International Journal of Thermofluids 21 (2024) 100590

Contents lists available at ScienceDirect

International Journal of Thermofluids

journal homepage: www.sciencedirect.com/journal/international-journal-of-thermofluids

Progress in artificial intelligence-based visual servoing of autonomous unmanned aerial vehicles (UAVs)

Muaz Al Radi [a], Maryam Nooman AlMallahi [b], Ameena Saad Al-Sumaiti [a,f], Concetta Semeraro [c,d], Mohammad Ali Abdelkareem [c,e,*], Abdul Ghani Olabi [c]

[a] Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi, United Arab Emirates
[b] Department of Mechanical and Aerospace Engineering, United Arab Emirates University, Al Ain City P.O. Box 15551, United Arab Emirates
[c] Sustainable Energy & Power Systems Research Centre, RISE, University of Sharjah, P.O. Box 27272, Sharjah, United Arab Emirates
[d] Department of Industrial and Management Engineering, University of Sharjah, PO Box 27272, Sharjah, United Arab Emirates
[e] Chemical Engineering Department, Minia University, Elminia, Egypt
[f] Advanced Power and Energy Center, Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi, United Arab Emirates

ARTICLE INFO

Keywords:
Unmanned aerial vehicles
Visual servoing
Artificial intelligence
Artificial neural networks
Fuzzy logic
Reinforcement learning

ABSTRACT

Unmanned aerial vehicles (UAVs) have attracted massive attention in many engineering and practical applications in the last years for their characteristics and operation flexibility. For the UAV system, suitable control systems are required to operate appropriately and efficiently. An emerging control technique is visual servoing utilizing the onboard camera systems for inspecting the UAV's environment and autonomously controlling the UAV's operation. Artificial intelligence (AI) techniques are widely deployed in the visual servoing of autonomous UAV applications. Despite the increasing research in the field of AI-based visual control of UAV systems, comprehensive review articles that showcase the general trends and future directions in this field of research are limited. This work comprehensively examines the application and advancements of AI-enhanced visual servoing in autonomous UAV systems, covering critical control tasks and offering insights into future research directions for enhancing performance and applicability which is limited in the current literature. The paper first reviews the application of intelligent visual servoing systems for autonomously executing various UAV control tasks, including 3D UAV positioning, aerial and ground object following, obstacle avoidance, and autonomous landing. Second, the research progresses in applying AI techniques in the visual servoing of autonomous UAV systems are discussed and analyzed. Finally, future directions and critical research gaps for further improving the performance and applicability of intelligent visual servoing systems are included.

1. Introduction

Turing and Haugeland denoted the birth of Artificial Intelligence (AI) in the book "Computing Machinery and Intelligence", published in 1950 [1]. AI leverages computers and machines to mimic the human mind to problem-solving and decision-making capabilities [2]. Nearby environments are perceived, and measures are taken to assure the maximum success rate of the desired goal [3]. AI has the potential to provide some of the most powerful and disruptive technologies of the twenty-first century [4]. Self-driving vehicles [5], robotic assistants [6], automated Unmanned Aerial Vehicles (UAVs) [7], automated disease diagnosis [8–10], dentronics [11], renewable energy [12,13], water and energy, and smart cities [14,15] are the main AI applications fields including machine learning, deep learning, robotics, computer vision and expert systems techniques [16]. Specifically, computer vision has been attracting attention in recent years due to its wide variety of applications and recent achievements in different fields [17,18]. AI offers multiple advantages and has been effectively utilized in various industries, including image and speech recognition, autonomous vehicles, and computer vision. In particular, nowadays, UAVs frequently employ artificial intelligence techniques [19]. UAVs are flying robots employed for different purposes, including surveillance [20], traffic monitoring [21], agriculture [22], firefighting [23], videography [24], industrial inspection [25], and civil infrastructure monitoring [26]. The main components of a UAV are an airframe, a propulsion system, and a navigation system [27]. They include a wide range of aircraft design configurations and supporting tools that help with a variety of applications.

* Correspondence author at: Sustainable Energy & Power Systems Research Centre, RISE, University of Sharjah, P.O. Box 27272, Sharjah, United Arab Emirates.
E-mail address: mabdulkareem@sharjah.ac.ae (M.A. Abdelkareem).

https://doi.org/10.1016/j.ijft.2024.100590

Available online 28 January 2024

List of abbreviations

AI	Artificial Intelligence
ANNs	Artificial Neural Networks
AR	Autoregressive
CamShift	Continuously Adaptive Mean Shift
CNN	Convolutional Neural Networks
DETR	Detection Transformers
DOF	Degrees-Of-Freedom
DRL	Deep RL
DRL-IBVS	Deep Reinforcement Learning Image-Based Visual Servoing
EKF	Extended Kalman Filter
FCMs	Fuzzy Cognitive Maps
FFNN	Feedforward Neural Networks
GCS	Global Concept Score
IBVS	Image-Based Visual Servoing
IMU	Inertial Measurement Unit
L-IBVS	IBVS With Lyapunov Candidate Function
LSPI	Least-Squares Policy Iteration
mAP	Mean Average Precision
MDPs	Markov Decision Processes
MMP	Mechatronic Multi-criteria Profile
NHL	Nonlinear Hebbian Learning
PBVS	Position-Based Visual Servoing
PID	Proportional-Integral-Derivative
PSO	Particle Swarm Optimization
QP	Quadratic Programming
R-CNN	Regional-CNN
RGB-D	Red-Green-Blue Depth
RL	Reinforcement Learning
RMSE	Root Mean Square Error
SA-TQ-IBVS	Simulated Annealing-Truncated Q-Learning-IBVS
SAP	Search Area Proposal
UAVs	Unmanned Aerial Vehicles
ViT	Vision Transformer
YOLO	You-Only-Look-Once

Although this technology is not new, it has only recently begun to meet traditional business requirements by providing a cheaper, faster, and better alternative to full-size aircraft. UAVs powered by AI greatly depend on computer vision. This technology enables UAVs to detect objects while flying and collect and record information [28,29]. Computer vision has achieved high-performance onboard image processing-based systems using neural networks [30,31]. Computer vision refers to the discipline of visually analyzing and comprehending real-time information using computer systems or machines [32]. This system resembles the visual capability of living organisms, particularly humans. This science simulates the machine to extract, analyze, and comprehend an image containing real-world information by visualizing it the way humans do. Picture capture, image preprocessing, feature extraction, object recognition, and data extraction from images are the fundamental phases of computer vision [33]. Remote sensing, medical image processing, precision agriculture, satellite imaging, military, and other industries may benefit from computer vision. Thus, computer vision plays a major role in autonomous facilities. The main aspects of UAVs technologies are path planning [34], navigation control, landing control, mapping and localization and target detection/tracking, which are already mature and well covered. However, when constructing autonomous UAV platforms for unstructured situations, assuring safe flying in crowds [35], avoiding passing over people [36], or emergency landing capabilities in the event of a failure is still an afterthought [37].

There have been several recent works that overview and analyze the recent trends in the application of several AI techniques in different types of engineering applications. For example, a recent study by Juan and Valdecantos [38] reviewed the application of Artificial Neural Networks (ANNs) in different problems of high randomness level related to ocean engineering, such as predicting the wave heights, water level, and wave power, and proved the extreme potential of such AI-based techniques in comparison with traditional modelling and prediction methods. Another study by Lu et al. [39] presented a comprehensive review of the recent application of ANN models as a Machine Learning-based approach for the prediction of buildings energy, and concluded with a thorough guideline for choosing the correct ANN model architecture and mitigating various training and implementation issues for efficient and accurate modeling of buildings energy. In another work, Hussain et al. [40] discussed the recent trends in the application of Computer Vision techniques for detecting defects Photovoltaic cells by employing electroluminescence-based imaging and concluded the recent focus on Convolutional Neural Networks (CNN)-based frameworks for the analysis of such imagery. By reflecting on these works, AI-based techniques have shown extreme potential in a wide variety of analysis, prediction, modelling, and optimization tasks in all types of engineering sectors and are still emerging to fill a wider range of applications that were challenging to use classical techniques.

From the standpoint of computer vision, this work reviews the research progress in intelligent systems for UAVs visual servoing. There has been growing research interest in the application of Artificial Intelligence (AI) methods and Visual Servoing control techniques that is extensively performed in recent years, as can be seen in Fig. 1 which was generated using recent data from Google Scholar. Although there is growing research in AI-based visual control for UAV systems, there is a scarcity of extensive review articles that outline the overall trends and prospective developments in this area. This study aims to fill this important gap by providing an in-depth examination of the applications and progress of AI-driven visual servoing in autonomous UAVs, addressing key control tasks, and presenting valuable perspectives on potential future research, which are currently underrepresented in existing literature. For better comprehension, the paper is divided into the following sections. Section 2 presents a background on visual servoing considering two types of visual servoing, image based visual servoing and position based visual servoing. Section 3 reviews and discusses AI techniques for visual servoing of UAV systems, including neural networks, fuzzy logic, reinforcement learning, and hybrid artificial intelligence systems. A technical comparison is presented in Section 4. Future research directions and recommendations for improving the technology are discussed in Section 5. Fig. 2 shows the overall taxonomy of this work and progress done on AI-based visual servoing control methods for UAV systems.

As AI techniques have attracted huge research attention in the last years, a great number of articles were published in the last years for all types of research problems. Among these problems is robotic control on which many AI techniques were utilized for automating the robotic control for various tasks. This review article has three main goals. The first goal is to review the recent application of AI techniques, including neural networks, fuzzy logic, and reinforcement learning, for the visual servoing control of autonomous UAVs. The second is to discuss the current literature in terms of the current trends regarding the application of AI techniques in visual servoing control of autonomous UAVs. Finally, the third is to provide future research directions on the current research gaps for further improving the current technology. This paper is focused on intelligent control techniques that were applied to control autonomous UAVs while performing various tasks.

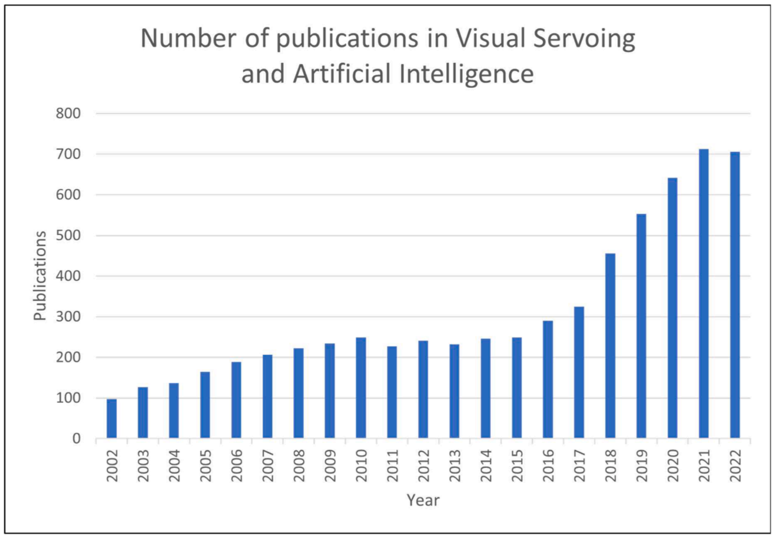

Fig. 1. The increasing number of publications in Visual Servoing and Artificial Intelligence from 2002 to 2022 (Data obtained using Google scholar advanced search: allintitle: "Visual servoing" OR "Artificial Intelligence" OR "Artificial Neural Network" OR "Fuzzy logic" OR "Reinforcement learning").

