



Article CA-LSTM: An Improved LSTM Trajectory Prediction Method Based on Infrared UAV Target Detection

Zhaoyang Dang, Bei Sun *, Can Li, Shudong Yuan, Xiaoyue Huang and Zhen Zuo

College of Intelligent Science, National University of Defense Technology, Changsha 410072, China * Correspondence: sunbei08@nudt.edu.cn

Abstract: In order to improve the UAV prevention and control capability in key areas and improve the rapid identification and trajectory prediction accuracy of the ground detection system in anti-UAV early warnings, an improved LSTM trajectory prediction network CA-LSTM (CNN-Attention-LSTM) based on attention enhancement and convolution fusion structure is proposed. Firstly, the native Yolov5 network is improved to enhance its detection ability for small targets of infrared UAVs, and the trajectory of UAVs in image space is constructed. Secondly, the LSTM network and convolutional neural network are integrated to improve the expression ability of the deep features of UAV trajectory data, and at the same time, the attention structure is designed to more comprehensively obtain time series context information, improve the influence on important time series features, and realize coarse–fine-grained feature fusion. Finally, tests were carried out on a homemade UAV infrared detection dataset. The experimental results show that the algorithm proposed in this paper can quickly and accurately identify infrared UAV targets and can achieve more accurate predictions of UAV flight trajectories, which are reduced by 9.43% and 23.81% in terms of *MAPE* and *MSE* indicators compared with the native LSTM network (the smaller the values of these evaluation indexes, the better the prediction results).

Keywords: UAV prevention and control; LSTM; target identification; trajectory prediction



Citation: Dang, Z.; Sun, B.; Li, C.; Yuan, S.; Huang, X.; Zuo, Z. CA-LSTM: An Improved LSTM Trajectory Prediction Method Based on Infrared UAV Target Detection. *Electronics* **2023**, *12*, 4081. https:// doi.org/10.3390/electronics12194081

Academic Editor: Shiho Kim

Received: 1 September 2023 Revised: 22 September 2023 Accepted: 24 September 2023 Published: 28 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

1. Introduction

As a product of the development of unmanned system technology, UAV flight equipment has great application value in civilian and military fields such as geological exploration, regional search, and battlefield reconnaissance. With the popularization and development of unmanned aerial equipment, the application scope of UAVs has gradually expanded from early low-altitude shooting and low-load transportation to high-altitude reconnaissance, signal link jamming, precision strikes, and other complex tasks [1]. The "black flying" behavior of UAVs has seriously threatened public safety and even national security, and the demand for UAV detection and counter-measures is becoming more and more urgent [2].

Due to its low flight trajectory, slow movement speed, small size, and other characteristics, the amount of signal echo data generated by the UAV and the resolution of the time domain signal are extremely low, and the feature feedback information is difficult to capture [3]. Even if a UAV flies at a relatively close distance from the ground monitoring platform, the UAV can easily break through the signal detection area by changing the trajectory with its own flexible flight control ability and achieve covert penetration and rapid penetration [4]. Due to the deep dimension of photoelectric detection image data, rich feature information, and high degree of spatial mapping, UAVs have more advantages in judging the target category. In addition, photoelectric detection combined with trajectory prediction technology can provide accurate positioning data for laser weapon pointing and orientation, and realize the continuous killing of targets, which has great application value [5]. However, due to the influence of the complex low-altitude airspace environment and the limitations of the detection means, in the case of conventional detection methods, the non-cooperative UAV data information obtained is very small [6], and UAV trajectory prediction, as an important part of UAV prevention and control, can make full use of the target historical trajectory information obtained in the target detection stage to predict the location of the target in the next stage, which can provide effective information for further target intention recognition and provide data support for target threat assessment, disturbance intent, and other types of situation recognition [7].

Based on this, we first improved the native Yolov5 network to improve the recognition ability of small UAV targets and better obtain the historical flight trajectory of UAVs. A CNN-LSTM network structure based on the attention mechanism is designed, which improves the expression ability of the deep features of UAV flight trajectory data through convolutional neural networks, more comprehensively obtains time series context through attention modules, improves the influence on important time series features, and realizes coarse–fine-grained feature fusion so as to characterize the time series data more comprehensively. In addition, we also construct a typical UAV infrared detection dataset, including three scenarios: conventional sky background, urban occlusion background, and complex mountain background. Finally, the proposed algorithm is tested on this dataset, and the test results show that the algorithm can achieve more accurate predictions of UAV flight trajectories in photoelectric detection scenarios, provide effective data support for further target intent recognition, and effectively improve UAV prevention and control capabilities in key areas.

