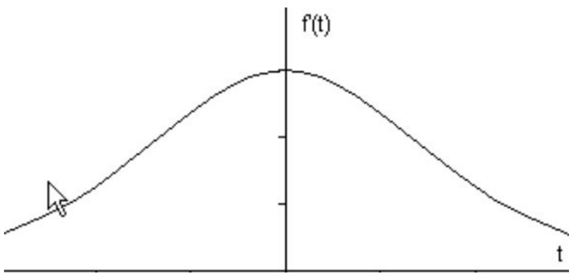
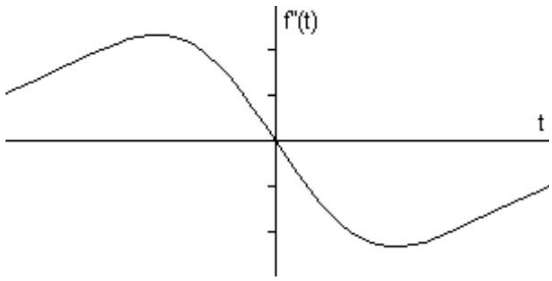


Fig4. *Test result for feature detection in house captured video[10]*





The derivative displays a peak at the center of the corner in the original signal, which is a feature of "gradient filter" edge detection filters such as Sobel method. If the gradient increases a specific threshold-value, the pixel's location is identified as an edge's location. Edges are characterized by higher pixel intensity values than their neighbouring pixels. Therefore, comparing the gradient value to the threshold value allows for the detection of an edge whenever the gradient surpasses the threshold. Additionally, when the first derivative reaches its peak, the second derivative becomes zero. Hence, the Laplacian method can be used to detect edges by identifying zeros in the second derivative of the signal. The second derivative of the signal is illustrated below[8].

2.1.5. Optical Flow: Optical flow is a method in computer vision that is utilized to monitor the movement of objects in an image or video. This technique involves examining the alterations in pixel intensities between successive frames of a video to calculate the apparent motion of objects present in the scene.

Optical flow can be used to solve various computer vision problems, such as object tracking, activity recognition, and video stabilization. The resulting optical flow field is a dense map of vectors, where each vector represents the motion of a pixel in the scene between consecutive frames. The direction and magnitude of the vector indicate the direction and speed of the motion.

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
https://doi.org/10.5281/zenodo.14065813 Cosmos Impact Factor-5.86

There are various techniques used to estimate optical flow, such as the Lucas-Kanade technique, HornSchunck method, and the Farneback method. These techniques differ in their assumptions about the motion field and the cost functions used to estimate the flow vectors[9].

Optical flow has numerous applications in different fields, such as robotics, autonomous navigation, and sports analysis. For example, optical flow can be used in self-driving cars to estimate the motion of surrounding vehicles and pedestrians, which is crucial for safe navigation. Additionally, optical flow can be used to monitor and avoid all obstacles in real time, enabling the drone to fly safely and autonomously in complex environments. Overall, optical flow is a powerful tool for drone navigation and has the potential to revolutionize the way drones are used in a vast variety of applications, from search and rescue operations to agriculture and delivery services [10].

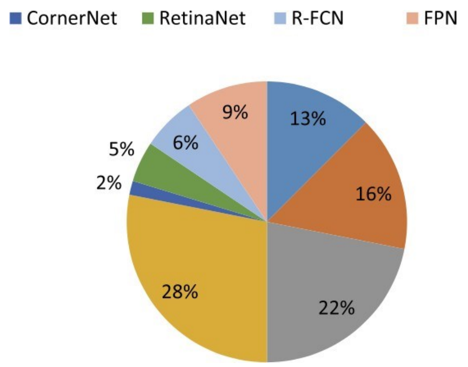


Fig5. *The distribution of deep learning papers in the UAV field, categorized by the type of deep learning technology used.[14].*

2.2. Deep learning-based object detection

Deep learning-based object detection and video monitoring is a CV(computer vision) method that aims to detect and locate objects within an image or video frame using DNN(deep-neural-networks). The approach is able to perform object detection tasks with a great accuracy, and has become increasingly popular in recent years due to the growing availability of large datasets and the computational power required for training DNN.

