

## Research Article

# Multiperspective Light Field Reconstruction Method via Transfer Reinforcement Learning

Lei Cai <sup>1</sup>, Peien Luo <sup>2</sup>, Guangfu Zhou <sup>2</sup>, Tao Xu <sup>1</sup> and Zhenxue Chen <sup>3</sup>

<sup>1</sup>School of Artificial Intelligence, Henan Institute of Science and Technology, Xinxiang 453003, China

<sup>2</sup>School of Information Engineering, Henan Institute of Science and Technology, Xinxiang 453003, China

<sup>3</sup>School of Control Science and Engineering, Shandong University, Jinan 250061, China

Correspondence should be addressed to Lei Cai; cailei2014@126.com

Received 9 July 2019; Revised 18 December 2019; Accepted 13 January 2020; Published 31 January 2020

Academic Editor: Cornelio Yáñez-Márquez

Copyright © 2020 Lei Cai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Compared with traditional imaging, the light field contains more comprehensive image information and higher image quality. However, the available data for light field reconstruction are limited, and the repeated calculation of data seriously affects the accuracy and the real-time performance of multiperspective light field reconstruction. To solve the problems, this paper proposes a multiperspective light field reconstruction method based on transfer reinforcement learning. Firstly, the similarity measurement model is established. According to the similarity threshold of the source domain and the target domain, the reinforcement learning model or the feature transfer learning model is autonomously selected. Secondly, the reinforcement learning model is established. The model uses multiagent (i.e., multiperspective) Q-learning to learn the feature set that is most similar to the target domain and the source domain and feeds it back to the source domain. This model increases the capacity of the source-domain samples and improves the accuracy of light field reconstruction. Finally, the feature transfer learning model is established. The model uses PCA to obtain the maximum embedding space of source-domain and target-domain features and maps similar features to a new space for label data migration. This model solves the problems of multiperspective data redundancy and repeated calculations and improves the real-time performance of maneuvering target recognition. Extensive experiments on PASCAL VOC datasets demonstrate the effectiveness of the proposed algorithm against the existing algorithms.

## 1. Introduction

Light field [1, 2] is the parametric representation of a 4D light radiation field that contains the position and direction information in space. In other words, the light field contains all images of the same object taken at different positions and different angles.

The light field can be used as a feature library of the target images. The multiperspective light field is a complete light field represented by multiple perspectives. The multimodal fusion of the captured target information is performed through the cooperation mechanism among different perspectives, and finally, the dataset of the multiperspective light field is obtained. With the rapid development of light field reconstruction technologies, the target recognition employing the light field feature library has been widely

studied in the fields of computer vision, pattern recognition, and image processing. In the military field, the target recognition is applied to counterterrorism, missile guidance, search and rescue, and target monitoring in the airspace. In the industrial field, the target recognition is employed to the robot navigation, industrial part detection, and assembly line. In the civil field, the target recognition is applied to biomedicine, intelligent transportation, and motion tracking. Modal perception in the area of the computer is a channel that uses computer technology to imitate human beings and make machines connect with the physical world. Environment segmentation is to use computer technology to segment the acquired environment according to different categories. Finally, the formal representation and description of the segmented image are called feature representation. With the popularization of artificial intelligence

technology, many learning methods have been widely investigated, for example, reinforcement learning [3], transfer learning [4], and antagonistic learning [5]. The successful application of artificial intelligence technology in the field of image recognition is no more than deep learning [6]. In recent years, researchers have trained image samples through a convolution neural network [7] and other deep learning models, which improve the accuracy of target recognition.

However, the complex computing process seriously affects the real-time performance and accuracy of target recognition. To solve the problem, this paper proposes the multiperspective light field reconstruction method based on transfer reinforcement learning, as shown in Figure 1.

The transfer reinforcement learning algorithm is free from the constraints of large data and large sample training. The transfer learning and reinforcement learning are selected independently according to similarity thresholds based on the samples of tags with limited source domains.