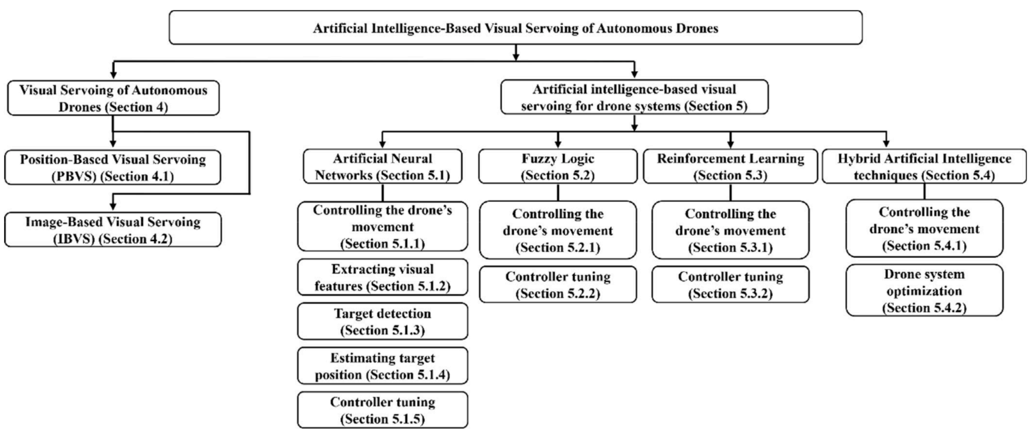

Fig. 2. Taxonomy of AI-based visual servoing of UAV systems.

2. Research methodology

With the increasing number of AI techniques that are developed and proposed in literature, AI has attracted huge research attention, and many articles were published in the last years for solving all types of problems. In the current work, several criteria were applied to narrow the used research articles. The databases from which the articles were collected include MDPI, IEEE, Springer, and ScienceDirect. The search was focused on the key words "Visual servoing control" and "vision-based control" along with the key words "Artificial Intelligence", "Artificial Neural Networks", "Fuzzy logic", and "Reinforcement learning". This search resulted in a huge number of articles which was narrowed by the inclusion of articles in which the AI technique was used for the control process and for analyzing the visual data obtained from the on-board camera stream. The collected papers were grouped based on the implemented AI technique and then were sub-grouped based on the task that the visual-servoing control system was used for. From the collected articles, it was noticed that Artificial ANNs attracted the most attention among the other techniques.

3. Visual servoing of autonomous UAVs

The process of "servoing" is an automated control method that involves comparing the current state of the robotic system to a desired state and generating a feedback control signal [41]. Visual servoing control is a type of servoing control that utilizes visual characteristics of the target object, captured via an onboard or external camera, to direct the control actions [41,42]. Visual servoing can be divided into two main categories. The first is Position-Based Visual Servoing (PBVS), which employs pictures of the target object and its known geometrical dimensions along with the geometrical dimensions of objects in the observed environment to guide the control process [43]. To determine the goal position, relevant information is extracted and processed from the camera's captured images . The second category of visual servoing is called Image-Based Visual Servoing (IBVS), which eliminates the need for position estimation used in the first type by utilizing visual features extracted from the incoming video stream to directly influence the control actions [41]. This section overviews the two main categories of visual servoing techniques: PBVS and IBVS, as shown in Fig. 3.

3.1. Position-based visual servoing (PBVS)

PBVS control systems use the video stream obtained from the UAV's onboard camera for extracting visual features about the observed scene and then use these features to predict the relative pose difference between the UAV system's current pose and the required pose [43]. Fig. 3 (a) shows the general structure of PBVS control systems. For such techniques to be useable, the geometric shape and the dimensions of the

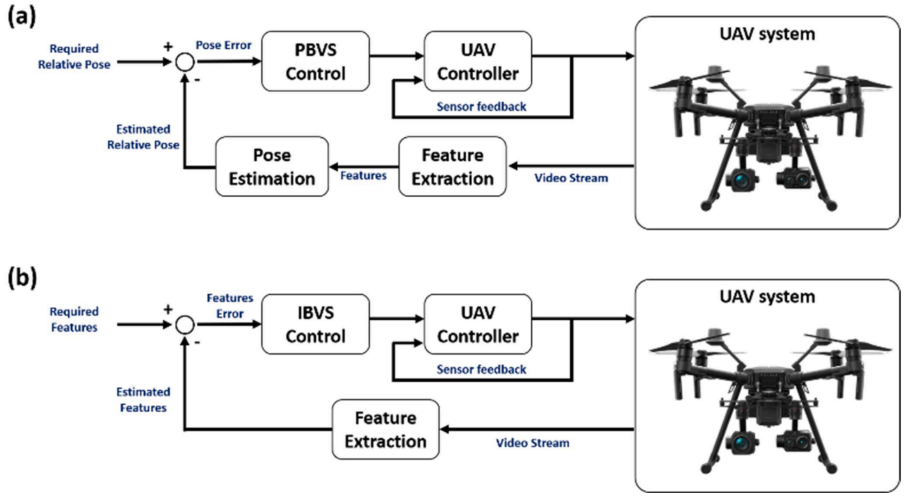

Fig. 3. Schematics of the two types of visual servoing: (a) Position-Based Visual Servoing (PBVS) and (b) Image-Based Visual Servoing.

reference target object must be known beforehand as well as the used camera's properties and types of image features that will be utilized [44]. The obtained video stream first goes to the feature extraction stage, where various image analysis techniques, such as line detectors [45,46], edge detectors [47,48], color analyzers, and Convolutional Neural Networks (CNNs) [49], are utilized to extract useful visual features. The extracted features are passed to the pose estimation stage that uses the features and the object's geometry to find the relative pose of the camera regarding the target object [50]. The PBVS controller uses the relative pose difference to send control commands to the UAV to move the UAV in the direction that lowers the estimated pose difference. The servoing loop continues until the pose error is minimized which occurs when the UAV reaches the desired position. In the case of UAV missions, the target object can be a landing platform or any other object with known geometrical shape and dimensions [51].

3.2. Image-based visual servoing (IBVS)

IBVS control is similar to PBVS in using the camera's video stream to extract visual features about the scene but ignores the pose estimation stage of PBVS systems [41,52]. Fig. 3(b) shows the general structure of IBVS control systems. In IBVS, the values of the extracted features are used directly as a feedback signal to guide the visual servoing process as they are compared to desired feature values to provide an error signal, thus implicitly representing the pose error [53,54]. The IBVS controller uses the error in image features to send actuation commands to the UAV system so that the observed features move to their required positions in the image plane. For example, features that can be used for controlling the UAV movement include the corners of the platform of landing and the center of the tracked object [55]. IBVS control stops when the feature's error converges to zero, corresponding to the UAV system reaching the required position [56].

4. Artificial intelligence-based visual servoing for UAV systems

In various tasks, artificial intelligence techniques are used in literature to design visual servoing control systems for controlling UAVs, including 3D UAV positioning, aerial object following, ground object following, obstacle avoidance, and autonomous landing. The AI techniques include ANNs, Fuzzy logic, Reinforcement learning, and Hybrid AI techniques. Several AI-based techniques, such as Artificial Neural Networks (ANN), Fuzzy logic, Reinforcement Learning (RL) and hybrid AI Techniques, are used in the vision-based control of UAV systems. This section reviews the application of these techniques in various UAV operation tasks. Each technique is discussed in a specific sub-section.

4.1. Artificial neural networks (ANN)

The artificial neural network (ANN) works as a parallel processing network, where each of the cells in the brain is interconnected and provides the brain with the ability to analyze, learn, and recall [57]. Each of these cells is linked to the next cell, providing the brain with the ability to analyze, recall, learn, and think [58,59]. The ability of neural networks to express complex input/output relationships is perceived with major significance. A typical ANN system has multiple layers, with the input layer being the first, followed by several hidden levels, and lastly, an output layer. The input layer is trained and weighted by the neurons in the hidden layer during the learning phase. The output layer provides the desired target, as shown in Fig. 4. This system undergoes an extensive training process to learn how to reflect the input patterns in the actual system [60,61].

Fig. 4 shows the schematic diagram of the multilayer feedforward structure. The efficiency of ANN algorithm strongly depends on the network parameters set, such as the hidden layer's amount of neurons, the momentum constant, and the learning rate [62]. However, when increasing the number of neurons in the hidden layer, the training time of the ANN will increase and could result in an overfitting issue which influences the ANN's ability to generalize [63]. On the other hand, using a low number of neurons could result in the inaccuracy of the output where the system would be inadequate to learn and alight the input and output data [64]. The number of neurons is generally calculated as the difference between the number of input and output data [65]. The momentum constant ensures that the system does not become caught in local minima [66]. The learning rate is used during the training phase to control the weight size and bias until the predicted output matches the intended objective [67]. ANNs can be used in various visual servoing-based robotic systems applications. ANN-based detection networks, such as You-Only-Look-Once (YOLO) [68] and Detection Transformers (DETR) [69], can be used for detecting the followed object using the video stream and guide the visual control process. Also, Convolutional Neural Networks (CNNs) can be used for extracting visual features that can be used during the control process. ANNs can be implemented with conventional controllers, such as Proportional-Integral-Derivative (PID) controllers for intelligent tuning of the control gains. Finally, the ANN is applied to estimate useful parameters that can be implemented in control, such as estimating the target's position and distance.

4.1.1. Controlling the UAV's movement

Ananthakrishnan et al. [70] developed an offline backpropagation neural network controller for visual servoing during landing. The quadrotor control system is designed to land correctly on a marker

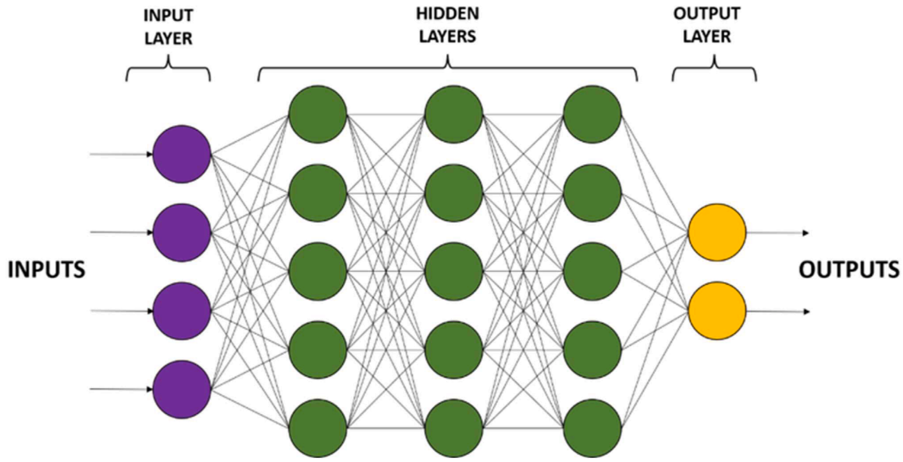

Fig. 4. Schematic of the structure of an artificial neural network.

platform within the time and distance constraints. The training yields a network weight that may be applied in two ways. First, as a starting weight for an online learning neural network that adjusts to the environment or for precision landing. Second, the network weights, particularly the training data, were more universal. Regardless of the challenges and the environment, the quadrotor managed to land and orient itself. The quadrotor landed over the landing station despite its starting location and orientation while obtaining the correct orientation. Corrections to orientation, position, and altitude are done simultaneously rather than separately.