2. Related Work

UAV prevention and control technology mainly refers to a series of means such as detection, tracking, prediction, identification, coordination, and disposal of low-altitude airspace targets [8]. As an important part of UAV prevention and control technology, UAV trajectory prediction technology needs to make full use of the information obtained in the target detection and tracking link so as to provide real-time decision support for subsequent identification intentions and collaborative disposal [9].

2.1. UAV Target Recognition Algorithm

In order to solve the problem of environmental clutter and disturbance clutter, the researchers shifted the center of gravity of the target feature acquisition method from active radar echo signals to passive imaging photoelectric signals [10]. Although the detection distance is closely related to the photoelectric acquisition equipment, it can ensure accurate detection of targets in high-resolution images, and the anti-electromagnetic interference ability is strong [11]. The rapid development of deep learning technology also greatly improved the detection, recognition, and positioning performance of small UAV targets [12]. However, because UAVs are typical examples of low, slow, and small targets, their flight altitude is low, their speed is slow, their degree of freedom is high, and their detectable area is small, and there are a large number of birds in the air whose size and flight speed are similar to those of UAVs; these characteristics mean that the traditional aerial threat detection system is no longer applicable in UAV target detection, and it is necessary to develop an anti-UAV system for the corresponding characteristics of UAVs so as to achieve the real-time detection of invasive UAVs [13].

In order to effectively detect UAV targets based on the characteristics of "low, slow, and small" targets, Xue [14] proposed a UAV target recognition method based on a convolutional neural network and trained it on a small sample UAV dataset. The PGAN proposed by Li [15] generates high-resolution feature representations for detecting small targets, making the feature representations of small targets similar to those of large targets, thereby improving the feature representation capabilities of small targets; Cheng [16] uses dilated convolutions instead of conventional neural networks. The traditional convolution in the network expands the receptive field of the network for feature extraction and reduces the number of redundant calculations [17].

Zhao [18] introduced the YOLO target detection algorithm into the field of target detection, improved it on the basis of the YOLOv3 algorithm, and proposed the ST-YOLO algorithm for UAV target detection, which combined the TopHat transform with binary trees for detection of small targets for small target UAVs. However, due to the performance limitation of the YOLOv3 algorithm, the recognition effect of low-altitude UAVs is not ideal [19]. Ma [20] adds a residual network and multi-scale prediction on the basis of YOLOv3, uses the K-means clustering algorithm on the dataset of low-altitude UAVs to obtain the optimal anchor box, and fuses the residual network and the original network on the basis of the original YOLOv3 network to obtain a new O-YOLOv3 network, which makes the network easier to train and has a good recognition effect. Zheng et al. [21] propose an enhanced adaptive feature pyramid network target detection algorithm that alleviates the differences between receptive fields and semantic information at different levels by sharing convolution, and it solves the problem of weakening feature expression in cross-scale feature fusion. Zhao et al. [22] designed a novel C3-PANet neural network structure on the basis of YOLOv7, which, combined with the C3 structure, increased the receptive field, and optimized the size of the anchor frame by using the K-means++ clustering method, and they designed an anchor frame more suitable for the detection of the infrared small targets of UAVs, which improved the detection efficiency on the infrared dataset. Xu et al. [23] proposed a new infrared small target detection algorithm, SSD-ST, on the basis of the SSD network framework, which removes the low-resolution layer in the network and enhances the high-resolution layer, making it more suitable for the detection of infrared small targets.

2.2. UAV Trajectory Prediction Algorithm

Prediction is based on historical moment data for estimating the future moment state of the process because the UAV flight speed is fast and unpredictable, and the urban low-altitude background building occlusion is more serious. So, it is necessary to rely on the prediction method to predict its motion state to provide support for accurate countermeasure prevention and control [24]. Common academic prediction methods can be divided into association rules, Markov models, neural networks, and other methods [25].