The main contributions of this paper are as follows:

- (1) This paper introduces reinforcement learning and transfer learning algorithms into the area of light field reconstruction. According to the similarity measure model, the reinforcement learning and transfer learning are autonomously chosen.
- (2) The reconstruction method for the multiperspective light field via transfer reinforcement learning effectively reduces the calculation of similar data in the reconstruction process, shortens the time of light field reconstruction, and improves the real-time performance of target recognition.
- (3) This paper conducts extensive experiments to evaluate the performance of the proposed method in target recognition. The experimental result shows that this method is superior to existing methods.

The remainder of this paper is organized as follows: Section 2 provides an overview of the related literature. Section 3 describes the multiperspective representation and source-domain establishment of the light field, the transfer reinforcement learning algorithm, and the reconstruction method for the multiperspective light field based on transfer reinforcement learning. The simulation experiments are discussed in Section 4. Finally, Section 5 concludes this paper.

## 2. Related Works

The complex computing process of big data affects the real-time performance and accuracy of data processing. Many researchers have proposed different optimization algorithms. Literature [8] proposes a novel deep learning framework for attribute prediction in the wild. This framework not only outperforms the state-of-the-art with a large margin but also reveals valuable facts on learning face representation. Literature [9–11] proposes the reinforcement learning model, which matches behavioral data to explain the observed strategic behavior. The results show that this method is more effective than similar methods. The

effect of transfer learning depends to some extent on common factors between learning materials. Literature [12–14] eliminates the training needs of similar learning by learning the transferable structure of scalable image recognition tasks, which significantly reduces the training parameters and the error rate and shortens the training time. In the area of the light field, it is difficult to obtain a complete light field. The huge amounts of data and complex scenes have been obstacles to build light fields.

Many researchers have proposed the concept of light field reconstruction, using the sparsity of the light field to reconstruct the unknown information of the targets from locally known information. Cai et al. [15] proposed a target recognition method based on multiperspective reconstruction, through which the light field data from multiple perspectives are fused and the consistency of the average of all observations is obtained. The experiment shows the robustness and reliability of solving large-scale complex problems with the collaboration of multiple agents. The literature [16–18] uses the features of wavelet transform and multiresolution analysis to effectively suppress the window effect on reconstruction results.

The reconstruction efficiency is improved. In [19], light field imaging is used to solve underwater imaging problems under low-intensity lighting conditions. By using depth convolutional neural networks, the problem of scattering in the light field is effectively solved. By combining light field imaging with structured illumination to perform multiperspective depth measurements, many light field reconstruction methods are proposed [20–23]. The flexible calibration strategy is accordingly designed to determine the mapping coefficients for each light field ray, enabling efficient 3D reconstruction. The angle superresolution method [24–26] is used to capture the light field by using a sparse camera array. It makes use of the compressive sensing reconstruction to collect the dense light field.

Target recognition is an important application in the field of computer vision. Literature [4] proposes a combination method for target tracking with deep learning and preference learning. This method solves the problem of target position and size change during target tracking. It effectively finds out the target object in each frame of the video. Literature [16] introduces light field reconstruction technology to the field of target recognition. The method uses sampled images to reconstruct the light field image and finally tests its validity in the field of target recognition.

To sum up, many researchers have made useful improvements in optimizing algorithms for light field reconstruction. Moreover, they successfully introduced deep learning into light field reconstruction. However, there are still problems: (1) duplicate processing of large amounts of data is constraining real-time performance of target recognition and (2) it is still a challenge to obtain the required data for reconstructing the complete light field.

## 3. Proposed Method

*3.1. Multiperspective Representation of the Light Field.* The Lego Bulldozer of the light field library of the Stanford

