

# UAV Land Classification Method Based on Federated Learning

Kun Yao\*, Xinyue Cao

Hohai University, Changzhou, 213022, China

**Abstract:** In the application of unmanned aerial vehicle (UAV), land classification is an important field. Image recognition technology based on deep learning is a new method to solve the problem of land classification in recent years. In this paper, we propose a land classification method based on federated learning (FL) which uses Fedavg-Adam algorithm. This method enables UAVs to learn land models using training data distributed in their local database. By aggregating local computational updates of the land classification model, a shared global model is constructed. UAVs can collectively benefit from global models without sharing datasets and protect sensitive information collected by UAVs. To obtain good classification performance, we further propose an improved CNN network. The experimental results show that the improved CNN network suitable for federated learning has the good performance considering the influence of time and accuracy. In the RSSCN7 Dataset, the FedAvg-Adam algorithm converges in the 43rd rounds with an accuracy rate of 82.71%. Compared with FedAvg and FedAdam, the final accuracy of the three is similar, but the latter two converge at 56 rounds and 114 rounds respectively, and Fedavg-Adam has the fastest speed, which proves the superiority of our method.

**Keywords:** UAVs, Land Classification, CNN, Federated Learning

## 1. Introduction

By using cameras and embedded systems for computer vision, UAVs can be useful in applications such as surveillance, aerial photography and ground target detection [1]. Compared with ground detection, aerial detection has lower cost and higher efficiency because of its wide field of vision, low use cost and strong environmental adaptability.

In recent years, the rapid development of deep learning in the field of computer vision provides a new technical means for remote sensing image scene classification, target recognition, image segmentation and other fields [2]. Computer vision based on deep learning has great commercial and application potential in object recognition and detection. For example, in the field of vehicle behavior prediction on the road and strategic bombing target detection in the military, realizing the autonomous execution of UAVs can effectively reduce costs and improve efficiency.

But the computing and power resources of the UAVs are limited. The traditional deep learning (DL) auxiliary scheme uploads raw data directly to the server. This brings huge communication resource loss and the risk of data privacy leakage. Compared with traditional centralized machine learning, federated learning (FL) is an effective solution without uploading raw data.

## 2. Related Work

Xi Gong et al. [3] proposed a TLMoE classification model based on transfer learning, which makes full use of the characteristics of the full connection layer and the convolution layer. Yingqiong Peng et al. [4] proposed an improved image classification model for improved CNN, which replaced Softmax classifier with SVM and had a good accuracy in the application of fruit fly classification. Cong Tan et al. [5] proposed a lightweight FL model using YOLOV3-Tiny, which has good data convergence and accuracy while ensuring data privacy.

### 3. Deep Learning Model

#### 3.1. CNN

CNN usually includes convolution layer, pooling layer and full connection layer [6]. We established a CNN network consisting of two convolutional layers, two pooling layers and three full connection layers. CNN can have good generalization ability through local connection and weight sharing, and the convolution layer with nonlinear operation is as follows:

$$x_j^l = f \left( \sum_{i=1}^N x_i^{l-1} * k_{ij}^l + b_j^l \right) \quad (1)$$

where the matrix  $x_i^{l-1}$  is the  $i$  th feature graph of layer  $(l-1)$ ,  $x_j^l$  represents the  $j$  th feature graph of the current layer  $l$ , and  $N$  denotes the number of input feature graphs.  $k_{ij}^l$  and  $b_j^l$  are randomly initialized, then adjusted by back propagation.  $f(\bullet)$  is a nonlinear activation function, and  $*$  is the convolution operation.

#### 3.2. Improved CNN

Batch Normalization (BN), makes data out of order and accelerates neural network convergence, which has positive application in CNN [7]. We use BN layer to improve the established CNN network. The improved network, which shows in Figure 1, contains two convolutional layers, two pooling layers, two BN layers and two full connection layers. And in this process, we use ReLU as the activation function to add some nonlinear factors to the neural network.

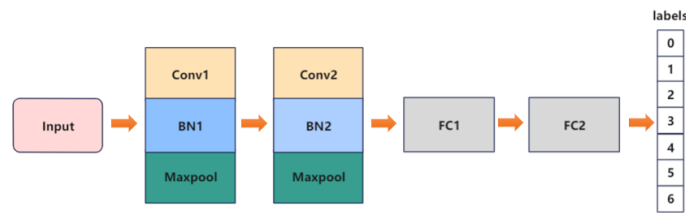


Figure 1: Improved CNN network structure

#### 3.3. ResNet 18

When there are too many layers of neural network, model degradation will occur [8], resulting in unsatisfactory results. ResNet 18, which shows in Figure 2, solves this problem by using residual. The network directly transmits input  $x$  and the proposed residual ( $F(x)$ ) as output result ( $H(x)$ ). The relationship can be expressed as:  $H(x) = F(x) + x$ . When  $F(x) = 0$ , it is called identity mapping.