The prediction method based on association rules first needs to construct association rules from the data and then perform trajectory prediction using the sequence matching method. The process of mining association rules can be divided into two stages: the first stage is to find all frequent itemsets from the data collection, and the second stage is to generate association rules from these high-frequency project groups [26]. Agrawal first proposed the concept of association rules and gave the corresponding mining algorithm AIS, but due to its limited performance, it has not been widely used on a large scale. In 1994, they proposed the theory of itemset-lattice spaces, and based on the above two theorems, they proposed the famous Apriori algorithm [27]. In subsequent developments, the partition algorithm optimizes the dependence of the Apriori algorithm on frequent database scans; uses hash functions to compress transaction sets; efficiently generates frequent itemsets through hash tables; and uses the characteristics of hash tables to prune the contents of the database, reduce transaction itemsets, and reduce database scanning.

The Markov model calculates the probability of a certain location point moving to other location points by constructing a state-transition matrix, inputting the current position into the constructed matrix, and determining the next position according to the maximum probability so as to obtain the prediction result [28]. The first-order Markov model has only one position-transition probability matrix. In order to obtain more comprehensive information, Yang [29] designed a high-order Markov model that uses the state information of n positions to predict the position of the next moment, improving the prediction accuracy. Qiao [30] proposed a trajectory prediction algorithm, HMTP, based on hidden Markov, which improved the problem of fast and difficult-to-predict speeds of moving objects, so that it could predict the continuous trajectory of objects. This type of research is based on the assumption that the position information at the current moment is only related to the previous moment,

and the obtained solutions are all local optimal solutions. However, high-order Markov calculations are complex and are not suitable for real-time prediction needs.

With the continuous development of neural networks, there are now a variety of neural networks that are used in the trajectory prediction of low, slow, and small UAVs, and they have achieved good results. Tan [31] et al. used a genetic algorithm to optimize the weight and structural parameters of the BP neural network, and they established a maneuvering target prediction model with high prediction accuracy. Zhang [32] proposed a fighter air combat trajectory prediction model based on an Elman neural network. When predicting the trajectory of a UAV, Yang [33] proposed a prediction model based on Bi-LSTM, which inputted parameters such as target pose and distance into the model, outputted the position of the next moment, and achieved the prediction of the UAV flight trajectory.

In summary, a single UAV prevention and control method struggles to meet the UAV prevention and control needs in key areas and faces many difficulties in practical applications. The effective combination of UAV target detection technology and trajectory prediction technology can solve the problem of low-altitude airspace. The complexity of the environment and the limitations of some prevention and control technologies have greatly improved the UAV prevention and control capabilities of the ground system and achieved accurate prediction of target flight trajectories in complex low-altitude backgrounds.

3. System Composition

The trajectory prediction network for UAV prevention and control designed in this paper considers how to more accurately complete the prediction of the UAV flight trajectory at the next moment in an actual setting. The overall implementation framework of the algorithm is shown in Figure 1.



Figure 1. Trajectory prediction network overall implementation framework.

3.1. UAV Target Recognition Method

In order to improve the extraction ability of small target features, we propose an improved object detection network based on Yolov5. As shown in Figure 2, we introduce the Convolutional Attention Module (CBAM) on top of the native Yolov5 network, embedding it at layer 23 of the network, between the Concat module and the C3 module. CBAM consists of two main components: the spatial attention module and the channel attention module, in which the spatial attention structure learns the mask through the residual network to suppress the background information of the original feature map. Channel attention enables the network to use global information to selectively enhance or suppress tasks between channels, generating different weight coefficients to capture the interrelationships between channels. The weights obtained by the attention mechanism are multiplied by the input feature map for adaptive feature refinement, while CBAM, as a lightweight general-purpose module, can be seamlessly integrated into existing CNN architectures with negligible overhead.



Figure 2. UAV object detection network framework.

3.2. UAV Trajectory Prediction Method

The UAV trajectory prediction method used in this paper is mainly composed of convolutional neural network long short-term memory network, and attention mechanism. The mathematical symbols involved are shown in Table 1.

