Drone for Intelligent Traffic Monitoring: Current Status and Future Trends

Hongbin Cao, Zongkun Wu, and Wenshuai Yu

Abstract With a large number of observation objects and diverse scene environments, limited fixed monitoring points and views make the traffic monitoring a complex and difficult task. UAVs, as mobile and agile vehicles, provide an aid for the dynamic implementation of traffic monitoring. Similar to many other application areas, the development of intelligence is mainly driven by deep learning. This paper reviews how UAVs can use deep learning methods for dynamic traffic monitoring. In particular, for detection and recognition, which fall under the umbrella of computer vision techniques, several high-performance methods are brought together, typically represented by the Yolo family of algorithms. For navigation and localization, deep learning methods can be combined into planning and scheduling as an overall control scheme for monitoring a swarm of drones. As for fusion and analysis, the detection results will be abstracted into real-time traffic flow states. It will also be integrated with the weather, time, and other related information, and heterogeneous data processing will be performed using deep learning methods to achieve intelligent analysis and decision-making. In addition, the UAV simulation platform is introduced to address the problem of insufficient actual training data. The types, performance, and typical application cases of traffic monitoring drones are also summarized and explained. With the continuous development of technology, the three aspects involved in dynamic traffic monitoring will form a more functional and cohesive closed-loop system with UAVs as the driving platform and intelligent processing algorithms as the core technology.

Keywords Drone · Traffic monitoring · Deep learning · Intelligent system

---

H. Cao · Z. Wu · W. Yu (✉)
College of Civil and Transportation Engineering, Shenzhen University, No. 3688, Nanhai Avenue, Nanshan District, Shenzhen 518060, China
e-mail: ywsh@szu.edu.cn

W. Yu
Guangdong Laboratory of Artificial Intelligence and Digital Economy (Shenzhen), Kerun Building, Guangming District, Shenzhen 518060, China

S. Li (ed.), *Computational and Experimental Simulations in Engineering*, Mechanisms and Machine Science 146, https://doi.org/10.1007/978-3-031-44947-5_88

1 Introduction

Most current traffic monitoring methods use fixed sensor devices such as fixed cameras, induction loops, or radar sensors for real-time monitoring of traffic conditions. But this faces the problems of damage not easy to repair, high maintenance costs, and, most importantly, the traffic information collected by these fixed equipment sensor devices is only for the specific area of installation, which is a passive collection method. The monitoring information contained in this way cannot accurately find the root cause of traffic congestion.

Uncrewed aerial vehicles (UAVs), also called drones, are used in various applications due to their mobility, flexibility, and low cost [1]. It is this flexibility and mobility that makes them particularly suitable for applications in the field of transportation, which have been implemented in various ways, such as traffic monitoring and surveillance, accident monitoring, and road safety [1–3].

This paper divides UAV-based traffic monitoring technology into three main aspects: navigation and localization, detection and identification, and fusion and analysis. Intelligent traffic monitoring has been extensively developed in recent years, and its intelligence is mainly fused with deep learning. Especially for detection and recognition, with the tremendous success of the AlexNet deep neural network proposed by Krizhevsky [4] in 2012 for image classification tasks, deep learning based on a convolutional neural network (CNN) has rapidly emerged. CNN-based architectures in target detection have been developed to different degrees in terms of accuracy and speed, which is impressive. R-CNN and YOLO series are computer vision's commonly used target detection algorithms. Detection and recognition are vital parts of implementing dynamic traffic monitoring. Traffic targets detected by UAVs can be well used for navigation and positioning of UAVs to achieve accurate monitoring of specific areas; picture data or video data of traffic conditions collected by UAVs can be transformed into real-time traffic flow status and prediction of future traffic flow through data fusion and analysis. Regarding navigation and positioning, the current collection of traffic information mainly relies on fixed sensor devices installed on the road to collect. Still, this method contains traffic information only for the specific area where the device is installed. Given this, using UAVs with high mobility, flexible type, and high altitude overhead characteristics for traffic monitoring is more effective than traditional ground monitoring methods. Then navigation and positioning play a non-negligible role, and most UAVs use GPS for navigation and positioning. The vigorous development of deep learning methods in recent years has made it possible to fuse them with the onboard cameras carried by UAVs to use vision for navigation and localization. Especially for traffic monitoring, if a UAV detects a vehicle blockage, it can use imagination to locate and gather to the blocked area for a better view of the traffic conditions on the ground. In terms of fusion and analysis, with the deployment of intelligent transportation systems and the development of sensor technologies in recent years, UAVs are gradually being used in traffic monitoring, and it has become possible to collect multi-source spatiotemporal traffic data from different platforms. These spatiotemporal data provide data support for traffic prediction, and how to analyze and make decisions intelligently on these data has become the focus of scholars' attention, in which data-driven traffic flow prediction methods are proposed, and the most representative model is the deep learning-based method.

This paper is organized as follows. Chapter 2, Navigation and Positioning, introduces current UAV navigation and positioning methods and future trends. Chapter 3, Detection and Identification, overviews current deep learning-based target detection algorithms and discusses future directions. Chapter 4, Fusion and Analysis, introduces the commonly used data fusion methods and focuses on the current state and future trends of deep learning-based traffic flow prediction analysis. Chapter 5, UAVs, provides a general classification of the types of UAVs and application cases and introduces commonly used UAV simulation platforms, and Chap. 6 concludes.