4.1.2. Extracting visual features

Choosing an appropriate set of features is critical since it directly affects the performance of any visual servoing approach. Saxena et al. [71] intend to understand visual feature representations appropriate for servoing tasks in an unstructured environment and unfamiliar settings, inspired by modern advancements in the performance of data-driven approaches on identification and localization tasks. When information on camera settings and scene geometry are not available a priori, the authors [72] proposed an end-to-end learning-based technique for visual servoing in various circumstances. A convolutional neural network is trained on color images with synchronized camera postures to achieve the results. Next, the authors used a CNN model to assess the relative camera change needed to achieve the intended picture posture in image space. The network's learned visuomotor representations transfer effectively across varied settings. The proposed approach was tested experimentally in various scenarios under different illumination levels, and the results proved the reliability and robustness of the developed system while not requiring any knowledge about the scene's geometry and camera parameters.

A similar study by Harish et al. [73] proposed a two-pronged strategy for visual servoing where optical flow is used as a visual property predicted by a deep neural network and an interaction matrix. These flow characteristics are systematically combined with depth estimations supplied by another neural network for guiding the visual servoing process. The authors suggested two designs to forecast the scene's depth at the present pose. Lastly, an interaction matrix comparable to traditional IBVS merges the depth and visual attributes. This aids in the development of a stable and precise visual servoing system. The results demonstrated convergence over 3 m and 40° mile while maintaining the exact placement of less than 2 cm and 1°, whereas previous techniques failed to converge for most cases over 1.5 m and 20°. The techniques proposed by Collewet et al. [74] and Saxena et al. [71] both diverge compared to both of the presented techniques (flow-depth based and depth-network based) that are capable of converging without oscillation. Also, Harish et al. [73] tested the technique on an aerial robot in a real-world environment. Without additional retraining or fine-tuning, the technique generalized to unexpected settings, resulting in accurate and robust servoing performance at 6 degrees of freedom positioning issues with even large camera transformations.

4.1.3. Target detection

UAV target detection is a new research trend. Chen et al. [75] utilized an IBVS tracking controller with lower tracking error and lower control effort requirements than conventional IBVS. The system is based on a target motion pattern derived from quadratic programming (QP) and bounding box features generated by a YOLO detector. The tracked object always remains within the image. Three simulations were run to compare the proposed controller's efficacy and performance to a traditional image-based visual servoing controller. The three simulations indicated three important features of the proposed control method. First, the proposed controller required lower control effort than traditional controllers while ensuring better control performance. Second, the generated control command value always stays in the UAV's acceptable range and does not surpass it. Third, the target is followed correctly, and its features are always located in the observed frame, in contrast to conventional IBVS where such localization is not guaranteed.

4.1.4. Estimating target position

One of the main challenges involved in the field is estimating the UAV's target position. For instance, Ramon-Soria et al. [76] developed a system that successfully grabbed the required object by an unmanned aerial vehicle for outdoor applications. ANN is applied to obtain necessary information about the object's position. The proposed algorithm consists of a Convolutional Neural Network (CNN) using a random sample consensus algorithm to detect the object and locate its required aerial manipulator commands using Extended Kalman Filter (EKF). An RGB-D camera was mounted on the UAV's body to extract the observed scene's color view and depth value. The system accurately estimated the required pose and moved the manipulators' joints to achieve the desired tasks. It also switched to another target grasp when the current target grasp could not be reached. The findings of several experiments prove that the efficiency of the proposed system is correctly performing intervention tasks under various environments, different illumination conditions, and various object shapes.

Similarly, Durdevic and Ortiz-Arroyo [77] developed a visual servoing control system on a stereo camera and two deep ANNs for positioning the UAV during a wind turbine inspection mission. The two ANNs were accurately used to measure the position of a wind turbine in a 3D space. Two YOLOv2 detector networks were used to estimate the observed object's distance using video streams from two 2D cameras and epipolar-plane image analysis. Some obstacles were considered first for the sensor to be implemented. This approach allowed for maintaining programmatic control over an item and the initial bounding box, which will be manually at the start. The virtual camera's motions might be monitored to automatically compute the position and size of the new bounding boxes. This data was collected and used to generate a labelled training set. The results revealed that when the camera is stationary or moving at a slow speed, the system successfully calculated the position of the examined wind turbine.

A comprehensive vision-based real-time deep target tracking system was presented by Kassab et al. [78] to command the UAV to follow a ground target. The utilized IBVS system allowed a flying UAV to estimate the distance to the target and control the UAV's orientation. The built tracking system consists of two deep neural networks: the approaching and the chasing networks. Both networks' detection accuracy was confirmed, and a series of real-world flight data demonstrated the suggested system's effectiveness in target tracking and following and both deep networks reached significant levels of accuracy. The summarized overall mean Average Precision (mAP) of the chasing networks for each ground target's orientation was more than 90 % for most cases. The proposed vision-based system was tested for 4 min during a flight. The findings indicated the efficiency of the proposed system in detecting and following the ground target.

Huang et al. [79] developed a limited image-based visual servoing control method for the shipboard landing issue of unmanned helicopters. An autoregressive (AR) model was used to find the ship's pitch and roll motion to identify the best landing time. This control technique showed high accuracy and robustness during the landing on a ship-mounted moving landing platform. The modified Chebyshev neural network predicted uncertainties such as the ship's linear acceleration and translational perturbation, while an adaptive rule adjusts the rotating disturbances. The controller requires only a vision sensor and an inertial measurement unit (IMU) to function. Finally, simulations are used to prove that the proposed shipboard landing control technology is capable despite the various disturbances. The comparison of the expected and actual motions revealed that the AR model technique has a high level of prediction accuracy. The simulation results demonstrated that the proposed constrained IBVS controller could achieve autonomous shipboard landing with high robustness and accuracy and ensure visual target visibility in the presence of a 6-DOF (degrees-of-freedom) moving ship, unknown relative linear velocity, disturbances, and measurement noises.

4.1.5. Controller tuning

Lopez-Franco et al. [80] developed an ANN-based proportional integral derivative PID controller that adapts to changes in mass and moment of inertia based on the vision sensors. PID controller helps create an IBVS control that knows the position of the UAV by utilizing a velocity vector as a reference to control the hexacopter position. This integration necessitates close coordination between control algorithms, models of the system to be controlled, sensors, hardware and software platforms, and well-defined interfaces, as well as the design of different processing stages with their respective communication architectures to allow for real-time implementation. Experiments were carried out on the Asctec Firefly on-board computer to illustrate the utility of the sensor integration and control method in handling these difficulties on a high nonlinear system with noisy sensors such as cameras. Both simulation and experiment results are included. This technique achieved lower Root Mean Square Error (RMSE) and average absolute deviation values than conventional PID.

4.2. Fuzzy logic

Fuzzy systems are information processing structures based on fuzzy approaches where using conventional sets theory and binary logic is impossible or hard [81]. This approach is suitable for many practical applications to process inaccurate information [82]. The terms fuzzy system, fuzzy model, a system based on fuzzy rules, fuzzy controller, and fuzzy associative memory are used interchangeably depending on the application type [83,84]. Their key feature is the representation of symbolic information in fuzzy conditional (if-then) rules [85]. Fuzzy logic has evolved as a valuable tool for controlling and steering systems, complicated industrial processes, domestic and entertainment devices, other expert systems, and applications such as SAR data classification [72]. These notions were invented as between values, which can mathematically be formulated and processed by computers [86]. Fuzzy systems, which have their roots in ancient Greek philosophy, are alternative to traditional notions of set membership and logic [87]. An example of a fuzzy set or crisp set would be the middle values between 0 and 1 or values between the minimum and maximum of the fuzzy set [88]. Thus, fuzzy logic provides a novel approach to control or classification [87]. Rather than describing how the system works, the fuzzy system focuses on what the system should perform [89]. One can focus on addressing the problem rather than attempting to represent the system [74] mathematically. On the other hand, the fuzzy technique necessitates adequate expert knowledge to formulate the rule base, the combining of the sets, and the defuzzification [90]. Fuzzy logic systems can be applied in visual sevoing control systems either by tuning the gains of another controller, such as a PID controller, or for directly generating control commands for the UAV.

4.2.1. Controlling the UAV's movement

Olivares-Mendez et al. [91] developed a fuzzy logic-based visual servoing system for controlling UAV's transitional and yaw movements. The control system aimed to follow the desired object while keeping the UAV at a constant distance away from it and maintaining it at the center of the image plane. Continuously Adaptive Mean Shift (CamShift) algorithm was utilized to detect the tracked object in the captured images based on their color, and a red balloon was used as a sample tracked object. After visual detection of the tracked object, two fuzzy logic controllers were used to give control actions to the UAV based on taken features from the vision system. The first fuzzy logic controller was used to control the UAV's yaw and the second fuzzy logic controller was used to act on the UAV's pitch. The UAV followed the object for around two minutes while keeping the object at a safe distance and the center of the camera's focus. The obtained results indicated the robustness of the proposed system against lighting variations in the visual systems and any uncontrolled weather changes.

Correspondingly, Olivares-Mendez et al. [92] proposed a fuzzy logic-based visual servoing method for autonomous object avoidance of a quadrotor UAV. The control method uses image data obtained using the quadrotor's front camera and extracts the location of the obstacle using a color detection algorithm called CamShift. Once the obstacle is in the UAV's range of vision, a fuzzy logic controller gives yaw commands to the quadrotor UAV so that the obstacle is either to the left or right of the UAV's front. The fuzzy logic controller takes two inputs: the distance between the obstacle's location in the image and the center of the picture and the difference between the last two distance values. The fuzzy logic controller's single output is the yaw command sent to the quadrotor to avoid the obstacle. When the method was experimentally applied, the quadrotor could quickly react to the detection of the obstacle and easily avoid it. The system enabled the quadrotor to react fast to obstacle detection and avoid it.

Similarly, Olivares-Mendez et al. [93] employed two control strategies for a see-and-avoid task by micro UAV. The fuzzy logic technique was utilized in both approaches. Each controller receives information from the image processing front-end, which recognizes and tracks targets in the environment. Visual data is then used to execute autonomous avoidance using a visual servoing technique. The optimized controller achieved a RMSE value below 3°, thus verifying the excellent performance of the vision-based system.

An efficient fuzzy visual servoing system for obstacle avoidance employing an unmanned aerial vehicle was demonstrated by Olivares-

Mendez et al. [94]. The authors outlined the image processing front end and visual servoing technique for heading control using fuzzy logic. The controller aims to generate the vehicle's intended yaw commands focused on the target's position in the picture plane. There are three inputs and one output on the controller. The output is the yaw instruction required by the vehicle to maintain the object at the appropriate relative bearing. Using the cross-entropy theory, the gains are optimized for the controllers. ROS-Gazebo 3D simulation was used to execute the optimization process. The use of cross-entropy methods successfully estimated the optimal gains, and it achieved great results when tested in real application flights. Root-Mean-Squared Error (RMSE) metric was used to evaluate the controller's behavior. The small RMSE values were obtained to verify the excellent behavior of the optimized controller.

Huang et al. [55] studied the application of a visual servoing system with a multi-level fuzzy logic controller for autonomous landing of a micro-UAV. The visual servoing system relies on two cameras' images located on the bottom and front of the micro-UAV. First, autonomous control strategies were used to enable the micro-UAV to find the landing platform's location by following a specific flying strategy. After the landing platform is located, the multi-level fuzzy logic controller is utilized to guide the landing process. The first layer fuzzy logic controller is used to retain the center of the landing platform at the image center of the bottom camera. However, the first layer fuzzy logic controller alone resulted in significant oscillations when approaching the landing platform due to the exerted force by the reflected air flow. Thus, a second layer fuzzy logic controller was applied to the first controller to fine-tune the control command actions. The combination of the two controllers resulted in significantly lower overshoot in the "x" and "y" error.