Symbol	Variable/Parameter	
И	Input weight	
b	Bias term	
f_t	Forgetting gate parameter	
g_t	Nonlinear mapping of hyperbolic tangent functions	
e_{tj}	Alignment score	
h_i	The hidden state of the encoder at moment <i>j</i>	
α_{tj}	Attention weights	
c_t	Context vector	
S_t	The hidden state of Decoder at time t	
y_t	Conditional probability output	

Table 1. List of variables/parameters.

3.2.1. Long Short-Term Memory

LSTM evolved from recurrent neural networks (RNNs), which control the accumulation speed of information by introducing gating, selectively adding new information, and forgetting part of the previously accumulated information. So, it is more sensitive to complex time-dimensional information, which can effectively solve the problem of gradient explosion or gradient diffusion in the training process of RNNs, and which is conducive to learning long-term dynamic information. Its LSTM network structure unit is shown in Figure 3. This structural unit adds a priori knowledge input gate, forget gate, output gate, and input modulation gate to the hidden layer, through which the signal processing between layers and the input signal at a certain moment can be more transparent.

$$i_t = s(U_{xi}x_t + W_{hi}h_{t-1} + b_i)$$
(1)

$$f_t = s(U_{xf}x_t + W_{hf}h_{t-1} + b_f)$$
(2)

$$o_t = s(U_{xo}x_t + W_{ho}h_{t-1} + b_o)$$
(3)

$$g_t = s(U_{xc}x_t + W_{hc}h_{t-1} + b_c)$$
(4)

where *U* represents the input weight, *W* represents the recursive permission of the implied layer at time t - 1, and *b* is the bias term. The LSTM gate structure is controlled by the nonlinear activation function, $s(x) = (1 + e^{-x})^{-1}$ represents the Sigmoid function, and $y(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ represents the hyperbolic tangent function. Moreover, *i*_t determines which information to keep, *f*_t determines which information to remove, *o*_t determines which part of the memory cell *c*_t will output, and *g*_t is a nonlinear mapping by hyperbolic tangent functions.



Figure 3. LSTM network structure unit.

3.2.2. Attention Mechanism

In order to improve the ability to extract key features in flight trajectory data, we introduce the attention mechanism into the LSTM network, focusing on the more critical points of the current prediction task among many trajectory points, reducing attention to other points, filtering out irrelevant information to solve the problem of information overload, and improving the efficiency and accuracy of prediction task processing [34]. The LSTM-Attention structural framework is shown in Figure 4.



Figure 4. LSTM-Attention network framework.

In the attention mechanism, the *i*-th target word context vector c_i is weighted according to the hidden vector of each source word; the encoder part uses a bidirectional recurrent neural network as the encoder; and the hidden state of *t* at each moment is obtained by forward and backward stitching. Introducing the attention mechanism in the decoder, the specific implementation process of the *t*-moment decoder is shown in Table 2.

Table 2. The specific implementation process of the attention mechanism.

Attention Mechanism

Step 1: Calculate the degree to which the information at moment *j* of the input sequence affects the decoding of the output sequence at time *t*, called the alignment score, denoted e_{tj} , where a is the function s_{t-1} learned by the fully connected network is the hidden state of the decoder at time t - 1, and h_j is the hidden state of the encoder at moment *j*:

$$e_{tj} = a\left(s_{t-1}, h_j\right) \tag{5}$$

Step 2: Calculate the attention weights of the input sequence *j* to the output sequence *t* moment, denoted as α_{tj} , its calculation formula is as follows:

$$a_t(s) = \operatorname{align}\left(h_t, \overline{h}_s\right) = \frac{\exp\left(h_t, \overline{h}_s\right)}{\sum\limits_{s'} \exp\left(\operatorname{score}\left(h_t, \overline{h}_{s'}\right)\right)}$$
(6)

Step 3: Calculate the context vector of the decoder at time t, denoted as c_t , where the attention c_t of step t is weighted by the hidden state h_t of the decoder step t and each hidden state h_t in the encoder:

$$c_t = \sum_{j=1}^{l} \alpha_{tj} h_j \tag{7}$$

Step 4: Calculate the hidden state S_t of the decoder at time t:

$$S_t = f(S_t - 1, y_{t-1}, c_t)$$
(8)

