Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

A Review on Deep Learning Techniques for Unmanned Aerial Vehicles Navigation

Divyam Prajapati[1], Dr. Shyamal S. Virnodkar[2], Varsha P. Gaikwad[3], Dr. Sangita B. Nemade[4], Mrunali Desai[5], Sanika S. Virnodkar[6]

[1]*Bachelor of Technology, Department of Computer Engineering, K. J. Somaiya Institute of Technology, Sion, Mumbai, India.*
[2]*Associate Professor, Department of Computer Engineering, K. J. Somaiya Institute of Technology, Sion, Mumbai, India.*
[3]*Assistant Professor, Information Technology Department, Government College of Engineering, Aurangabad, Maharashtra, India.*
[4]*Assistant Professor, Computer Engineering Department, Government College of Engineering and Research, Avasari, Pune, India.*
[5]*Assistant Professor, Computer Engineering Department, K. J. Somaiya Institute of Technology, Sion, Mumbai, India.*
[6]*Bachelor of Technology, Department of Chemical Engineering, Bract's Vishwakarma Institute of Technology, Pune, India.*
E-mail: divyam.prajapati@somaiya.edu

Over the past decade, there has been a significant advancement in the research and development of Unmanned Aerial Vehicles (UAVs) and Deep Learning (DL) methods, which, when used in conjunction with path planning algorithms, help UAVs achieve autonomy. UAVs have been flown autonomously using a variety of techniques, but modern deep learning techniques such as Convolutional Neural Networks (CNN), Imitation Learning (IL), and Deep Reinforcement Learning (DRL) are some of the techniques that have recently been favored. Different approaches like path planning, localization, and obstacle avoidance are used. Multiple approaches using CNN, RL, and IL were put out in this area. The majority of the literature on the subject since 2020 is reviewed in this study. The goal is to compare different methods used in the autonomous navigation of a UAV. Different existing algorithms and methods have been discussed and explored in the category of deep learning.

Keywords: Convolutional Neural Networks, Deep Learning, Deep Reinforcement Learning, Flight.

1. Introduction

A type of aircraft capable of autonomous flight in real-time without any human intervention is referred to as a UAV. There are various types of UAVs, such as single rotor, multi-rotor, fixed-wing, etc., depending on the purpose it is used for Mohsan, Khan, Noor, Ullah, & Alsharif (2022). The applications of these UAVs have recently increased in delivery, remote sensing, wireless networking, photography, military, mining, agriculture, 3D mapping of unknown locations (ex: collapsed buildings), and many more Ahmed, Mohanta, Keshari, & Yadav (2022); X. Jiang et al. (2022); Messina & Modica (2020); Park & Choi (2020). Being able to locate itself autonomously and in real-time and navigate its way is one of the critical requirements for a UAV in such applications. This study focuses on exploring the different ways using which UAVs can navigate around. The different techniques for UAV navigation can mainly be divided into 3 main sections: GPS-based navigation, vision-based navigation, and inertial navigation Lu, Xue, Xia, & Zhang (2018). GPS is a popular and traditional option because it is inexpensive and provides a high degree of accuracy, but it comes with a few

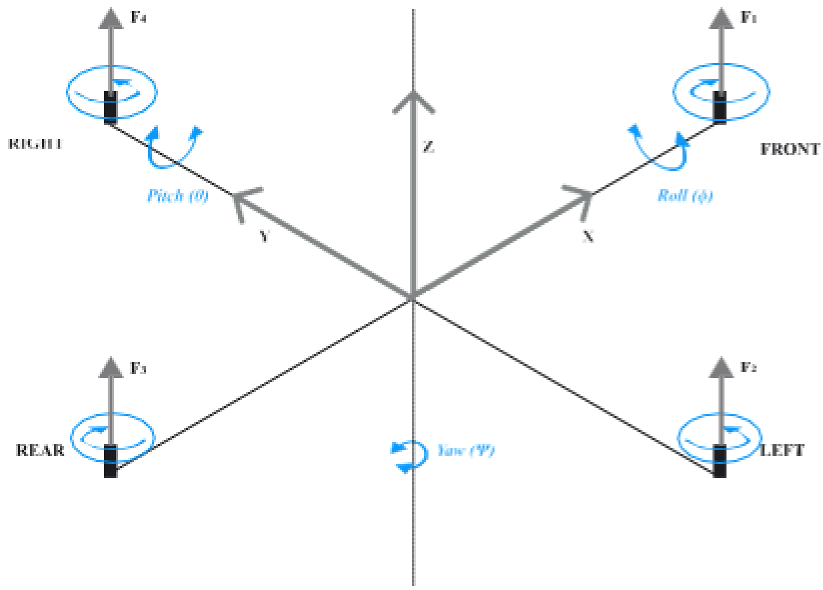

Figure 1. Inertial Navigation Frame denoting the roll, pitch, and yaw axes of a UAV.

It can only tell the UAV where it is at that particular moment, then there is GPS loss, sensor degradation, spoofing of events, etc. This means that GPS-based navigation does not provide "a continuous map" of an area that could be produced using a vision-based navigation system. Hence, integrating vision-based navigation or visual lidar odometry along with GPS is a decent option because it can work in all kinds of environments and doesn't require a lot of computing power Dissanayaka, Wanasinghe, De Silva, Jayasiri, & Mann (2023). Also, it works better in bad weather conditions than GPS does. Thus, Vision-based systems provide high levels of accuracy but they also come with their limitations on what they can and cannot detect. If an obstacle is in their way or thrown at them, they cannot react appropriately in that scenario. To overcome the limitations of Vision-based navigation researchers came up with the concept of Inertial Navigation Shao et al. (2020). It is a very good option because it can fly autonomously without crashing into obstacles while achieving great speed. There are multiple advantages of using Inertial Navigation systems over the previous traditional techniques of navigation, one of the main ones is that it does not require any kind of input from any external or artificial source for getting the input data, as all the sensors are embedded on the UAV itself, so there is no need for connecting the UAV over an unsafe network where there might be the slightest possibility of getting the data altered. Various sensors (gyroscopes, accelerometers, and magnetometers) that provide altitude information for the platform are the core material needed for Inertial Navigation as all the decisions for questions like "where to go next?", "What will be the next optimal move?" is taken based on that information, so they can be imagined as oxygen for UAVs. The main component that is responsible for collecting the data and reporting the changes in altitude, velocity, experienced gravitational forces, and also the changes in orientation of the UAV by using pitch, yaw, and roll (as shown in Figure 1) is given by Inertial Measure Unit (IMU).

IMU mainly comprises Accelerometers, Gyroscopes, and Magnetometers, which are used for reducing the noise from the input data. For further reduction in the noise, there are two approaches first by using the Kalman filter and the other is to use machine learning algorithms like neural networks. Based on the UAV's previous state, measurements, and control inputs, the Kalman filter is a mathematical technique that is used to estimate the state of the UAV, such as the position and velocity of a moving item, on noisy data over time. The UAV's position and orientation can be estimated more precisely by using the Kalman filter in combination with other navigation techniques, such as GPS. For instance, the Kalman filter can be fused with GPS and accelerometer readings, to correct any errors or drift in the accelerometer readings using the more accurate GPS readings for UAV navigation. Another approach is based on machine learning which has been more explored in the previous decade. In contrast to Kalman filters, these algorithms learn from training data, which comprises inputs and outputs, to generalize to previously unknown data. Overall, the trade-off between accuracy, computing efficiency, and robustness as well as the particular requirements of the UAV navigation application will determine whether to use a Kalman filter, machine learning, or a hybrid of both.