2 Navigation and Positioning

2.1 Current Status

Traditional traffic monitoring methods only target areas where specific equipment is installed and are reactive monitoring. UAV-based traffic monitoring can reach the target area quickly according to the wishes of traffic managers and is a kind of active monitoring. UAV navigation technology is crucial in advancing the target area by UAV.

Various navigation techniques have been proposed so far, divided into three main categories: satellite, inertial, and vision-based navigation [5, 6] (see Fig. 1).

Satellite navigation relies mainly on GNSS and is received and processed by UAVs to navigate according to pre-planned routes. Still, once the GPS signal is blocked or interfered with, it can significantly lead to inaccurate UAV navigation and positioning. Gyagenda et al. [7] have pointed out that GNSS is susceptible to spoofing,

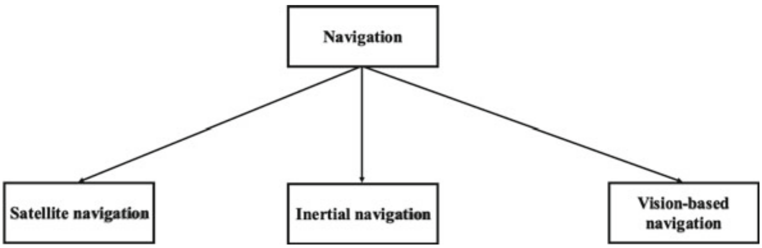

Fig. 1 Navigation techniques

interference, and environmental effects. Vanegas et al. [8] showed that GPS navigation could be limited or unavailable in dense outdoor environments, i.e., GNSS rejection interference.

Inertial navigation, where accelerometer and gyroscope measurements obtain the position and orientation of the device, is not dependent on the external environment and can be used for its navigation in case of loss of GPS, but only for short periods. The error increases over time due to the presence of cumulative mistakes and the presence of noise and drifts in the sensor itself, leading to deviations in navigation [9].

From the above, it is easy to find that both inertial navigation and satellite navigation have certain shortcomings, so navigation techniques based on GPS denial conditions have been widely studied [10–12].

Based vision navigation, UAV vision navigation is mainly a navigation technique that uses the combination of cameras carried by UAVs and computer vision technology to capture the external environment using cameras, obtain visual information, and extract and match features through computer vision algorithms based on pre-stored global or local maps to serve as the basis for UAV flight. Lu et al. [5] reviewed vision-based UAV navigation, pointing out that vision sensors can obtain richer visual information and have a solid anti-interference capability, showing significant advantages in UAV navigation. Zhang et al. [13] formulated the navigation problem as a tracking problem and solved it using particle filters to construct a vision-based UAV navigation method. Mittal et al. [14] proposed a vision-based autonomous navigation and landing system that uses depth information from stereo cameras to sense the surrounding environment and generate low-resolution maps for faster UAV navigation. Kumar et al. [15] proposed a novel approach to achieve UAV navigation using only the camera carried by the UAV and PID controller, applied edge detection and homography techniques to extract key features, and obtained good robustness. Su et al. [16] introduced the image feature extraction algorithm of SIFI and ORB to UAV vision navigation to estimate the pose and position of the camera by matching feature points and images.

Due to the wave of artificial intelligence and deep learning, autonomous UAV navigation and deep understanding are becoming more and more closely integrated. Amer et al. [17] used visual input from the UAV's onboard camera for visual localization to achieve autonomous navigation based on the use of a deep convolutional neural network and a regressor to predict the steering angle of the UAV along a specified path. Pandhy et al. [18] used video information extracted from the UAV's onboard camera by extracting video information, classified it by convolutional neural network and controlled the UAV flight commands by the classification results to achieve autonomous navigation in indoor environments without GPS signals. Chindhe et al. [19] introduced the mainstream UAV navigation methods, pointing out that vision-based navigation techniques have a very high potential. Using a combination of UAV on-board cameras and deep learning can ensure Huang et al. [20] combined deep reinforcement learning with UAV navigation, designed a deep neural network, proposed a

Q-learning learning strategy for UAV navigation, and obtained superior UAV navigation performance. Guo et al. [21] developed a deep reinforcement learning framework for autonomous UAV navigation in complex environments, dividing the navigation task into two subtasks, each solved by a DRL network of LSTM.

Regarding positioning, UAV positioning refers to determining the position and attitude information of the UAV in space, etc. By positioning the UAV position to make more accurate control of UAV navigation further, navigation can, in turn, feedback to positioning, UAV navigation, and setting combined. In the drone application, traffic monitoring is essential; it can reach the desired target monitoring area through accurate positioning, hovering, and shooting to collect traffic data.

The most commonly used positioning methods are based on inertial sensors and GNSS, which are usually used in combination with each other to constrain the cumulative error of INS through GNSS. Still, GNSS faces the problem of signal loss, which makes it much less reliable. Based on such issues, researchers have started to use vision-based positioning to compensate for the shortcomings of GNSS and INS.

There are two main approaches for vision-based UAV localization: relative vision localization and absolute vision localization.

Relative to visual localization, popular methods are visual odometry (VO) and simultaneous localization and map building (SLAM) [22]. Visual odometry is an estimation of the camera motion based on the captured images, and by sampling a specific time interval, the movement of the moving object at each time interval can be estimated. Its motion estimates can be used for navigation purposes and with information from other data sources such as GPS, inertial sensors, etc. [23]. SLAM is a simultaneous localization and map-building technique that updates its bit estimation and map-building from previous motion trajectories and current sensor data. It performs this process to optimize it to obtain more accurate positioning continuously. It has been used more often in the autonomous navigation of UAVs [24, 25].