4.2.2. Controller tuning

Touil et al. [95] developed a fuzzy logic approach to gain adjustment of PD controllers to control the quadrotor's direction and position. Unlike traditional visual servoing approaches, Touil et al. [95] computed the proportional controller gain by smoothing the difference between the desired and actual feature vector values. The ineffective management of this gain impacts the tracking of features in the field of vision. Compared to the classical visual servoing scheme, which requires a known gain value, which is chosen heuristically, causing the controller to fall victim to the velocity convergence problem. A fuzzy logic controller is used to change an adaptive gain value. The suggested controller uses the error between feature vectors and their derivatives as inputs. The suggested controllers, used in both the inner and outer loops, provide numerous benefits in dealing with uncertainty in trajectory tracking control. MATLAB simulations indicate that the new strategy outperforms the traditional method. The comparison findings also reveal that it outperforms the traditional PD controller in case of disturbance.

Another study by Singh and Anvar [96] proposed an image-based detection and tracking algorithm for the application of UAV platforms. The developed system uses an IBVS control technique that utilizes Viola-Jones method and Blob analysis. The goal of the work was to design a visual control technique for a UAV to land safely on a target. Fuzzy logic was used to calibrate a PID controller for guiding the UAV's movement throughout the vertical takeoff and landing mission. Initially, some over-shooting and fluctuations occurred in the step response before PID calibration. After calibrating each motor's controller, the system reached its critically damped response and achieved system stability.

4.3. Reinforcement learning

Reinforcement learning (RL) is an artificial intelligence subfield where an intelligent agent learns a policy for reaching the required system state starting from the initial state without prior knowledge of the environment [97]. An agent is an intelligent entity that observes its environment and acts on it autonomously, such as an autonomous robot. The goal of RL is finding the best possible policy, which is a sequence of possible control actions, to reach the goal state starting from the initial state [97]. RL works through a reward-based structure where desirable actions by the agent are rewarded, and undesirable actions are punished [98]. During the agent's operation, the agent explores its environment through trial-and-error and learns from the reward values that it observes [99]. The learning algorithm will then try to move more toward the regions of the state space where higher reward values are obtained, and vice versa [99]. As this operation continues and after many learning iterations, the agent will gain better performance in the required task and collect higher cumulative rewards [99]. However, an important aspect to be taken into consideration is the tradeoff between exploration and exploitation, where the agent could either exploit the regions of the state space where high reward values are guaranteed or explore the previously unreached state for a chance of achieving an even higher cumulative reward value [100]. The environments, where RL is employed, are commonly modeled as Markov Decision Processes (MDPs) where a state space S, an actions space A, a transition function $T(s,a,s')$, a reward function $R(s,a,s')$, and a learning discount factor γ are used to model the environment; however, neither the transition function nor the reward function are known beforehand [101]. The RL algorithm will either learn their values throughout the learning process for fining the optima policy or will directly find the optimal policy without estimating their values [102]. RL is utilized in several ways in visual servoing such as learning a policy that relates the visual features error to velocity commands and learning the values of visual control gains [103].

4.3.1. Controlling the UAV's movement

Sampedro et al. [104] discussed the application of an Image-Based Visual Servoing controller that employs a deep reinforcement learning technique named Deep Deterministic Policy Gradients (DDPG) for visually controlling the movement of an autonomous UAV following a ground target. The trained control algorithm could be transferred easily from simulation to real-world applications. The performance of the proposed system was compared to two well-established IBVS techniques: classical IBVS [53] and a partitioned IBVS [42]. When the three techniques were compared in a simulated environment, all techniques showed excellent results in visual object tracking. However, when these techniques were evaluated in a real experiment, the proposed reinforcement learning based IBVS system showed superiority in terms of exceptional tracking performance and ease of switching the algorithm from simulation to real-world application. The other two techniques required enormous parameter tuning to be implemented correctly in a real UAV system while the proposed reinforcement learning based IBVS system did not require such tuning even though it was trained using a different UAV model compared to the UAV system. The proposed reinforcement learning based IBVS showed noticeably lower feature errors compared to the two other techniques.

A study by Akhloufi et al. [105] proposed two visual control methods for controlling a UAV to track another UAV in a "UAV pursuit-evasion" scenario. The first method implements deep reinforcement learning that uses the video stream from the UAV's front camera to predict the required control actions for following the other UAV. Several experiments were carried out to investigate the effectiveness of the developed approaches. Outdoor experiments indicated the effectiveness of this method in tracking moving UAVs; however, this method showed limited tracking performance when the tracked UAV moved in a very high speed due to the long processing time required. The second method used a deep learning detector named YOLO-v2 combined with Search Area Proposal (SAP) to detect the goal UAV's position in the video stream and predict its position in the next video frame. The utilization of SAP improves the control system's performance in detecting far away UAVs and only occupying a small area in the obtained video frames. However, a serious limitation of this method is the low detection performance when the target UAV is taking most of the obtained frame. This is because of the working principle of CNNs that search for features in small windows throughout the frame. However, the UAV could efficiently follow the target throughout the mission. The use of Vision Transformer (ViT) networks has been investigated in [106] in such object detection systems due to their interesting property of extracting large-scale features from the obtained frames and thus can be a potential solution for this limitation.

Shaker et al. [107] presented a method that allowed fast and accurate learning of the landing process through reinforcement learning-based visual servoing control system for controlling the landing of a UAV system. A reinforcement learning method named Least-Squares Policy Iteration (LSPI) was used for learning a control policy from the visual features obtained from the UAV's bottom camera. This method does not require precise tuning of its parameters prior to the learning process as it does not require a learning rate parameter and all other parameters can be learnt throughout the training process. Processing of video frames obtained from the bottom camera was carried out using shape and color analysis through three main steps. The proposed system was examined in a simulated environment, and it allowed for fast and accurate learning of the landing process. However, validation of this system in an experimental UAV system while considering more noise parameters is required.

Research conducted by Shinde et al. [108] created a DRL-IBVS (deep reinforcement learning image-based visual servoing) controller for aerial robots. The model maps the bounding box errors of the target to linear velocity commands using deep deterministic policy gradient DRL with CNN (YOLOv2). The noise in detection is reduced via a multi-object Bayes filter, and the prediction model uses a correlation tracker to estimate the target prior. The authors utilized a reinforcement learning-based strategy to control the follower UAV to address the control challenges that emerge in traditional methods due to limitations in interaction matrix estimates. It was observed that the applied control effort is initially high because of the higher feature error, but it decreases with time as the error decreases. Also, the controller is always taking remedial action to reduce the number of faults. The model was able to follow the target and keep the identified UAV's location in the picture frame within a limited bound around the frame's center.

4.3.2. Controller tuning

Shi et al. [109] utilized Q-learning RL with bootstrap aggregation (bagging) algorithm to tune the controller's gains, generating velocity commands. A bagging method that utilizes time-varying image errors was used instead of the Moore-Penrose pseudo-inverse for finding the required inverse kinematics for the visual servoing. This technique significantly reduces the effects of image noise and overcomes the quadrotor's under-actuated dynamics. The developed Bagging-IBVS (B-IBVS) technique was compared to three other works in the literature which are C-IBVS [110], A-IBVS [53], and V-PID [111], and the findings showed the superiority of the proposed method in terms of convergence speed and noise rejection. Also, incorporating Q-learning results revealed excellent stability and a fast convergence rate. The collective system showed excellent stability and fast convergence rate compared to the B-IBVS without the reinforcement learning-based adaptation and thus, it is highly practical and efficient. However, the UAV's movement showed noticeable oscillations when the target's speed was higher than 0.6 m/s, which must be addressed in a future as it is a significant limitation to the possible applications of this technique.

Another study by Shi et al. [112] proposed an adaptive IBVS technique for controlling a quadcopter's position during hovering. The implemented technique named Simulated Annealing-Truncated Q-learning-IBVS (SA-TQ-IBVS) employs reinforcement learning and simulated annealing metaheuristic for achieving fast, accurate, and stable visual control of the quadcopter. Truncated Q-learning reinforcement learning technique was used for learning a policy for adaptively estimating the servoing gain for the IBVS throughout the quadcopter's mission. Simulated annealing algorithm was used for selecting the control actions throughout the mission while balancing the trade-off between exploration and exploitation, thus speeding up the convergence of the learning process and minimizing the chance of physical damage to the quadrotor during learning. Several experimental and simulation tests were conducted to test the proposed system. The simulated annealing-truncated Q-learning technique was compared to 0-greedy truncated Q-learning and 0.2-greedy truncated Q-learning in relation to the number of time steps in the obtained solution. The results showed that the simulated annealing-truncated Q-learning method achieved faster convergence speed and better stability. Finally, the proposed IBVS system was compared to other IBVS techniques: C-IBVS [53] and VPID-IBVS [113]. The experimental results illustrated that the proposed system converges faster and shows lower oscillation next to the goal position.

A similar study by Shi et al. [114] proposed a hybrid visual servoing system that combines Q-learning reinforcement learning algorithm with fuzzy logic for visual servoing of a quadrotor UAV tracking ground objects. The proposed IBVS control system uses a decoupled controller design where two decoupled controllers are used for visual servoing, one for linear motion and another for angular motion. The decoupled controller architecture enables the correction of any visual distortion induced by the quadrotor's roll and pitch. A reinforcement learning algorithm named Q-learning was used to improve the control system's stability and convergence rate and to adjust the two controllers' servoing gain. The fuzzy logic controller takes the changing rate of the features' error as its input and returns the Q-learning algorithm's learning rate as its output. Simulations of the proposed method were carried out followed by experimental verification. The proposed visual servoing system was compared to two other systems which are Conventional-IBVS [115] and IBVS with Lyapunov candidate function (L-IBVS) [116]. The hybrid visual servoing system achieved fast and accurate UAV control as it illustrated better control results and a simpler structure.

5. Hybrid artificial intelligence techniques

Various AI techniques were hybridized in literature for the design of visual servoing control systems for UAVs. In this section, the hybrid artificial intelligence techniques that were applied in literature are reviewed.

5.1. Controlling the UAV's movement

Amirkhani et al. [117] proposed Fuzzy Cognitive Maps (FCMs) for IBVS of quadrotors to follow ground objects to control the UAV's movement. The proposed control method relies on FCMs, which combine the basic characteristics of ANNs and fuzzy logic systems and thus integrate the learning ability of neural networks with the causality provided by the fuzzy relations. First, the authors used an untrained FCM-based controller for following the moving target. The untrained controller provided acceptable object tracking but better tracking performance was still accessible. Thus, the FCM controller was trained using the Nonlinear Hebbian Learning (NHL) algorithm and used again for visual servoing. The obtained simulation results show a significant improvement in the tracking performance after the training. Finally, the effect of disturbing torques and forces on the trained controller were evaluated to assess the robustness of the controller. The quadrotor could efficiently and robustly follow the target and correct its control signal for different disturbances. The controller showed accurate target following despite the target's height and yaw data unavailability.

5.2. UAV system optimization

A multi-criteria fuzzy decision system can help in quadrotor design, a very complex and composite process. Mohebbi et al. [118] presented fuzzy integral based neural network for UAV system optimization. A novel mechatronic multi-criteria profile (MMP) for concept evaluation in mechatronic system design was presented, and the assessment approach was thoroughly investigated. The MMP is made up of five primary design criteria: intelligence, reliability, complexity, flexibility, and affordability, and it may be included in an automated design procedure. A global concept score (GCS) has been computed based on the measured MMP for each idea, and different aggregation approaches to facilitate concept evaluation, selection, and customization. The UAV system was designed for a visual servoing task. The optimal system performance was verified in a simulation by matching four image features to four specified image plane positions and measuring the camera frame velocity of the target features in systems based on using two proposed concepts.