This review article will go into detail about the various DL algorithms such as CNN, DRL, and IL. The majority of these findings are from research published in 2020 and later, with an emphasis on navigation, path planning, obstacle avoidance, and localization for UAVs. The review is organized in the following manner: Section 2 is divided into three parts 2.1, 2.2, and 2.3, subsection 2.1 gives a brief idea about various methods used in CNN for navigation, path planning, localization, and obstacle detection and avoidance, whereas subsection 2.2 gives the idea of DRL techniques in a detailed manner, and subsection 2.3 gives details about the IL approach. Next, Section 3 is an extensive comparison between these CNN, DRL, and IL techniques. Finally, in Section 4 conclusions are drawn along with future work.

2. Autonomous Navigation

There is a plethora of frequently used machine learning and deep learning approaches for attaining autonomy some of which include Deep Autoencoders (DAE), Deep Belief Networks

(DBN), CNN, IL, and DRL. Out of these, DBN and DAE are not the first choices that pop up in the researchers' minds, as numerous weights need to be trained using the raw inputs during the training process using supervised learning Ren, Huang, & Gabbar (2022). While on the other hand, CNN reduces the number of weights using the concepts of weight sharing and local receptive field during the training phase. The main drawback of utilizing a CNN is that it may prevent us from obtaining the global maxima during the training phase and instead lead us into the path of local minima. The most popular methodology among these is DRL, which uses the policy gradient algorithm or the Q-Learning approach to obtain the highest rewards depending on the agent's behaviors. Hence, in the context of UAV navigation, machine learning techniques, such as DRL, IL, and CNN are the most prominent choices, and these techniques can be used to perform tasks such as navigation, path planning, obstacle avoidance, and localization. As depicted in Figure No. 2 each branch is a different subsection of this Section 2 and will dive deeper into these concepts individually and see what different architectures are being used by the researchers in these areas.

2.1 Deep Convolution Neural Networks (DCNN)

DCNN is a DL algorithm that typically works with multidimensional input arrays of images and has a multi-layered architecture. All the layers of CNN are assigned learnable biases and weights so that they are easily distinguishable from one another. It overcomes the drawbacks of the Feed-Forward Neural Network (FFNN) by using partially connected layers and the weight-sharing capabilities of CNN, also it can work well on high-resolution images as well. There are various types of pre-trained CNN architectures available that can be used. Some of them are VGGNet, ResNet, Inception, Xception, GoogLeNet, ZFNet, DenseNet, InceptionResNet, MobileNet, and many others K. He, Zhang, Ren, & Sun (2015); Howard et al. (2017); Simonyan & Zisserman (2014); Szegedy, Ioffe, Vanhoucke, & Alemi (2017); Szegedy et al. (2014).

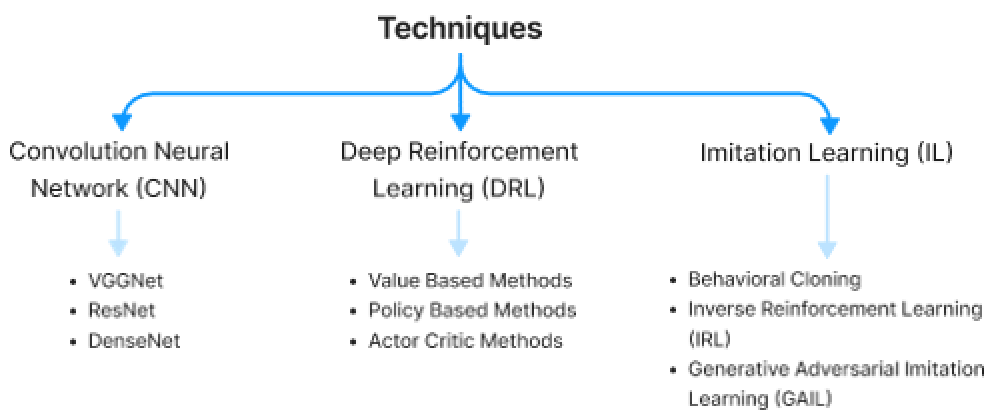

Figure 2. Different techniques under deep learning for autonomous navigation of a UAV.

Let's now go over some UAV navigational methods that have been used in recent years using the CNN approach. Beginning with a novel hybrid aerodynamic quadrotor model called NeuroBEM that mixes blade-element-momentum (BEM) with a neural network and performs better than both learned residual dynamics and first-principles BEM theory Bauersfeld, Kaufmann, Foehn, Sun, & Scaramuzza (2021). Gaussian processes can also improve the quadrotors' nominal dynamics. So the training of Gaussian processes on the data from previously recorded flights to forecast the acceleration error of the nominal model given its current velocity is given in Torrente, Kaufmann, Föhn, & Scaramuzza (2021).

By combining FFNN with the Model Predictive Control (MPC) framework, it is possible to forecast a quadrotor's whole translational dynamics, therefore overcoming FFNN's disadvantages B. Jiang et al. (2022). Software mapping approaches that enable visual navigation are investigated, along with a closed-loop end-to-end DNN-based visual navigation engine for autonomous nano-drones by Palossi et al. (2019).

Visual odometry is the method of estimating the distance traveled by determining equivalent odometry data from successive camera pictures. A deep learning solution for this visual odometry is proposed which validates that CNN is capable of predicting camera motion when it receives only optical flow as input Pandey, Pena, Byrne, & Moloney (2021). Dronet, a CNN that can safely pilot a UAV across city streets, forecasts both the likelihood of a collision and the appropriate steering angle for the unmanned vehicle. This CNN is integrated with a fully linked regression to forecast the steering angles of UAVs Amer, Samy, Shaker, & ElHelw (2021). Some approaches use style transfer between artificial and actual training pictures that may also be done with the help of GANs (Generative Adversarial Networks). A path planner resembling the "A*" method in a heuristic sense, but unlike "A*" it does not call for the storage of frontier nodes in memory is proposed where this planner finds collision-free pathways based on the relative locations of identified items or obstacles Tullu, Endale, Wondosen, & Hwang (2021).

Localization is another key concept that is used by UAVs for navigation purposes. A continual learning Simultaneous Localization and Mapping (SLAM) system is proposed in which SLAM was performed with a drone in a challenging and visually ambiguous warehouse environment W. Chen et al. (2022). Using mapping, a UAV can explore unknown environments and deliver scaled exploration maps. Such a 3D map-ping method in real-time that combines deep learning and SLAM on a budget-friendly UAV is given in Steenbeek & Nex (2022). Suggestions on improving how the UAV images are put together and processed for accurately flying over ungrounded woodland routes are given in MENFOUKH, TOUBA, KHENFRI, & GUETTAL (2020). An automatic obstacle avoidance system for UAVs to fly safely in indoor/outdoor environments is proposed by Dai et al. (2020).