The main idea behind DNN based object detection is to train a neural network to identify patterns and features that are indicative of objects within an image or video frame. There are several popular techniques for performing object detection with deep learning, including region-based detection including convolutional neural networks (R-CNNs), and single-shot detection methods (SSDs) like You Only Look Once (YOLO).

2.2.1. Region based object detection: It is a technique in deep learning based object detection that uses a two-stage approach to identify targeted objects in images. In the first stage, a set of region proposals are generated that has a very high possibility to contain objects. These region proposals are generated using a region proposal network (RPN) that scans the image at multiple scales and identifies potential object locations [11]. The RPN outputs a set of bounding boxes with corresponding objectness scores indicating the likelihood of each box containing an object.

In the second stage, the region proposals are refined and classified by a Region-Based-Convolutional-Neural-Network (RCNN). The RCNN takes each proposed region as input and outputs a label and a more accurate bounding rectangles for the object within these region.

The Region-based Convolutional Neural Network (RCNN) usually consists of a layer known as the Region of Interest (ROI) pooling layer. Its purpose is to extract a feature vector of a predetermined length from each region proposal. followed by one or more fully connected layers that perform classification and bounding box regression [14].

Methods for detecting objects based on regions, like faster R-CNN and cascade R-CNN, mask R-CNN have attained excellent performance on object detection benchmarks like COCO and PASCAL VOC. These techniques are extensively used in diverse applications, including object tracking, autonomous driving, and surveillance.

The Faster RCNN is introduced as a deep learning based object detection method. It consists of a pre-trained CNN for feature extraction, followed by two more sub-networks that are trained. The Region- Proposal-Network (RPN) generates object proposals, while the second subnetwork predicts the object's class. The main difference between faster R-CNN and other region-based detectors is that RPN is added at the last convolutional layer, allowing for real-time frame rates without the need for selective search[13]. Additionally, Faster R-CNN outperforms other region-based detectors in terms of mAP and allows for single-stage training of both classification and regression. Feature Pyramid Network (FPN) is another method for generating multi-scale feature representations at high resolution levels, which can improve object detection in multiple scales. The Deep Drone framework proposed by Han, Shen, & Liu (2016) [12] uses a CNN for object detection and achieves fast and accurate results.

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
https://doi.org/10.5281/zenodo.14065813 Cosmos Impact Factor-5.86

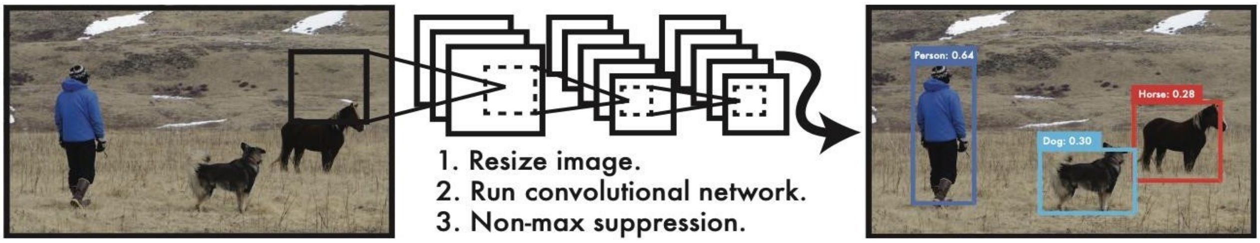


Fig6. *Working with the YOLO Detection System & resizing the input image to 448 × 448, processing of image with a single convolutional network, and applying a confidence threshold to the resulting detections [18].*

2.2.2. Single Shot Object detection: Although region-based detection methods deliver impressive accuracy, their computational speed is not optimal. In contrast, single-shot detection methods offer faster processing times and require less memory than region-based methods. These methods use the concept of "multibox" to identify multiple objects in a single shot. They achieve higher efficiency and accuracy by eliminating the need for bounding box proposals, which is a requirement for RCNN. Instead, they use a convolution filter to gradually predict the classes of the objects and its location offsets[15].