Step 5: Calculate the conditional probability output y_t of the decoder at time t, take the attention c_t and the hidden state h_t of the original decoder at step t as input, obtain \tilde{h}_t at step t and output \hat{y}_t , where g is the function of fully connected network learning:

$$p(y_t \mid y_1, \dots, y_{t-1}, X) = g(y_{t-1,St}, c_t)$$
(9)

3.2.3. Trajectory Prediction Model Architecture

As shown in Algorithm 1, the implementation process of the trajectory prediction network is shown. We introduce the attention module and convolutional fusion module described above on the basis of the native LSTM network. Among them, the attention module and the convolution fusion module can enable the LSTM network to obtain time series context information more comprehensively and realize the coarse-granular fusion of time series data features. The model uses one hidden layer, each layer consisting of 64 neurons, using Adam to update and correct the parameters of the model, and the model uses Rule as the activation function of neurons. The improved trajectory prediction network structure is shown in Figure 5.

Algorithm 1: CA-LSTM trajectory prediction network.

Known:

The model uses one hidden layer with 64 neurons per layer.

The specific input sequence X^i is denoted as $\{x_1^i, x_2^i, \cdots, x_s^i\}$.

Step:

S1: The sequence X^i passes through the CNN module to obtain convolutional features, and the output is W_{CNN} .

S2: The output W_{CNN} is used as the input to one node of LSTM.

S3: The attention module operates on the input variables from LSTM by querying the variable q, and calculates the attention distribution weights of the input variables.

S4: Follow the above steps to adjust the input sequence forward to complete the forward propagation process and obtain the predicted output value p.

S5: Combined with the real values, backpropagation solves and updates the model parameters, cycling through the above steps until the algorithm converges.



Figure 5. CA-LSTM network structure.

4. Experiment

4.1. Experiment Setting

In order to maintain the consistency of the comparative experiments, all experimental steps were performed in the same experimental environment. The specific experimental environment configuration is shown in Table 3. At the same time, in order to ensure the authenticity and validity of the experimental results, this paper constructed some datasets for UAV target detection and trajectory prediction under typical infrared detection backgrounds.

Table 3. Experimental environment configuration.

Environment Configuration	Version
CPU	Intel Core i9-12900H @ 2.50 GHz
GPU	NVIDIA GeForce RTX 3070
Operating system	Windows 11
Programming language	Python 3.9

4.1.1. Experimental Environment

The test computer used for this experiment was the Intel Core i9-12900H @ 2.50 GHz, and the laptop GPU was the NVIDIA GeForce RTX 3070. The operating system is Windows 11, and the programming language is Python 3.9.

4.1.2. Dataset Building

The infrared image can detect thermal radiation; ignore lighting conditions; penetrate fog, smoke, and other occlusions; use long visual distances; and work around the clock. So, this paper uses the photoelectric pod to carry out infrared imaging of the UAV target, sets three datasets for different environments to construct scenes (conventional sky background, urban occlusion background, and complex mountain background), and constructs the historical movement trajectory of the UAV target in the time domain according to the change trajectory of the corresponding UAV target in a certain time domain. The UAV infrared detection dataset is shown in Figure 6.



Figure 6. UAV infrared detection dataset.

4.2. Comparative Experiment

In order to prove the effectiveness of the proposed algorithm, the target measurement network and the trajectory prediction network were subjected to ablation experiments in which the object detection and ablation experiment set all the networks to not input the pre-trained model; the number of network training parameter iterations was set to 100; the batch size was set to 16; the initial learning rate was set to 10^{-4} ; and the specific parameter settings were shown in Table 4. The time step of the trajectory prediction ablation experiment was selected as 16, the optimizer selected Adam, the learning rate was dynamically adjusted, and the specific parameter settings were shown in Table 5.

Table 4. Object detection ablation experiment parameter setting.

Parameter	Value
Training epoch	100
Batch size	16
Initial learning rate	10^{-4}

Table 5. Trajectory prediction ablation experiment parameter settings.