Table 1. Comparisons of a few approaches using convolution neural networks for navigation. These models are compared based on their metrics and limitations.

Ref.	Model	Metrics	Limitations
Chhikara, Tekchan-Dani, Kumar, Chamola, & Guizani (2020)	DCNN-GA MAPE: 1.9587	MSE: 0.0082 MAE: 0.0243	There are several lights-ing difficulties in the images of the dataset because it was generated using the onboard Camera in the corridor. This The Corridor dataset explains why The quadrotor performs better. Indoors than outdoors. The The drone's battery life is about 30 minutes, and maneuvering it in crowded spaces is a challenging

Padhy, Ahmad, Verma, Bakshi, & Sa (2021)	DenseNet-201 MAE (T): 1.79144 MAE (R): 3.6442	MSE (T): 0.12383 MSE (R): 0.0828	The Parrot drones' unpre-predictable and irregular state es-time prevents it from us-ing location feedback in the path planning, which occasionally affects the overall navigation Process. The fact that other factors, such as control and State estimations have an im-. pact on actual UAV flying in real-world situations should be emphasized in addition to Their network prediction.
Pham, Sarabakha, Odnoshyvkin, & Kayacan (2022)	PencilNet MAE Ed: 0.031 MAE Eθ: 0.018 MAE FN: 20.5	MAE Ec: 0.017	By taking into consideration the drifts and deteriorated state estimation performance in dark illumination, which Have a detrimental impact on the perception system's ac-Curacy, the system's performance may be enhanced. Mance may be enhanced.
Safa et al. (2022)	DLSC - QBS	MAE (L): 0.588 MAE (M): 0.1572	Aisles are where I llumina-on conditions abruptly shift, making feature matching un-Predictable. The end of the Aisles is where it performs the Worst.
Arshad et al. (2022)	Drone-STM-RENet	RMSE: 0.48	One drawback is that drones' agile dynamics aren't properly used In contrast to ear-lier CNN-based controllers, it is therefore not possible to Assign the robot a particular Goal to pursue.

There is also a tabular comparison in Table 1 between different CNN techniques for UAV navigation concerning the models used with their metrics and limitations. So, after going through all these different CNN methodologies, using CNN for UAV navigation gives a wide range of challenges — understanding the inner workings of the model, incorporating spatial information into the model, and dealing with overfitting the training data. However, adopting CNN approaches provides us with a competitive advantage since they perform exceptionally well on image recognition tasks, can readily be trained on big datasets, and can learn and extract significant features from the input data. There are a few disadvantages to utilizing CNN as well. First, it may not work well with sequential data or time series data. Second, it may have trouble with tiny datasets or insufficient training data. Finally, training CNN may be computationally expensive.

2.2 Deep Reinforcement Learning (DRL)

DRL is a fusion of RL with DL. It is beneficial for handling dynamically changing environments. It primarily contains two components: an Agent and an Environment. An Agent interacts with an Environment and attempts to influence that environment Arulkumaran, Deisenroth, Brundage, & Bharath (2017). In terms of UAV navigation, the UAV is our Agent and the Environment is the surroundings of the UAV in which it is maneuvering. So as the UAV navigates, it employs a trial-and-error method that results in rewards and penalties. The following are the primary three categories into which DRL techniques are divided:

2.2.1 Value-Based Methods: These methods, which primarily use Q-Learning and SARSA (State Action Reward State Action), such as Deep Q-Networks (DQN), are based on the idea of the value function.

2.2.2 Policy-Based Methods: These methods, which are based on the idea of a policy function, are primarily represented by REINFORCE and Proximal Policy Optimization (PPO), such as Trust Region Policy Optimization (TRPO).

2.2.3 Actor-Critic Methods: These methods integrate approaches that are value-based and policy-based. For instance, Asynchronous Advantage Actor- Critic(A3C) teaches both the actor and critic roles concurrently.

One of the key objectives of this deep neural network is the development of adaptive systems that are capable of experience-driven learning in the real world. DRL is an independent Markov Decision Process (MDP) based on the mathematical framework for experience-driven learning that enables the modeling of almost any complicated environment H. Jiang, Wang, Yau, & Wan (2020). Figure 3 gives us a more precise idea about the flow of the mathematical framework of MDP used by DRL. The MDP is composed of the following five tuples: $< S, A, R, p, \gamma >$: S — a set of states, A — a set of actions, R — reward function, p — transition function, γ — discounting factor.

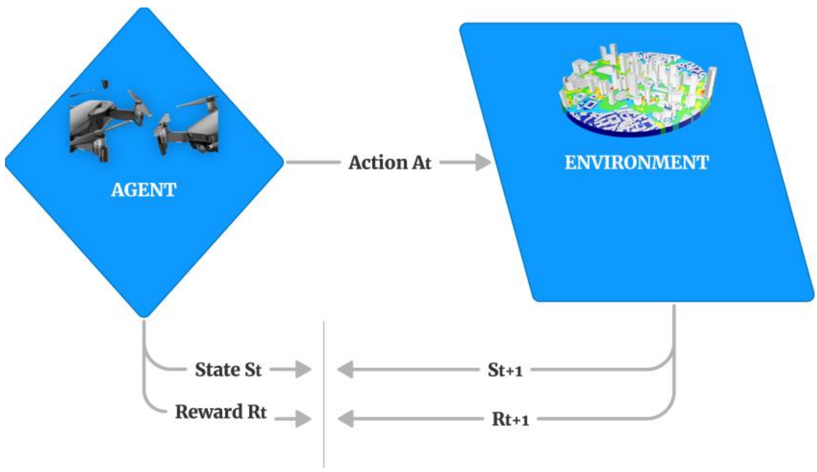

Figure 3. Diagrammatic representation of reinforcement learning structure.

The mathematical framework of DRL typically comprises the following:

a Set of states ‘S’ - The distribution of the starting state is represented as $p(s_0)$.

The final state is represented as S_T.

b A set of actions A.

c Policy $\pi\theta(s_{t+1}|s_t, a_t)$ that maps state(s) and action(a) at a particular time(t) onto a distribution of states at time t + 1 using parameter set Θ.

d A reward function with each transition $R(s_t, a_t, s_{t+1})$.

Some of the DRL techniques that have been implemented in the past few years for UAV navigation consider the issue of UAV navigation in a highly dynamic environment as an MDP, where the problem is broken down into three easier subproblems and successfully solved using the created Layered-RQN. The problem is formulated as a PO-MDP and a deep reinforcement learning method, PPO (Proximal Policy Optimization) is then applied to solve the problem Tong et al. (2021). To produce a digital twin of the location, a high-fidelity rendering engine is combined with 3D scans of the actual environment T. Lee, Mckeever, & Courtney (2022).