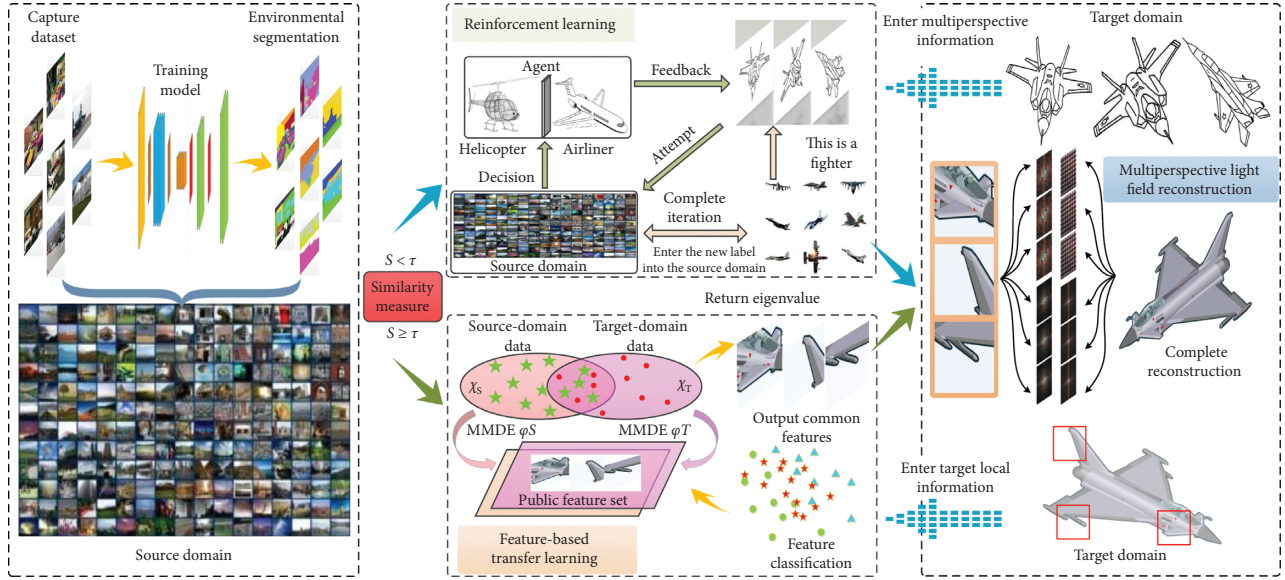


FIGURE 1: Overall flow chart of the multiperspective light field reconstruction method via transfer reinforcement learning.

University is represented as a weight graph  $G = (v, \varepsilon)$ , a given image. The node  $v$  is the pixels of the image, the light field domain  $\varepsilon$  is the choice of the neighborhood structure, and the weight of the domain is  $e_{ij} \in \varepsilon$ . Define each view field threshold as  $C$ , then the minimum generalization function  $E(C)$  is

$$E(C) = \int_C -|\nabla I(C(\varepsilon))|^2 + \alpha|C_\varepsilon(\varepsilon)|^2 + \beta|C_{\varepsilon\varepsilon}(\varepsilon)|^2 d\varepsilon, \quad (1)$$

where  $\nabla I$  is the standard that the threshold value is in the gray gradient area,  $\alpha$  is a random parameter, and  $\beta$  is a weighting function. Specific parameters are added to the threshold, i.e., minimizing functional calculation of the segmentation smooth approximation  $u$  of the gamma function  $I$ . Let  $I: \Omega \rightarrow \mathbb{R}$  be the gray value input map in each light field.

$$\min_{\Omega_i} \left\{ \frac{1}{2} \sum_{i=0}^k \text{per}_g(\Omega_i; \Omega) + \sum_{i=0}^k \int_{\Omega_i} f_i(x) dx \right\} \bigcup_{i=0}^k \Omega_i. \quad (2)$$

In this paper, the convex representation of image segmentation is introduced into multiperspective light field reconstruction [27]. The region  $\Omega_i$  in the above formula is represented by the labeling function  $u: \Omega \rightarrow \{0, \dots, k\}$ . The  $k$  binary functions  $\theta(x) = (\theta_1(x), \dots, \theta_k(x))$  are equivalent to the multilabel function. Finally, the marker function  $u$  is recovered from these sequential functions:

$$u(x) = \sum_{i=1}^k \theta_i(x). \quad (3)$$

Therefore, the light field model is divided into many perspectives, as shown in Figure 2.

