



ARTICLE

Broad Federated Meta-Learning of Damaged Objects in Aerial Videos

Zekai Li¹ and Wenfeng Wang^{2,3,4,5,6,*}

¹School of Aerospace Engineering, Xi'an Jiaotong University, Xi'an, 710049, China

²Research Institute of Intelligent Engineering and Data Applications, Shanghai Institute of Technology, Shanghai, 201418, China

³Research Center for Ecology and Environment of Central Asia, Chinese Academy of Sciences, Urumqi, 830011, China

⁴Applied Nonlinear Science Lab, Anand International College of Engineering, Jaipur, 391320, India

⁵Department of Visual Engineering, International Academy of Visual Art and Engineering, London, CR2 6EQ, UK

⁶Institute of Intelligent Management and Technology, Sino-Indian Joint Research Center of AI and Robotics, Bhubaneswar, 752054, India

*Corresponding Author: Wenfeng Wang. Email: wangwenfeng@sit.edu.cn

Received: 31 December 2022 Accepted: 21 March 2023 Published: 03 August 2023

ABSTRACT

We advanced an emerging federated learning technology in city intelligentization for tackling a real challenge—to learn damaged objects in aerial videos. A meta-learning system was integrated with the fuzzy broad learning system to further develop the theory of federated learning. Both the mixed picture set of aerial video segmentation and the 3D-reconstructed mixed-reality data were employed in the performance of the broad federated meta-learning system. The study results indicated that the object classification accuracy is up to 90% and the average time cost in damage detection is only 0.277 s. Consequently, the broad federated meta-learning system is efficient and effective in detecting damaged objects in aerial videos.

KEYWORDS

Fuzzy; learning system; mixed-reality; 3D-reconstructed; oblique photography

1 Introduction

Unmanned aerial vehicles have irreplaceable advantages in large-scale monitoring areas to achieve effective video scheduling and risk prediction [1–3]. However, data privacy is still a problem in debate [4]. Federated learning can present a possible solution [5], where each UAV can run the same and complete model and there is no communication between the UAVs. Technologies are already mature in object detection and classification, but the precision understanding and learning of damaged objects in aerial videos are still great challenges [6]. The ultimate reason is that there are significant differences in the features of damaged objects in the real world and even for the same object, the damaged degree and features can also be very different in different situations [7]. Considering the influences of COVID-19 and the necessity to extend the applications of unmanned intelligent technologies, this is an urgent problem to be solved [8,9].



In the traditional oblique photography technology, the methods of machine learning and the related theories of federated learning are rarely applied, and most of them stay at the level of three-dimensional modeling [9–12]. After using the UAV to complete the 3D modeling, the problem of image classification and recognition can be solved with some effective machine learning system, especially the theory of federated learning can be further introduced. Meta-learning and fuzzy broad learning system (FBLs) have been proven to be effective for image recognition [13–22]. Integration with these two learning systems can also further develop the federated learning theory.

Therefore, the major objectives of this study are (1) to introduce the federated learning theory in the treatment of UAV photography, which improves the texture methodology and the massive video-integrated mixed reality technology [23,24], and (2) to further develop the theory of federated learning through an integration with the method of meta-learning and FBLs to segment the 3D modeling virtual-real fusion video as a data set, which realizes image classification and recognition with high accuracy. The organization of the article is as follows. Section 2 uses a five-eye UAV system for oblique photography and builds a 3D model. The mixed reality of surveillance video and 3D model will be carried out and a broad federated meta-learning system will be constructed. Section 3 uses the theory of learning the damaged objects in aerial videos. The broad federated meta-learning system will be used to recognize and classify images, respectively, and its performance in damaged degree analyses is compared with traditional CNN networks. The conclusions and some next priorities are presented in Section 4.

2 Theoretical Analyses

As shown in Fig. 1, the five-view camera is used from orthogonal and four oblique directions to shoot the target objects, where the camera views the target from orthogonal and four oblique directions object to shoot [18].

Date treatment for federated learning in the position orientation system is shown in Fig. 2. A local image transformation is performed for mapping the mean and variance of the selected window to the rest of the image so that the mean and variance of the selected area are balanced, and the selected window area is moved in a certain order to filter the image globally. The selected window areas can be moved in a certain order to filter the picture globally [19–21].

The broad federated meta-learning processes are as follows. First, data were collected by the UAVs according to Fig. 1 and inertially calculated according to Fig. 2. The well-treated data from each UAV were transformed for privacy protection and collected by the server for broad learning and meta-learning. The UAVs can download the latest model from the server in turn. Second, each user trained the model with local data and uploaded the encryption gradient to the server. Integrating with the inertial calculations, the server can aggregate the gradient information for updating model parameters. After then, the server returns the updated model parameters to each UAV. Finally, all the UAVs updated their models to learn about damaged objects in aerial videos. The users can be people from different regions. They have absolute control over associated UAVs and data. Each user can stop computing and participate in communication at any time. In meta-learning and broad learning systems, the server acquires the candidate worker nodes and then controls the selected nodes by updating the model parameters and nodes values.

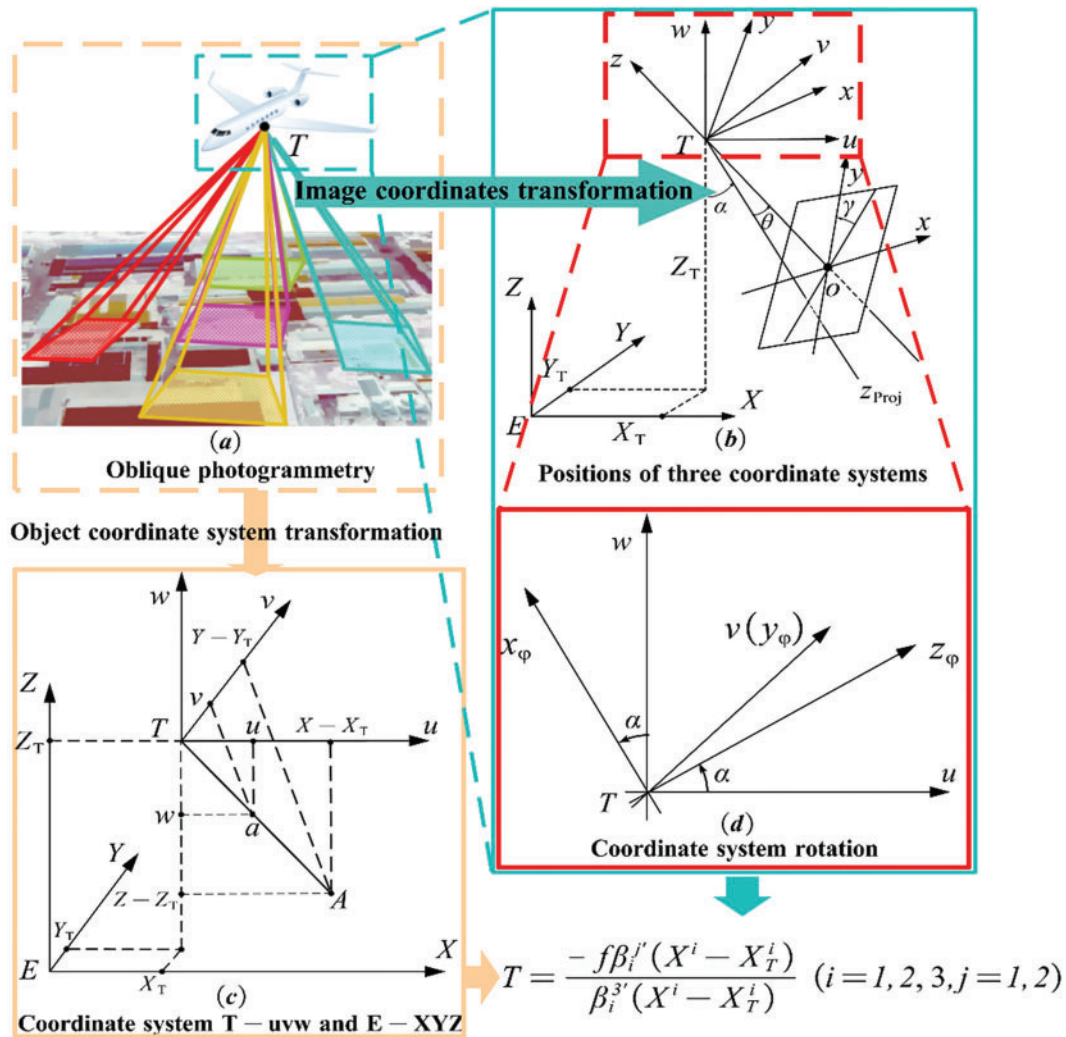


Figure 1: Mechanisms for data acquisition through the oblique photography technology—(a) the x-axis and y-axis of the image space coordinate system; (b) the angle between the projection and the image plane [18]; (c) the coordinates of the image points obtained from the central projection relationship; (d) the coordinated rotations to find the ground points (Image encryption display)