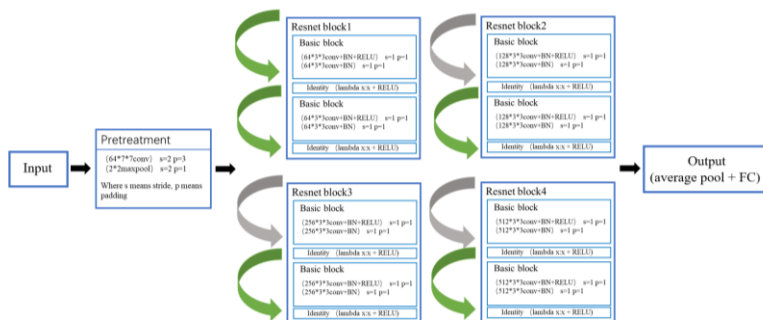


Figure 2: ResNet 18 network structure

## 4. Establishment of Federated Learning Model

### 4.1. Description of the task

In exploration tasks, terrain discrimination and classification are the core work, and the exploration range and flight altitude of UAV are affected by ground service center (GFC). The GFC needs to be at a position with significant height, such as the top of a mountain, to reduce communication costs and expand exploration coverage. As showed in Figure 3, UAVs equipped with cameras traverse the detection area, classifying targets as they arrive, and GFC coordinated UAV swarm exploration model can adopt federated learning to solve the problem. Specifically, each UAV updates the model locally, and then transfers the model parameters obtained by training to GFC, which aggregates the parameters and returns them to the UAVs.

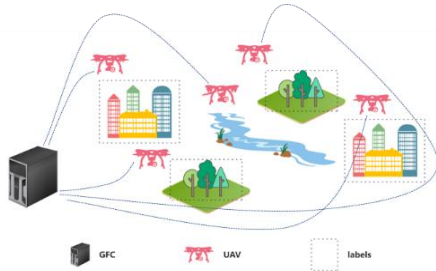


Figure 3: A multi-UAV system designed for performing image classification tasks

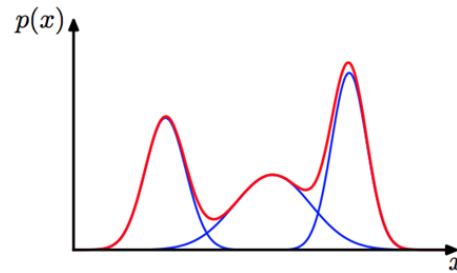


Figure 4: An example of Dirichlet distribution

### 4.2. Dataset

RSSCN7 Dataset contains 2800 images of 7 types of typical scenes. The scene images in the dataset include meadows, forests, farmland, parking lots, residential areas, industrial areas, and rivers and lakes [9]. The high-altitude images in the dataset are of great value in this application scenario because they contain different seasons and weather conditions and UAV-like perspectives.

### 4.3. Heterogeneous Data

In practical applications, the data obtained by the server from UAV is diverse due to factors such as geography, time and UAV state. This means that the data overlap degree of different UAVs is very small. In order to simulate this scenario, we adopt the method of dividing non-IID data sets according to Dirichlet distribution [10]. It is assumed that there is a similarity which is the data conforms to a mixed distribution between the data of each UAV. This mixed distribution can be intuitively understood in Figure 4. When the  $t$  th UAV generates a sample point ( $x$ ), select a component from  $n$  composition samples. The possibility that the sample point ( $x$ ) comes from the  $t$  th UAV is described as  $p(x|\theta_t)$ :

$$p(x|\theta_t) = \sum_{n=1}^N z_m p(x|\theta_n) \quad (2)$$

Where  $z_{tn}$  is a hidden variable in the  $t$  th UAV, which represents the probability that data in the  $t$  th UAV come from composition  $n$ .

### 4.4. Loss Function in UAVs

In this study, the Cross-entropy error function [11] is used as the loss function to conduct experiments on the three networks, so as to select the network suitable for FL task. The loss function is expressed as:

$$Q = -\frac{1}{N} \sum_{n=1}^N (y_n \log_2(o_n) + \tilde{y}_n \log_2(\tilde{o}_n)) \quad (3)$$

Where  $\tilde{y}_n = 1 - y_n$ ,  $\tilde{o}_n = 1 - o_n$ ,  $\tilde{o}_n = 1 - o_n$ ,  $N$  is the number of batch training samples,  $y$  is the true label, and  $o$  is the actual output.