A similar study by Mohebbi et al. [117] developed a fuzzy logic-based Particle Swarm Optimization (PSO) for the multidisciplinary design and optimization of a visually controlled quadrotor UAV. Different aspects of the quadrotor's design were considered, such as the structure, system dynamics, propellers' aerodynamics, flight control, and stereo camera based IBVS system. A unified performance index was derived using a fuzzy-based method to serve as a multi- objective function for multicriteria optimization. Before implementing PSO, the unified objective function was obtained using fuzzy measures and a multicriteria fuzzy aggregation method to obtain the optimal design parameters. The fuzzy measures were used to represent the importance of each criteria, and the different objectives were combined using the Cascade Choquet integration-based aggregation method. The quadrotor's IBVS system takes image data from two stereo cameras and uses it for feature extraction. The performance of the optimized quadrotor system was evaluated in a visual servoing experiment where the UAV is required to follow a moving object on the ground and then intercept it by landing on it. The obtained results were compared to that of a commercial micro UAV, the AR. UAV 2. The results showed that the optimized quadrotor system achieved enhanced performance compared to the commercial UAV. Furthermore, the performance of the optimized UAV system was further improved by expanding the number of optimization iterations and the swarm size of the PSO algorithm. The further optimized quadrotor system achieved better visual servoing performance as the visual feature errors converged faster to zero.

6. Summary and discussion

Many studies have been published in the literature on artificial intelligence approaches in the visual servoing control of autonomous UAVs. Table 1 summarizes the progress made in implementing various types of AI techniques, such as Artificial Neural Networks, Fuzzy logic, Reinforcement learning, and hybrid AI techniques to design visual servoing systems for autonomous UAVs. The developed visual servoing systems were used in various control tasks such as 3D UAV positioning, aerial and ground moving objects following, obstacle avoidance, aerial manipulation, and autonomous landing. Moreover, various types of vision systems were applied in the literature. These visual sensing systems include conventional RGB cameras [70], stereo vision cameras [118], RGB-Depth (RGB-D) cameras [73], event cameras [51], and thermal cameras [51]. Among these visual sensors, RGB cameras were the most used in literature. This is due to their low cost and easy-to-process output data format that can be easily processed using artificial intelligence data processing paradigms, such as Convolutional Neural Networks (CNN) [75,119]. However, other types of visual sensors are of extreme importance and could be utilized with high potential for the control of UAV systems. For example, event cameras offer a significant advantage in dynamic and challenging lighting conditions due to their high temporal resolution and low latency, enabling more accurate and efficient capture of fast-moving objects and reducing motion blur compared to traditional cameras which could positively influence the controller performance. Another important type of visual sensors is thermal cameras which when mounted on UAVs provide the distinct advantage of enabling night-time and low-visibility operations, as they can detect heat signatures and thermal anomalies, making them invaluable various applications ranging from search and rescue missions to surveillance in various environmental conditions. These types of cameras were used in much less frequency in literature and thus further investigation of such sensors in UAV control is recommended. Regarding ANNs, the types of ANNs that were used in UAV visual servoing systems include Feedforward neural networks (FFNN), Convolutional Neural Networks (CNN), and Chebyshev neural networks. The most used ANN type is convolutional neural networks that were used for extracting visual features [119], detecting the required object [120], or estimating the target's position [77]. This is attributed to their excellent ability in automatically detecting and learning hierarchical patterns in visual data, making them highly effective for image recognition and processing tasks, such as facial recognition, object detection, and autonomous vehicle navigation. However, various types of ANN techniques for image processing were not investigated previously in the literature. These techniques include regional-CNN (R-CNN) which offers a significant advantage in object detection tasks by combining region proposals with CNN features, enabling precise localization and classification of multiple objects within an image [121,122], Siamese networks which are particularly advantageous for tasks that involve comparing and differentiating between pairs of visual inputs, such as in face verification or tracking, due to their unique architecture which efficiently learns to distinguish subtle differences and similarities by processing two image inputs simultaneously [123,124], and Vision Transformers (ViT) which leverage the power of self-attention mechanisms to process images as a sequence of patches, offering an advantage in capturing long-range dependencies and complex patterns in visual data, which enhances performance in tasks like image classification and object detection, particularly in large-scale and diverse datasets [69,125,126]. Regarding fuzzy logic, most of the works used simple fuzzy logic controllers either for generating control commands for the UAV or tuning another controller's parameters that are used in-parallel, such as a Proportional-Integral-Derivative (PID) controller. Fuzzy logic can be effectively incorporated with other AI models to control UAVs, providing a robust hybrid approach that combines the interpretability and rule-based reasoning of fuzzy systems with the learning capabilities of AI, resulting in enhanced decision-making and adaptability in complex and uncertain environments. About Reinforcement Learning (RL), various conventional RL and deep RL (DRL) algorithms were implemented in the literature. These algorithms were either used for adaptively tuning the parameters of a controller or for learning a policy for mapping visual features to control commands, thus they were combined with other controllers or feature extraction methods. Incorporating RL with other AI models for controlling UAVs offers a distinct advantage, as it enables the system to learn optimal control strategies through trial and error in low risk simulated environments, enhancing adaptability and efficiency in real-world applications, such as autonomous navigation and obstacle avoidance. Finally, hybrid AI systems were used to control the UAV's movements or optimize the design of a UAV system for visual servoing tasks. Hybridizing two AI models for controlling UAVs combines the strengths of different approaches, such as the precision of supervised learning with the adaptability of reinforcement learning, leading to more robust and versatile control systems capable of handling a wider range of scenarios and tasks with improved accuracy and efficiency.

7. Conclusions and future research directions

This work reviews the progress in applying artificial intelligence-based techniques in visual servoing of UAV systems. This study seeks to bridge the gap of limited comprehensive review articles by thoroughly reviewing the use and progress of AI in visual servoing for autonomous UAVs, focusing on essential control tasks, and highlighting areas for future research that are not adequately covered in the current literature. First, background on visual servoing control systems and the

Table 1
Progress on applying artificial intelligence techniques in visual servoing control of autonomous UAVs.

AI type	AI technique	AI role	UAV control task	Visual servoing type	Vision system	Performance Issues	Ref
Artificial neural networks	Feed-Forward Neural network	Controlling UAV's translational and yaw movements	Autonomous landing	IBVS	RGB camera	• The UAV system could land autonomously in the desired orientation regardless of its initial conditions	[70]
	Convolutional Neural network (FlowNet2 and Depth Network)	Controlling UAV's translational and yaw movements	3D positioning	IBVS	RGB-D camera	• The network showed convergent results for 3 m and 40° of pose error. • The technique's efficiency was validated in a UAV visual servoing experiment.	[73]
	Convolutional Neural Network	Extract visual features	Moving object following	PBVS	RGB camera	• The UAV system achieved the required pose despite the large initial transformations in the camera's pose and the large variation in appearance between the objects in the same category.	[119]
	Convolutional Neural Network (FlowNet)	Extract visual features, Estimate relative camera transformation	3D positioning	IBVS	RGB-D camera	• The approach did not require knowledge of scene geometry and camera parameters. • The system showed robust performance in non-homogeneous illumination and different scene textures. • The obtained camera pose error was around 4 %.	[71]
	Convolutional Neural Network (YOLO)	Target detection	Moving object following	IBVS	RGB camera	• The system achieved lower tracking error and lower control effort requirement compared to conventional IBVS. • The control effort is always within the UAV's capability. • The tracked object always remains within the image.	[75]
	Convolutional Neural Network (YOLOv3)	Target detection	Moving object following	IBVS	RGB camera	• The system could detect multiple UAVs in various environments. • Kalman Filter secured reliable and steady visual information. • The detector network achieved a Precision value of 93 %, a Recall value of 82 %, and a mAP value of 80.18 %.	[120]
	Convolutional Neural Network (AlexNet)	Target detection	3D positioning	IBVS	RGB camera	• The UAV could detect the wind turbine, approach it, and stop at a safe distance autonomously.	[127]
	Convolutional Neural Network	Estimating target position	Aerial manipulation	PBVS	RGB-D camera	• CNN significantly increased the speed of pose estimation. • The system performed well on low light and on light conditions. • The system successfully grabbed the required object. • A RMSE value of less than 0.2 m was achieved.	[76]
	Convolutional Neural Network (YOLOv2)	Estimating target position	3D positioning	PBVS	Stereo vision depth camera	• The system estimated the position of the inspected wind turbine successfully when the camera is stationary or moving slowly.	[77]
	Convolutional Neural Network	Estimating distance to the target, Control the UAV's orientation	Ground object following	IBVS	RGB camera	• The two deep neural networks achieved high accuracy values.	[78]
	Deep Neural Networks with the modified relay feedback test	Real-time hardware identification and controller parameter tuning	3D positioning	PBVS	RGB, Event camera, Thermal camera	• The system accounted for delay dynamics and tuned the control parameters in real time. • The system showed resilience against external disturbances. • Using the RGB camera the system showed a percentage overshoot of 7.85 % and a Rise Time Error of 13.4 %.	[51]
	Neural network based PID control	PID controller parameter tuning and dealing with nonlinearities and uncertainties	Moving object following	IBVS	RGB-D camera	• The ANN based PID controller adapts to changes in mass and moment of inertia. • The technique achieved lower RMSE and Average absolute deviation values compared to conventional PID	[80]
	Chebyshev neural network	Controlling UAV's translational movements	Autonomous landing	IBVS	RGB camera	• The control technique showed high accuracy and robustness during landing on a ship-mounted moving landing platform.	[79]

(*continued on next page*)

Table 1 (*continued*)