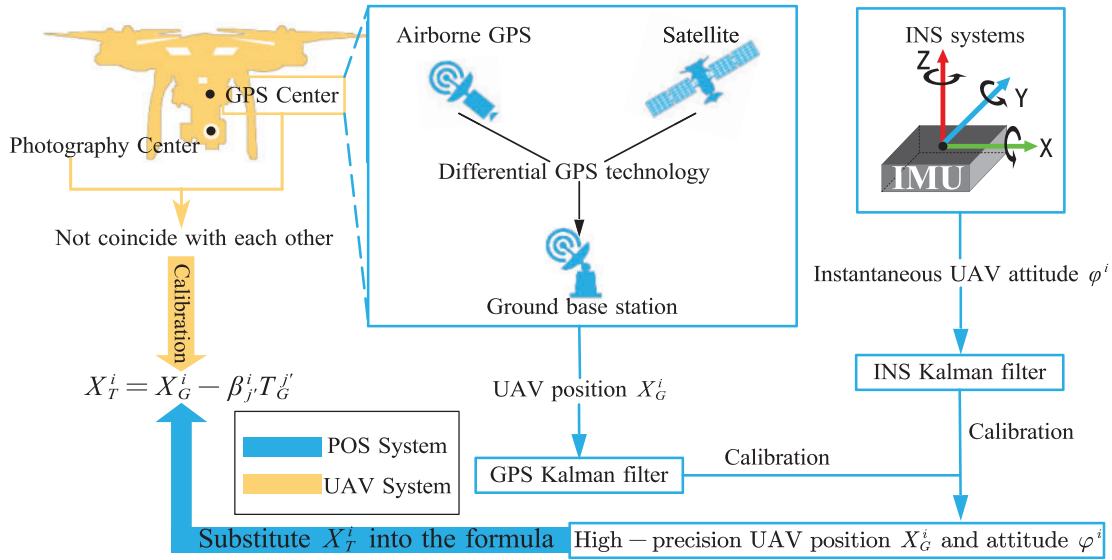


Figure 2: Data treatment for federated learning in the position orientation system, which consists of an on-board GPS (Global Positioning System) and an INS (Inertial Navigation System), which is composed of the UAV position information obtained by differential GPS technology and the IMU (Inertial Measurement Unit) captured by the INS

3 Implementation of the Theory

3.1 Mixed Reality from 3D Reconstruction

For mixed reality from 3D reconstruction, we first perform feature extraction and match on the 3D point cloud formed by aerial triangulation using ORB, then construct Digital Terrain Model (DTM) by multi-visual matching, then process the surveillance video sequences, and finally integrate the virtual and real fusion based on the surveillance video images and DTM model to achieve lossless zooming of local areas and scene recognition with classification [20], as shown in Fig. 3.

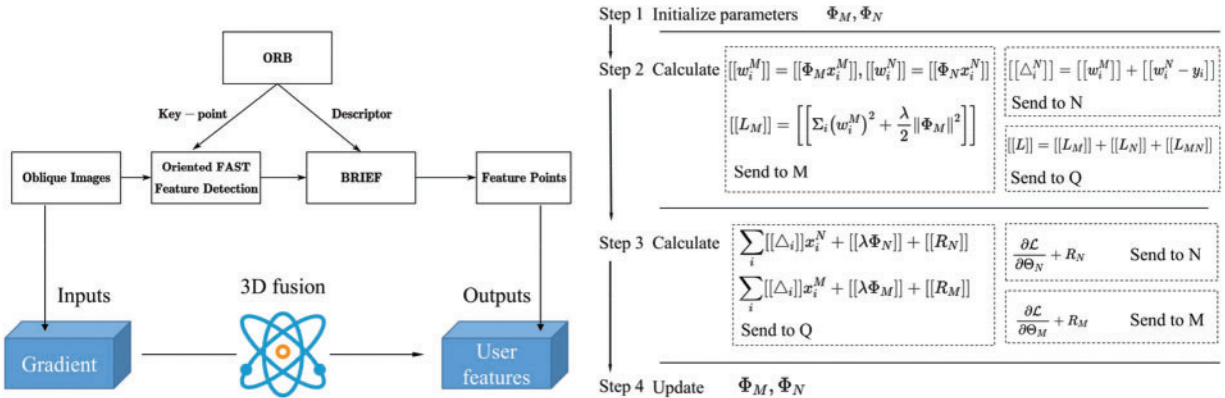


Figure 3: Flow chart of ORB key-point detector and federated learning [25,26]. The ORB (Oriented FAST and Rotated BRIEF) method is adopted in this paper to extract feature points in multi-view matching for the 3D model fusion and fast secondary processing of the aerial videos

3.2 Approach to Learning of Damaged Objects

Utilizing the subsystems with fuzzy rules for federated learning of damaged objects, we define the activation level (fire strength) and membership function corresponding to fuzzy sets with defuzzified outputs. We choose the Gaussian membership function and initialize the coefficients in the polynomial function of the first-order fuzzy system. Applying the k-means method to divided the data into classes and derive K clustering centers, the centers of Gaussian membership functions are initialized by the values of clustering centers.

The intermediate output matrix for each training sample is calculated without aggregation. During federated learning of damaged objects, the output matrix of the i^{th} fuzzy subsystem for all the training samples can be derived for all the training samples. The output matrix of intermediate output can be calculated from the output matrix of the enhancement layer and the output matrix of each fuzzy subsystem for all the training samples. The fuzzy subsystems are constructed on certain fuzzy rules—defining the activation level (fire strength) and membership function corresponding to fuzzy sets and the defuzzified outputs can be interpreted by finite IF-THEN rules, as follows:

$$\text{If } x_{s1} \text{ is } A_{k1}^i \text{ and } x_{s2} \text{ is } A_{k2}^i \dots \text{ and } x_{sM} \text{ is } A_{kM}^i, \text{ then } z_{sk}^i = f_k(x_1, x_2, \dots, x_M), \quad k = 1, 2, \dots, K_i. \quad (1)$$

where A_{ki}^i is a fuzzy set, x_{sj} is the system input ($j = 1, 2, \dots, M$), and K_i is the number of fuzzy rules in the i th fuzzy subsystem.

Taking f_k^i to be a polynomial function of x and using the Gaussian membership function, this learning system can obtain the first-order fuzzy model with a first-order polynomial f_k^i [14,16].

The meta-learning system can be defined as a procedure to process the dataset D and make predictions \hat{y}^* from the known factors x^* [27]. As a particular function that maps x^* to \hat{y}^* , f can be linked with the meta-learning methods [28–31]. Exactly, the goal of such learning process is to recover a function from data, including meta-learning classifiers, regression, and policies for effective prediction and control [32]. The learning function f is said to be consistent if [33]

$$\lim_{|D| \rightarrow \infty} f(D, x_i^*) \rightarrow y_i^*, \quad \forall (x_i^*, y_i^*). \quad (2)$$

Hence the broad federated meta-learning of damaged objects can be formulated as a large-scale task classification. We improve the traditional gradient descent to update the weights of the meta-learning function f , which can be represented as

$$\hat{f}(\cdot; \theta) = f_{out} \left(\left(\prod_{i=1}^N W_i \right) \varnothing(\cdot; w_j^l, \theta_{f_l}^l, \theta_b^l); \theta_{out}^l \right) \quad l = 1, 2, 3, \dots, n, \quad (3)$$

where l represents the l^{th} task, $\varnothing(\cdot; w_j^l, \theta_{f_l}^l, \theta_b^l)$ and $f_{out}(\cdot, \theta_{out}^l)$ are respectively the input and output of the meta-learning system, $\prod_{i=1}^N W_i$ is the weight matrices.

The considered scenarios are to detect damaged objects in aerial videos. A real-time detection should be expected and hence, the communication cost of federated learning can be majorly focused on time. The roles of the two integrated learning systems are different—the fuzzy broad learning system serves for differentiate the damaged objects from other ones and meta-learning system serves for objects recognition. The meta-learning system works at first and presents the necessary basis for the broad fuzzy learning system. We also utilized CNN to classify the data that participated in the federated learning processes to ensure their independence. To minimize the influences of feature differences in learning the damaged objects in aerial videos, meta-learning as a “learn to learn” process is further utilized, as shown in Fig. 4.