A RL PPO-based learning algorithm is proposed by Chikhaoui, Ghazzai, & Massoud (2022), allowing the UAV to navigate in 3D environments. Nowadays DRL is also used for the navigation of a swarm of drones so that they can move from source to destination without colliding with one another. Another approach uses the FFCN algorithm which is trained to control a swarm of drones, which offers innovative fault tolerance by building dynamic patterns and avoiding collisions Raja, Baskar, Dhanasekaran, Nawaz, & Yu (2021). An RL framework for the UAV swarm flocking problem is presented by Yan, Bai, Zheng, & Guo (2020). Localization can also be implemented using DRL. A UAV visual localization approach uses deep learning characteristics from satellite images Hou et al. (2020). The localization is achieved in a Global Navigation Satellite System (GNSS) denied environment. A network called Dragon (DRGN) is proposed that is spatially and temporally aware, with a well-designed GAT-FANET communication structure based on network architecture search, and a memory unit with Gate Recurrent Units (GRU) to support long-term memory Ye, Wang, Chen, Jiang, & Song (2022). Deep Deterministic Policy Gradient (DDPG) and Q-learning (QL) are two RL approaches that are used to teach the UAV to interpret its surroundings and provide efficient scheduling to carry out a data-gathering task. To collect data from geographically dispersed ground sensor nodes throughout a particular geographic region in a predefined amount of time, a self-taught UAV is used as a flying mobile unit. The QL is designed to pick the sequence of nodes to visit to shorten the data-collecting time, while the DDPG is intended to automatically decide the optimal path to travel in an area with obstacles Bouhamed, Ghazzai, Besbes, & Massoud (2020b).

To locate pathways that avoid collisions while yet allowing numerous cellularly linked UAVs to connect to Ground Base Stations (GBSs) in the existence of a dynamic jammer, the UAV route navigation issue is thought to be a sequential choice issue. Thus, a signal-to-interference-plus-noise ratio (SINR) mapping with an offline Temporal Difference (TD) learning technique is suggested for the RL agent X. Wang, Gursoy, Erpek, & Sagduyu (2021). A DRL-QiER model, a DRL solution with a special quantum-inspired experience replay (QiER) framework, is designed to enable the UAV to select the ideal flying path inside each time frame by transforming the navigation issue into an MDP. The suggested DRL-QiER system offers an improved trade-off between sampling priority and variety by using the Grover-iteration-based amplitude amplification method and connecting the significance of an experienced transition to its associated quantum bit (qubit). Numerical data demonstrate the efficacy and superiority of the suggested DRL-QiER solution when compared to different sample baselines. Also, an effort was made to cut down on UAV flight routes, estimated outage length, and the weighted sum of time cost Li & Aghvami (2022).

DDQN (Dueling Deep Q Network) is trained for path coverage planning and data harvesting

Theile, Bayerlein, Nai, Gesbert, & Caccamo (2021). Instead of considering the trained model as a "black box," a new DRL strategy for model explainability is given that allows huge maps to be fed directly into the convolutional layers of the DRL agent L. He, Aouf, & Song (2021). Small quadcopter UAVs can avoid collisions by seeking a way through crowd-spaced 3D objects with the use of a 3D vision cone-based reactive navigation algorithm. This proposed method is evaluated in MATLAB utilizing a range of 3D obstacle scenarios to demonstrate its applicability. It is also compared against two other 3D navigation algorithms Ming & Huang (2021).

There are also some tabular comparisons between several DRL strategies for UAV navigation in Table 2 concerning the trajectory types, models utilized, their success rates, and their limitations. Though there are so many unique DRL methodologies, using DRL for UAV navigation offers several challenges, including dealing with partial observability, balancing exploration and exploitation, overcoming the curse of dimensionality, and factoring existing information into the training process. Additionally, employing DRL has several benefits — It offers an end-to-end learning method, learns from interactions with the dynamic environment, manages high dimensional and noisy observations, and is employed in a range of control tasks. The use of DRL has some drawbacks, including slower computation, high training costs, difficulties with convergence and stability, a potential need for large amounts of training data, and sensitivity to hyperparameter selections.

2.3 Imitation Learning (IL)

IL is a machine learning approach that includes learning from examples or demonstrations of experts rather than through interacting with the environment in a trial-and-error manner.

Table 2. Comparisons of some methods using deep reinforcement learning. Different parameters like the "Trajectory type" are used for training, their success rates, and a few drawbacks of the approaches.

Authors	Method	Static (%)	Dynamic (%)	Notes
Ramezani Dooraki & Lee (2022)	SCAN	98%	82%	1) RGB-D camera noises 2) No obstacle avoidance 3) Not applicable on swarm
D. Wang, Fan, Han, & Pan (2020)	ORCA 3D-S	100%		The two-stage RL policy may oscillate when the scenario changes rapidly due to shifting obstacles.
	ORCA 3D-L	100%		Only considers observation at the current moment.
Y. Chen, González-Prelcic, & Heath (2020)	Object Detection + DQN	Flight distance: 200m	OD+DQN: 78%	
	CAD2RL	Flight distance: 1200m	OD+DQN: 40%	
C. Wang, Wang, Wang, & Zhang (2020)	LwH POfD DDPGfD A3C	CAD2RL: 40%	SenAvo-Pri: 97%, NaivePri: 96%	When baseline decay rate and standard deviation are calibrated, LwH shows resistance.
			SenAvo-Pri: 76%, NaivePri: 41%	
			SenAvo-Pri: 0%, NaivePri: 0%	

Imitative learning may be applied to UAV navigation to help drones travel effectively and securely by taking cues from experienced pilots or pre-planned flight patterns. In this approach, the UAV picks up navigation skills by watching the actions of a human expert or another autonomous agent, so it is well-suited for obstacle avoidance and path planning, but it is restricted to the quality of the observations and the variety of the UAV's surroundings. The different techniques in IL, include:

i Behavioral Cloning (BC): The model learns to imitate the expert's behavior by supervised learning on the expert's demonstration data.

ii Inverse Reinforcement Learning (IRL): The model learns to imitate the expert's behavior by finding a reward function that best explains the expert's actions.

iii Generative Adversarial Imitation Learning (GAIL): The model learns to imitate the expert's behavior by competing against a discriminator that tries to distinguish between the expert's actions and the model's actions.

An end-to-end neural network controller, a controller based on linear regression, and a variational autoencoder (VAE) based controller are all trained using data aggregation in virtual environments. With a greater range of travel in both training and testing, VAE performs better than the other two Wei, Liang, Michelmore, & Kong (2022). Another approach is where an encoder function similar to DroNet which is similar to an equivalent 8-layer Resnet is used which takes the current Image as an input and provides output in the form of a latent compressed vector. Two dense layers and sixtranspose convolutional layers are used for the gate decoder and the image decoder respectively. The goal of training a total of 5 policies was to reduce the expected difference of D distance between our expert and control policy across the observed states Bonatti, Madaan, Vineet, Scherer, & Kapoor (2020).