The researchers suggested a deep learning-based model for object detection in drone images of a particular class. The images were captured using Parrot AR Drone 2, and the data was processed on a connected PC via WiFi. To overcome the computational burden of region-based algorithms, they used SSD (Single Shot Detector) for object detection. The resulting output was fed to a PID (Proportional Integral Derivative) controller, which tracked the objects in a 3D plane comprising x, y, and z axes. This approach outperformed other methods in terms of computational time also accuracy, making it ideal for realtime applications [15].

2.2.3 You Only Look Once : Ross Girshick, Joseph Redmon, Santosh Divvala, and Ali Farhadi developed YOLO in 2016 as an object detection system. YOLO is a deep learning algorithm that can be used to detect objects in real time videos and images quickly and accurately. Instead of taking a sliding window approach, YOLO computes object detection as a regression problem and estimates bounding rectangles and class probabilities directly from full image in a single neural network.

At present, detection systems utilize classifiers in order to conduct object detection. The process involves using a classifier in question and testing it at different locations and the scales within an images.The Deformable parts models (DPM) and similar systems uses a sliding window technique where the classifier is applied at regular intervals throughout the image[16].

Newer methods such as R-CNN start by creating possible bounding boxes in an image using region proposal techniques. Then, a classifier is applied to these boxes and post-processing is done to enhance the accuracy of the boxes by removing redundant detections and scoring them based on other objects in the scene [17]. These methods involve many components that need to be trained separately, making them difficult and time consuming to optimise.

The creators of YOLO have redefined the problem of object detection as a single regression-problem that estimate the class probabilities and bounding rectangles coordinates of multiple objects in an image directly from the pixels. YOLO only needs one pass over an image to predict the location of all objects in the image. This makes YOLO simpler, faster, and more accurate than traditional methods of object detection. YOLO can run on a new image at test time, without the need for a complex pipeline or individual component training, which makes it very fast. The base network can run at 45 fps and the fast version can run at more than 150 fps. According to research, YOLO has the capability to process video streams in real-time with a latency of under 25 milliseconds. It has been shown to outperform other real-time systems with more than twice the mean average precision, as documented in [18].

2.3. Real-time tracking

Currently, there are many research initiatives aimed at creating dependable cloud-based robotic applications for the future. These initiatives can be classified into two main categories: Cloud Robotics Systems and Drone-based Systems. The author's introduced some of the algorithms and demonstrated the implementation of a cloud-based system called Robot Cloud. The purpose was to take advantage of the flexibility, re-usability, and extensibility offered by cloud based robot systems. They built a prototype of Robot Cloud using Service Oriented Architecture (SOA) and deployed it on Google App

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
https://doi.org/10.5281/zenodo.14065813 Cosmos Impact Factor-5.86

Engine(GAE) [19]. The researchers created an openEASE system that enables robots and researchers to remotely solve complex mental parallel problems using the cloud's vast storage and computational resources. They incorporated learning algorithms into the system and provided the robot with suggested solutions for dealing with situations [20]. The text [21] explains that a Cloud Robotics Middleware has been introduced that permits the transfer of storage and computations from robots to the cloud. This is considered to be an initial implementation of cloud robotics systems, similar to the works mentioned previously.

The problem of insufficient computing power on robots for modelling Simultaneous Localisation and Mapping (SLAM) tasks [22], which involves creating a map of the robot's surroundings. To address this issue, they suggest a software framework called Cloudroid, which is designed to deploy robotic packages to the cloud for cloud as services. They also conduct tests to evaluate the framework's performance in dynamic and resource-limited environments, with a focus on request response time. New system was introduce then which is called, Context Aware Cloud Robotics (CACR) [23] which includes decision-making capabilities for industrial robots like automated guided vehicles. The system's design incorporates cloud based application for simultaneous localisation and mapping, and the researchers highlight energy proficiency and cost savings as the major advantage of using the cloud-based approach.