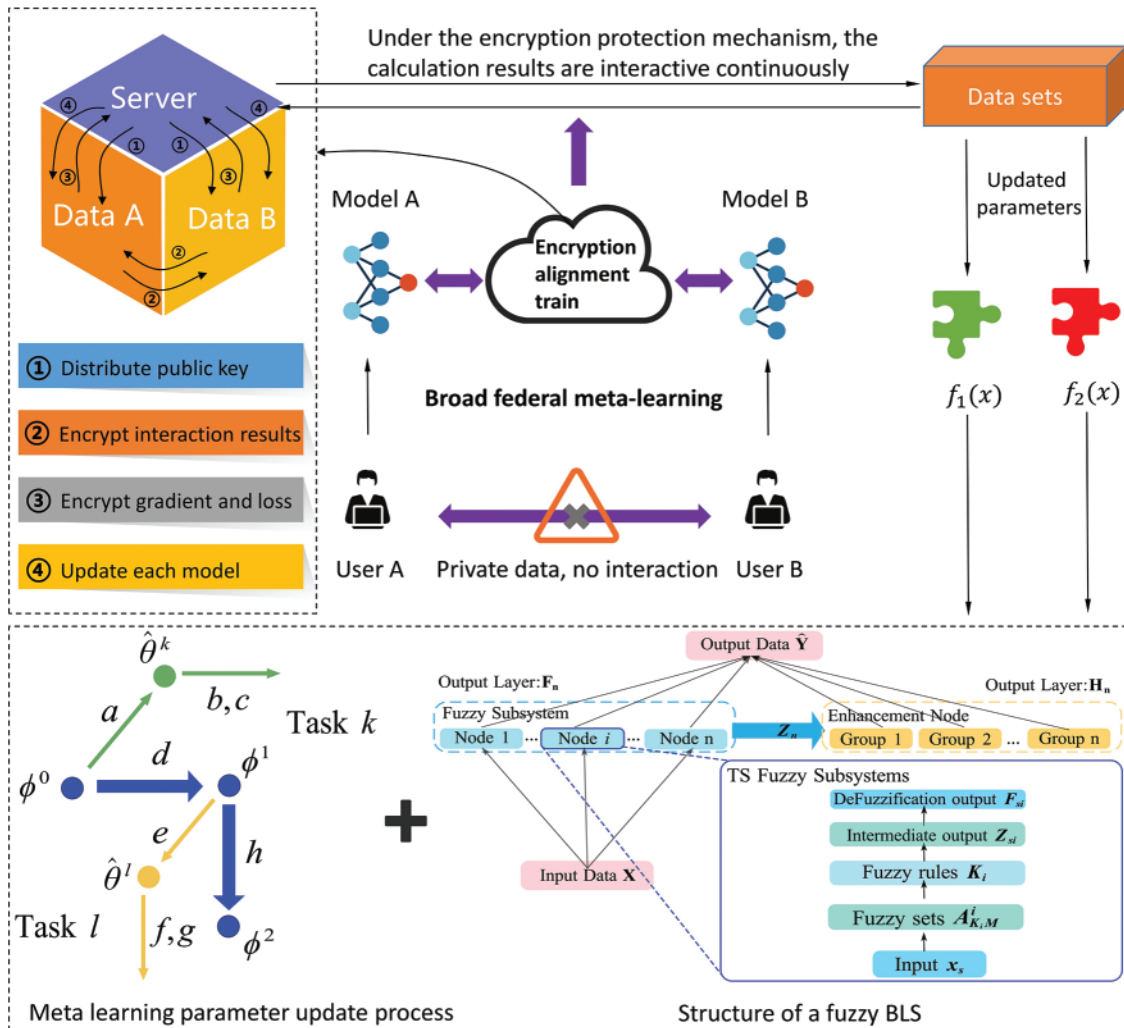


Figure 4: Broad federated meta-learning processes, including the structure of FBLs and meta-learning parameter updating process—the first green arrow is parallel to the second green arrow from left to right, which means that as the federated gradients in the same direction are updated simultaneously in approaches to learn damaged objects in aerial videos

Specifically, for each client, the latest global model parameters are first downloaded from the central server. The client then trains its own CNN model using its local data for 50–100 epochs and calculates the gradient of the FBLs model relative to the binary cross-entropy loss function of its local data. The client then uploads this local model gradient information to the central server. Once the central server has received all the uploaded gradient information from the clients, it aggregates the gradients using the Federated Averaging algorithm and updates the weights and bias parameters of the global model. Finally, the central server sends the updated global model weights and bias parameters to the clients, who use their local optimizers to update their local model parameters and minimize their local loss function.

Upload gradient information of local model parameters, which is the local data loss function gradient according to the classification model trained by each client, to a central server.

For the cross entropy function, we have:

$$loss_k = - \sum_{i=1}^s \sum_{j=1}^c y_{ij} \ln \hat{y}_{ij} \quad (k = 1, 2, 3 \dots, n), \tag{4}$$

where $loss_k$ represents the calculation result of the k^{th} client's cross entropy, y_{ij} represents the real value of the j^{th} category label in the i^{th} sample, $\hat{y}_{ij} \in [0, 1]$ represents the predicted value of the model for the j^{th} category in the i^{th} sample, s and c are respectively the number of samples and the number of categories.

For the binary model of whether the target is damaged, we take $c = 2$, then $y_{i2} = 1 - y_{i1}$, $\hat{y}_{i2} = 1 - \hat{y}_{i1}$, thus we can simplify the above equation as:

$$loss_k = - \sum_{i=1}^s [y_{i1} \ln \hat{y}_{i1} + (1 - y_{i1}) \ln (1 - \hat{y}_{i1})] \quad (k = 1, 2, 3 \dots, n), \tag{5}$$

for the loss of image category, we take the label value y to 1, so we have:

$$loss_k = - \sum_{j=1}^s \ln(\hat{y}_j) \quad (k = 1, 2, 3 \dots, n). \tag{6}$$

In FBLS, $\hat{y}_i = f(\omega) \in [0, 1]$, where $\omega = (\omega_1, \omega_2, \omega_3 \dots, \omega_l) (l = 1, 2, 3 \dots, m)$ is for the weight of the l^{th} middle layer, m represents FBLS in the middle layer to the total number of nodes, then we have:

$$loss_k = - \sum_{i=1}^s \ln(\hat{y}_i(\omega)) = - \sum_{i=1}^s J_i(\omega) = J_k(\omega) \quad (k = 1, 2, 3 \dots, n). \tag{7}$$

For the update of the k^{th} client weight parameter ω_k , we have:

$$\omega_k^{t+1} = \omega_k^t - \alpha \frac{\partial J_k(\omega_k^t)}{\partial \omega_k^t} \quad (k = 1, 2, 3 \dots, n), \tag{8}$$

where t is the number of iterations, α represents the learning rate, $\frac{\partial J_k(\omega_k^t)}{\partial \omega_k^t}$ is the gradient information of fbfs loss function, represents the update of the t^{th} weight parameter of the k^{th} client, which is denoted as $\Delta \omega_k^t$.

With respect to the central server, we aggregate $\Delta \omega_k^t$ according to the Federated algorithm and return the obtained parameters to the client for parameter iteration and optimization in accordance with Eq. (8).

These devices participate in federated learning through online communication. Considering the unstable situations and the difference in their computing capabilities, we perform the system on each class of objects, respectively. This reduced the number of communications so that they can work together and finish learning tasks simultaneously. To make it possible to quickly learn new tasks based on the acquisition of existing “knowledge”, the system is trained by feeding it many support sets to achieve good recognition in the testing set and for new-shot learning, it learns with an alternative strategy—learns from small samples and constantly correcting the model parameters to achieve good recognition and classification in a new testing set. The key point of broad federated meta-learning is to find the meta-learner that can be adapted to the new damaged objects through the updating data. The key point is not to find the mapping between input and output by computing the loss function to better implement the mapping between features and labels (requiring the input of too many data sets for

training and validation), but to learn autonomously with only a small amount of data. Consequently, the broad federated meta-learning system can learn damaged objects quickly and accurately with new, small amounts of data.

4 From Theory to Practices

4.1 Damaged Objects in Federated Learning

As shown in Fig. 5 (left), the meta-learning system achieves lossless enlargement and screen recognition of the details on 3D-reconstructed data sets with damaged objects. We divided the degree of lawn into four categories according to the percentage of damaged areas from 0% to 100% with a 25% interval width. After adding morphological operators to extract the breakage and pre-training the algorithm by introducing the lawn damage instances of the 4 categories of damage in 3D-reconstructed data sets, we use the improved algorithm to classify Chaparral, Baseball Field, Golf course, Meadow, in the reality data sets. The classification results are shown in Fig. 5 (right).