Parameter	Value
Time step	16
Optimizer	Adam
Learning rate	Dynamic adjustment

4.2.1. Object Detection Network

In order to test the detection performance of the improved Yolov5 network proposed in this paper on the small target of infrared UAVs and to test the target detection accuracy of the CBAM attention module on the infrared UAV detection dataset, all parameter settings, experimental software, and hardware configurations are made to be the same during the experimental training process to ensure the reference of the comparative experiment. The comparative experimental results are shown in Table 6 from the experimental results using the improved Yolov5 network proposed in this paper. The mAP@0.5 and mAP@0.5:0.95 are increased by 3.5% and 1.6%, respectively, which can effectively cope with the detection of small targets of infrared UAVs, achieve the rapid and accurate identification of UAV targets, and construct the historical motion trajectory of UAVs in this time domain according to the changing trend of the spatial position center of the image of the corresponding UAV in a certain time domain. The UAV trajectory acquisition results are shown in Figure 7.

 Table 6. Comparison of detection results of different target detection algorithms.

mAP@0.5	mAP@0.5:0.95
0.871	0.503
0.906	0.519
	mAP@0.5 0.871 0.906



Figure 7. UAV trajectory acquisition.

4.2.2. Trajectory Prediction Network

This section compares the performance of the three time series prediction algorithms of native LSTM, CNN-LSTM, and CA-LSTM proposed by us, based on the time series data of UAV flight trajectories obtained by the target detection network. In order to optimize the accuracy of the prediction results, the time step is selected as 16, the learning rate is automatically and dynamically adjusted, and the optimization function is Adam. In this experiment, the Mean Absolute Percentage Error and Mean Square Error are selected from several algorithm evaluation indicators common to continuous time series data to evaluate the model's performance.

Mean Absolute Percentage Error (*MAPE*). *MAPE* is the expected value of the relative error loss, which is the percentage of absolute error and true value.

$$MAPE(\mathbf{y}_{i}, \hat{\mathbf{y}}_{i}) = \frac{100\%}{n} \sum_{i=1}^{n} |\frac{\hat{\mathbf{y}}_{i} - \mathbf{y}_{i}}{\mathbf{y}_{i}}|$$
(10)

Mean Squared Error (*MSE*). *MSE* corresponds to the expectation of squared (quadratic) error.

$$MSE(y_{i}, \hat{y}_{i}) = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_{i} - y_{i})^{2}$$
(11)

In the above two evaluation indicators, y_i refers to the actual value and \hat{y}_i refers to the predicted value of the model. The smaller the value of the two evaluation indicators, the better the prediction effect.

As shown in Figure 8, the prediction results of the three trajectory prediction models are compared, and it can be roughly seen from the prediction curve that the prediction effect of the proposed CA-LSTM trajectory prediction model is better than that of CNN-LSTM and LSTM. The specific evaluation results are shown in Table 3.



Figure 8. Comparison of prediction results of different trajectory prediction models.

From Table 7, it can be confirmed that in this experiment, the prediction effect of the model can be depicted as follows: CA-LSTM > CNN-LSTM > LSTM. The prediction accuracy of the CA-LSTM trajectory prediction network proposed in this paper is 1.3% higher than that of CNN-LSTM, 9.43% higher than LSTM, 18.71% lower than CNN-LSTM, and 23.81% lower than LSTM on MSE indicators. Considering the real-time and accurate requirements of UAV trajectory prediction, CA-LSTM can meet the needs of accurate prediction and complete the trajectory prediction task in the preset scenario of this experiment.

Table 7. Evaluation of prediction results for different trajectory prediction n	ıode	e	ls
--	------	---	----

Method	MAPE	MSE
LSTM CNIN I STM	0.1044	3.2470
CA-LSTM (ours)	0.0231	3.0089

In the photoelectric detection scenario, the two-dimensional motion trajectory of the UAV obtained through photoelectric detection cannot reflect the real location information of the UAV target and there is also a lack of height data. However, when combined with the visual servo system of the photoelectric pod, it can detect and monitor the UAV target more intuitively and better perform the UAV prevention and control task. This section uses this detection background to set the mission scene and predict the historical motion trajectory of the UAV target in the time domain obtained from the photoelectric detection image. As shown in Figure 9, the UAV trajectory prediction algorithm proposed in this paper also has good trajectory prediction results for UAV photoelectric detection trajectories and can still maintain good trajectory prediction accuracy for the influence of unfavorable factors such as lens shake and environmental occlusion in the actual detection process, which can better improve the UAV prevention and control ability of ground systems in the field of machine vision.