Absolute visual localization is used for localization by matching the current view acquired by the UAV with the graphical view established from the reference data. The benefit of this approach is that it reduces the effect of cumulative error, i.e., drift, in relative visual localization. Still, it requires a large amount of data and matching data, and with the rise of deep learning, image matching can be well combined in absolute visual localization.

In summary, different UAV navigation technologies are proposed so traffic managers can choose the appropriate navigation technology for other environmental conditions and guide the UAV to the designated target area for traffic monitoring. For example: in outdoor open and good GPS conditions, a combination of inertial navigation and satellite navigation technology can be selected to achieve the UAV flight plan. But in areas where GPS signals are disturbed, such as underground tunnels, culverts, etc., the UAV cannot achieve accurate navigation and positioning, thus affecting traffic monitoring. Vision-based technology makes up for this shortcoming, and more and more scholars are beginning to study vision-based navigation and positioning.

2.2 Future Trends

In terms of navigation and positioning, more and more vision-based technologies will be proposed in the future, and the following trends will emerge:

UAV clustering. In the face of the increasingly complex traffic environment and UAV clustering, research has been continuously brought up in recent years; in the future, a UAV network can be deployed for traffic monitoring through the UAV cluster [21, 26] in at least two neighboring UAVs to communicate and share each other sensing information, improve the sensing range of each UAV and report on traffic conditions, and further improve the overall navigation through clustering capability.

Vision-based autonomous navigation is a key area of advancement in artificial intelligence, especially for Unmanned Aerial Vehicles (UAVs). This technology merges environmental sensing with vision-based navigation, offering a reliable alternative in areas with poor GPS signal. Such advancements are crucial for intelligent traffic monitoring and represent a significant step towards more sophisticated autonomous flight systems.

Image feature extraction. Thanks to the development of the GPU, deep learning algorithms can be well applied to the autonomous navigation of UAVs, and convolutional neural networks can be used to feature extraction of images to achieve more accurate matching of environmental views, thereby improving the accuracy of UAV navigation and positioning.

3 Detection and Identification

3.1 Current Status

Traditional traffic monitoring mainly relies on equipment such as detectors and cameras to detect traffic flow, vehicle speed, traffic operation conditions, etc. However, there are disadvantages such as limited traffic monitoring range, practical only for the area where specific equipment is installed, and requiring a lot of manual processing methods that cannot achieve intelligent monitoring.

The use of UAV airborne cameras combined with advanced target detection algorithms in computer vision can achieve intelligent detection and identification of traffic data, such as traffic flow, traffic composition of motor vehicles, non-motorized vehicles, and pedestrians, detection of traffic accidents and other data, creating great convenience for subsequent analysis and mining of traffic data, while due to the high mobility of UAVs and the characteristics of high altitude overlooking, they can monitor a larger At the same time, due to the increased mobility and high altitude view of the UAV, it can monitor a more extensive range and respond more flexibly to the traffic situation of unexpected situations.

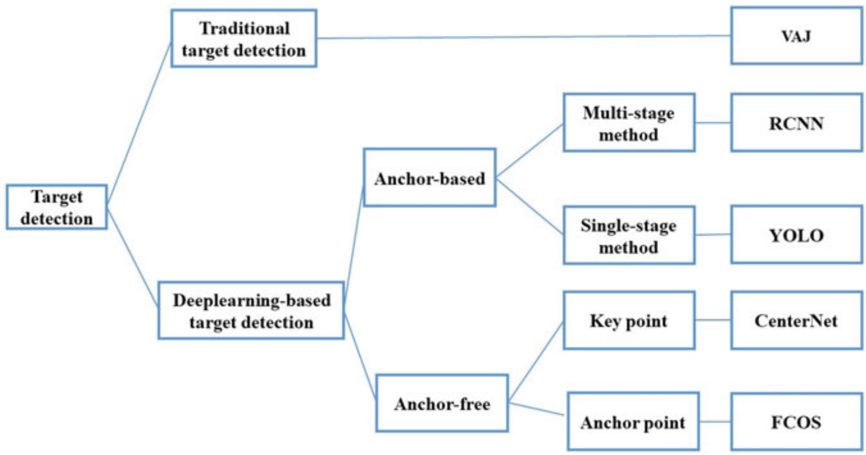

Fig. 2 Target detection method

Target detection is a significant core direction in computer vision, which requires two tasks classification and localization, i.e., determining whether there is an object of interest on a pair of images, and, if so, giving the class and location of all things of interest. The classification task is to determine whether an image contains objects matching the desired category when it is acquired and, if so, to output a series of labels with confidence scores to indicate the probability that the target object is present in the input image. The role of the target localization task is mainly to determine the location and range size of the target category in the input image and output the range frame of the object, i.e., target recognition.

Target detection is mainly divided into traditional target detection algorithms and deep learning-based target detection algorithms (see Fig. 2).

Traditional target detection is performed by manually adding features and classifiers and setting sliding windows of different scales and aspect ratios to traverse the entire image. This “brute force” approach includes all possible locations of the target, but its disadvantages are apparent: no targeting, high computational time and complexity, and redundant windows.

Deep learning-based target detection methods are mainly divided into anchor-based and anchor-free. Among them, anchor-frame-based methods are divided into single-stage and multi-stage forms.