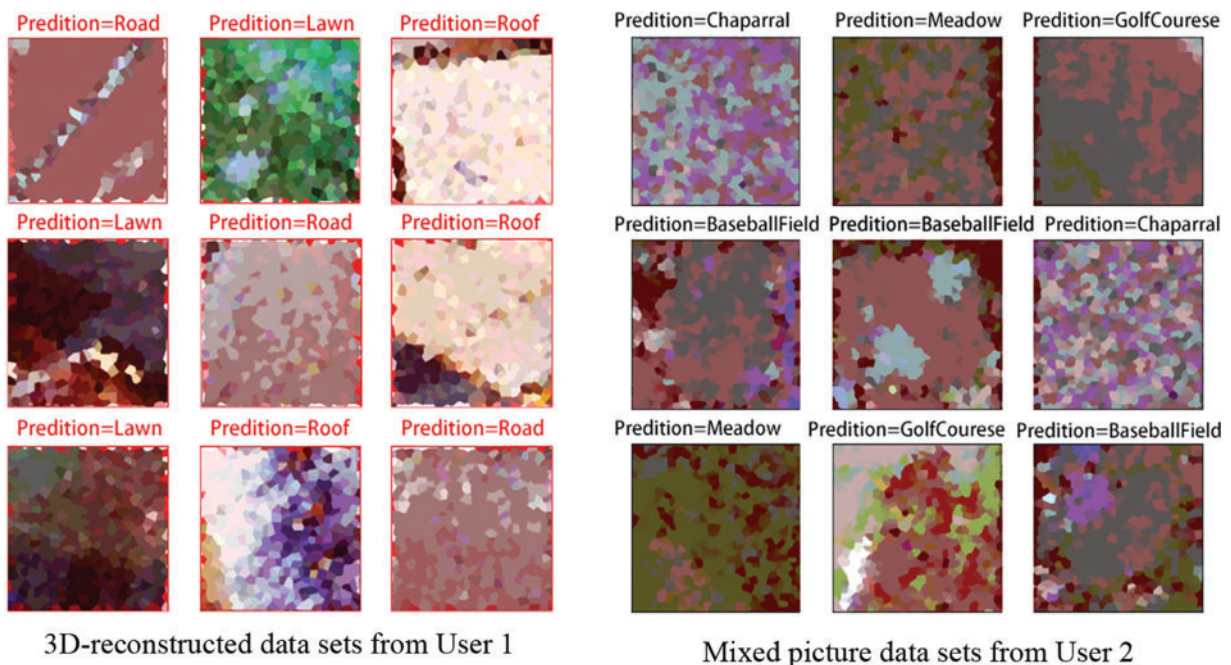


Figure 5: Meta-learning of objects from 3D-reconstructed data sets and mixed picture data sets. The mixed picture data sets for Chaparral, Baseball Field, Golf course, and Meadow were manually processed to ensure that the percentage of damaged areas was in a quartile interval from 0% to 100% with a 25% interval width (Image encryption display)

The training results of the broad federated meta-learning system are represented in a confusion matrix with the actual model, as shown in Figs. 6 and 7. The performance of the 3D-reconstructed data from the oblique photography of the UAV was compared with that of the mixed data set of aerial videos segmentation pictures. For 3D-reconstructed data, damaged objects are divided into three categories according to lawn, road and roof. For the mixed picture data sets, we used the same classification as before, setting up Chaparral, Baseball Field, Golf course, and Meadow as four different degrees of deterioration.

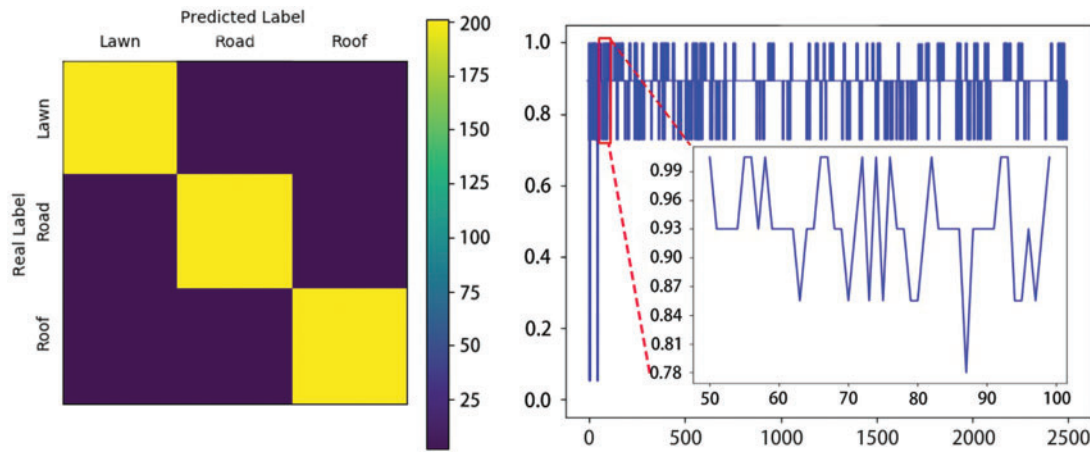


Figure 6: Real label vs. predicted label confusion matrix—the 3D-reconstructed data from the oblique photography of the UAV and the surveillance footage are processed according to the video, and the images are extracted frame by frame, and each image is divided into sixteen equal parts

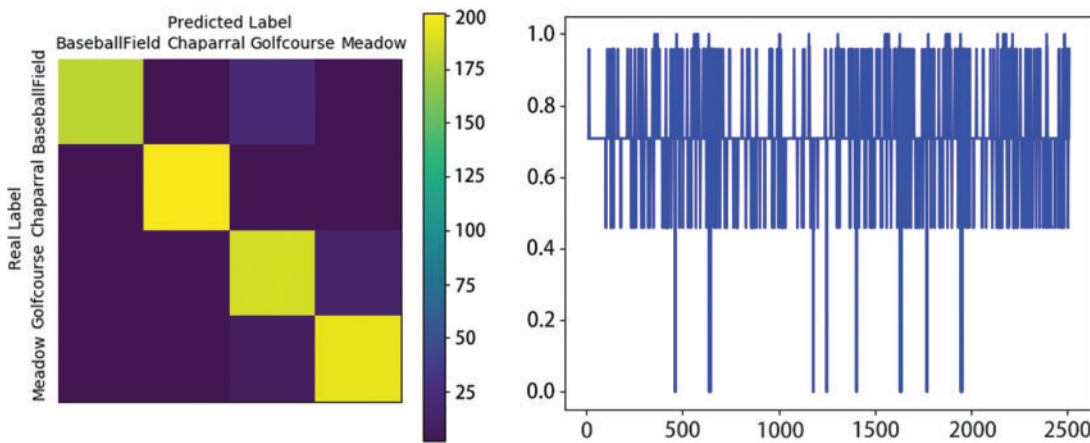


Figure 7: Real label vs. predicted label confusion matrix—the mixed picture set of aerial videos segmentation are processed according to the video, and the images are extracted frame by frame, and each image is divided into sixteen equal parts

In the figure above, the federated learning model is trained relatively quickly on the 3D-reconstructed data sets, with the accuracy of the model essentially stabilising at 90% after 50 steps. The picture on the right shows that only a small number of damaged targets are incorrectly identified and the model performs well overall.

In the above figure, the classification accuracy of meta-learning on mixed picture data sets is not as high as that in Fig. 6, and its classification accuracy is volatile and does not converge. As shown in Figs. 8c and 8d, for the classification of mixed picture data sets, we can achieve over 90% accuracy by using conventional CNN.