Figure 9. Prediction results of UAV photoelectric detection trajectory.

5. Conclusions

In this paper, an LSTM trajectory prediction network based on attention enhancement and convolutional fusion structure improvement is proposed, which can improve the network's ability to obtain time series context information. Convolutional fusion structure can effectively use the deep features in UAV trajectory data, which can effectively improve the prediction accuracy of UAV trajectory data by the network. We also constructed a typical infrared drone target dataset based on photoelectric detections, and we trained and tested on this dataset with a 9.43% reduction in MAPE and a 23.81% reduction in MSE compared to the native LSTM network.

As an important part of UAV prevention and control technology, UAV trajectory prediction technology has rapid identification and accurate prediction capabilities that are prerequisites for realizing UAV intent identification and threat assessment. Trajectory prediction technology can be effectively combined with photoelectric detection platforms to improve UAV prevention and control capabilities in key areas. In future research work, we will consider how to solve the problems of UAV trajectory data noise and discontinuity in the environment of low-altitude mountains, urban buildings, other obstructions, and complex backgrounds, so as to provide real-time decision-making support for subsequent UAV intent identification and collaborative disposal.

Author Contributions: Conceptualization, B.S. and S.Y.; Methodology, Z.D., B.S., X.H. and Z.Z.; Software, Z.D.; Validation, C.L.; Writing—original draft, Z.D.; Writing—review & editing, B.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under Grants 52101377.

Data Availability Statement: The data presented in this study are openly available in https://github. com/Dang-zy/SIDD.git.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Liu, F.; Gao, C.; Chen, F.; Meng, D.; Zuo, W.; Gao, X. Infrared Small-Dim Target Detection with Transformer under Complex Backgrounds. *arXiv* 2021, arXiv:2109.14379.
- 2. Guo, S.S. Development status of anti-drone technology and products. *Mil. Dig.* 2016, 377, 36–39.
- 3. Wei, Q.Y. Research on Low, Slow and Small Target Passive Detection Based on External Radiation Source. Master's Thesis, Nanjing University of Science and Technology, Nanjing, China, 2020. [CrossRef]
- 4. Ma, Y. Research on infrared weak and small target detection technology based on SSD framework. *Laser Infrared* **2021**, *51*, 1342–1347.
- 5. Hui, W.Y. A dataset of small and weak aircraft target detection and tracking in infrared images under ground/air background. *China Sci. Data* **2020**, *5*, 291–302.
- 6. Zuo, Z.; Tong, X.; Wei, J.; Su, S.; Wu, P.; Guo, R.; Sun, B. AFFPN: Attention Fusion Feature Pyramid Network for Small Infrared Target Detection. *Remote Sens.* **2022**, *14*, 3412. [CrossRef]
- Li, M.; Ning, D.J.; Guo, J.C. CNN-LSTM model based on attention mechanism and its application. *Comput. Eng. Appl.* 2019, 55, 20–27.
- Qiao, S.J.; Shen, D.; Wang, X.; Han, N.; Zhu, W. A Self-Adaptive Parameter Selection Trajectory Prediction Approach via Hidden Markov Models. *IEEE Trans. Intell. Transp. Syst.* 2015, 16, 284–296. [CrossRef]
- 9. Wu, Q.G. Research on Mission-Oriented Cooperative UAV Trajectory Prediction Technology. Master's Thesis, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2019.
- 10. Qu, X.T.; Zhuang, D.C.; Xie, H.B. Low, slow and small" drone detection method. Command. Control Simul. 2020, 42, 128–135.
- 11. Zhang, M.; Zhang, R.; Yang, Y. ISNet: Shape Matters for Infrared Small Target Detection. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 19–24 June 2022; pp. 867–876.
- 12. Li, X.Y. Application of UAV countermeasure technology and equipment in low-altitude airspace management. *China Secur.* **2023**, 202, 31–36.
- 13. Jiang, C.C.; Ren, H.H. Object detection from UAV thermal infrared images and videos using YOLO models. *Int. J. Appl. Earth Obs. Geoinf.* 2022, 112, 1569–8432. [CrossRef]
- 14. Xue, S. Image recognition method of anti UAV system based on convolutional network. Infrared Laser Eng. 2020, 49, 250–257.
- Li, J. Perceptual generative adversarial networks for small object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1222–1230.
- 16. Cai, W. Detection of weak and small targets in infrared image under complex background. J. Appl. Opt. 2021, 42, 643–650.
- 17. Wu, X.; Li, W.; Hong, D.; Tao, R.; Du, Q. Deep Learning for Unmanned Aerial Vehicle-Based Object Detection and Tracking: A survey. Geosci. *Remote Sens.* 2022, *10*, 91–124. [CrossRef]
- Zhao, Y. Research on UAV Target Detection Method Based on Infrared and Visible Light Image Fusion. Master's Thesis, Harbin Institute of Technology, Harbin, China, 2019; pp. 23–35.
- 19. Arafat, M.Y.; Alam, M.M.; Moh, S. Vision-Based Navigation Techniques for Unmanned Aerial Vehicles: Review and Challenges. Drones 2023, 7, 89. [CrossRef]
- 20. Ma, Q. Low-altitude UAV detection and recognition method based on optimized YOLOv3. *Laser Prog. Optoelectron.* **2019**, *56*, 279–286.
- Zheng, L.; Peng, Y.P. Infrared Small UAV Target Detection Algorithm Based on Enhanced Adaptive Feature Pyramid Networks. IEEE Access 2022, 10, 115988–115995. [CrossRef]
- 22. Zhao, X.; Xia, Y.; Zhang, W.; Zheng, C.; Zhang, Z. YOLO-ViT-Based Method for Unmanned Aerial Vehicle Infrared Vehicle Target Detection. *Remote Sens.* 2023, *15*, 3778. [CrossRef]
- Xu, X.; Sun, Y.W. A Novel Infrared Small Target Detection Algorithm Based on Deep Learning. In Proceedings of the 4th International Conference on Advances in Image Processing (ICAIP'20), Chengdu, China, 13–15 November 2020; Association for Computing Machinery: New York, NY, USA, 2021; pp. 8–14. [CrossRef]
- 24. Li, H.; Deng, L. Enhanced YOLO v3 tiny network for real-time ship detection from visual image. *IEEE Access* 2021, 9, 16692–16706. [CrossRef]
- 25. Yu, K. Research on Infrared Target Detection and Tracking Methods in Anti-UAV Systems. Master's Thesis, National University of Defense Technology, Changsha, China, 2017.