The multiagent (i.e., multiperspective) method is used to perceive the environment for establishing the source domain, which is also called the instinctive database. It sets the basic action ability for the agent so that the agent is capable of environment interaction and trial-and-error action without

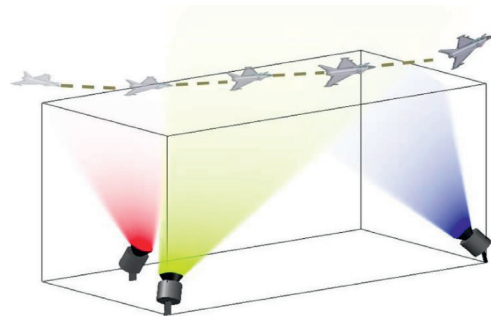


FIGURE 2: Multiperspective light field 3D representation model. The method uses red, yellow, and blue visible regions to represent the visual range of the multiperspective light field, respectively. In the process of moving the target, the visual range of each perspective and the collaborative relationship between multiple perspectives are embodied.

training. The multiagent method integrates the collected environmental information and finally completes the establishment of the source domain. According to the content of the environment, the perceptual environment is modeled to train the image. After the image dataset training is completed, the image information is subject to modal analysis. The image is segmented into different categories based on its threshold [28–31]. At present, most segmentation methods are based on pixel-based SoftMax classification. The relationship among pixel points is described by a binary function. Through the relationship, the same labels are assigned to similar pixels and different labels are assigned to pixels with larger gaps so that the image boundary is effectively segmented.

$$\psi_p(x_i, x_j) = u(x_i, x_j) \sum_{m=1}^M W^m k_G^m(f_i, f_j). \quad (4)$$

$u_q(x; W)$  represents the score of the pixel  $x$  belonging to the category  $q$ . The probability of the pixel class contained in the image is output by the SoftMax function:

$$S(x) = p(q|x, W) = \frac{\exp(u_q(x; W))}{\sum_{K=1}^K \exp(u_q(x; W))}, \quad (5)$$

where  $S(x)$  represents the result of a light field image segmentation from a single perspective. The light field images captured from multiple perspectives are normalized, and the feature information fusion  $u_r$  from multiple perspectives is completed:

$$u_r = [S(x)_1, S(x)_2, \dots, S(x)_N]^T. \quad (6)$$

To add a label to the split body, the label set is called the source domain, as shown in Figure 3. The segmentation method proposed in this paper allows for a wider range of images and lays a good foundation for enhancing the ability of light field reconstruction.

**3.2. Transfer Reinforcement Learning Algorithm.** Duplicate processing of large amounts of data is constraining real-time performance of target recognition. In this paper, the transfer reinforcement learning algorithm is proposed to implement an agent autonomous selection learning model with the ability of analysis, judgment, decision-making, and execution. The target samples captured by the multiple perspectives are traversed in the source domain, and the corresponding learning strategies are selected according to the thresholds of the similar metric models. If the similarity is less than the threshold, reinforcement learning is performed. The decision is adjusted by a small amount of known information and the environment interaction feedback, and finally, the unlabeled sample is iterated with the labeled sample. The new label samples are added to the source domain, and the sample size of the source domain is continuously expanded to improve the environmental cognitive ability and scene-understanding ability. If the similarity is greater than or equal to the threshold, the target-domain sample is directly determined by the feature transfer learning method. This method effectively reduces the number and time of interactions of target-domain samples in the source domain and ensures the real-time performance of target recognition.

### 3.2.1. Establishment of the Similarity Measurement Model.

The perceptual hash algorithm is a mapping of key data in a digital image into a short length sequence. According to the response of machine vision to images in different environments, the perceptual hash algorithm relies on the similarity of the scene to give the corresponding hash value [32]. In this paper, the similarity measures of the perceptual hash algorithm are used to autonomously select reinforcement learning and transfer learning. The specific methods are as follows:

- (1) The observation image captured by the above multiperspective  $u(x)$  is preprocessed, and the image size is adjusted to  $256 * 256$ .

- (2) The feature extraction is performed on the multiperspective image  $u(x) = \{u | x = 1, 2, \dots, n\}$ , and the feature vectors  $R = \{R_1, R_2, \dots, R_n\}$  are obtained, where  $R_i$  represents a feature point vector in the graph. The compression of the feature matrix is achieved by summing the multiperspective feature matrices:

$$H(i) = \sum_{x=1}^n R_{i,x}, \quad 1 \leq i \leq 128. \quad (7)$$

In this paper, the feature clustering analysis is used to quantify  $H$ , and the image information is mapped to 0 or 1 according to the feature clustering threshold  $t$ . Finally, the hash value  $H$  is obtained.