#### 4.5. BN Layer in Local Training

BN layer's work on local UAV can be described as follows:

$$\hat{x}_i = \frac{z_i - \mu(z_i)}{\sigma(z_i)} \quad (4)$$

$$\text{BN}(\hat{x}_i) = \gamma_i \hat{x}_i + \beta_i \quad (5)$$

Where  $z_i$  is a neuron across a batch,  $\mu(z_i)$ ,  $\sigma(z_i)$  are the mean and standard deviation in  $z_i$ , and  $\gamma_i$ ,  $\beta_i$  are parameters learned in local UAV training.

#### 4.6. The Optimization Goal

We designed the following optimization goal to allow UAVs benefit from global model while building different local models [12]:

$$\min L = \sum_{k=1}^K \frac{x_k}{x} l_k(M_k) \quad (6)$$

$$M_k = (\Omega_1 \cdots \Omega_{i_1}, P_{k_1}, \Omega_{i_1+1} \cdots \Omega_{i_m}, P_{k_m}, \Omega_{i_m+1} \cdots \Omega_j)$$

Where  $x_k$  is the number of samples on UAV  $k$ ,  $x$  is the total number of samples on all UAVs,  $l_k$  is the loss function on UAV  $k$ ,  $\Omega$  consists of Federated model parameters,  $P$  consists of parameters in BN layer, and  $M_k$  is the model on UAV  $k$ .

#### 4.7. FL Training Algorithms

FedAvg is one of the most basic and widely used algorithms in federated learning. It first requires the UAV to perform local model initialization and train the model using local data for SGD, which uses gradient descent to update model parameters. The UAV then sends the model parameters to the central server, which aggregates them by Average. Average takes the average of all the parameters. Finally, the server returns the resulting aggregation parameters to the UAV, which continues training with the returned parameters. Repeat these steps.

Other improved algorithms include FedAdam and FedAvg-Adam [13]. Adam update local UAV model parameters according to the mean of the gradient and the uncentered variance. Table 1 introduces these training algorithms.

*Table 1: LocalUpdate, GlobalModelUpdate and GlobalOptimUpdate Used by the FedAvg, FedAdam, and FedAvg-Adam FL Training Strategies*

Optimisation Strategy	LocalUpdate	GlobalModelUpdate	GlobalOptimUpdate
FedAvg	SGD	Average	-
FedAdam	SGD	Adam	-
FedAvg-Adam	Adam	Average	Average

### 5. Experiments

All the experiments were carried out under the Pytorch framework on a workstation with NVIDIA GeForce GTX1080Ti and Inter×CoreI5-3479 CPU×3.20 GHz and RAM of 12.0GB.

Training parameters are set as follows: Input the image size as 224\*224, the learning rate as 0.001, batch size as 8, and select SGD as the optimization algorithm. The results obtained are shown in the following table 2:

(Acc is the convergence accuracy obtained by training, N is the epochs required to achieve the

convergence accuracy, and  $t$  is the average time required by an epoch in the experiment,  $T$  is the total time, calculated as  $t * N$ )

Table 2: The Results Obtained by Experiments

Type	Acc	N	t	T
CNN	65.85%	39	22.3s	869.7s
CNN+BN	72.02%	20	23.1s	462.0s
ResNet 18	86.43%	83	80.4s	6673.2s

We assume a system of 20 UAVs and that half of all UAVs participate in training in each round.

Set the batch size on the local UAV to 10, the learning rate to 0.1, the UAV's local iteration to 3, and total communication rounds to 200. The results obtained are shown in Figure 5.