The primary target of this study is to integrate robots with the cloud and offer task-oriented services while ensuring a high quality of service. RoboCloud presents a cloud service with a specified mission and controllable resources that are determined based on predictable behaviour. The effectiveness of the proposed approach is assessed by analyzing the quality of service variables, such as latency, of a cloud service that offers cloud focused object monitoring.

2.3.1 Drone Based System: Several initiatives have aimed to combine drones with cloud computing and the Internet of Things (IoT). One such initiative is the Internet-of-Drones model proposed by Gharibi et al. [24]. Their model comprises three primary networks - air traffic control network, cellular network, and the internet - which provide generic service for various UAV applications such as delivery, surveillance, search and rescue, among others. However, the article only presents a theoretical framework for the IoD without any concrete implementation or realization of the proposed architecture. In contrast, our research presents a validated architecture for the IoD along with a real-world implementation and experimentation.

The research paper [25] introduces a cloud robotics platform, FLY4SmartCity, that uses ROS as the base. The platform architecture consists of essential features that allow the creation of drone instances as nodes, which are managed by the platform manager for management event and planning. During events, the service manager provides services, while the rule manager handle the actions. Ermacora et al. [22] presented another paper on a cloud robotics platform designed for monitoring and based on ROS. The platform utilizes cloud computing to offload data and computational capabilities. The architecture is layered and includes services that use APIs provided by applications utilizing drone capabilities and adaptation. Drones generally serve as the physical layer of the architecture.

Dronemap Planner is a cloud-based system that utilizes the Internet-of-Drones (IoD) concept to enable users to control and manage multiple autonomous drones. A mission, such as visiting a set of waypoints, is initiated by the user through the cloud. Virtual UAVs are then created and mapped to physical UAVs using a service-oriented approach based on REST or SOAP Web services. Once a mission request is received, the selected UAVs carry out the mission and send real-time data to the cloud service, which then stores, processes, and forwards the synthesized results back to the user. An overview of the Dronemap Planner's architecture is provided below.[25].

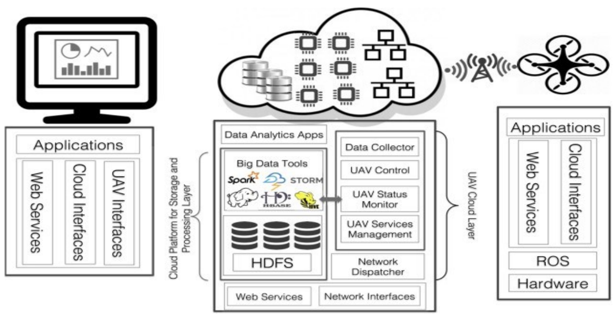


Fig7. *Drone-map system architecture[21]*

• **The UAV layer** : This layer provides users with services by making system resources available to them. The UAV layer facilitates communication with hardware through the use of the Micro Air Vehicle Link (MAVLink) and the Robot Operating System (ROS) communication protocol. ROS is widely used middleware that is useful in developing robotics applications, and the Micro-Air-Vehicle-Link (MAVLink) communication protocol helps in message exchange between drones and ground stations using different transport protocols such as TCP, UDP, and USB ,Telemetry. When ROS and MAVLink are used

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
https://doi.org/10.5281/zenodo.14065813 Cosmos Impact Factor-5.86

together, developers can have an interface that allows them to control and monitor drones at a high-level without having to program or interact with the hardware directly.