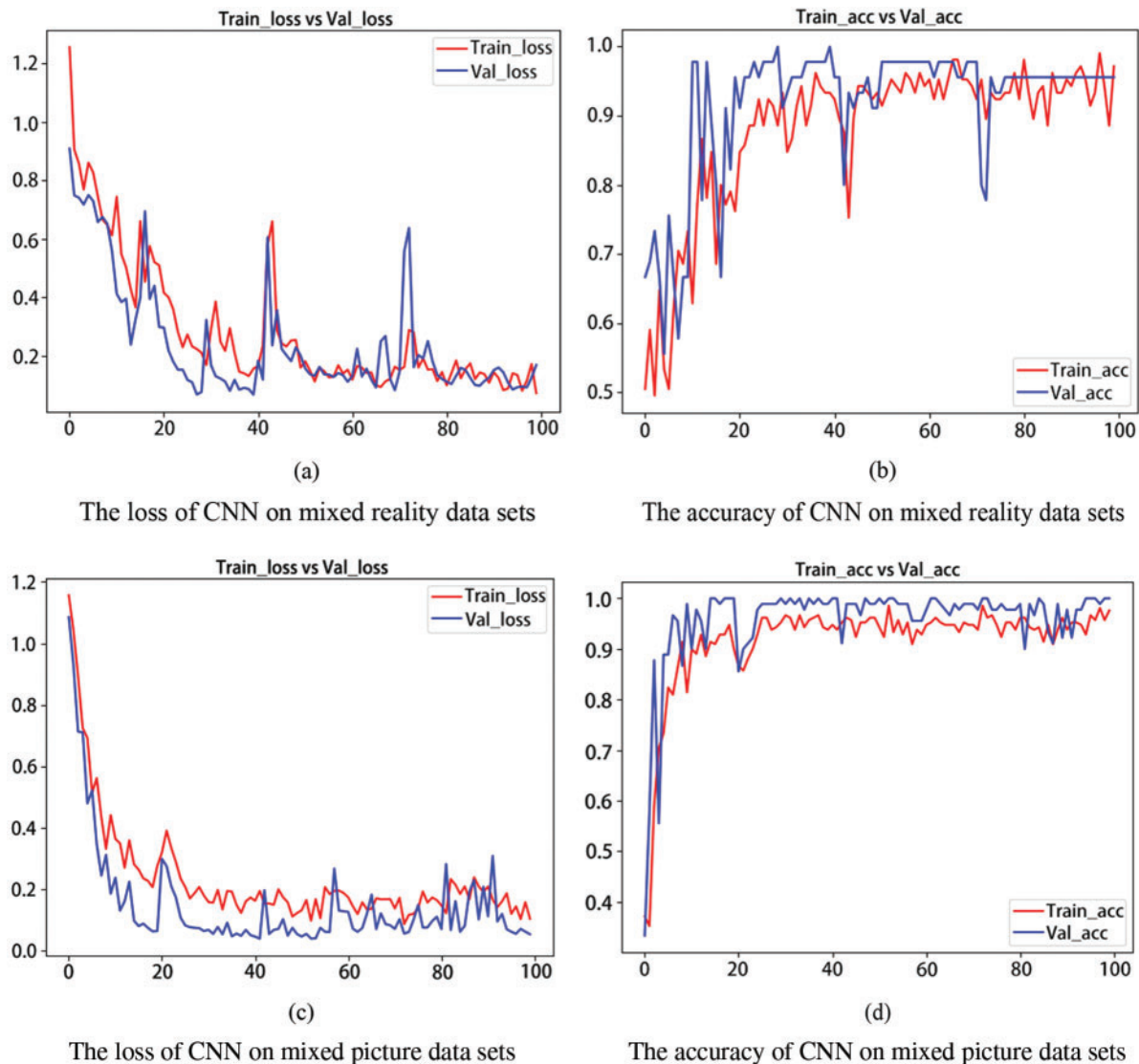


Figure 8: Performance of a non-federated learning system—CNN network for the comparison of the loss function of the training set and the validation set (left) and the comparison of the accuracy of the training set and the validation set (right)

4.2 Performance of Federated Learning System

As shown in Fig. 8, on 3D-reconstructed data sets, the traditional CNN network is unstable in both training and verification phases compared to meta learning, while on mixed picture data sets, convergence speed and final accuracy are superior to meta learning final results. The calculated loss values and accuracies on the training set loss function `train_loss` and the validation set loss function `val_loss` are shown respectively.

For federated meta-learning, the divided dataset is also divided into a training set and a testing set. In contrast with the non-federated learning system, the loss function of the federated meta-learning system (used for both training and testing sets of the model, respectively) converged very quickly

under the assumption of small samples. The CNN network did not converge under the same data input conditions, but in some cases the accuracies of the federated meta-learning system recognition fluctuated more. For an effective recognition and classification of damaged objects, it is necessary for the federated meta-learning system to be further integrated with FBLs, as shown in Fig. 9.

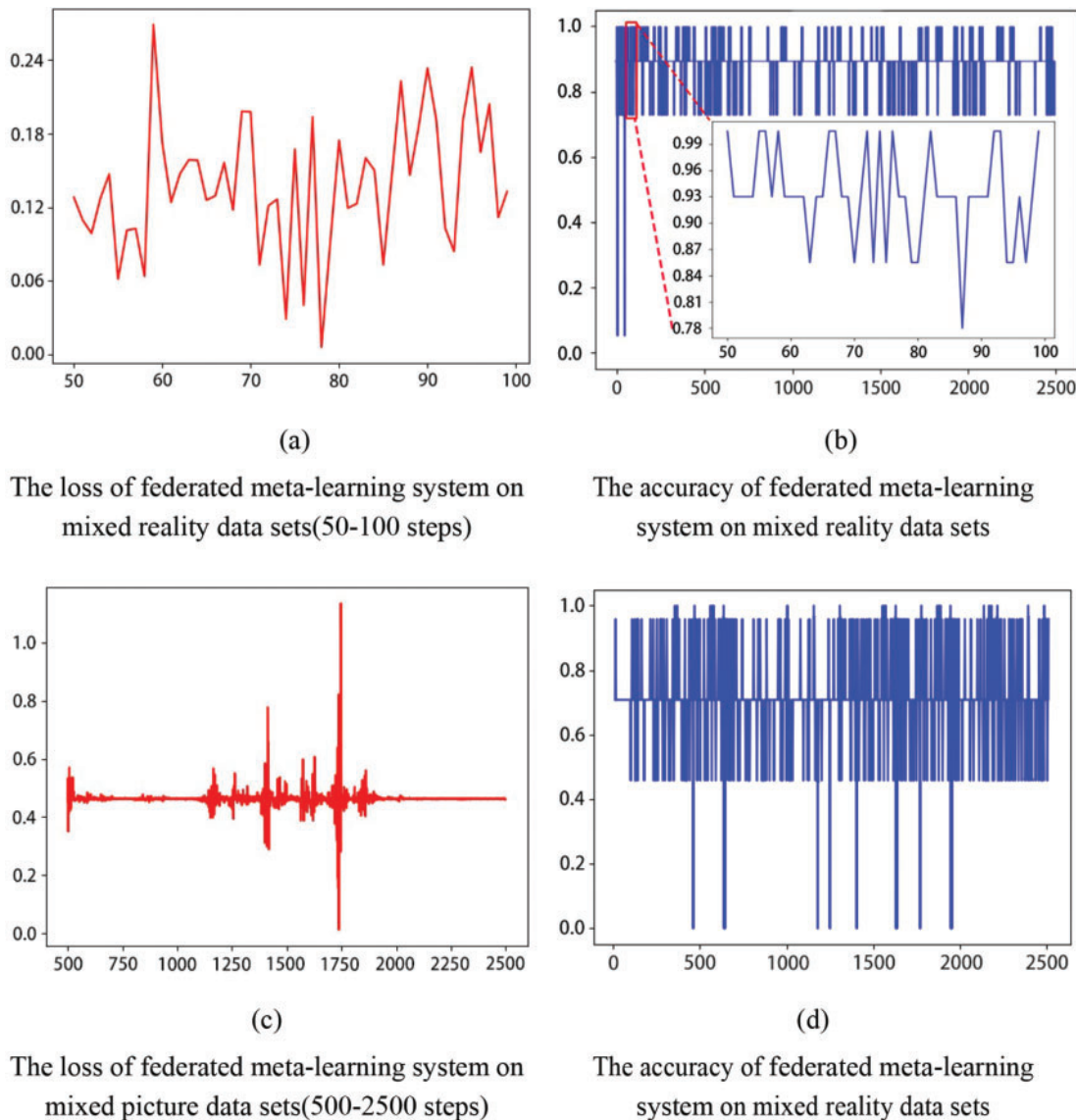


Figure 9: Performance of the federated meta-learning system with the comparison of the loss function of the training set and the validation set (left) and the comparison of the accuracy of the training set and the validation set (right). Due to the large initial loss of the system, the details after loss convergence will be compressed. Therefore, the figure starts from the step whose loss is less than 1

After integrated with FBLs, the broad federated meta-learning system become more efficient, as shown in Table 1.

Table 1: Performance of the broad federated meta-learning system from two users—in comparisons of training accuracy, training time, testing accuracy and testing time

| Damaged objects | Training accuracy | | Training time | | Testing accuracy | | Testing time | |
|---|-------------------|-------|---------------|-------|------------------|--------|--------------|-------|
| | User1 | User2 | User1 | User2 | User1 | User2 | User1 | User2 |
| Initial broad federated meta-learning system | | | | | | | | |
| Roof | 96.3% | 90.3% | 1.669 s | 1.663 | 97.8% | 91.5% | 0.233 s | 0.252 |
| Road | 71.5% | 70.5% | 1.408 s | 1.502 | 55.5% | 55.6% | 0.274 s | 0.274 |
| Lawn | 71.8% | 68.9% | 1.535 s | 1.524 | 76.4% | 60.37% | 0.221 s | 0.225 |
| Performance in server | 78.2% | | 1.563 s | | 72.9% | | 0.250 s | |
| Improved broad federated meta-learning system with morphological operators and damage instances | | | | | | | | |
| Roof | 96.3% | 90.3% | 1.669 s | 1.664 | 97.8% | 91.5% | 0.233 s | 0.252 |
| Road | 85.5% | 80.2% | 1.765 s | 1.740 | 95.5% | 80.5% | 0.269 s | 0.280 |
| Lawn | 82.8% | 76.3% | 1.865 s | 1.831 | 90.6% | 84.6% | 0.295 s | 0.301 |
| Performance in server | 85.2% | | 1.745 s | | 90.8% | | 0.277 s | |

Through actual observation, we believe that the damaged area can be determined as roof damage if the proportion of the damaged area reaches more than 10%, while for the classification of roads and lawns, we adjust the threshold to 25%. It can be seen from [Table 1](#) that the identification accuracy of roof damage is high, while the identification effect of road and lawn is poor. It can be seen from the analysis that most of the roofs are rusty, and the color difference between the damaged area and the surrounding area is obvious, while it is difficult to accurately identify whether the damaged area is damaged due to the close color between the damaged area and the surrounding area. Through further experiments, we found that the extraction of crack areas can better extract the damaged areas and improve the accuracy of identifying the damaged areas of roads and lawns by comparing them with specific instances of roads and lawns. As shown in the [Fig. 10](#), for the extraction of damaged areas, we further segment the image, introduce the threshold we set, convert the image set of the objects into a binary image, and then use the erosion and expansion operators in morphology to process the binary image, and compare it with the damaged examples in the example to judge the damage of the road.