Another method is Imitation Learning with Indirect Intervention (I3L) which replicates situations in which the learning agent is watched over by an expert (human) and receives assistance via a communication channel. In I3L, an "advisor" interacts with the agent during both training and testing time, in contrast to standard IL approaches where the "teacher" only interacts with the agent during training. This approach combines BC with IL and it is called Behavior Cloning under Intervention (BCUI) Nguyen, Dey, Brockett, & Dolan (2019). BC is an approach used in IL wherein the goal of the model is to accurately replicate the activities of an expert rather than gaining high performance at a specific activity. The approximate BC algorithm initially trains a loss neural network and predicts how distant an action is from the expert's action in a certain state. It then uses the updated loss network to compute the loss for the imitator Lowman, McClellan, & Mullins (2021).

A technique that combines a DRL model and an IL model is called Imitation Augmented Deep Reinforcement Learning (IADRL). The IL model is influenced by GAIL, which is based on GANs, whereas DRL is based on the PPO network with an actor-critic architecture. The latest PPO-based generator is combined with a TRPO-based generator, which also acts as the DRL model's policy Zhang et al. (2020). A Robust Model Based Imitation Learning framework (RMBIL) is provided, in which an end-to-end differentiable tracking control problem developed based on Model-Based Imitation Learning (MBIL) is implemented using the NDI algorithm. To approximate dynamical differences, the Neural ODE model is employed, which backpropagates through a black box ODE solver utilizing the adjoint sensitivity approach to solve the initial-value issue Lin, Li, Zhou, Wang, & Meng (2021).

An alternative approach concentrates on RL-based anomaly identification and rectification. To identify test system anomalies, the corresponding goal weights and purpose of the real-time observed test system are learned. Normal behaviors teach us regular control of attention and objective weight. From the observed behaviors, it is also possible to learn the objective purpose and the equivalent objective function weight. Learning this equivalent weight to the normal for the abnormal system and by resisting the impact of the potential noise anomalies, the suggested algorithm eliminates the anomaly and causes the abnormal actions to have the normal goal Lian et al. (2022).

There are also a few approaches that use IL for controlling swarms which are discussed now. The first technique describes a strategy for coordinating UAV operations in scattered circumstances with insufficient observations; this job is seen as a multi-agent partially observable Markov game. The following UAVs aim to repeatedly mimic the leader UAV’s movements until the desired precision is obtained. The Line Of Sight (LOS) communication range is the main topic of the study. A belief and policy component is included in the proposed Belief Policy Interrelated Imitation (BPII) algorithm B. Yang, Ma, & Xia (2022). The second method takes into account the deployment of many UAVs in a large metropolis to carry out activities related to monitoring urban traffic. A unified machine learning model is trained using the Federated Learning (FL) framework suggested with the help of the leader UAV. An MDP is used to formulate the issue. The objective function is reformulated using the GAIL model with Earth Mover Distance (EMD) to achieve accurate swarm control. To improve the UAV’s imprecise external guidance, a Self-Imitation Learning (SIL) model is employed to take advantage of the distinction between past beneficial experiences and imitation deviations. Mean Squared Error (MSE) is used to reduce the errors that occurred during training B. Yang, Shi, & Xia (2022).

An E2EIL (End-to-end Incremental Learning) lane-keeping and collision-checking model is proposed, which is combined with MPC to achieve real-time control. The basic objective is to learn an E2EIL cost map that is “generalizable”. On top of this, MPC in image space is carried out using a real-time produced agent-view costmap K. Lee, Vlahov, Gibson, Rehg, & Theodorou (2021). This is another method that chooses the optimum course of action to navigate the system using the RL policy. Accurately determining the values of each action under multiple conditions is the main issue here. An expert policy is created that retrains the network to make educated guesses about certain activities. Two output signals are created, the first of which provides the target velocity’s tuning action and the second provides the navigation velocity’s tuning action Hua & Fang (2022).

A policy that maps inputs given by a human expert toward an objective by avoiding obstacles is proposed by Candare & Daguil (2020). CNN (3D), LSTM-RNN, and CNN (2D) are used for the policy representation of the desired controller behavior, each policy was trained at 100 epochs.

A technique that uses an Artificial Neural Network (ANN) consists of one hidden layer (35 neurons) and one output layer (4 neurons) is used to map higher-level features to end effector control values. Various methods like the Monte Carlo DAgger, Sequential DAgger, GNC DAgger, and Moving Window DAgger are applied to design the opti- mal Guidance Navigation Control system (GNC) by comparing with the expert GNC policy Shukla, Keshmiri, & Beckage (2020).

Using IL for UAV navigation presents a variety of difficulties, including developing and gathering expert demonstrations, selecting the best algorithm for a given job, and including exploration and unpredictability in the learned policy. Additionally, using IL offers various advantages, including the ability to use already-existing expert knowledge, shorter training times than DRL techniques, and less risk of overfitting compared to supervised learning techniques. IL has a variety of limitations, such as difficulty with partial observability, the potential inability to learn novel behavior that deviates from the expert, and the need for fine-tuning for practical applications.

3 Comparative Study

UAVs frequently use DRL, CNN, and IL techniques for navigation. But each of these techniques has its pros and cons for its usage, also each has a unique feature which cannot be found in others. Like for instance processing visual data from the UAV cameras allows CNNs to navigate by using image recognition and classification. On the other hand, DRL allows the UAV to make decisions based on reward-based systems and involves learning through trial-and-error interactions with the dynamic complex environment, while the IL approach uses expert demonstration and learns from it. So, the CNNs analyze visual input more quickly and effectively, but DRL makes decisions more flexibly and adjusts to changing situations more effectively, and IL learns and adapts more quickly and accurately. In contrast, CNNs could have trouble adapting to new contexts and need to be retrained, similarly, in the case of IL, it may require more data if the environment's complexity increases, whereas DRL can learn with more experience. The main aspects of UAV navigation include localization, path planning, and obstacle avoidance. To determine a comparative study, we will discuss all three areas of UAV navigation for each approach.

(1) Path Planning – Starting with path planning as it is the most basic and most important component of any autonomous navigation system it is the process of determining the best path between a starting point and an endpoint, finding a course that is both safe and effective. In the case of CNN, the decisions are based on the visual input given by the camera hence in a rapidly changing environment there can be challenges faced. Whereas in the case of IL, it is guided by learning from expert demonstrations, so it works effectively in surroundings that are structured and have a clear path, but it can struggle in environments that are not structured and do not have a well-defined path. As for DRL, it is more effective in terms of adaptability and path planning as it works on trial and error so maneuvering in complex environments is comparatively easy, but the training time of a DRL model is very high.

(2) Obstacle Avoidance – It is the technique of avoiding collisions while the UAV travels to a goal, it is also one of the important aspects of UAV navigation. Let's start with IL, it may be able to handle simple stable obstacles but it struggles when the count of obstacles increases or in diverse environments. DRL techniques can also handle simple obstacles using trial and error, but integrating some real-time-based decision-making may require a lot of training for a greater amount of time. Finally, CNNs are considered to be the best when it comes to collision/ obstacle avoidance as they use visual information as their input, so they can be

trained to avoid obstacles. The only drawback of CNN is that it has to be trained in different scenarios to be able to avoid obstacles effectively in complex environments.