• **Cloud services layer** : The cloud services layer is in charge of deploying cloud services utilizing three component sets: remote computation, communication interactions, and cloud-based storage. In the cloud, data from UAVs is stored, including data on environmental variables, mission data, localization parameters, and time-stamped transmitts data streams, such as images and sensor data. The data is stored in a distributed file system, like HBase or HDFS (Hadoop Distributed File System),depending on the specific application requirements. The usage of distributed file system storage allows for extensive batch processing through tools such as Hadoop Map/Reduce. This system offers both realtime and batch processing of the data. In the case of real-time data streams, the cloud operates on incoming data to detect crucial events or threats that demand prompt action or to execute dynamic computation in a distributed environment. On the other hand, for batch processing the incoming data which is retained in HDFS and is analyzed at a later stage.

In addition, the system offers cloud-based remote computation, encompassing resource-intensive algorithms for data analysis and image processing. Furthermore, it supports Map/Reduce jobs that run on Hadoop, enhancing processing speed and boosting system efficacy. Additionally, data analytics algorithms is executed on voluminous sets of stored data.

The third component of Cloud Services layer comprises communication interfaces. The system enables interactions via network interfaces and web services. The network interfaces rely on server-side network sockets that receive JSON serialized messages sent from UAVs. Meanwhile, web services enable clients to manage drone missions and parameters. Both SOAP and REST web services are utilized to provide end-users and client applications with varied means of controlling and monitoring drones. Network interfaces are mainly utilized to manage continuous streams, while web services are employed to issue commands to drones and retrieve cloud-based data.

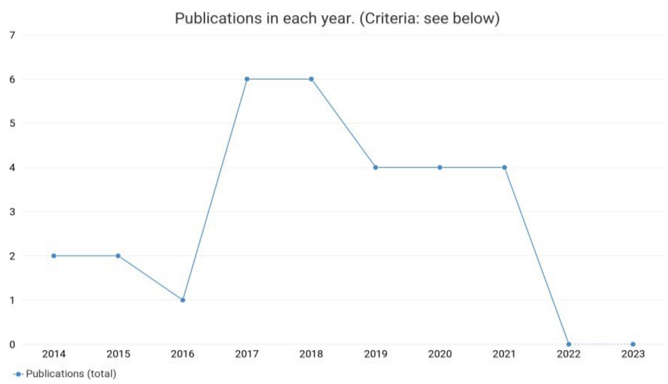


Fig9. *Year wise selected papers*

• **Client-Layer**: The layer described here provides interfaces for drone application developers and end-users alike. End-users can access the cloud services layer and UAV layer via drone map client side web applications running on the client layer. This allows users to register multiple UAVs, modify mission parameters based on cloud results, remotely monitor and control the UAVs and their missions. The front-end interface includes features to connect/disconnect, utilise physical UAVs and their services, configure, control a mission and keep track of UAV parameters. For developers, the client layer provides many Api in various programming languages to facilitate the development of drone applications.

3. BIBLIOMETRIC ANALYSIS

To identify the research papers and in the field of real-time object detection and video monitoring in drone system , we used the Dimensions.ai database to retrieve all publications related to real-time detection in Drone system using the following search query:

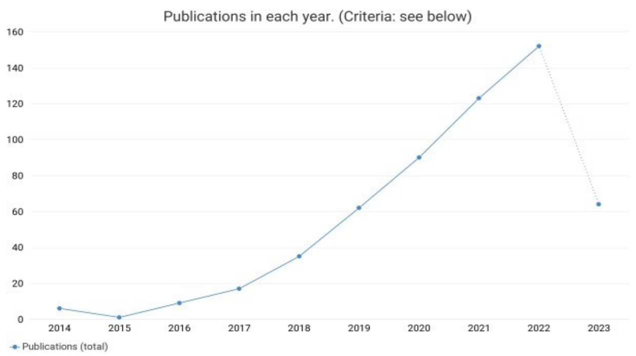


Fig8. Year wise distribution of query result

Exclusion Criteria:
Research paper older than 2014 are excluded and 31 papers are included.