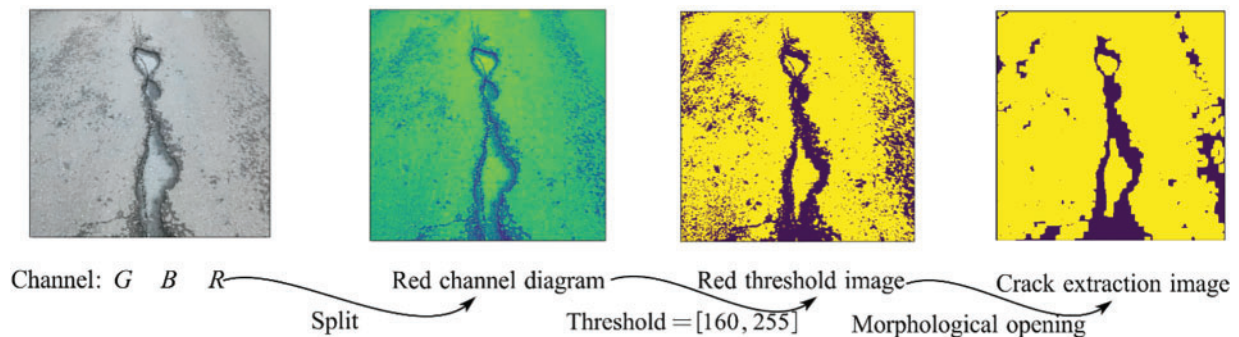


Figure 10: Introducing thresholds for crack extraction by morphological operators

As shown in [Table 1](#), the average accuracy of multiple classification of roof, road and grass is 90%, and the accuracy of damage identification of rusty roof directly reaches 96.3%. After adding damaged cases for pre-training and morphological operator extraction of damaged parts, the testing accuracy of judging whether the grass and road are damaged is above 90%. For the calculation of the area proportion of the damaged part, the accuracy of CNN classification converges at 90%, no matter the mixed reality data sets or the mixed picture set of aerial video segmentation images. The classification of damaged objects and the detection of the damaged area show good effect on the two types of data sets and the mixed data set, and the accuracy is higher than 90%. In comparisons of the performance on two different data sets, we found that meta-learning system works much better in the 3D-reconstructed data from the oblique photography of the UAV and the mixed reality than in the mixed picture set of aerial videos segmentation. This also highlighted the necessity to introduce CNN further as an alternative system for integrating with federated learning. We also met some overlapping data features in federated learning processes and in order to reduce overlapping samples, we introduced necessary thresholds to match the information and improve the model capability, as shown in [Figs. 11](#) and [12](#).

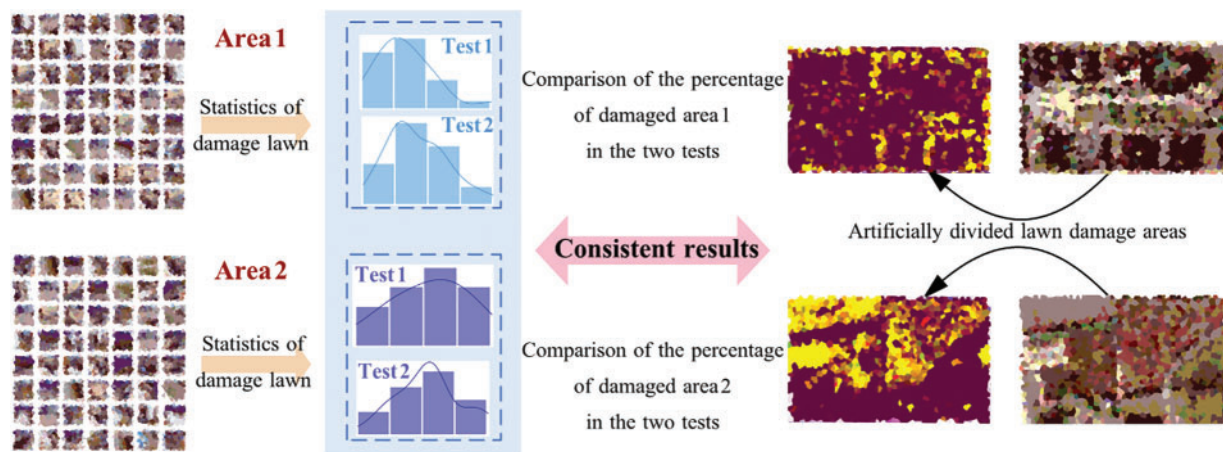


Figure 11: Improved the model capability—learning of damaged green belts as an example (Image encryption display)

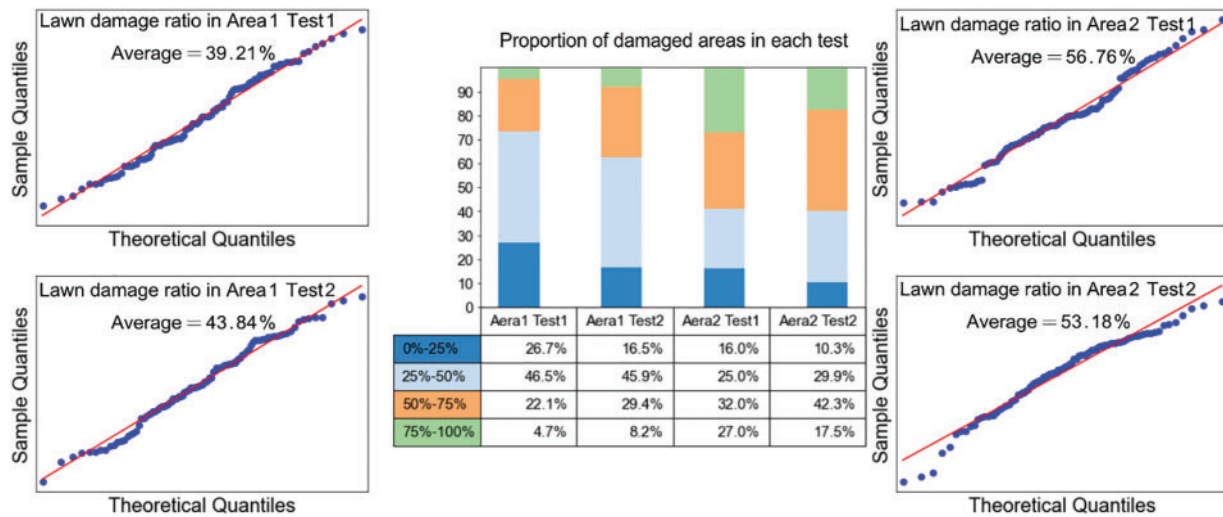


Figure 12: Learning results of four green areas with the improved model

Comparative analysis was adopted in the lawn damage ratio analysis. The residential area was divided into two parts manually and aerial sampling was conducted on the two parts, respectively. To be specific, we view the three-dimensional model established by UAV aerial photography in the two regions, respectively, according to the specific route, and carry out statistical analysis on it. It can be seen from the histogram that the green area roughly follows the rule of normal distribution. In this experiment, we calculated the lawn damage ratio by using the two-dimensional numerical map of the complete aerial map of manually divided areas. That is, the real lawn damage ratio could be obtained by calculating the area proportion of damaged lawn in the two-dimensional map. Finally, the experimental data are consistent with the average of the normal distribution of lawn damage ratio obtained by statistical analysis of the tilting camera model. Therefore, it is reasonable to calculate the lawn damage ratio of green belts by aerial photography method.

As can be seen from the figure above, after the video shot according to the specific route has been divided, the lawn damage ratio in the picture has been tested by Quantile-Quantile Plots, and the distribution of lawn damage ratio in the two areas is basically consistent with the expected curve of the normal red distribution. Therefore, we can use the normal distribution to calculate the proportion of damage of the whole damaged lawn. The area of the damaged lawn was manually divided and the proportion of damaged lawn area in the two areas was statistically analyzed using the binary graph as the true value. That is, the true value of damaged lawn rate in Area1 and Area2 was 38.35% and 55.73%, respectively. After four tests, the error of the average lawn damage ratio was 5.76%, indicating that the broad federated meta-learning system could accurately calculate the proportion of damaged areas.

According to the lawn damage image recognition classification method shown in 4.1, we classified the damage areas into 4 categories by the preliminary screening of the damage areas calculated by the algorithm and fine screening of the manually classified damage area proportions and sent these 4 categories of data to meta-learning and CNN for multi-classification, which can yield the above results. Overall, the CNN was more accurate in classifying the proportion of lawn area damaged, while meta-learning was faster but less robust.