(3) Localization – The process of localization involves figuring out the UAV's position and orientation. Beginning with DRL, using this method UAVs can be taught to estimate the it's location based on sensor inputs and past information. In the case of CNN algorithms, they are adept at establishing the position and orientation from visual data but may suffer if there is very little visual data available or the quality of the visual data is not up to the mark. Whereas IL algorithms have a heavy reliance on expert demonstrations and hence may have difficulty with localization.

Overall, while all three IL, CNN, and DRL techniques can be used for navigation along with path planning, obstacle avoidance, and localization, each method has its challenges depending on the environment and has its advantages and disadvantages. Several different types of CNN and DRL algorithms that are used for UAV navigation are compared in Table 3. The choice of which technique to use depends on the requirements and constraints of the application.

Table 3. Various CNN and DRL algorithms for UAV navigation are compared. The "Input" and "Output" columns list the types of information used as input for the UAV and what is returned as the output from each technique, respectively. The "Real-time" columns show whether each technique has been operated in real-time or not.

Ref.	Technique	Algorithm	Dataset Input	Output	Real-time
Mumuni & Amuzu (2022)	CNN (visual odometry)	Cowan-GGR	KITTI, Camera, State Information	Depth Maps with Object Boundaries	YES
S. Singh et al. (2022)	DRL (Actor-Critic)	CACLA		Surge Angle (75 to -75)	NO
Qiu, Jin, Lv, & Zheng (2022)	Lidar navigation system based on ArUco-LIO.	ArUco	Custom Camera, State Information, and Lidar	Pose in the real-world frame	YES
Zhao, Yang, Zhang, Yan, & Yue (2022)	DRL (Policy-Based)	PPO		Control Signals	NO
Bouhamed, Ghazzai, Besbes, & Massoud (2020a)	DRL (Actor-Critic)	DDPG		Control Signals	NO
Shin, Kang, & Kim (2019)	DRL (Value-Based)	DD-DQN	RGB and Depth Map	Control Signals	NO
Santos, Matos-Carvalho, Tomic, Beko, & Correia (2022)	RNN + LSTM	GTRS	Custom State Information	Control Signals	NO
A. Singh & Jha (2021)	DRL (Actor-Critic)	Safe-MADDPG		Control Signals	NO
Xue & Chen (2022)	DRL (Actor-Critic)	FRSVG (0)		Control Signals	NO
Mughal, Khokhar, & Shahzad (2021)	CNN (visual odometry) + GPS		Custom-labeled template images, Geo-tagged ortho mosaics, GPS	Coordinates: Latitude (ϕ), Longitude (λ)	YES
Yang, Zhang, Liu, & Song (2020)	DRL (Value-Based)	DDQN-PER	Raw pixels from the camera and state information	Control Signals	NO
Abedin, Munir, Tran, Han, & Hong (2020)	DRL (Value-Based)	UAV-BS (DQN with replay memory)	State Information	Control Signals	NO

4 Conclusion And Future Work

The most widely used methods for autonomous UAV navigation are RL and DL. Researchers are looking into ways to make UAV navigation systems more autonomous and less reliant on human input as a result of the growing demand for autonomous UAV navigation. This includes creating algorithms that can learn and adjust to dynamic environments and circumstances without requiring human involvement and reduce the collision probability. A growing demand for cutting-edge navigation systems that can guarantee safe and reliable UAV operations will arise as UAVs become more common in a variety of industries, including agriculture, logistics, and emergency response. In the end, improvements in DL and the combination of several navigation methods to produce more autonomous and adaptable systems will define the future of UAV navigation. These developments will significantly contribute to making UAVs an even bigger part of our daily lives.

The future of UAV navigation is rapidly evolving with advances in DRL, IL, IRL, and CNN. As there were some limitations when using CNN and DRL approaches, there was IL to solve some of its issues. Some new methods have been emerging in the past few years in the area of IL and DRL, some of these approaches have a hybrid combination of these techniques like Hybrid Deep Reinforcement and Imitation Learning (HDRIL), Deep Reinforcement Learning with Expert Demonstrations (DRLED) and Meta-Reinforcement Learning (MetaRL). All these methods combine different features of DRL and IL algorithms for accurate navigation and path planning along with avoiding obstacles and mapping. So, there is no single approach that can solve this problem so all the new techniques are either some evolutionary approaches or some hybrid of the existing approaches which can be used in a complex dynamic environment. Finally, specific techniques used for UAV navigation may vary based on the requirements and constraints of the specific application.

References

1. Abedin, S. F., Munir, M. S., Tran, N. H., Han, Z., & Hong, C. S. (2020). Data freshness and energy-efficient uav navigation optimization: A deep reinforcement learning approach. IEEE Transactions on Intelligent Transportation Systems, 22 (9), 5994–6006.
2. Ahmed, F., Mohanta, J., Keshari, A., & Yadav, P. S. (2022). Recent advances in unmanned aerial vehicles: A review. Arabian Journal for Science and Engineering , 1–22.
3. Amer, K., Samy, M., Shaker, M., & ElHelw, M. (2021). Deep convolutional neural network based autonomous drone navigation. In Thirteenth international conference on machine vision (Vol. 11605, pp. 16–24).
4. Arshad, M. A., Khan, S. H., Qamar, S., Khan, M. W., Murtza, I., Gwak, J., & Khan, A. (2022). Drone navigation using region and edge exploitation-based deep cnn. IEEE Access, 10 , 95441–95450.
5. Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforce- ment learning: A brief survey. IEEE Signal Processing Magazine, 34 (6), 26-38. doi:
6. Bauersfeld, L., Kaufmann, E., Foehn, P., Sun, S., & Scaramuzza, D. (2021). Neurobem: Hybrid aerodynamic quadrotor model. arXiv preprint arXiv:2106.08015 .
7. Bonatti, R., Madaan, R., Vineet, V., Scherer, S., & Kapoor, A. (2020). Learning visuomotor policies for aerial navigation using cross-modal representations. In 2020 ieee/rsj interna- tional conference on intelligent robots and systems (iros) (pp. 1637–1644).