Inclusion Criteria:

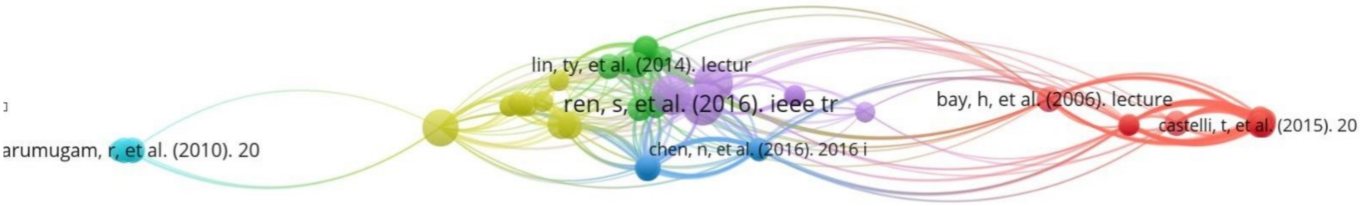
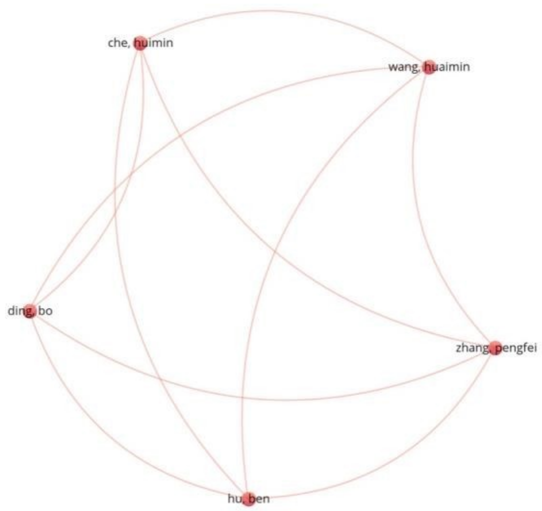


Fig10. Co-Citation Network

Co-Citation analysis using cited sources as the unit of analysis. The condition is for a source to have at least been cited 2 times. This filtered out 69 sources from 566, as displayed in Figure 10.



Co-Authorship analysis using Authors as the unit of analysis. The condition is for the author to have at least 2 documents. Out of 101 authors, 22 met this threshold. Out of 22, 5 formed the largest connected set (cluster).

Fig 11. Co-Authorship Network

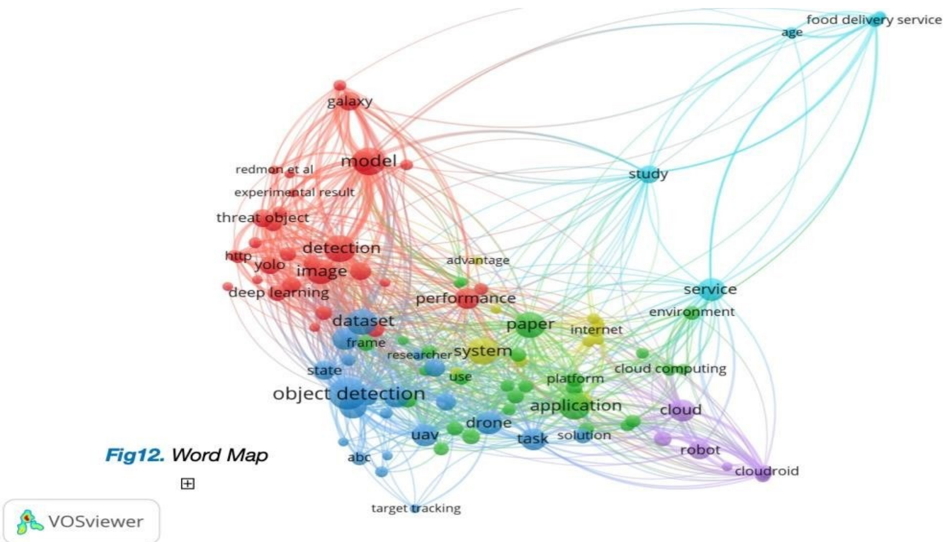


Fig12. Word Map

Occurrences of words was counted from Title and Abstract Fields, using the Full Counting method. All common words with a minimum occurrence of 3 were chosen, giving a total of 105 words.