4.3 Outstanding Questions and Next Research Priorities

Based on the virtual-real fusion model obtained in oblique photography, the image is extracted and segmented frame by frame. This provides sufficient data for federal learning [34–36]. At the same time, in order to verify the damage recognition accuracy of the model on UAV aerial images without virtual-real fusion, we also added the real aerial video images, forming a mixed data set of virtual reality and reality fusion. This data is in contrast to 3D reconstruction data in Federated Learning. In the decision-making part of federated learning, we use the meta-learning method to conduct multi-classification of segmented images to realize the detection of roofs, lawns and roads [37–40]. Further, for the results of multiple classifications, the fuzzy width learning method is used to distinguish the damage of the three, and the classification of the damage degree of the roof is realized directly. Then, according to the characteristics of road and lawn image damage, the damage detection algorithm is designed, and damage samples are introduced to pre-train the model, so as to improve the detection accuracy of damage [41–45]. In the second part, based on the 3D-reconstructed data, an improved broad federated meta-learning system with morphological operators and damage instances was used to calculate the proportion of the damaged area of regional lawns. And the damage degree is divided into four categories, respectively input meta-learning model and CNN training model. At the same time, we also verified the feasibility of classifying damaged images on realistic data sets for federation learning [46–50]. We define 3D-reconstructed data sets and Mixed picture data sets from User1 and User2, respectively. From Figs. 8 to 9 and Table 1, we weighed the training results of different users according to the number of pictures to measure the overall performance of the server. In the end, the average testing accuracy is 90.8% and the average testing time is 0.277 s.

In image classification, we use the mixed data set of virtual reality and real image segmentation to carry out experiments and adopt the method of meta-learning and fuzzy broad learning system to classify images. The Federal Learning Process [51] classifies images into three categories: roofs, lawns, and roads. Firstly, the modeling images obtained in oblique photography were segmented, manually marked, and sent into the meta-learning model for learning, and the classification results were obtained to verify the accuracy. In the damage section, we assessed the damage to roofs, roads and lawns. Secondly, we select the corresponding image set in the meta-learning multi-classification, use convolution to filter the damaged image set, and roughly determine the size of the damaged area. Then, according to the set damaged area threshold, the image is divided into lossless and damaged. These two kinds of images are then sent to the fuzzy width learning network for training and final classification. For damaged roofs, the color difference between the damaged area and the surrounding area is larger. The size of the damaged region can be determined by direct convolution filtering. For roads and lawns, morphological operators and damage instances need to be introduced to extract cracks. Then, we manually produced four data sets with different damage ratios and used the damaged area ratio of images as damage examples to train the improved broad federated meta-learning system. Finally, we applied the model to calculate the damage ratio of the two regions in the data set that was segmented after shooting according to specific routes and verified the damage ratio calculated by the model.

From the federated learning results, we can see that our algorithm for judging the damaged area and the algorithm for calculating the proportion of the damaged lawn area also have high accuracy in the real data set. Next, research priorities to future improve the federated learning system for detecting damaged objects can be described as follows. How to check the damage rate according to a uniform route to ensure learning accuracy? How to define the damage degree of different objects without a loss in the model performance? The federated learning accuracy of the accurate identification element of the damaged area of lawn is poor, which is not as good as the traditional CNN method and needs to be

further improved. Possible solutions are as follows [52–56]. For federated learning of the proportion of damaged lawn areas, we take the model of virtual and real fusion as the basis, view the model according to a certain route and extract the picture frame by frame and segment it [52]. First, calculate the greening rate, then use the algorithm in the previous part to extract the collection of damaged lawn images, and calculate the proportion of damaged lawn areas, and divide the damage degree into four categories, which are sent into the meta-learning and CNN training model. Finally, we made a data set with fixed damaged area using the public aerial photography data set, and verified the feasibility of classifying the damaged image according to the proportion of damaged area in the quality aerial photography image. In the calculation of the overall greening rate, first of all, we divide the observation area into two parts. In these two parts, we view and record the 3D model obtained by the combination of oblique photography and video images according to a certain route, and segment it frame by frame. Through convolution filtering, we identified the green area of the extracted picture, calculated the greening rate, and carried out statistical analysis on it. The experimental observation shows that its distribution follows the normal distribution, and the greening rate myopia obtained from the manual division of two parts of the observation area and accurate calculation of the green area, that is, the average value can be used to calculate the greening rate of the aerial photos of specific routes.

5 Conclusions

The traditional UAV mapping system has a low degree of automation and relies on manual mapping for damaged objects, which requires the measurement of plenty of control point position coordinates. Broad federated meta-learning systems can make full use of GPS positioning and inertial navigation in the role of IMU units in accurate positioning, providing accurate initial values for the error equation. The recognition accuracy is high, and the convergence speed is fast, but the accuracy rate can be improved. In the subsequent research, we can try to combine the advantages of the two to form an image classification and recognition system with fast convergence speed and high accuracy rate at the same time.

Funding Statement: This research was funded by the Strategic Priority Research Program of Chinese Academy of Sciences (XDA20060303) and the National Natural Science Foundation of China (41571299) and the High-Level Base-Building Project for Industrial Technology Innovation (1021GN204005-A06).

Availability of Data and Materials: All the data utilized to support the theory and models of the present study are available from the corresponding authors upon request.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Rho, S., Bharnitharan, K. (2014). Interactive scheduling for mobile multimedia service in M₂M environment. *Multimedia Tools and Applications*, 71(1), 235–246.
2. Moore-Ede, M., Heitmann, A., GuTTKuHN, R., Trutschel, U., Aguirre, A. et al. (2004). Circadian alertness simulator for fatigue risk assessment in transportation: Application to reduce frequency and severity of truck accidents. *Aviation, Space, and Environmental Medicine*, 75(3), A107–A118.
3. Owolabi, H. A., Bilal, M., Oyedele, L. O., Alaka, H. A., Ajayi, S. O. et al. (2018). Predicting completion risk in PPP projects using Big Data Analytics. *IEEE Transactions on Engineering Management*, 67(2), 430–453.

4. Ch, R., Srivastava, G., Gadekallu, T. R., Maddikunta P. K., R., Bhattacharya, S. (2020). Security and privacy of UAV data using blockchain technology. *Journal of Information security and Applications*, 55, 102670.
5. Li, Z., Sharma, V., Mohanty, S. P. (2020). Preserving data privacy via federated learning: Challenges and solutions. *IEEE Consumer Electronics Magazine*, 9(3), 8–16.
6. Chowdhury, T., Murphy, R., Rahneemoonfar, M. (2022). RescueNet: A high resolution UAV semantic segmentation benchmark dataset for natural disaster damage assessment. arXiv: 2202.12361.
7. Nasser, N., Fadlullah, Z. M., Fouda, M. M., Ali, A., Imran, M. (2022). A lightweight federated learning based privacy preserving B5G pandemic response network using unmanned aerial vehicles: A proof-of-concept. *Computer Networks*, 205(3), 108672. <https://doi.org/10.1016/j.comnet.2021.108672>
8. Aabid, A., Parveez, B., Parveen, N., Khan, S. A., Zayan, J. M. et al. (2022). Reviews on design and development of unmanned aerial vehicle (drone) for different applications. *Journal of Mechanical Engineering Research*, 45(2), 53–69.
9. Mishra, R. (2022). A Comprehensive review on involvement of drones in COVID-19 pandemics. *ECS Transactions*, 107(1), 341–351. <https://doi.org/10.1149/10701.0341ecst>
10. Abouzaid, L., Elbiaze, H., Sabir, E. (2022). Agile roadmap for application-driven multi-UAV networks: The case of COVID-19. *IET Networks*, 11(6), 195–206. <https://doi.org/10.1049/ntw2.12040>
11. Alsamhi, S. H., Almalki, F. A., Afghah, F., Hawbani, A., Shvetsov, A. V. et al. (2021). Drones' edge intelligence over smart environments in B5G: Blockchain and federated learning synergy. *IEEE Transactions on Green Communications and Networking*, 6(1), 295–312. <https://doi.org/10.1109/TGCN.2021.3132561>
12. Epstein, R. A. (2022). The private law connections to public nuisance law: Some realism about today's intellectual nominalism. *The Journal of Law, Economics & Policy*, 17, 282.
13. Snell, J., Swersky, K., Zemel, R. (2017). Prototypical networks for few-shot learning. *Advances in Neural Information Processing Systems*, 30, 4080–4090.
14. Feng, S., Chen, C. L. P. (2018). Fuzzy broad learning system: A novel neuro-fuzzy model for regression and classification. *IEEE Transactions on Cybernetics*, 50(2), 414–424. <https://doi.org/10.1109/TCYB.2018.2857815>
15. Poullis, C., You, S. (2008). Photorealistic large-scale urban city model reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, 15(4), 654–669. <https://doi.org/10.1109/TVCG.2008.189>
16. Chen, C. P., Liu, Z. (2017). Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. *IEEE Transactions on Neural Networks and Learning Systems*, 29(1), 10–24. <https://doi.org/10.1109/TNNLS.2017.2716952>
17. Sugeno, M., Takagi, T. (1993). Fuzzy identification of systems and its applications to modelling and control. *Readings in Fuzzy Sets for Intelligent Systems*, 15(1), 387–403.
18. Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A. et al. (2019). Towards federated learning at scale: System design. *Proceedings of Machine Learning and Systems*, 1, 374–388.
19. Li, T., Sahu, A. K., Talwalkar, A., Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60. <https://doi.org/10.1109/MSP.2020.2975749>
20. Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W. et al. (2021). A survey on federated learning. *Knowledge-Based Systems*, 216(1), 106775. <https://doi.org/10.1016/j.knosys.2021.106775>
21. Schenk, T. (1997). Towards automatic aerial triangulation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(3), 110–121. [https://doi.org/10.1016/S0924-2716\(97\)00007-5](https://doi.org/10.1016/S0924-2716(97)00007-5)
22. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M. et al. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, 14(1–2), 1–210. <https://doi.org/10.1561/22000000083>
23. Li, L., Fan, Y., Tse, M., Lin, K. Y. (2020). A review of applications in federated learning. *Computers & Industrial Engineering*, 149(5), 106854. <https://doi.org/10.1016/j.cie.2020.106854>