$$h(i) = \begin{cases} 1, & R_{i,x} \geq t, \\ 0, & R_{i,x} < t. \end{cases} \quad (8)$$

We use the Hamming distance to judge the similarity among images from multiple perspectives according to the above hash value. Let the multiperspective image be  $h_n$ , and then the Hamming distance between the image and the feature environment information  $h_0$  is  $D = h_n - h_0$ . Then, the similarity model of multiperspective images is

$$S = \frac{i - D}{i}. \quad (9)$$

Finally, this paper sets the similarity threshold to  $\tau$ . When the image feature value is less than  $\tau$ , the model determination features are not similar and the reinforcement learning algorithm is selected; when the image feature value is greater than or equal to  $\tau$ , the model determination features are similar and the transfer learning algorithm is selected.

**3.2.2. Reinforcement Learning Algorithm.** Reinforcement learning should try to make a judgment, adjust the previous behavior in the process of interacting with the environment, and then complete the recognition of the target sample of continuous iteration. It is assumed that each step of reinforcement learning has corresponding observations, and the action that can be performed is supported by a small number of label samples in the source domain. The implementation of each step requires a combination of previous actions and observations to take action [33]. Then, the specific steps of reinforcement learning are given below. To simplify the description, the following symbols are defined:

- (1)  $x_t \in R^{m*n}$  is the corresponding observation image when the target and environment interact at step  $t$  ( $t = 1, 2, \dots, T$ ).
- (2)  $a_t \in \Lambda$  is the action performed by the observation  $x_t$ , where  $\Lambda$  is the set of all behaviors under the reinforcement learning rule.
- (3)  $r_t$  is the feedback obtained after performing an action  $a_t$  under the observation  $x_t$ .

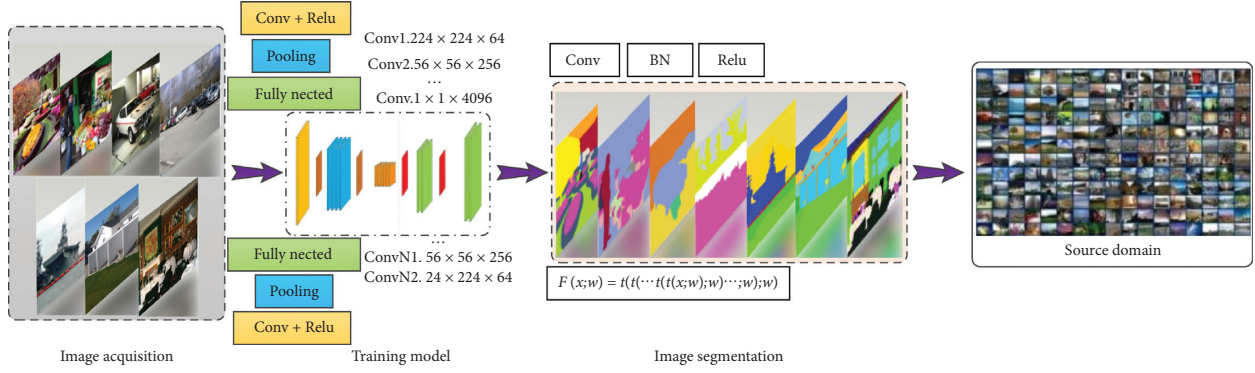


FIGURE 3: Establishment of the source domain. The acquired image is trained to obtain the state after the image is segmented, and the interpretation of the environment is output.

$$R_t = \sum_{t'=t}^T \gamma^{(t'-t)} \cdot r_{t'}, \quad (10)$$

where  $R_t$  is the collection of all feedback obtained from step  $t$  to the end time and  $\gamma$  is the supervisory budget representing the limitation of computing power and time. In addition, the state  $s$  at a certain time  $t$  is  $s_t = (x_1, a_1, \dots, x_{t-1}, a_{t-1}, x_t)$ . Then, the main idea of reinforcement learning is to realize the optimization learning of action state function based on the iteration of  $Q$ -learning:

$$\begin{cases} Q_{k+1}(s_t, a_t) = Q_k(s_t, a_t) + \alpha_k \cdot \delta_k, \\ \delta_k = r_{t+1} + \gamma \cdot \max_{a' \in \Lambda} Q_k(s_{t+1}, a') - Q_k(s_t, a_t), \end{cases} \quad (11)$$