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
https://doi.org/10.5281/zenodo.14065813 Cosmos Impact Factor-5.86

4. CONCLUSION

In conclusion, real-time object detection and video monitoring are critical functionalities in modern drone systems that enable them to perform a wide range of tasks, from surveillance and monitoring to search and rescue operations. These capabilities are made possible by using advanced computer vision algorithms and deep learning models that can process data in real-time and detect objects with high accuracy.

Traditional computer vision algorithms are good approach for detecting objects in a static image but it becomes a bit slower for 40fps to 60fps videos and it becomes quite difficult to detect objects in a live stream, whereas deep learning algorithms like SSD, YOLO, R-CNN and Faster R-CNN are quite efficient for detecting objects in a high frame rate videos but there is a big problem just because of size of drones, since these technologies are quite complex and requires a lot of computation it is difficult to provide such a high end system on a drone hence the cloud connecting drone is introduced and all these complex computations are computed on dedicated server machine and then the result is sent back.

Moreover, cloud connection of drones is becoming increasingly important as it enables data storage, computation, and communication capabilities that are crucial for many drone applications. Storing data in the cloud enables drones to engage in extensive batch processing using tools like Hadoop Map/Reduce. Meanwhile, real-time processing is capable of detecting pressing events and threats that demand prompt action.

The cloud also provides remote computation and communication interfaces for both end-users and developers, enabling them to control and monitor the drones remotely and develop applications that leverage the drones' capabilities. Overall, the integration of real-time object detection, video monitoring, and cloud connectivity is a powerful combination that can enable drones to perform complex tasks with high efficiency and accuracy.

5. REFRENCES

[1] Soo, Sander. "Object detection using Haar-cascade Classifier." *Institute of Computer Science, University of Tartu* 2.3 (2014): 1-12.

[2] Nagabhushana, S. "Introduction in Computer Vision and Image Processing." *New Age International (P) Ltd. Publishers, New Delhi* (2005): 3.

[3] Lovell, Nathan, and Vladimir Estivill-Castro. "Color classification and object recognition for robot soccer under variable illumination." *Robotic Soccer*. IntechOpen, 2007.

[4] T. M. Inc., “Train a Cascade Object Detector,” [Online]. Available: http://www.mathworks.se/help/ vision/ug/train a cascade object detector.html#btugex8. 2014

[5] Felzenszwalb, Pedro F., et al. "Object detection with discriminatively trained part-based models." *IEEE transactions on pattern analysis and machine intelligence* 32.9 (2009): 1627-1645.

[6] Banharnsakun, Anan, and Supannee Tanathong. "Object detection based on template matching through use of best-so-far ABC." *Computational intelligence and neuroscience* 2014 (2014): 7-7.

[7] Brunelli, Roberto. Template matching techniques in computer vision: theory and practice. John Wiley & Sons, 2009.

[8] Maini, Raman, and Himanshu Aggarwal. "Study and comparison of various image edge detection techniques." *International journal of image processing (IJIP)* 3.1 (2009): 1-11.

[9] Lezki, Hazal, et al. "Joint exploitation of features and optical flow for real-time moving object detection on drones." *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. 2018.

[10] Berker Logoglu, K., Lezki, H., Kerim Yucel, M., Ozturk, A., Kucukkomurler, A., Karagoz, B., Erdem, E., Erdem, A.: Feature-based efficient moving object detection for low-altitude aerial platforms. In: The IEEE International Conference on Computer Vision (ICCV) Workshops. (Oct 2017)

[11] Anitha Ramachandran[a], Arun Kumar Sangaiah, “A review on object detection in unmanned aerial vehicle surveillance” International Journal of Cognitive Computing in Engineering 2 (2021) 215–228

[12] Han, Song, William Shen, and Zuozhen Liu. "Deep drone: Object detection and tracking for smart drones on embedded system." *URL https://web. stanford. edu/class/cs231a/prev_projects_2016/deepdrone-object__2_. pdf* (2016).