AI type	AI technique	AI role	UAV control task	Visual servoing type	Vision system	Performance Issues	Ref
Fuzzy logic	Fuzzy logic controller	Tuning the controller's gains	Ground object following	IBVS	RGB camera	• The system guaranteed the visibility of the ground platform despite the various disturbances. • The proposed controller showed improved results compared to conventional PD controller.	[95]
	Fuzzy logic controller	Controlling UAV's translational and yaw movements	Moving object following	IBVS	RGB camera	• The proposed system allowed the UAV to detect the desired object in an angle of 33° and a distance of 8 m.	[128]
	Fuzzy logic controller	Controlling UAV's translational and yaw movements	3D positioning	PBVS	RGB camera	• The proposed controller kept the UAV hovering in the required position and showed better performance compared to the onboard internal controller.	[129]
	Fuzzy logic controller	Controlling UAV's translational and yaw movements	Moving object following	IBVS	RGB camera	• The system showed robustness against lighting variations in the visual systems and any uncontrolled weather changes.	[91]
	Fuzzy logic controller	Controlling UAV's yaw movements	Obstacle avoidance	IBVS	RGB camera	• the system allowed the quadrotor to quickly react to detection of the obstacle and avoid it.	[92]
	Fuzzy logic controller	Controlling UAV's yaw movements	Obstacle avoidance	IBVS	RGB camera	• The controller showed excellent behavior indicated by quick response and the small error.	[94]
	Fuzzy logic controller	Controlling UAV's yaw movements	Obstacle avoidance	IBVS	RGB camera	• The controller showed excellent obstacle avoidance performance indicated by a low error value in the yaw control of around 3°.	[93]
	Fuzzy logic controller	Controlling UAV's altitude	Autonomous landing	PBVS	RGB camera	• The system allowed the UAV to land autonomously even when the camera sees only 30 % of the landing platform.	[130]
	Multi-level fuzzy logic controller	Controlling UAV's translational movements	Autonomous landing	IBVS	RGB camera	• The multi-level fuzzy logic control application results in significantly lower overshoot in the position error.	[55]
	Fuzzy logic based PID controller	Controlling UAV's translational movements	Moving object following	PBVS	RGB camera	• The developed technique allowed for efficiently following the required path. • The implemented adjustable pan-tilt camera is suitable for inspection tasks requiring an adjustable view. • RMSE values of 0.0704 m, 0.1109 m, and 0.0663 m were achieved for the x, y, and z directions.	[131]
Reinforcement learning	Q-learning RL with fuzzy logic	Tuning the controller's gain	Ground object following	IBVS	RGB camera	• The hybrid visual servoing system was able to achieve fast and accurate control of the UAV.	[114]
	Q-learning RL with bootstrap aggregation (bagging) algorithm	Tuning the controller's gains, generating velocity commands	Ground object following	IBVS	RGB camera	• The proposed system showed excellent convergence speed and noise rejection. • Incorporating Q-learning results in excellent stability and fast convergence rate.	[109]
	Truncated Q-learning RL with Simulated Annealing	Estimating the servoing gain, selecting the control actions	3D positioning	IBVS	RGB camera	• The proposed system converges faster and shows lower oscillation next to the goal position compared to other techniques in literature	[112]
	Least-Squares Policy Iteration (LSPI) RL	Map visual features error to velocity commands	Autonomous landing	IBVS	RGB camera	• The proposed method allowed for fast and accurate learning of the landing process	[107]
	Deep Deterministic Policy Gradient DRL with CNN (YOLOv2)	Map target's bounding box errors to linear-velocity command	Moving object following	IBVS	RGB camera	• The system could track the target and keep the detected object in the image frame and within a limited bound around the frame's center.	[108]
	REINFORCE DRL with CNN (YOLO v2)	DRL: predict the required control action, CNN: Predict the target's position in the next frame.	Moving object following	IBVS	RGB camera	• The UAV could efficiently follow the target throughout the mission.	[105]
	Deep Deterministic Policy Gradient DRL	Map visual features error to velocity commands	Ground object following	IBVS	RGB camera	• The system showed exceptional tracking performance compared to other techniques. • The trained control algorithm could be transferred easily from simulation to real world even when using a different UAV than the model in the simulation.	[104]
Hybrid AI Techniques	Fuzzy cognitive maps	Controlling UAV's translational and yaw movements	Ground object following	IBVS	RGB camera	• The controller could provide accurate target following despite the unavailability of height and target's yaw data.	[132]

(continued on next page)

Table 1 (*continued*)

AI type	AI technique	AI role	UAV control task	Visual servoing type	Vision system	Performance Issues	Ref
						• The quadrotor could robustly follow the target and correct for different disturbances.	
	Fuzzy cognitive maps	Controlling UAV's translational and yaw movements	Ground object following	IBVS	RGB camera	• The system could successfully follow the moving ground target.	[133]
	Fuzzy integral-based neural network	UAV system optimization	Moving object following	IBVS/ PBVS	Stereo vision RGB camera	• The visual servoing-based UAV system was optimized. • The optimal system uses IBVS architecture, a PID controller, and a stereo vision camera. • The optimal system performance was verified in a simulation.	[118]
	Fuzzy-based Particle Swarm Optimization	UAV system optimization	Ground object following	IBVS	Stereo vision RGB camera	• The system achieved higher performance after the optimization process. • The optimized system showed lower tracking error and a smoother path.	[117]

used artificial intelligence methods is included. Next, the progress done on applying artificial intelligence-based techniques in visual servoing systems of UAV systems was reviewed and discussed. The AI techniques discussed in the literature include Artificial Neural Networks, Fuzzy logic, Reinforcement learning, and hybrid AI techniques. Visual servoing was applied for several control tasks, including 3D positioning, aerial and ground object following, obstacle avoidance, autonomous landing, and aerial manipulation. The current progress on AI-based visual servoing for autonomous UAVs used the AI techniques for several roles, such as controlling the UAV's movement, tuning the controller's parameters, feature extraction, target position estimation, and system design optimization.

However, some critical research gaps need to be investigated further to improve the performance of such systems. The gaps are summarized as follows:

The currently available progress on AI-based visual servoing of UAVs varies largely in terms of control task, application environment, and implemented hardware. This results in difficulty in comparing the performance of different proposed techniques and, in most cases, quantitative results for comparing different methods are either hard to access or inapplicable. Thus, general benchmarking methods for comparing different techniques under a fixed simulation environment are extremely important. It is essential to establish a standardized set of metrics for performance evaluation. These metrics should not only focus on accuracy and efficiency of the UAV controller but also consider factors like energy consumption, robustness to environmental changes, and operational safety, thereby providing a comprehensive view of the system's capabilities.

For further research and improvement of the system's performance, it is suggested to hybridize the system with other object detection algorithms and deep learning-based computer vision systems to enhance the system's performance. Additionally, incorporating adaptive learning models with AI techniques could further refine the accuracy of UAV navigation and control. These models would allow UAV systems to learn and adapt in real-time to various environmental conditions, which is a critical aspect often overlooked in static simulation environments. Moreover, integrating real-time data analytics with existing AI models would offer immediate system performance feedback, allowing for more agile adjustments. Collaboration with fields like robotics and human-machine interaction can provide innovative approaches to address the existing challenges, which is highly recommended.

In the current literature, publicly available datasets on visual servoing are extremely limited and are insufficient for designing and evaluating the performance of proposed techniques. Most of the current research uses private datasets generated specifically for the required task and are ungeneralizable. The development of visual servoing datasets for training AI techniques for different tasks, such as pose estimation and control action generation, under various control tasks, such as autonomous landing and target following, is required to make research on this topic more approachable. To address this gap, there is a serious need for the creation of publicly available comprehensive datasets that encompass a wide range of scenarios, environmental conditions, and task-specific challenges. Such datasets should include high-quality, annotated data that can help in developing more robust and adaptable visual servoing systems. In addition, standardization of the dataset formats and annotations will enable researchers to easily use and contribute to these resources to further accelerate research in these directions.

The current progress on AI-based visual servoing control focused on applying specific UAV control tasks, such as obstacle avoidance, autonomous landing, and target following. However, the work that discussed the full design of UAV-based missions is very limited. Such designs would include taking-off, reaching the required position, performing the required task, and returning safely to the landing station. Such works will overall improve the practicality and applicability of visual servoing controlled UAV systems for various types of complex tasks. For example, autonomous UAV systems are a promising option for monitoring and inspection missions. However, the design of visual servoing-guided UAV systems for inspection tasks was not discussed much in literature and is highly recommended. The design procedure shows huge potential for a wide variety of mechatronics applications and similar studies should be carried out for other types of UAVs using an experimental setup instead of computer simulations.

Various types of vision sensors, such as depth cameras, event cameras, thermal cameras, and multispectral cameras, can be used in UAV systems performing various tasks. However, the progress in implementing such camera types for producing visual signals for guiding the visual servoing for UAVs is extremely limited and requires further research. Expanding the research in this domain could involve developing models and techniques focusing on the unique characteristics of each sensor type, thereby optimizing their effectiveness in UAV applications. For instance, integrating depth cameras for visual servoing-guided 3D mapping and obstacle avoidance, or utilizing thermal cameras for vision-based search and rescue operations in low-visibility conditions, could significantly enhance the UAV capabilities. Furthermore, there's a need for exploring the potential of fusing data from multiple sensor types to create a more comprehensive and reliable visual feedback system. This multimodal approach could lead to more robust and versatile UAV systems, capable of operating in a wider range of environments and conditions. Additionally, addressing the challenges of sensor calibration, real-time data processing, development of hardware-friendly algorithms, and efficient integration with UAV control systems is crucial for advancing the use of these vision sensors in practical applications.