8. Bouhamed, O., Ghazzai, H., Besbes, H., & Massoud, Y. (2020a). Autonomous uav navigation: A ddpg-based deep reinforcement learning approach. In 2020 ieee international symposium on circuits and systems (iscas) (pp. 1–5).
9. Bouhamed, O., Ghazzai, H., Besbes, H., & Massoud, Y. (2020b). A uav-assisted data collection for wireless sensor networks: Autonomous navigation and scheduling. IEEE Access, 8 , 110446–110460.
10. Candare, R. J., & Daguil, R. C. (2020). Deep imitation learning for safe indoor autonomous mi- cro aerial vehicle navigation. In 2020 ieee 12th international conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (hnicem) (p. 1-6). doi:
11. Chen, W., Shang, G., Ji, A., Zhou, C., Wang, X., Xu, C., . . . Hu, K. (2022). An overview on visual slam: From tradition to semantic. Remote Sensing , 14 (13), 3010.
12. Chen, Y., Gonz´alez-Prelcic, N., & Heath, R. W. (2020). Collision-free uav navigation with a monocular camera using deep reinforcement learning. In 2020 ieee 30th international workshop on machine learning for signal processing (mlsp) (pp. 1–6).
13. Chhikara, P., Tekchandani, R., Kumar, N., Chamola, V., & Guizani, M. (2020). Dcnn-ga: A deep neural net architecture for navigation of uav in indoor environment. IEEE Internet of Things Journal , 8 (6), 4448–4460.
14. Chikhaoui, K., Ghazzai, H., & Massoud, Y. (2022). Ppo-based reinforcement learning for uav navigation in urban environments. In 2022 ieee 65th international midwest symposium on circuits and systems (mwscas) (pp. 1–4).
15. Dai, X., Mao, Y., Huang, T., Qin, N., Huang, D., & Li, Y. (2020). Automatic obstacle avoidance of quadrotor uav via cnn-based learning. Neurocomputing , 402 , 346–358.
16. Dissanayaka, D., Wanasinghe, T. R., De Silva, O., Jayasiri, A., & Mann, G. K. (2023). Review of navigation methods for uav-based parcel delivery. IEEE Transactions on Automation Science and Engineering .
17. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition.
18. arXiv. Retrieved from https://arxiv.org/abs/1512.03385 doi:
19. He, L., Aouf, N., & Song, B. (2021). Explainable deep reinforcement learning for uav autonomous path planning. Aerospace science and technology , 118 , 107052.
20. Hou, H., Xu, Q., Lan, C., Lu, W., Zhang, Y., Cui, Z., & Qin, J. (2020). Uav pose estimation in gnss-denied environment assisted by satellite imagery deep learning features. IEEE Access, 9 , 6358–6367.
21. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., . . . Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 .
22. Hua, H., & Fang, Y. (2022). A novel learning-based trajectory generation strategy for a quadrotor. IEEE Transactions on Neural Networks and Learning Systems, 1-12. doi:
23. Jiang, B., Li, B., Zhou, W., Lo, L.-Y., Chen, C.-K., & Wen, C.-Y. (2022). Neural network based model predictive control for a quadrotor uav. Aerospace, 9 (8), 460.
24. Jiang, H., Wang, H., Yau, W.-Y., & Wan, K.-W. (2020). A brief survey: Deep reinforcement learning in mobile robot navigation. In 2020 15th ieee conference on industrial electronics and applications (iciea) (p. 592-597). doi:
25. Jiang, X., Sheng, M., Nan, Z., Chengwen, X., Weidang, L., & Xianbin, W. (2022). Green uav communications for 6g: A survey. Chinese Journal of Aeronautics, 35 (9), 19–34.
26. Lee, K., Vlahov, B., Gibson, J., Rehg, J. M., & Theodorou, E. A. (2021). Approximate inverse reinforcement learning from vision-based imitation learning. In 2021 ieee international conference on robotics and automation (icra) (pp. 10793–10799).
27. Lee, T., Mckeever, S., & Courtney, J. (2022). Generating reality-analogous datasets for autonomous uav navigation using digital twin areas. In 2022 33rd irish signals and systems

conference (issc) (pp. 1–6).
28. Li, Y., & Aghvami, A.-H. (2022). Intelligent uav navigation: A drl-qier solution. In Proc. ieee int. conf. commun.(icc) (pp. 1–6).
29. Lian, B., Kartal, Y., Lewis, F. L., Mikulski, D. G., Hudas, G. R., Wan, Y., & Davoudi, A. (2022). Anomaly detection and correction of optimizing autonomous systems with inverse reinforcement learning. IEEE Transactions on Cybernetics, 1-12. doi:
30. Lin, H., Li, B., Zhou, X., Wang, J., & Meng, M. Q.-H. (2021). No need for interactions: Robust model-based imitation learning using neural ode. In 2021 ieee international conference on robotics and automation (icra) (pp. 11088–11094).
31. Lowman, C. A., McClellan, J. S., & Mullins, G. E. (2021). Imitation learning with approximated behavior cloning loss. In 2021 ieee/rsj international conference on intelligent robots and systems (iros) (pp. 8921–8927).
32. Lu, Y., Xue, Z., Xia, G.-S., & Zhang, L. (2018). A survey on vision-based uav navigation.
33. Geo-spatial information science, 21 (1), 21–32.
34. MENFOUKH, K., TOUBA, M. M., KHENFRI, F., & GUETTAL, L. (2020).
35. Optimized convolutional neural network architecture for uav navigation within unstructured trail. In 2020 1st international conference on communications, control systems and signal processing (ccssp) (pp. 211–214).
36. Messina, G., & Modica, G. (2020). Applications of uav thermal imagery in precision agriculture: State of the art and future research outlook. Remote Sensing , 12 (9). Retrieved from https://www.mdpi.com/2072-4292/12/9/1491 doi:
37. Ming, Z., & Huang, H. (2021). A 3d vision cone based method for collision free navigation of a quadcopter uav among moving obstacles. Drones, 5 (04), 134.
38. Mohsan, S. A. H., Khan, M. A., Noor, F., Ullah, I., & Alsharif, M. H. (2022). Towards the unmanned aerial vehicles (uavs): A comprehensive review. Drones, 6 (6). Retrieved from https://www.mdpi.com/2504-446X/6/6/147 doi:
39. Mughal, M. H., Khokhar, M. J., & Shahzad, M. (2021). Assisting uav localization via deep contextual image matching. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing , 14 , 2445-2457. doi:
40. Mumuni, F., Mumuni, A., & Amuzuvi, C. K. (2022). Deep learning of monocular depth, optical flow and ego-motion with geometric guidance for uav navigation in dynamic environments. Machine Learning with Applications, 10 , 100416.
41. Nguyen, K., Dey, D., Brockett, C., & Dolan, B. (2019, June). Vision-based navigation with language-based assistance via imitation learning with indirect intervention. In Proceedings of the ieee/cvf conference on computer vision and pattern recognition (cvpr).
42. Padhy, R. P., Ahmad, S., Verma, S., Bakshi, S., & Sa, P. K. (2021). Localization of unmanned aerial vehicles in corridor environments using deep learning. In 2020 25th international conference on pattern recognition (icpr) (pp. 9423–9428).
43. Palossi, D., Loquercio, A., Conti, F., Flamand, E., Scaramuzza, D., & Benini, L. (2019). A 64-mw dnn-based visual navigation engine for autonomous nano-drones. IEEE Internet of Things Journal , 6 (5), 8357–8371.
44. Pandey, T., Pena, D., Byrne, J., & Moloney, D. (2021). Leveraging deep learning for visual odometry using optical flow. Sensors, 21 (4), 1313.
45. Park, S., & Choi, Y. (2020). Applications of unmanned aerial vehicles in min- ing from exploration to reclamation: A review. Minerals, 10 (8). Retrieved from https://www.mdpi.com/2075-163X/10/8/663 doi:
46. Pham, H. X., Sarabakha, A., Odnoshyvkin, M., & Kayacan, E. (2022). Pencilnet: Zero-shot sim-to-real transfer learning for robust gate perception in autonomous drone racing. IEEE Robotics and Automation Letters, 7 (4), 11847–11854.
47. Qiu, Z., Lin, D., Jin, R., Lv, J., & Zheng, Z. (2022). A global aruco-based lidar navigation

system for uav navigation in gnss-denied environments. Aerospace, 9 (8), 456.