[13] Wang, Xiaoliang, et al. "Fast and accurate, convolutional neural network based approach for object detection from UAV." *IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2018.

[14] Subash, K. V. V., Srinu, M. V., Siddhartha, M., Harsha, N. S., & Akkala, P. (2020). Object detection using Ryze Tello drone with help of mask-RCNN. In Proceedings of the 2020 2nd *international conference on innovative mechanisms for industry applications (ICIMIA)*

[15] Rohan, Ali, Mohammed Rabah, and Sung-Ho Kim. "Convolutional neural network-based real-time object detection and tracking for parrot AR drone 2." *IEEE access* 7 (2019): 69575-69584.

[16] Felzenszwalb, Pedro F., et al. "Object detection with discriminatively trained part-based models." *IEEE transactions on pattern analysis and machine intelligence* 32.9 (2009): 1627-1645.

[17] Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.

[18] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[19] Du, Zhihui, et al. "Robot cloud: Bridging the power of robotics and cloud computing." *Future Generation Computer Systems* 74 (2017): 337-348.

[20] A. K. Bozcuogˇlu and M. Beetz, ''A cloud service for robotic mental simulations, ''in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 2653–2658.

[21] C.Huang, L.Zhang, T.Liu, and H.Y.Zhang,''A control middle ware for cloud robotics, ''in *Proc. IEEE Int. Conf. Inf. Autom. (ICIA)*, Aug. 2016, pp. 1907–1912.

[22] B. Hu, H. Wang, P. Zhang, B. Ding, and H. Che, ''Cloudroid: A cloud framework for transparent and QoS-aware robotic computation outsourcing, ''in *Proc. IEEE 10th Int. Conf. Cloud Comput. (CLOUD)*, Jun. 2017, pp. 114–121.

[23] Wan, Jiafu, et al. "Context-aware cloud robotics for material handling in cognitive industrial Internet of Things." *IEEE Internet of Things Journal* 5.4 (2017): 2272-2281.

[24] Gharibi, Mirmojtaba, Raouf Boutaba, and Steven L. Waslander. "Internet of drones." *IEEE Access* 4 (2016): 1148-1162.

[25] Bona, Basilio. "'Advances in human robot interaction for cloud robotics applications." *Polytech. Univ. Turin, Turin, Italy, Tech. Rep* (2016).

[26] Sien, Jonathan Phang Then, King Hann Lim, and Pek-Ing Au. "Deep learning in gait recognition for drone surveillance system." *IOP Conference Series: Materials Science and Engineering*. Vol. 495. No. 1. IOP Publishing, 2019.

[27] Hwang, Jinsoo, and Hyunjoon Kim. "Consequences of a green image of drone food delivery services: The moderating role of gender and age." *Business Strategy and the Environment* 28.5 (2019): 872-884.

[28] Kyrkou, Christos, et al. "DroNet: Efficient convolutional neural network detector for real-time UAV applications." *2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2018.

[29] Nuijten, Rik JG, Lammert Kooistra, and Gerlinde B. De Deyn. "Using unmanned aerial systems (UAS) and object-based image analysis (OBIA) for measuring plant-soil feedback effects on crop productivity." *Drones* 3.3 (2019): 54.

[30] Kyrkou, Christos, and Theocharis Theocharides. "EmergencyNet: Efficient aerial image classification for drone-based emergency monitoring using atrous convolutional feature fusion." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13 (2020): 1687-1699.

[31] Kim, Hangeun, et al. "Development of a UAV-type jellyfish monitoring system using deep learning." *2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*. IEEE, 2015.