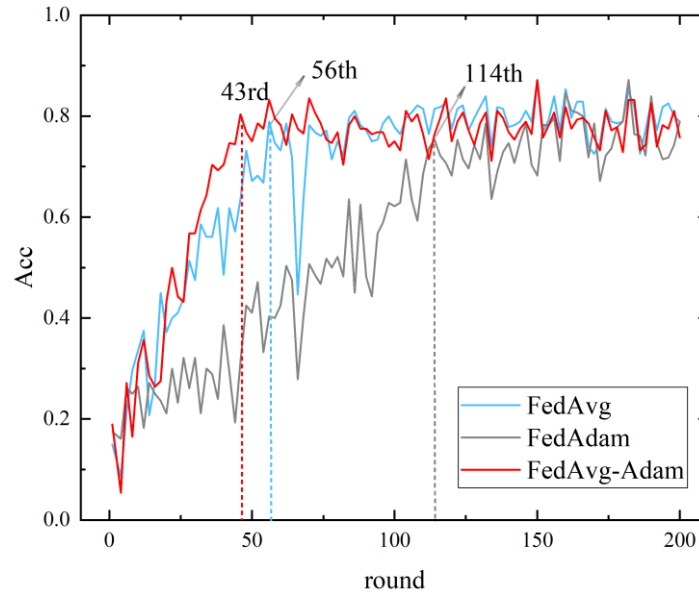


Figure 5: Results of federated learning under the three algorithms

It can be found that the final accuracy of FedAvg-Adam and FedAdam is similar and slightly higher than that of FedAvg. FedAvg-Adam converges in the 43rd round, FedAvg converges in the 56th round, and FedAdam converges in the 114th round. So FedAvg-Adam works best in this application scenario.

## 6. Conclusion

In this paper, we propose a land classification method based on federated learning. This method uses Fedavg-Adam algorithm, which enables UAVs to learn land models using training data distributed in their local database. UAVs can collectively benefit from the shared classification model aggregating local computational updates of the land classification model without sharing datasets and protect sensitive information. To obtain good classification performance, we further propose an improved CNN network for UAVs' local train. By adding BN layer, the performance of improved CNN considering the influence of time and accuracy is better than CNN and ResNet 18. In the RSSCN7 Dataset, the FedAvg-Adam algorithm converges in the 43rd rounds with an accuracy rate of 82.71%. Compared with FedAvg and FedAdam, the final accuracy of the three is similar, but the latter two converge at 56 rounds and 114 rounds respectively, and Fedavg-Adam has the fastest speed, which proves the superiority of our method.

## References

- [1] M. Bhaskaranand, a. J. (n.d.). Low-complexity video encoding for UAV reconnaissance and surveillance. *Proc. IEEE Military Communications Conference (MILCOM)*, pp. 1633-1638, 2011.
- [2] Dong, Q.; Zou, Q. Visual UA V detection method with online feature classification. In *Proceedings of the IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chengdu, China, 15–17 December 2017; pp. 429–432.

- [3] Gong Xi, et al. "A hybrid expert classification model based on transfer learning for scene classification of high-resolution remote sensing images" [J]. *Acta Optica Sinica*, 2019, 41(23):19-31.
- [4] Peng, Yingqiong, et al. "CNN-SVM: a classification method for fruit fly image with the complex background." *IET Cyber-Physical Systems: Theory & Applications* 5.2 (2020): 181-185.
- [5] Tan, Cong, et al. "A Lightweight Underwater Object Detection Model: FL-YOLOV3-TINY." *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, 2021.
- [6] SHI Chao-jun et al. Automatic Classification Method of Star Spectra Data Based on Convolutional Neural Network[J]. *Spectroscopy and Spectral Analysis*, 2018, 39(04): 1312-1316.
- [7] P. K. Mudrakarta, M. Sandler, A. Zhmoginov, and A. Howard, "K for the price of 1: Parameter efficient multi-task and transfer learning," in *Proc. Int. Conf. Learn. Representations*, 2019.
- [8] Z. Hao, F. Ren, X. Kang, H. Ni, S. Lv and H. Wang, "Classification of Steel Strip Surface Defects Based on Optimized ResNet18," *2021 IEEE International Conference on Agents (ICA)*, 2021, pp. 61-62.
- [9] Q. Zou, L. Ni, T. Zhang and Q. Wang, "Deep Learning Based Feature Selection for Remote Sensing Scene Classification," in *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 11, pp. 2321-2325, Nov. 2015. Available: <https://hyper.ai/datasets/5440/>
- [10] J. Zhang et al., "Adaptive Federated Learning on Non-IID Data with Resource Constraint," in *IEEE Transactions on Computers*. 2020.
- [11] Q. Wang, H. Wang and Y. Yan, "Reconstruction of CT images based on cross-entropy method," *2010 IEEE Instrumentation & Measurement Technology Conference Proceedings*, 2010, pp. 1510-1514.
- [12] J. Mills, J. Hu and G. Min, "Multi-Task Federated Learning for Personalised Deep Neural Networks in Edge Computing," in *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 3, pp. 630-641, 1 March 2022.
- [13] D. Leroy, A. Coucke, T. Lavril, T. Gisselbrecht and J. Dureau, "Federated Learning for Keyword Spotting," *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 6341-6345.