Multi-stage method, which generates a series of candidate frames as samples, and then classifies the models by convolutional neural networks, standard algorithms by the RCNN family (R-CNN, Fast R-CNN, Faster R-CNN, and Mask R-CNN). Zhang et al. [27] used UAV-based video and deep learning methods for traffic analysis to alleviate urban traffic congestion problems using Mask RCNN for recognition and tracking. Wang et al. [28] studied the application of deep learning in UAVs, using an embedded system installed in UAVs for image recognition and target detection recognition, and applied the Faster R-CNNN target detection algorithm to achieve target detection and UAV flight control finally.

Single-stage method, characterized by a fast, direct target frame localization problem, is transformed into a regression problem without generating candidate frames, the most representative algorithms YOLO series [29–32] (YOLOv1, YOLOv2, YOLOv3, and YOLOv4, etc.).YOLO, full name You Only Look Once [29], all roughly follow the following algorithmic ideas: Firstly, the input image is divided into S $\times$ S grids, and each grid is responsible for detecting the targets falling into it, using a non-great suppression algorithm to filter the redundant targets, and finally outputting the borders of the contained marks, the position information of the localization, and all the class confidence levels. Tan et al. [33] take the starting point that the aerial images of UAVs are affected by the flight height and speed, etc., and the photos present smaller targets, and improve the YOLOv4 algorithm and propose the YOLOV4_Drone model to enhance the ability to detect small targets. Luo et al. [34] analyzed the shortcomings of current detection methods based on the fact that UAV aerial images have a large number of minor marks and complex backgrounds and proposed the YOLOD structure based on YOLOv4 to obtain better average accuracy than YOLOv4. Sun et al. [35] improved YOLOv3 to improve the accuracy of small target detection for the characteristics of small size and high density of images collected by UAVs.

Some scholars also compared the representative algorithms of the above two methods in experiments and concluded that YOLO is faster in terms of speed. Gupta et al. [36] proposed using UAV cameras to collect visual information on traffic status using a deep learning object detection model for traffic monitoring by comparing Faster R-CNN, YOLO v3, and v4 for the shortcomings of traditional traffic detection methods. Through experiments, YOLOv4 is considered to have more muscular efficiency and better adaptability. Benjdira et al. [37] proposed using aerial drone images to accurately detect and identify cars for traffic monitoring purposes in the context of drones being increasingly used for traffic monitoring. Analysis and two state-of-the-art target algorithms, Faster R-CNN and YOLOv3, point out that their accuracies are comparable, but the processing speed of YOLOv3 is better than that of Faster R-CNN.

Anchor free uses the idea of semantic segmentation to solve the target detection problem. Compared with anchor-based algorithms, it does not require the artificial design to design anchor-related hyperparameters, which avoids the calculation of IoU between a large number of anchor frames and real frames and saves computational costs. Current anchor-free detection algorithms are mainly divided into two categories: keypoint-based and semantic map-based. The core idea of CornerNet is to represent the target area as a pair of key points; the upper left corner point and the lower right corner point, i.e., to predict the pair of key points while establishing correspondence with each other. The semantic map based on FCOS as an example proposes a pixel-by-pixel fully convolutional target detection network, which combines the single-stage and no-anchor-frame algorithms to predict the offset between the current pixel and the corresponding target centroid, introduces more positive samples, eliminates low-quality prediction results to accelerate model convergence, and has a higher accuracy of the prediction frame.

3.2 Future Trends

In summary, deep learning-based target detection has been widely used under the UAV-based traffic monitoring system because of the great advantage of convolutional neural networks in image processing, which significantly eases the difficulty of traffic data collection for traffic managers and planners. In the future, UAV-based detection and identification technology will present the following trends:

YOLO algorithm. YOLO-based target detection algorithm is widely used due to the fast speed of YOLO, is incredibly applicable to traffic monitoring, can quickly detect and identify traffic monitoring data, which also produces a new technology—YOLO-based drone technology.

Lightweight network. Traffic is a complex composition, and most of the existing research is on the detection and analysis of cars or pedestrians, with the increase in the number of detection will lead to the detection algorithm requiring increased computing power and the hardware resources of the UAV platform is limited, the future can prune the model or design a more lightweight network structure to better deploy to the embedded system of little computing power carried by the UAV.

Small target detection. UAV aerial images will cause such characteristics as small image targets and high density, and the YOLO algorithm has certain shortcomings for small target objects; there are already many scholars using the existing YOLO target detection algorithm to improve it so that it can be better applied to the detection of small target objects, and this trend will continue in the future.

Label assignment. The process of distinguishing between positive and negative samples when training a target detection network to assign and learning targets to the network. This is important in target detection and is the key to improving speed and accuracy. An imbalance between positive and negative samples will make the model produce undesirable results, which is worth investigating in the future and can be applied in target detection networks using dynamic adaptive label assignment methods, such as ATSS and OTA.

Anchor free. Most of the current UAV-based target detection algorithms are based on the anchor frame, no anchor frame can avoid the impact of hyperparameter settings based on the anchor frame, saving computational costs and more balanced positive and negative samples, while maintaining high accuracy, the future can use CenterNet, FCOS, and other algorithms without anchor free for target detection.