where  $\alpha_k$  is the learning rate,  $s_t$  and  $a_t$  are the states and actions corresponding to step  $t$ , respectively, and  $\delta_k$  is the time difference.  $a'$  is the action that  $\Lambda$  can perform in  $s_{t+1}$ . The state action value function under the optimal control strategy is obtained as the number of iterations tends to infinity, and the optimal execution strategy in each state is summarized. Finally, the strategy for selecting the optimal executable action in a certain state is to maximize the expected value:

$$Q^*(s, a) = E_{s' \sim \xi} \left( r + \gamma \cdot \max_{a'} Q^*(s', a') \mid s, a \right), \quad (12)$$

where  $\xi$  is the environment,  $s'$  is the state after  $s$  performs an action  $a$ , and  $a'$  is all possible actions of the state  $s'$ . The above formula uses the expectation to analyze the state action value function and uses  $Q(s, a, \theta) \approx Q^*(s, a)$  to estimate the state action function. In reinforcement learning, the  $Q$ -network implements parameter updation by minimizing the objective function.

$$L_k(\theta_k) = E_{s, a \sim \rho(\cdot)} [y_k - Q(s, a, \theta_k)]^2. \quad (13)$$

Here,  $\rho(s, a)$  is the probability distribution of the state  $s$  and behavior  $a$ , and  $y_k$  is the target output corresponding to the  $k$ th iteration and is given by

$$y_k = E_{s' \sim \xi} \left( r + \gamma \cdot \max_{a'} Q(s', a', \theta_{k-1}) \mid s, a \right). \quad (14)$$

Let all subjects evaluate the results. If the objective function obtained by the subject under various constraints is smaller, the subject should receive a larger reward:

$$R'(s'_k, s'_{k+1}, a'_k) = \begin{cases} \frac{W}{f_{\text{Best}}}, & (s'_k, a'_k) \in y_k, \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

where  $f_{\text{Best}}$  represents the fitness function of the optimal state at the  $k$ th iteration and  $W$  is a positive constant; the smaller the value of the objective function, the larger the reward value. In this paper, based on the experience of previous learning, the prompt for the next target environment interaction is  $e_t$ . Assume that the number of steps at the end of reinforcement learning is  $N$  and the set of reviews for experience is  $D = [e_1, e_2, \dots, e_N]$ . We feed  $D$  back to the source domain, increasing the amount of data of known samples in the source domain. Therefore, the behavior state action value in the transfer reinforcement learning algorithm is changed to  $\text{TRL}(s, a, \theta) \rightarrow \text{TRL}(\varphi(s), a, \theta)$ .  $X$  is the feature learning in the transfer reinforcement learning algorithm, so the state of the step  $t$  in the transfer reinforcement learning is  $s_{t+1} = (s_t, a_t, x_{t+1})$ . Finally, the modified equation for the interaction of the target environment in reinforcement learning is obtained:

$$\begin{cases} D \rightarrow \bar{D} = [\bar{e}_1, \bar{e}_2, \dots, \bar{e}_N], \\ \bar{e}_t = (\varphi(s_t), a_t, r_t, \varphi(s_{t+1})) \end{cases} \quad (16)$$

The red box represents the target information captured from multiple perspectives. The local information captured from each perspective is preliminarily judged as a helicopter or an airliner, and the target information and the environment keep interacting to conclude that the target is a fighter. We will input new label samples into the source domain after reinforcement learning, continuously update the source-domain sample data, expand the label sample size, and improve the environmental cognitive ability and scene-understanding ability, as shown in Figure 4.