- Wang, Y.H. A brief discussion on how drones and anti-drone technologies complement each other in the 5G era. *China Secur.* 2020, 175, 78–81.
- 27. Zhang, D.D. Research on the development status and anti-swarm strategy of UAV swarms abroad. *FlyingMissiles* **2021**, 438, 56–62. [CrossRef]
- Lu, J.X. Short-term load forecasting method based on CNN-LSTM mixed neural network model. *Autom. Electr. Power Syst.* 2019, 43, 131–137.
- 29. Yang, J.; Xu, J.; Xu, M. Predicting next location using a variable order Markov model. In Proceedings of the 5th ACM SIGSPATIAL International Workshop on GeoStreaming, Dallas, TX, USA, 4 November 2014; pp. 37–42.
- 30. Wei, Y.; You, X. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognit.* **2016**, *58*, 216–226. [CrossRef]
- 31. Tan, W.; Lu, B.C.; Huang, M.L. Neural networks combined with genetic algorithms are used for track prediction. J. Chongqing Jiao Tong Univ. (Nat. Sci. Ed.) 2010, 29, 147–150.
- Zhang, T.; Guo, J.L.; Xu, X.M.; Jun, W. Fighter Air Combat Trajectory Prediction Based on Elman Neural Network. *Flight Mech.* 2018, 36, 86–91.
- Yang, R.N.; Yue, L.F.; Song, M. UAV Trajectory Prediction Model and Simulation Based on Bi-LSTM. Adv. Aeronaut. Eng. 2020, 11, 77–84. [CrossRef]
- 34. Teng, F.; Liu, S.; Song, Y.F. Bi-LSTM-Attention: A Tactical Intent Recognition Model for Air Targets. Aviat. Weapons 2021, 28, 24–32.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.