4 Data Fusion and Analysis

4.1 Current Status

The most important thing to achieve intelligent traffic monitoring is that thoughtful analysis and decision-making can be reached from the monitored traffic data. With the development of communication technology, the massive use of sensors, and the application of drones, traffic data can be collected and analyzed from multiple sources, presenting characteristics such as heterogeneity, diversity, large scale, and high complexity. This requires data fusion from various sources to achieve complementary goals, reduce the problems of singularity and insufficient data generalization from a single source, and better understand the monitored data.

There are three mainstream data fusion methods: statistical, probabilistic, and artificial intelligence [38].

Probabilistic methods: data fusion based on the Kalman filter is one of the frequently used algorithms, Kalman filter can achieve more accurate information from noisy sensor data, and this method is well suited to handle multi-sensor heterogeneous data [39]. Wang et al. [40] proposed a fusion method for heterogeneous traffic data based on the Kalman filter and Gaussian mixture model, using data from GPS and remote traffic sensors, which cover the accuracy of traffic state estimation. Byon et al. [41] used a single constraint time Kalman filter to estimate the current traffic state for multiple data sources. However, the Kalman filter can only accurately estimate models in linear scenarios [42].

Artificial intelligence methods: neural networks in deep learning are widely used in various prediction models that can be trained to produce predictions by extracting features from the data. It is now commonly used in traffic data fusion and prediction. However, this technique requires higher computing power equipment and data quality requirements. Peng et al. [43] proposed a model based on convolutional neural networks for extracting spatiotemporal features of traffic flows to predict urban passenger flows. Essien et al. [44] used extended and short-memory neural networks combined with traffic and weather data from road networks and employed a data-into-data exo data fusion technique to fuse traffic data with weather data to produce better data quality.

After the data is fused, the most critical thing is to use the data for deeper mining and analysis. Traffic analysis determines the current traffic condition and predicts the future traffic flow trend through historical and real-time traffic data to provide a basis for traffic management and planners to make decisions.

Among them, predicting traffic flow has been a popular topic of research in the field of transportation, and the methods of traffic flow prediction can be broadly divided into traditional methods, traditional machine learning methods, and deep learning-based methods.

Statistical methods: Based on previous observations, time series analysis is used to predict future values but cannot reflect the nonlinear characteristics of traffic flow.

Traditional machine learning methods: can learn nonlinear patterns from traffic data and external factors, can handle high-dimensional data, and capture complex nonlinear relationships, but their application is greatly hindered by artificially designed features based on them.

Deep learning-based approaches: apply neural networks end-to-end for training and show superior performance in traffic prediction. Researchers have proposed several architectures for deep learning, generally employing convolutional neural networks to extract spatial correlations of grid-structured data described by images or videos; graph convolutional networks and attention mechanisms have also been applied to process non-Euclidean spatial data, and LSTM has been used to model temporal dependencies. Ma et al. [45] proposed a deep learning-based daily traffic flow prediction method. The technique deployed a specific convolutional neural network (CNN) to extract daytime and intra-day traffic flow patterns. Wu et al. [46] proposed a deep neural network-based traffic flow forecasting model (DNN-BTF). The model uses convolutional and recurrent neural networks to mine traffic flow's spatial and temporal characteristics. Ma et al. [47] proposed a short-time traffic flow prediction model based on traffic flow temporal analysis combined with an improved extended short-term memory network (LSTM). The improved LSTM network connects the traditional long- and short-term memory network with a bidirectional long- and short-term memory network to overcome significant prediction errors effectively. However, most approaches use a hybrid deep learning framework that combines different deep learning techniques to obtain spatial and temporal correlations of traffic data. Shi et al. [48] proposed a novel attention-based period-time neural network (APTN) to model spatial, short-term, and long-term period dependencies. Zheng et al. [49] proposed a short-time traffic flow prediction model with an attention mechanism as a hybrid deep learning model. Zhang et al. [50] proposed a graph-based temporal attention framework, GTA, which considers spatial and temporal correlations to predict traffic flow based on data collected from multiple sensors. To address the difficulty of existing studies in making short-term traffic flow predictions for isolated points.

4.2 Future Trends

From the above, it is easy to see that deep learning-based approaches are becoming increasingly popular, can well fuse heterogeneous traffic data, and have significant superiority in traffic flow prediction. However, there are still sufficient research challenges, and future development directions are as follows:

Real-time forecasting. The purpose of real-time traffic forecasting is to perform data processing and traffic condition assessment quickly. However, as the scale of traffic data increases, this comes to the point where deeper models with more parameters are required to predict more accurate results, which leads to an increase in computing time and cannot ensure the requirement of real-time prediction. The proposed lightweight model improves the speed under the guarantee of sure accuracy, so an efficient, light neural network can be designed to improve the speed and achieve the purpose of real-time prediction.

Data set construction. UAV traffic data for acquisition collection is greatly affected by adverse weather conditions; obtaining clear pictures or video data is complex and non-periodic, making the training sample size smaller than in normal traffic conditions and more challenging to learn. The data determines the upper limit of the model, which can be optimized by some image or video processing techniques to ensure sufficient data.

Model interpretability. It is well known that a neural network is a black box; how each part of the network works and why it can play a powerful predictive performance has been a question worthy of deeper investigation. If its interpretability is somewhat solved, traffic flow prediction will achieve a more accurate forecast for a more extended period, thus promoting the further development of intelligent transportation.