CRediT authorship contribution statement

Muaz Al Radi: Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing. **Maryam Nooman AlMallahi:** Formal analysis, Writing – original draft, Writing – review & editing. **Ameena Saad Al-Sumaiti:** Data curation, Writing – original draft, Writing – review & editing. **Concetta Semeraro:** Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Mohammad Ali Abdelkareem:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing. **Abdul Ghani Olabi:** Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] A.M. Turing, J. Haugeland, Computing machinery and intelligence, Turing Test Verbal Behav. Hallmark Intell. (1950) 29–56.
[2] G. Von Krogh, Artificial intelligence in organizations: new opportunities for phenomenon-based theorizing, Acad. Manage. Discov. (2018).
[3] Y. Cao, et al., Adversarial sensor attack on lidar-based perception in autonomous driving, in: Proceedings of the 2019 ACM SIGSAC conference on computer and communications security, 2019, pp. 2267–2281.
[4] J. Bedford, J. Farrar, C. Ihekweazu, G. Kang, M. Koopmans, J. Nkengasong, A new twenty-first century science for effective epidemic response, Nature 575 (7781) (2019) 130–136.
[5] S. Manoharan, An improved safety algorithm for artificial intelligence enabled processors in self driving cars, J. Artif. Intell. 1 (02) (2019) 95–104.
[6] V.V. Unhelkar, et al., Human-aware robotic assistant for collaborative assembly: integrating human motion prediction with planning in time, IEEE Robot Autom. Lett. 3 (3) (2018) 2394–2401.
[7] T. Rakha, A. Gorodetsky, Review of unmanned aerial system (UAS) applications in the built environment: towards automated building inspection procedures using drones, Autom. Constr. 93 (2018) 252–264.
[8] S. Basaia, et al., Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks, NeuroImage: Clin. 21 (2019) 101645.
[9] J.M. Brown, et al., Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks, JAMA Ophthalmol. 136 (7) (2018) 803–810.
[10] A. Madani, J.R. Ong, A. Tibrewal, M.R. Mofrad, Deep echocardiography: data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease, NPJ Digit. Med. 1 (1) (2018) 1–11.
[11] J. Grischke, L. Johannsmeier, L. Eich, L. Griga, S. Haddadin, Dentronics: towards robotics and artificial intelligence in dentistry, Dent. Mater. 36 (6) (2020) 765–778.
[12] K.Y. Yap, C.R. Sarimuthu, J.M.-Y. Lim, Artificial intelligence based MPPT techniques for solar power system: a review, J. Modern Power Syst.Clean Energy (2020).
[13] M.A.E. Sattar, A. Al Sumaiti, H. Ali, A.A.Z. Diab, Marine predators algorithm for parameters estimation of photovoltaic modules considering various weather conditions, Neur. Comput. Appl. 33 (18) (2021) 11799–11819.
[14] A.C. Şerban, M.D. Lytras, Artificial intelligence for smart renewable energy sector in europe—smart energy infrastructures for next generation smart cities, IEEE Access 8 (2020) 77364–77377.
[15] Z. Ullah, F. Al-Turjman, L. Mostarda, R. Gagliardi, Applications of artificial intelligence and machine learning in smart cities, Comput. Commun. 154 (2020) 313–323.
[16] Y. Quiñonez, An overview of applications of artificial intelligence using different techniques, algorithms, and tools,, Latin Am. Women Res. Contrib. IT Field (2021) 325–347.
[17] A. Esteva, et al., Deep learning-enabled medical computer vision, NPJ Digit. Med. 4 (1) (2021) 1–9.
[18] V. Kakani, V.H. Nguyen, B.P. Kumar, H. Kim, V.R. Pasupuleti, A critical review on computer vision and artificial intelligence in food industry, J. Agric. Food Res. 2 (2020) 100033.
[19] J. Johnson, Artificial intelligence, drone swarming and escalation risks in future warfare, RUSI J. 165 (2) (2020) 26–36.
[20] A.V. Savkin, H. Huang, A method for optimized deployment of a network of surveillance aerial drones, IEEE Syst. J. 13 (4) (2019) 4474–4477.
[21] E. Barmpounakis, N. Geroliminis, On the new era of urban traffic monitoring with massive drone data: the pNEUMA large-scale field experiment, Transp. Research Part C Emerg. Technol. 111 (2020) 50–71.
[22] V. Puri, A. Nayyar, L. Raja, Agriculture drones: a modern breakthrough in precision agriculture, J. Stat. Manage. Syst. 20 (4) (2017) 507–518.
[23] M.S. Allauddin, G.S. Kiran, G.R. Kiran, G. Srinivas, G.U.R. Mouli, P.V. Prasad, Development of a surveillance system for forest fire detection and monitoring using drones, in: IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2019, pp. 9361–9363.
[24] Y.-A. Chen, et al., ARPilot: designing and investigating AR shooting interfaces on mobile devices for drone videography, in: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, 2018, pp. 1–8.
[25] M. Al Radi, H. Karki, N. Werghi, S. Javed, J. Dias, Vision-based inspection of flare stacks operation using a visual servoing controlled autonomous unmanned aerial vehicle (UAV), in: IECON 2022–48th Annual Conference of the IEEE Industrial Electronics Society, IEEE, 2022, pp. 1–6.
[26] H. Shakhatreh, et al., Unmanned aerial vehicles (UAVs): a survey on civil applications and key research challenges, IEEE Access 7 (2019) 48572–48634.
[27] M. Hassanalian, A. Abdelkefi, Classifications, applications, and design challenges of drones: a review, Progr. Aerospace Sci. 91 (2017) 99–131.
[28] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, M. Debbah, A tutorial on UAVs for wireless networks: applications, challenges, and open problems, IEEE Commun. Surveys Tutor. 21 (3) (2019) 2334–2360.
[29] S. Bang, H. Kim, Context-based information generation for managing UAV-acquired data using image captioning, Autom. Constr. 112 (2020) 103116.
[30] C. Kyrkou, G. Plastiras, T. Theocharides, S.I. Venieris, C. Bouganis, DroNet: efficient convolutional neural network detector for real-time UAV applications, in: 2018 Design, Automation & Test in Europe Conference & Exhibition (DATE), 2018, pp. 967–972.
[31] K. Gopalakrishnan, S.K. Khaitan, A. Choudhary, A. Agrawal, Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection, Constr. Build. Mater. 157 (2017) 322–330, 2017/12/30/.
[32] S.S. Chouhan, U.P. Singh, S. Jain, Applications of computer vision in plant pathology: a survey, Arch. Comput. Methods Eng. 27 (2) (2020) 611–632.
[33] V. Wiley, T. Lucas, Computer vision and image processing: a paper review, Int. J. Artif. Intell. Res. 2 (1) (2018) 29–36.
[34] T. Dudi, R. Singhal, R. Kumar, A. Al-Sumaiti, T.D. Do, Robust shortest path planning for aircraft using bounded region voronoi diagram, in: 2020 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), IEEE, 2020, pp. 1–6.
[35] E. Kakaletsis, et al., Computer vision for autonomous UAV flight safety: an overview and a vision-based safe landing pipeline example, ACM Compu. Surveys (CSUR) 54 (9) (2021) 1–37.
[36] B.H. Sababha, A. Al-mousa, R. Baniyounisse, J. Bdour, Sampling-based unmanned aerial vehicle air traffic integration, path planning, and collision avoidance, Int. J. Adv. Robot Syst. 19 (2) (2022) 17298806221086431.
[37] B. Ayhan, C. Kwan, Y.-B. Um, B. Budavari, J. Larkin, Semi-automated emergency landing site selection approach for UAVs, IEEE Trans. Aerosp. Electron. Syst. 55 (4) (2018) 1892–1906.
[38] N.P. Juan, V.N. Valdecantos, Review of the application of artificial neural networks in ocean engineering, Ocean Eng. 259 (2022) 111947.
[39] C. Lu, S. Li, Z. Lu, Building energy prediction using artificial neural networks: a literature survey, Energy Build. 262 (2022) 111718.
[40] T. Hussain, M. Hussain, H. Al-Aqrabi, T. Alsboui, R. Hill, A review on defect detection of electroluminescence-based photovoltaic cell surface images using computer vision, Energies (Basel) 16 (10) (2023) 4012.
[41] X. Sun, X. Zhu, P. Wang, H. Chen, A review of robot control with visual servoing, in: 2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), IEEE, 2018, pp. 116–121.
[42] F. Chaumette, S. Hutchinson, Visual servo control. II. Advanced approaches [Tutorial], IEEE Robot. Autom. Mag. 14 (1) (2007) 109–118.
[43] J. Haviland, F. Dayoub, P. Corke, Control of the final-phase of closed-loop visual grasping using image-based visual servoing, arXiv preprint (2020) *arXiv:2001.05650*.
[44] B. Thuilot, P. Martinet, L. Cordesses, J. Gallice, Position based visual servoing: keeping the object in the field of vision, in: Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292) 2, IEEE, 2002, pp. 1624–1629.
[45] C. Akinlar, C. Topal, EDLines: a real-time line segment detector with a false detection control, Pattern Recognit. Lett. 32 (13) (2011) 1633–1642.
[46] V.N. Nguyen, R. Jenssen, D. Roverso, LS-Net: fast single-shot line-segment detector, Mach. Vis. Appl. 32 (1) (2021) 1–16.
[47] C. Harris, M. Stephens, A combined corner and edge detector, in: Alvey vision conference 15, Citeseer, 1988, pp. 10–5244.
[48] R. Zecca, D.L. Marks, D.R. Smith, Symphotic design of an edge detector for autonomous navigation, IEEE Access 7 (2019) 144836–144844.