24. Zhou, Z., Meng, M., Zhou, Y. (2017). Massive video integrated mixed reality technology. *ZTE Technology Journal*, 23(6), 10–13.
25. Giese, M. A., Poggio, T. (1999). Synthesis and recognition of biological motion patterns based on linear superposition of prototypical motion sequences. *Proceedings IEEE Workshop on Multi-View Modeling and Analysis of Visual Scenes (MVIEW'99)*, pp. 73–80. Fort Collins, USA.
26. Kang, J., Xiong, Z., Niyato, D., Zou, Y., Zhang, Y. et al. (2020). Reliable federated learning for mobile networks. *IEEE Wireless Communications*, 27(2), 72–80. <https://doi.org/10.1109/MWC.001.1900119>
27. Brazdil, P. B., Soares, C., Costa, J. (2003). Ranking learning algorithms: Using IBL and meta-learning on accuracy and time results. *Machine Learning*, 50(3), 251–277. <https://doi.org/10.1023/A:1021713901879>
28. Giraud-Carrier, C., Vilalta, R., Brazdil, P. (2004). Introduction to the special issue on meta-learning. *Machine Learning*, 54(3), 187–193. <https://doi.org/10.1023/B:MACH.0000015878.60765.42>
29. Prudêncio, R. B. C., Ludermir, T. B. (2004). Meta-learning approaches to selecting time series models. *Neurocomputing*, 61(1), 121–137. <https://doi.org/10.1016/j.neucom.2004.03.008>
30. Zhang, K., Han, Y., Chen, J., Zhang, Z., Wang, S. (2020). Semantic segmentation for remote sensing based on RGB images and lidar data using model-agnostic meta-learning and partial Swarm optimization. *IFAC-PapersOnLine*, 53(5), 397–402. <https://doi.org/10.1016/j.ifacol.2021.04.117>
31. Kordík, P., Koutník, J., Drchal, J., Kovářík, O., Čepek, M. et al. (2010). Meta-learning approach to neural network optimization. *Neural Networks*, 23(4), 568–582. <https://doi.org/10.1016/j.neunet.2010.02.003>
32. Brewer, L., Horgan, F., Hickey, A., Williams, D. (2013). Stroke rehabilitation: Recent advances and future therapies. *QJM: An International Journal of Medicine*, 106(1), 11–25. <https://doi.org/10.1093/qjmed/hcs174>
33. Wang, W. F., Zhang, J., An, P. (2022). EW-CACTUs-MAML: A robust metalearning system for rapid classification on a large number of tasks. *Complexity*, 2022, 7330823.
34. Tutmez, B. (2009). Use of hybrid intelligent computing in mineral resources evaluation. *Applied Soft Computing*, 9(3), 1023–1028. <https://doi.org/10.1016/j.asoc.2009.02.001>
35. Lyu, L., Yu, H., Yang, Q. (2020). Threats to federated learning: A survey. arXiv: 2003.02133.
36. Huang, D. S., Wunsch, D. C., Levine, D. S., Jo, K. H. (2007). *Advanced intelligent computing theories and applications-with aspects of theoretical and methodological issues*. Switzerland: Springer Press.
37. Kwon, O. (2006). The potential roles of context-aware computing technology in optimization-based intelligent decision-making. *Expert Systems with Applications*, 31(3), 629–642. <https://doi.org/10.1016/j.eswa.2005.09.075>
38. Jolly, K. G., Ravindran, K. P., Vijayakumar, R., Kumar, R. S. (2007). Intelligent decision making in multi-agent robot soccer system through compounded artificial neural networks. *Robotics and Autonomous Systems*, 55(7), 589–596. <https://doi.org/10.1016/j.robot.2006.12.011>
39. Jabeur, K., Guitouni, A. (2009). A generalized framework for multi-criteria classifiers with automated learning: Application on FLIR ship imagery. *Journal of Advances in Information Fusion*, 4(2), 75–92.
40. Ghosh, A., Chung, J., Yin, D., Ramchandran, K. (2020). An efficient framework for clustered federated learning. *Advances in Neural Information Processing Systems*, 33, 19586–19597.
41. Brisimi, T. S., Chen, R., Mela, T., Olshevsky, A., Paschalidis, I. C. et al. (2018). Federated learning of predictive models from federated electronic health records. *International Journal of Medical Informatics*, 112(3–4), 59–67. <https://doi.org/10.1016/j.ijmedinf.2018.01.007>
42. Mammen, P. M. (2021). Federated learning: Opportunities and challenges. arXiv: 2101.05428.
43. Lee, K., Maji, S., Ravichandran, A., Soatto, S. (2019). Meta-learning with differentiable convex optimization. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10657–10665. Long Beach, USA.
44. Aledhari, M., Razzak, R., Parizi, R. M., Saeed, F. (2020). Federated learning: A survey on enabling technologies, protocols, and applications. *IEEE Access*, 8, 140699–140725. <https://doi.org/10.1109/ACCESS.2020.3013541>

45. Goetz, J., Malik, K., Bui, D., Moon, S., Liu, H. et al. (2019). Active federated learning. arXiv:1909.1264.
46. Chen, M., Yang, Z., Saad, W., Yin, C., Poor, H. V. et al. (2020). A joint learning and communications framework for federated learning over wireless networks. *IEEE Transactions on Wireless Communications*, 20(1), 269–283. <https://doi.org/10.1109/TWC.2020.3024629>
47. Nguyen, H. T., Schwag, V., Hosseinalipour, S., Brinton, C. G., Chiang, M. et al. (2020). Fast-convergent federated learning. *IEEE Journal on Selected Areas in Communications*, 39(1), 201–218. <https://doi.org/10.1109/JSAC.2020.3036952>
48. Li, B., Zhang, L. L. (1999). A new method for global optimal value estimation. *Journal of Shandong Institute of Mining & Technology (Natural Science Edition)*, 18(2), 107–109.
49. AbdulRahman, S., Tout, H., Ould-Slimane, H., Mourad, A., Talhi, C. et al. (2020). A survey on federated learning: The journey from centralized to distributed on-site learning and beyond. *IEEE Internet of Things Journal*, 8(7), 5476–5497. <https://doi.org/10.1109/JIOT.2020.3030072>
50. Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J. et al. (2021). Federated learning for internet of things: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 23(3), 1622–1658. <https://doi.org/10.1109/COMST.2021.3075439>
51. Cheng, K., Fan, T., Jin, Y., Liu, Y., Chen, T. et al. (2021). Secureboost: A lossless federated learning framework. *IEEE Intelligent Systems*, 36(6), 87–98. <https://doi.org/10.1109/MIS.2021.3082561>
52. Khan, L. U., Saad, W., Han, Z., Hossain, E., Hong, C. S. (2021). Federated learning for Internet of Things: Recent advances, taxonomy, and open challenges. *IEEE Communications Surveys & Tutorials*, 23(3), 1759–1799. <https://doi.org/10.1109/COMST.2021.3090430>
53. Charles, Z., Garrett, Z., Huo, Z., Shmulyian, S., Smith, V. (2021). On large-cohort training for federated learning. *Advances in Neural Information Processing Systems*, 34, 20461–20475.
54. Hasnat, M. A., Alata, O., Tremeau, A. (2016). Model-based hierarchical clustering with Bregman divergences and Fishers mixture model: Application to depth image analysis. *Statistics and Computing*, 26(4), 861–880. <https://doi.org/10.1007/s11222-015-9576-3>
55. Zhan, Y., Li, P., Qu, Z., Zeng, D., Guo, S. (2020). A learning-based incentive mechanism for federated learning. *IEEE Internet of Things Journal*, 7(7), 6360–6368. <https://doi.org/10.1109/JIOT.2020.2967772>
56. Pillutla, K., Kakade, S. M., Harchaoui, Z. (2022). Robust aggregation for federated learning. *IEEE Transactions on Signal Processing*, 70, 1142–1154. <https://doi.org/10.1109/TSP.2022.3153135>