5 Drones

5.1 Types of Drones

Due to the diversity of drones with their different specifications, sizes, shapes, etc., drones can be classified according to flight structure as fixed-wing, rotary-wing, etc. According to the size classification, they can be further divided into micro, light, small, and large drones. Multi-rotor drones are divided into tri-rotor, quad-rotor, hexa-rotor, etc. Quad-rotor drones (Fig. 3) are the most commonly used, with the characteristics of vertical landing, high mobility, simple and stable structure, low cost, and small size.

It is due to these characteristics of quadrotor UAVs that they are widely used in traffic monitoring systems. Barmpounakis et al. [51] applied DJI 7 Spreading Wings S900 hexacopter to extract natural trajectory data from aerial video of intersections and sidewalks, demonstrating the accuracy of UAV data collection. Gao et al. [52] proposed a new method based on quadrotor UAVs. Niu et al. [53] proposed a new vehicle trajectory detection method based on quadrotor UAVs, which can acquire vehicle trajectories and collect data such as vehicle speed for road traffic analysis. Niu et al. [53] proposed an aerial traffic monitoring system based on quadrotor UAVs to overcome the limitations of static cameras for traffic monitoring.

Fig. 3 Quad-rotor drones

In summary, UAVs have become the most promising new tool for collecting road traffic status data and conducting traffic monitoring in areas such as detecting vehicle speeding, detecting vehicles and giving vehicle paths, detecting road traffic accidents, etc.

5.2 Simulation Platform

UAV simulation platforms can help users perform various tests and verifications in a virtual environment to ensure that no one is as effective and safe as expected. In the real world, it often takes a lot of time and money to achieve the desired testing environment, especially for UAVs used in traffic monitoring; using simulation platforms for testing and verification is undoubtedly one of the best ways to accomplish this. For example, they are testing UAVs for autonomous navigation and obstacle avoidance, simulating their reliability and stability in different traffic environments, etc. Here are a few common UAV simulation platforms.

AirSim: an open-source UAV simulation platform developed by Microsoft, which can simulate different types of UAVs and various scenarios, and supports a variety of hardware and software platforms.

Gazebo: Open source simulation platform developed by OSRF (Open Source Robotics Foundation), which can simulate various robots and drones, widely used in ROS, and supports a variety of sensors and controllers.

DJI Simulator: A simulation platform developed by DJI that simulates different types of DJI drones and provides high simulation and accuracy in actual flight.

All the simulation platforms have advantages and disadvantages, and users can choose the right venue according to their needs when facing specific problems.

6 Conclusion

This paper divides the UAV-based traffic monitoring technology into three main aspects: navigation and localization, detection and identification, and fusion and analysis.

In terms of navigation, there are the following future trends:

(1) **Vision-based autonomous navigation**, UAVs using visual information to achieve autonomous navigation has become a trend that does not depend on signal interference, while visual information data has a rich planning path on the characteristics of the visual data and can be well combined with algorithms in deep learning to improve the accuracy of UAV navigation and achieve intelligent autonomous navigation.
(2) **UAV clusters**, through the UAV cluster share sensing information to improve the overall navigation capability.

In terms of detection and recognition, deep learning-based target detection methods are widely used in UAVs to achieve detection and recognition of specific traffic targets, providing for subsequent data fusion and analysis, presenting the following trends in the future:

(1) **Small target detection capability**, UAV aerial images will cause such characteristics as small image targets and high density, there are already many scholars based on the YOLO algorithm to improve, improve the small target detection capability, the future with the widespread use of UAVs, methods to improve the effectiveness of small target detection will be increasingly proposed.
(2) **Label assignment**, which is important in target detection, is the key to improving accuracy and speed, in the future can be used in the target detection network using dynamic adaptive tag assignment methods, such as ATSS and OTA.
(3) **In anchor free algorithm**, no anchor frame can avoid the impact of hyperparameter settings based on the anchor frame, saving computational costs and more balanced positive and negative samples, which can be applied to embedded devices such as UAVs in the future.

In terms of fusion and analysis, the methods in deep learning can well fuse traffic data characterized by heterogeneity, diversity, large scale, and high complexity, and combine with neural networks to achieve predictive analysis of traffic flow for intelligent analysis and decision-making. The following trends are presented in the future:

(1) **Real-time prediction**, in the future, an efficient lightweight neural network can be designed to improve the speed of predicting traffic flow and achieve the purpose of real-time prediction.
(2) **Data set construction**, the data determines the upper limit of the model and can be optimized by some image or video processing techniques to ensure sufficient data.
(3) **Model interpretability**, if we can know how the neural network is so effective in prediction, the traffic flow prediction will achieve a longer period of accurate prediction, thus promoting the further development of intelligent transportation, the study of its interpretation is of practical significance.

Through analyzing the current situation of these three main aspects and future trends, with UAVs as the driving platform and intelligent processing algorithms as the core technology, the three aspects involved in dynamic traffic monitoring will form a more functional and cohesive closed-loop system.