48. Raja, G., Baskar, Y., Dhanasekaran, P., Nawaz, R., & Yu, K. (2021). An efficient formation control mechanism for multi-uav navigation in remote surveillance. In 2021 ieee globecom workshops (gc wkshps) (pp. 1–6).
49. Ramezani Dooraki, A., & Lee, D.-J. (2022). A multi-objective reinforcement learning based controller for autonomous navigation in challenging environments. Machines, 10 (7), 500.
50. Ren, J., Huang, R., & Gabbar, H. (2022, 10). The current trends of deep learning in autonomous vehicles: A review. Journal of Engineering Research and Sciences, 1 , 56-68. doi:
51. Safa, A., Verbelen, T., Ocket, I., Bourdoux, A., Sahli, H., Catthoor, F., & Gielen, G. (2022). Learning to slam on the fly in unknown environments: A continual learning approach for drones in visually ambiguous scenes. arXiv preprint arXiv:2208.12997 .
52. Santos, R., Matos-Carvalho, J. P., Tomic, S., Beko, M., & Correia, S. D. (2022). Applying deep neural networks to improve uav navigation in satellite-less environments. In 2022 international young engineers forum (yef-ece) (p. 63-68). doi:
53. Shao, W., Dou, L., Zhao, H., Wang, B., Xi, H., & Yao, W. (2020). A visual/inertial relative navigation method for uav formation. In 2020 chinese control and decision conference (ccdc) (p. 1831-1836). doi:
54. Shin, S.-Y., Kang, Y.-W., & Kim, Y.-G. (2019). Automatic drone navigation in realistic 3d landscapes using deep reinforcement learning. In 2019 6th international conference on control, decision and information technologies (codit) (p. 1072-1077). doi:
55. Shukla, D., Keshmiri, S., & Beckage, N. (2020). Imitation learning for neural network autopilot in fixed-wing unmanned aerial systems. In 2020 international conference on unmanned aircraft systems (icuas) (p. 1508-1517). doi:
56. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv. Retrieved from https://arxiv.org/abs/1409.1556 doi:
57. Singh, A., & Jha, S. S. (2021). Learning safe cooperative policies in autonomous multi-uav navigation. In 2021 ieee 18th india council international conference (indicon) (p. 1-6). doi: Singh, S., Kishore, K., Dalai, S., Irfan, M., Singh, S., Akbar, S. A., . . . Yechangunja, R. (2022). Cacla-based local path planner for drones navigating unknown indoor corridors. IEEE Intelligent Systems, 37 (5), 32–41.
58. Steenbeek, A., & Nex, F. (2022). Cnn-based dense monocular visual slam for real-time uav exploration in emergency conditions. Drones, 6 (3), 79.
59. Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In Thirty-first aaai conference on artificial intelligence.
60. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S. E., Anguelov, D., . . . Rabinovich,
A. (2014). Going deeper with convolutions. CoRR, abs/1409.4842 . Retrieved from http://arxiv.org/abs/1409.4842
61. Theile, M., Bayerlein, H., Nai, R., Gesbert, D., & Caccamo, M. (2021). Uav path planning using global and local map information with deep reinforcement learning. In 2021 20th international conference on advanced robotics (icar) (pp. 539–546).
62. Tong, G., Jiang, N., Biyue, L., Xi, Z., Ya, W., & Wenbo, D. (2021). Uav navigation in high dynamic environments: A deep reinforcement learning approach. Chinese Journal of Aeronautics, 34 (2), 479–489.
63. Torrente, G., Kaufmann, E., F¨ohn, P., & Scaramuzza, D. (2021). Data-driven mpc for quadrotors. IEEE Robotics and Automation Letters, 6 (2), 3769–3776.
64. Tullu, A., Endale, B., Wondosen, A., & Hwang, H.-Y. (2021). Machine learning approach to real-time 3d path planning for autonomous navigation of unmanned aerial vehicle. Applied Sciences, 11 (10), 4706.
65. Wang, C., Wang, J., Wang, J., & Zhang, X. (2020). Deep-reinforcement-learning-based au-

tonomous uav navigation with sparse rewards. IEEE Internet of Things Journal , 7 (7), 6180–6190.

66. Wang, D., Fan, T., Han, T., & Pan, J. (2020). A two-stage reinforcement learning approach for multi-uav collision avoidance under imperfect sensing. IEEE Robotics and Automation Letters, 5 (2), 3098–3105.
67. Wang, X., Gursoy, M. C., Erpek, T., & Sagduyu, Y. E. (2021). Jamming-resilient path planning for multiple uavs via deep reinforcement learning. In 2021 ieee international conference on communications workshops (icc workshops) (pp. 1–6).
68. Wei, P., Liang, R., Michelmore, A., & Kong, Z. (2022). Vision-based 2d navigation of un- manned aerial vehicles in riverine environments with imitation learning. Journal of Intelli- gent & Robotic Systems, 104 (3), 47.
69. Xue, Y., & Chen, W. (2022). A uav navigation approach based on deep reinforcement learning in large cluttered 3d environments. IEEE Transactions on Vehicular Technology, 1-14. doi:
70. Yan, P., Bai, C., Zheng, H., & Guo, J. (2020). Flocking control of uav swarms with deep reinforcement leaming approach. In 2020 3rd international conference on unmanned systems (icus) (pp. 592–599).
71. Yang, B., Ma, C., & Xia, X. (2022). An interrelated imitation learning method for heteroge- neous drone swarm coordination. IEEE Transactions on Emerging Topics in Computing, 10 (4), 1704–1716.
72. Yang, B., Shi, H., & Xia, X. (2022). Federated imitation learning for uav swarm coordination in urban traffic monitoring. IEEE Transactions on Industrial Informatics.
73. Yang, Y., Zhang, K., Liu, D., & Song, H. (2020). Autonomous uav navigation in dynamic environments with double deep q-networks. In 2020 aiaa/ieee 39th digital avionics systems conference (dasc) (p. 1-7). doi:
74. Ye, Z., Wang, K., Chen, Y., Jiang, X., & Song, G. (2022). Multi-uav navigation for partially observable communication coverage by graph reinforcement learning. IEEE Transactions on Mobile Computing .
75. Zhang, J., Yu, Z., Mao, S., Periaswamy, S. C. G., Patton, J., & Xia, X. (2020). Iadrl: Imitation augmented deep reinforcement learning enabled ugv-uav coalition for tasking in complex environments. IEEE Access, 8 , 102335-102347. doi:
76. Zhao, X., Yang, R., Zhang, Y., Yan, M., & Yue, L. (2022). Deep reinforcement learning for intelligent dual-uav reconnaissance mission planning. Electronics, 11 (13), 2031.