[49] J. Gu, et al., Recent advances in convolutional neural networks, Pattern Recognit. 77 (2018) 354–377.
[50] K. Fathian, J. Jin, S.-G. Wee, D.-H. Lee, Y.-G. Kim, N.R. Gans, Camera relative pose estimation for visual servoing using quaternions, Rob. Auton. Syst. 107 (2018) 45–62.
[51] O.A. Hay, M. Chehadeh, A. Ayyad, M. Wahbah, M. Humais, Y. Zweiri, Unified identification and tuning approach using deep neural networks for visual servoing applications, arXiv preprint (2021) *arXiv:2107.01581*.
[52] F. Chaumette, S. Hutchinson, P. Corke, Visual servoing. Springer Handbook of Robotics, Springer, 2016, pp. 841–866.
[53] F. Chaumette, S. Hutchinson, Visual servo control. I. Basic approaches, IEEE Robot Autom. Mag. 13 (4) (2006) 82–90.
[54] M. Al Radi, H. Karki, N. Werghi, S. Javed, J. Dias, Autonomous inspection of flare stacks using an unmanned aerial system. Unmanned Aerial Vehicles Applications: Challenges and Trends, Springer, 2023, pp. 201–223.
[55] C.-M. Huang, M.-L. Chiang, T.-S. Hung, Visual servoing of a micro quadrotor landing on a ground platform, Int. J. Control Autom. Syst. 15 (6) (2017) 2810–2818.
[56] T.V. Venna, S. Patel, T. Sobh, Application of image-based visual servoing on autonomous drones, in: 2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA), IEEE, 2020, pp. 579–585.
[57] V. Eskov, V. Pyatin, V. Eskov, L. Ilyashenko, The heuristic work of the brain and artificial neural networks, Biophysics (Oxf) 64 (2) (2019) 293–299.
[58] O.I. Abiodun, A. Jantan, A.E. Omolara, K.V. Dada, N.A. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: a survey, Heliyon 4 (11) (2018) e00938.
[59] H.A. Prince, A. Ghosh, M.M.H. Siam, M.A.H. Mamun, AI predicts MHD double-diffusive mixed convection and entropy generation in hybrid-nanofluids for different magnetic field inclination angles by ANN, Int. J. Thermofluids (2023) 100383.
[60] R. Bala, D. Kumar, Classification using ANN: a review, Int. J. Comput. Intell. Res 13 (7) (2017) 1811–1820.
[61] F. Aljuaydi, Z. Khan, S. Islam, Numerical investigations of ion slip and hall effects on Cattaneo-Christov heat and mass fluxes in darcy-forchheimer flow of Casson fluid within a porous medium, utilizing non-fourier double diffusion theories through artificial neural networks ANNs, Int. J. Thermofluids 20 (2023) 100475.
[62] M. Saemi, M. Ahmadi, A.Y. Varjani, Design of neural networks using genetic algorithm for the permeability estimation of the reservoir, J. Petrol. Sci. Eng. 59 (1–2) (2007) 97–105.
[63] M. Adil, R. Ullah, S. Noor, N. Gohar, Effect of number of neurons and layers in an artificial neural network for generalized concrete mix design, Neur. Comput. Appl. (2020) 1–9.
[64] J.M. Alvarez, M. Salzmann, Learning the number of neurons in deep networks, Adv. Neural Inf. Process. Syst. 29 (2016).
[65] T. Vujicic, T. Matijevic, J. Ljucovic, A. Balota, Z. Sevarac, Comparative analysis of methods for determining number of hidden neurons in artificial neural network, in: Central european conference on information and intelligent systems, 2016, p. 219. Faculty of Organization and Informatics Varazdin.
[66] W.H. Bangyal, J. Ahmad, H.T. Rauf, Optimization of neural network using improved bat algorithm for data classification, J. Med. Imaging Health Inform. 9 (4) (2019) 670–681.
[67] A. Mathew, P. Amudha, S. Sivakumari, Deep learning techniques: an overview, in: International Conference on Advanced machine learning technologies and applications, Springer, 2020, pp. 599–608.
[68] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: unified, real-time object detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.
[69] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, S. Zagoruyko, End-to-end object detection with transformers, in: European conference on computer vision, Springer, 2020, pp. 213–229.
[70] U. Ananthakrishnan, N. Akshay, G. Manikutty, R.R. Bhavani, Control of quadrotors using neural networks for precise landing maneuvers. Artificial Intelligence and Evolutionary Computations in Engineering Systems, Springer, 2017, pp. 103–113.
[71] A. Saxena, H. Pandya, G. Kumar, A. Gaud, K.M. Krishna, Exploring convolutional networks for end-to-end visual servoing, in: 2017 IEEE International Conference on Robotics and Automation (ICRA, IEEE, 2017, pp. 3817–3823.
[72] R. Czabanski, M. Jezewski, J. Leski, Introduction to fuzzy systems. Theory and Applications of Ordered Fuzzy Numbers, Springer, Cham, 2017, pp. 23–43.
[73] Y. Harish, H. Pandya, A. Gaud, S. Terupally, S. Shankar, K.M. Krishna, Dfvs: deep flow guided scene agnostic image based visual servoing, in: 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2020, pp. 9000–9006.
[74] C. Collewet, E. Marchand, Photometric visual servoing, IEEE Trans. Robot. 27 (4) (2011) 828–834.
[75] C.-W. Chen, H.-A. Hung, P.-H. Yang, T.-H. Cheng, Visual servoing of a moving target by an unmanned aerial vehicle, Sensors 21 (17) (2021) 5708.
[76] P. Ramon-Soria, B.C. Arrue, A. Ollero, Grasp planning and visual servoing for an outdoors aerial dual manipulator,, Engineering 6 (1) (2020) 77–88.
[77] P. Durdevic, D. Ortiz-Arroyo, A deep neural network sensor for visual servoing in 3D spaces, Sensors 20 (5) (2020) 1437.
[78] M.A. Kassab, A. Maher, F. Elkazzaz, Z. Baochang, UAV target tracking by detection via deep neural networks, in: 2019 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2019, pp. 139–144.
[79] Y. Huang, M. Zhu, Z. Zheng, K.H. Low, Linear velocity-free visual servoing control for unmanned helicopter landing on a ship with visibility constraint, IEEE Trans. Syst. Man Cybern. Syst. (2021).
[80] C. Lopez-Franco, J. Gomez-Avila, A.Y. Alanis, N. Arana-Daniel, C. Villaseñor, Visual servoing for an autonomous hexarotor using a neural network based PID controller, Sensors 17 (8) (2017) 1865.
[81] K. Shihabudheen, G.N. Pillai, Recent advances in neuro-fuzzy system: a survey, Knowl. Based Syst. 152 (2018) 136–162.
[82] A. Sumaiti, S.R. Konda, L. Panwar, V. Gupta, R. Kumar, B.K. Panigrahi, Aggregated demand response scheduling in competitive market considering load behavior through fuzzy intelligence, IEEE Trans. Ind. Appl. 56 (4) (2020) 4236–4247.
[83] K. Salmi, H. Magrez, A. Ziyyat, A novel expert evaluation methodology based on fuzzy logic, Int. J. Emerg. Technol. Learn. 14 (11) (2019).
[84] L.A. Zadeh, R.A. Aliev, Fuzzy Logic Theory and applications: Part I and Part II, World Scientific Publishing, 2018.
[85] J.M. Alonso Moral, C. Castiello, L. Magdalena, C. Mencar, An overview of fuzzy systems,, Explain. Fuzzy Syst. (2021) 25–47.
[86] M. Kumar, L. Misra, G. Shekhar, Survey in fuzzy logic: an introduction, Int. J. Sci. Res. Dev 3 (6) (2015) 822–824.
[87] R. Bělohlávek, J.W. Dauben, G.J. Klir, Fuzzy Logic and mathematics: a Historical Perspective, Oxford University Press, 2017.
[88] J. Krejčí, Fuzzy set theory,, Pairwise Compar. Matrices Fuzzy Extens. (2018) 57–84.
[89] A. Fernandez, F. Herrera, O. Cordon, M.J. del Jesus, F. Marcelloni, Evolutionary fuzzy systems for explainable artificial intelligence: why, when, what for, and where to? IEEE Comput. Intell. Mag. 14 (1) (2019) 69–81.
[90] F. Sabahi, M.R. Akbarzadeh-T, Extended fuzzy logic: sets and systems, IEEE Trans. Fuzzy Syst. 24 (3) (2015) 530–543.
[91] M. Olivares, I. Mondragon, P. Campoy, L.Mejias Alvarez, C. Martinez, Aerial object following using visual fuzzy servoing, in: Proceedings of the 1st Workshop on Research, Development and Education on Unmanned Aerial Systems (RED-UAS 2011), 2011, pp. 61–70. Centro Avanzado de Tecnologas Aeroespaciales (CATEC).
[92] M.A. Olivares-Mendez, L. Mejias, P. Campoy, I. Mellado-Bataller, Quadcopter see and avoid using a fuzzy controller. Uncertainty Modeling in Knowledge Engineering and Decision Making, World Scientific, 2012, pp. 1239–1244.
[93] M. Olivares, L. Mejias Alvarez, P. Campoy, I. Mellado-Bataller, I. Mondragon, Uas see-and-avoid using two different approaches of fuzzy control, in: Proceedings of the 2012 International Conference on Unmanned Aircraft Systems (ICUAS'12), International Conference on Unmanned Aircraft Systems Association, Inc, 2012, pp. 1–9.
[94] M.A. Olivares-Mendez, P. Campoy, I. Mellado-Bataller, L. Mejias, See-and-avoid quadcopter using fuzzy control optimized by cross-entropy, in: 2012 IEEE International Conference on Fuzzy Systems, Ieee, 2012, pp. 1–7.
[95] D.E. Touil, N. Terki, A. Aouina, R. Ajgou, Intelligent image-based-visual servoing for quadrotor air vehicle, in: 2018 International Conference on Communications and Electrical Engineering (ICCEE), IEEE, 2018, pp. 1–7.
[96] G. Singh, A. Anvar, Investigating feasibility of target detection by visual servoing using UAV for oceanic applications, in: 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), IEEE, 2014, pp. 1621–1626.
[97] T. Zhang, H. Mo, Reinforcement learning for robot research: a comprehensive review and open issues, Int. J. Adv. Robot. Syst. 18 (3) (2021) 17298814211007305.
[98] S. Levine, Reinforcement learning and control as probabilistic inference: tutorial and review, arXiv preprint (2018) *arXiv:1805.00909*.
[99] A. OroojlooyJadid, D. Hajinezhad, A review of cooperative multi-agent deep reinforcement learning,, arXiv preprint (2019) *arXiv:1908.03963*.
[100] L. Canese, et al., Multi-agent reinforcement learning: a review of challenges and applications, Appl. Sci. 11 (11) (2021) 4948.
[101] T.W. Killian, S. Daulton, G. Konidaris, F. Doshi-Velez, Robust and efficient transfer learning with hidden parameter markov decision processes, Adv. Neural Inf. Process. Syst. 30 (2017).
[102] A. Wachi, Y. Sui, Safe reinforcement learning in constrained markov decision processes, in: International Conference on Machine Learning, PMLR, 2020, pp. 9797–9806.
[103] O.-M. Pedersen, E. Misimi, F. Chaumette, Grasping unknown objects by coupling deep reinforcement learning, generative adversarial networks, and visual servoing, in: 2020 IEEE international conference on robotics and automation (ICRA), IEEE, 2020, pp. 5655–5662.
[104] C. Sampedro, A. Rodriguez-Ramos, I. Gil, L. Mejias, P. Campoy, Image-based visual servoing controller for multirotor aerial robots using deep reinforcement learning, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2018, pp. 979–986.
[105] M.A. Akhloufi, S. Arola, A. Bonnet, Drones chasing drones: reinforcement learning and deep search area proposal, Drones 3 (3) (2019) 58.
[106] S. Khan, M. Naseer, M. Hayat, S.W. Zamir, F.S. Khan, M. Shah, Transformers in vision: a survey, arXiv preprint (2021) *arXiv:2101.01169*.
[107] M. Shaker, M.N. Smith, S. Yue, T. Duckett, Vision-based landing of a simulated unmanned aerial vehicle with fast reinforcement learning, in: 2010 International Conference on Emerging Security Technologies, IEEE, 2010, pp. 183–188.
[108] C. Shinde, R. Lima, K. Das, Deep reinforcement learning based dynamic object detection and tracking from a moving platform, in: 2019 Sixth Indian Control Conference (ICC), IEEE, 2019, pp. 244–249.

[109] H. Shi, K.-S. Hwang, X. Li, J. Chen, A learning approach to image-based visual servoing with a bagging method of velocity calculations, Inf. Sci. (Ny) 481 (2019) 244–257.

[110] A. Santamaria-Navarro, J. Andrade-Cetto, Uncalibrated image-based visual servoing, in: 2013 IEEE International Conference on Robotics and Automation, IEEE, 2013, pp. 5247–5252.

[111] Y. Kubota, Y. Iwatani, Dependable visual servo control of a small-scale helicopter with a wireless camera, in: 2011 15th International Conference on Advanced Robotics (ICAR), IEEE, 2011, pp. 476–481.

[112] H. Shi, L. Shi, G. Sun, K.-S. Hwang, Adaptive image-based visual servoing for hovering control of quad-rotor, IEEE Trans. Cogn. Dev. Syst. 12 (3) (2019) 417–426.

[113] K. Watanabe, Y. Yoshihata, Y. Iwatani, K. Hashimoto, Image-based visual PID control of a micro helicopter using a stationary camera, Adv. Robot. 22 (2–3) (2008) 381–393.

[114] H. Shi, X. Li, K.-S. Hwang, W. Pan, G. Xu, Decoupled visual servoing with fuzzy Q-learning, IEEE Trans. Industr. Inform. 14 (1) (2016) 241–252.

[115] R. Mahony, V. Kumar, P. Corke, Multirotor aerial vehicles: modeling, estimation, and control of quadrotor, IEEE Robot. Autom. Magaz. 19 (3) (2012) 20–32.

[116] D. Lee, T. Ryan, H.J. Kim, Autonomous landing of a VTOL UAV on a moving platform using image-based visual servoing, in: 2012 IEEE international conference on robotics and automation, IEEE, 2012, pp. 971–976.

[117] A. Mohebbi, S. Achiche, L. Baron, Integrated and concurrent detailed design of a mechatronic quadrotor system using a fuzzy-based particle swarm optimization, Eng. Appl. Artif. Intell. 82 (2019) 192–206.

[118] A. Mohebbi, S. Achiche, L. Baron, Multi-criteria fuzzy decision support for conceptual evaluation in design of mechatronic systems: a quadrotor design case study, Res. Eng. Des. 29 (3) (2018) 329–349.

[119] G. Kumar, H. Pandya, A. Gaud, K.M. Krishna, Pose induction for visual servoing to a novel object instance, in: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS, IEEE, 2017, pp. 2953–2959.

[120] A. Barisic, M. Car, S. Bogdan, Vision-based system for a real-time detection and following of UAV, in: in *2019 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED UAS)*, IEEE, 2019, pp. 156–159.

[121] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: towards real-time object detection with region proposal networks, Adv. Neural Inf. Process. Syst. 28 (2015).

[122] R. Girshick, Fast r-cnn, in: Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440–1448.

[123] K. Fu, D.-P. Fan, G.-P. Ji, Q. Zhao, J. Shen, C. Zhu, Siamese network for RGB-D salient object detection and beyond, IEEE Trans. Pattern Anal. Mach. Intell. (2021).

[124] L. Bertinetto, J. Valmadre, J.F. Henriques, A. Vedaldi, P.H. Torr, Fully-convolutional siamese networks for object tracking, in: European conference on computer vision, Springer, 2016, pp. 850–865.

[125] A. Dosovitskiy, et al., An image is worth 16x16 words: transformers for image recognition at scale, arXiv preprint (2020) *arXiv:2010.11929*.

[126] S. Khan, M. Naseer, M. Hayat, S.W. Zamir, F.S. Khan, M. Shah, Transformers in vision: a survey, ACM Comput. Surveys (CSUR) (2021).

[127] P. Durdevic, D.O. Arroyo, S. Li, Z. Yang, Uav visual servoing navigation in sparsely populated environments, IFAC-PapersOnLine (2022).

[128] M. Algabri, et al., Wireless vision-based fuzzy controllers for moving object tracking using a quadcopter, Int. J. Distrib. Sens. Netw. 13 (4) (2017) 1550147717705549.

[129] C.-H. Pi, V.B. Sheng, S. Cheng, A dual-loop approach with visual servoing fuzzy control for marker navigation quadcopter, in: Proceeding of the 4th IIAE International Conference on Intelligent Systems and Image Processing, 2016.

[130] M.A. Olivares-Méndez, I.F. Mondragón, P. Campoy, C. Martínez, Fuzzy controller for uav-landing task using 3d-position visual estimation, in: International Conference on Fuzzy Systems, Ieee, 2010, pp. 1–8.

[131] A. Wendel, M. Maurer, M. Katusic, and H. Bischof, *Fuzzy visual servoing for micro aerial vehicles*. na, 2012.

[132] A. Amirkhani, M. Shirzadeh, E.I. Papageorgiou, M.R. Mosavi, Visual-based quadrotor control by means of fuzzy cognitive maps, ISA Trans. 60 (2016) 128–142.

[133] A. Amirkhani, M. Shirzadeh, E.I. Papageorgiou, M.R. Mosavi, Fuzzy cognitive map for visual servoing of flying robot, in: 2016 IEEE international conference on fuzzy systems (FUZZ-IEEE), IEEE, 2016, pp. 1371–1376.