References

1. Gupta, H., Verma, O.P.: Monitoring and surveillance of urban road traffic using low altitude drone images: a deep learning approach. Multimed. Tools Appl. **81**, 19683–19703 (2022). https://doi.org/10.1007/s11042-021-11146-x
2. Zheng, C., Breton, A., Iqbal, W., Sadiq, I., Elsayed, E., Li, K.: Driving-behavior monitoring using an unmanned aircraft system (UAS). In: Duffy, V. (eds.) Digital Human Modeling. Applications in Health, Safety, Ergonomics, and Risk Management: Ergonomics and Health. DHM 2015. Lecture Notes in Computer Science, vol. 9185. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-21070-4_31
3. Barmpounakis, E.N., Vlahogianni, E.I., Golias, J.C.: Unmanned aerial aircraft systems for transportation engineering: current practice and future challenges. Int. J. Transp. Sci. Technol. **5**(3), 111–122 (2016). ISSN 2046-0430.https://doi.org/10.1016/j.ijtst.2017.02.001
4. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. Commun. ACM **60**(6) (2017)
5. Lu, Y., Xue, Z., Xia, G.S., et al.: A survey on vision-based UAV navigation. Geo-Spat. Inf. Sci. **21**(1), 21–32 (2018)
6. Rezwan, S., Choi, W.: Artificial intelligence approaches for UAV navigation: recent advances and future challenges. IEEE Access **10**, 26320–26339 (2022). https://doi.org/10.1109/ACCESS.2022.3157626
7. Gyagenda, N., Hatilima, J.V., Roth, H., et al.: A review of GNSS-independent UAV navigation techniques. Robot. Auton. Syst., 104069 (2022)
8. Vanegas, F., Gaston, K.J., Roberts, J., Gonzalez, F.: A framework for UAV navigation and exploration in GPS-denied environments. In: 2019 IEEE Aerospace Conference, Big Sky, MT, USA, pp. 1–6 (2019). https://doi.org/10.1109/AERO.2019.8741612

9. Woodman, O.J.: An Introduction to Inertial Navigation. University of Cambridge, Computer Laboratory (2007)
10. Schnipke, E., Reidling, S., Meiring, J., et al.: Autonomous navigation of UAV through GPS-denied indoor environment with obstacles, p. 0715. AIAA Infotech@ Aerospace (2015)
11. Cheng, C., Li, X., Xie, L., et al.: Autonomous dynamic docking of UAV based on UWB-vision in GPS-denied environment. J. Franklin Inst. **359**(7), 2788–2809 (2022)
12. Balamurugan, G., Valarmathi, J., Naidu, V.P.S.: Survey on UAV navigation in GPS denied environments. In: 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), pp. 198–204. IEEE (2016)
13. Zhang, J., Liu, W., Wu, Y.: Novel technique for vision-based UAV navigation. IEEE Trans. Aerosp. Electron. Syst. **47**(4), 2731–2741 (2011)
14. Mittal, M., Mohan, R., Burgard, W., et al.: Vision-based autonomous UAV navigation and landing for urban search and rescue. In: Robotics Research: The 19th International Symposium ISRR, pp. 575–592. Springer International Publishing, Cham (2022)
15. Kumar, J., Kandath, H., Agrawal, P.: Vision based UAV navigation through narrow passages. arXiv preprint arXiv:2303.15803 (2023)
16. Yue-Hua, S., Yuan-Ii, C.: Image feature extraction for vision-based UAV navigation. In: 2018 Chinese Automation Congress (CAC), Xi'an, China, pp. 1130–1134 (2018). https://doi.org/10.1109/CAC.2018.8623264
17. Amer, K., Samy, M., Shaker, M., ElHelw M.: Deep convolutional neural network based autonomous drone navigation. In: Proceedings of the SPIE 11605, Thirteenth International Conference on Machine Vision, 1160503, 4 Jan 2021
18. Padhy, R.P., Verma, S., Ahmad, S., et al.: Deep neural network for autonomous UAV navigation in indoor corridor environments. Proc. Comput. Sci. **133**, 643–650 (2018)
19. Chindhe, B., Ramalingam, A., Chavan, S., et al.: Advances in vision-based UAV manoeuvring techniques. In: International Symposium on Intelligent Informatics: Proceedings of ISI 2022, pp. 455–470. Springer Nature Singapore, Singapore (2023)
20. Huang, H., Yang, Y., Wang, H., Ding, Z., Sari, H., Adachi, F.: Deep reinforcement learning for UAV navigation through massive MIMO technique. IEEE Trans. Veh. Technol. **69**(1), 1117–1121 (2020). https://doi.org/10.1109/TVT.2019.2952549
21. Tong, G.U.O., Jiang, N., Biyue, L.I., et al.: UAV navigation in high dynamic environments: a deep reinforcement learning approach. Chin. J. Aeronaut. **34**(2), 479–489 (2021)
22. Couturier, A., Akhloufi, M.A.: A review on absolute visual localization for UAV. In: Robot. Auton. Syst. **135**, 103666 (2021)
23. Nister, D., Naroditsky, O., Bergen, J.: Visual odometry. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. Washington, DC, USA, p. I (2004). https://doi.org/10.1109/CVPR.2004.1315094
24. Li, R., Liu, J., Zhang, L., Hang, Y.: LIDAR/MEMS IMU integrated navigation (SLAM) method for a small UAV in indoor environments. In: 2014 DGON Inertial Sensors and Systems (ISS), Karlsruhe, Germany, pp. 1–15 (2014). https://doi.org/10.1109/InertialSensors.2014.7049479
25. Schmuck, P., Chli, M.: Multi-UAV collaborative monocular SLAM. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, pp. 3863–3870 (2017). https://doi.org/10.1109/ICRA.2017.7989445
26. Huang, H., Savkin, A.V., Huang, C.: Decentralized autonomous navigation of a UAV network for road traffic monitoring. In: IEEE Trans. Aerosp. Electron. Syst. **57**(4), 2558–2564 (2021)
27. Zhang, H., Liptrott, M., Bessis, N., Cheng, J.: Real-time traffic analysis using deep learning techniques and UAV based video. In: 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Taipei, Taiwan, pp. 1–5 (2019)
28. Wang, C., Zhao, R., Yang, X., et al.: Research of UAV target detection and flight control based on deep learning. In: 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), pp. 170–174. IEEE (2018)
29. Redmon, J., Divvala, S., Girshick, R., et al.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2016)

30. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7263–7271 (2017)
31. Redmon, J., Farhadi, A.: Yolov3: an incremental improvement. arXiv preprint arXiv:1804.02767 (2018)
32. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: Yolov4: optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934 (2020)
33. Tan, L., Lv, X., Lian, X., et al.: YOLOv4_Drone: UAV image target detection based on an improved YOLOv4 algorithm. Comput. Electr. Eng. **93**, 107261 (2021)
34. Luo, X., Wu, Y., Zhao, L.: YOLOD: a target detection method for UAV aerial imagery. Remote Sens. **14**, 3240 (2022)
35. Sun, W., Dai, L., Zhang, X., et al.: RSOD: Real-time small object detection algorithm in UAV-based traffic monitoring. Appl. Intell. **52**, 8448–8463 (2022). https://doi.org/10.1007/s10489-021-02893-3
36. Gupta, H., Verma, O.P.: Monitoring and surveillance of urban road traffic using low altitude drone images: a deep learning approach. In: Multimedia Tools and Applications, pp. 1–21 (2022)
37. Benjdira, B., Khursheed, T., Koubaa, A., et al.: Car detection using unmanned aerial vehicles: comparison between faster R-CNN and yolov3. In: 2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS), pp. 1–6. IEEE (2019)
38. El Faouzi, N.E., Leung, H., Kurian, A.: Data fusion in intelligent transportation systems: progress and challenges—a survey. Inf. Fusion **12**(1), 4–10 (2011)
39. Kashinath, S.A., et al.: Review of data fusion methods for real-time and multi-sensor traffic flow analysis. IEEE Access **9**, 51258–51276 (2021). https://doi.org/10.1109/ACCESS.2021.3069770
40. Wang, C., Zhu, Q., Shan, Z., Xia, Y., Liu, Y.: Fusing heterogeneous traffic data by Kalman Filters and Gaussian mixture models. In: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), Qingdao, pp. 276–281 (2014). https://doi.org/10.1109/ITSC.2014.6957704
41. Byon, Y.J., Shalaby, A., Abdulhai, B., et al.: Traffic condition monitoring with SCAAT Kalman filter-based data fusion in toronto, Canada. KSCE J. Civ. Eng. **23**, 810–820 (2019)
42. Wielitzka, M., Busch, A., Dagen, M., et al.: Unscented Kalman filter for state and parameter estimation invehicle dynamics. In: Kalman Filters-Theory for Advanced Applications, pp. 56–75. InTech, Rijeka (2018)
43. Peng, H., Wang, H., Du, B., Bhuiyan, M.Z.A., Ma, H., Liu, J., et al.: Spatial temporal incidence dynamic graph neural networks for traffic flow forecasting. Inf. Sci. **521**, 277–290 (2020)
44. Essien, A., Petrounias, I., Sampaio, P., Sampaio, S.: Improving urban traffic speed prediction using data source fusion and deep learning. In: 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), Kyoto, Japan, pp. 1–8 (2019). https://doi.org/10.1109/BIGCOMP.2019.8679231
45. Ma, D., Song, X., Li, P.: Daily traffic flow forecasting through a contextual convolutional recurrent neural network modeling inter-and intra-day traffic patterns. IEEE Trans. Intell. Transp. Syst. **22**(5), 2627–2636 (2020)
46. Wu, Y., Tan, H., Qin, L., et al.: A hybrid deep learning based traffic flow prediction method and its understanding. Transp. Res. Part C Emerg. Technol. **90**, 166–180 (2018)
47. Ma, C., Dai, G., Zhou, J.: Short-term traffic flow prediction for urban road sections based on time series analysis and LSTM_BILSTM method. IEEE Trans. Intell. Transp. Syst. (2021)
48. Shi, X., Qi, H., Shen, Y., et al.: A spatial–temporal attention approach for traffic prediction. IEEE Trans. Intell. Transp. Syst. **22**(8), 4909–4918 (2020)
49. Zheng, H., Lin, F., Feng, X., et al.: A hybrid deep learning model with attention-based conv-LSTM networks for short-term traffic flow prediction. IEEE Trans. Intell. Transp. Syst. **22**(11), 6910–6920 (2020)
50. Zhang, S., Guo, Y., Zhao, P., et al.: A graph-based temporal attention framework for multi-sensor traffic flow forecasting. IEEE Trans. Intell. Transp. Syst. (2021)

51. Barmpounakis, E.N., Vlahogianni, E.I., Golias, J.C., et al.: How accurate are small drones for measuring microscopic traffic parameters? Transp. Lett. **11**(6), 332–340 (2019)
52. Gao, H., Kong, S.L., Zhou, S., et al.: Automatic extraction of multi-vehicle trajectory based on traffic videotaping from quadcopter model. In: Applied Mechanics and Materials, vol. 552, pp. 232–239. Trans Tech Publications Ltd. (2014)
53. Niu, H., Gonzalez-Prelcic, N., Heath, R.W.: A UAV-based traffic monitoring system-invited paper. In: 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), pp. 1–5. IEEE (2018)