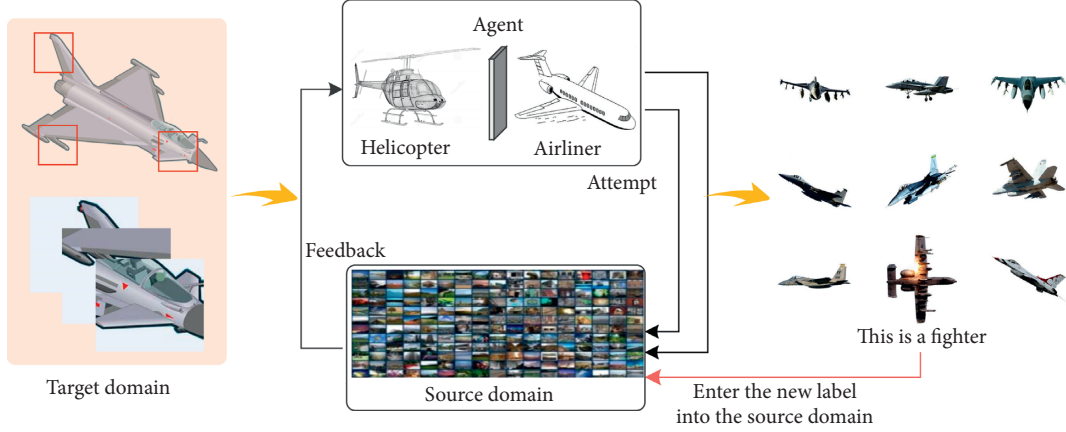


FIGURE 4: Reinforcement learning of unlabeled samples in the target domain.

3.2.3. *Transfer Learning Algorithm.* The multiperspective image is taken as the target domain. By comparing with the label samples in the source domain, the feature transfer method in the transfer learning is selected to recognize the environment of the multiperspective samples [34, 35]. Let the source domain after reinforcement learning be  $D_s = \{x_s, y_s\}$ . The information captured by multiple perspectives is the target domain  $D_T = \{x_{T_i}\}$ . As  $Y_s$  continues to increase, there will be more and more  $Y_T$  in  $D_T$  similar to  $Y_s$ . In order to avoid repeated calculation of similar feature data, this paper adopts PCA based on the transfer learning algorithm. It is known that  $\varphi$  is feature learning; then,  $\varphi(X_s)$  and  $\varphi(X_T)$  represent  $X_s$  and  $X_T$  data features, respectively. The unmarked data  $x_{T_i}$  in  $D_T$  are mapped to the feature space to obtain a new representation  $\varphi(X_{T_i})$ . In transfer learning, the optimal learning state is the minimum expected risk. Therefore, the optimal model for learning the target domain is proposed:

$$\theta = \arg \min_{\theta} \sum_{i=1} P(D_T) l(x, y, \theta), \quad (17)$$

where  $P(D_T)$  is the probability of edge distribution in the target domain and  $l(x_i, y_i, \theta)$  is the loss function. When  $P(D_s) \neq P(D_T)$ , the above model is optimized to enhance the generalization ability of the transfer learning target domain:

$$\begin{aligned} \theta &= \arg \min_{\theta} \sum_{i=1} \frac{P(D_T)}{P(D_s)} P(D_s) l(x, y, \theta), \\ &\approx \arg \min_{\theta} \sum_{i=1} \frac{P_T(x_{T_i}, y_{T_i})}{P_S(x_{S_i}, y_{S_i})} l(x_{S_i}, y_{S_i}, \theta). \end{aligned} \quad (18)$$

According to the probability distribution function of the source domain and the target domain, the maximum embedded space between them can be obtained:

$$\begin{aligned} \text{Dist}(\varphi(X_S), \varphi(X_T)) &= \left\| \frac{1}{n_S^2} \sum_{i,j=1}^{n_S} k(x_{S_i}, x_{S_j}) + \frac{1}{n_T^2} \right. \\ &\quad \cdot \left. \sum_{i,j=1}^{n_T} k(x_{T_i}, x_{T_j}) - \frac{2}{n_S n_T} \sum_{i,j=1}^{n_S n_T} k(x_{S_i}, x_{T_j}) \right\|. \end{aligned} \quad (19)$$

In order to obtain the optimal eigenvalues in the maximum embedded space, we use the kernel matrices  $K_S$  and  $K_T$  to represent the feature matrix  $K = \begin{bmatrix} K_{S,S} & K_{S,T} \\ K_{T,S} & K_{T,T} \end{bmatrix}$  on the source- and target-domain data. In this section,  $\tilde{W}$  of  $(n_S + n_T) \times m$  matrix transformation is used to map the kernel matrix to the  $m$ -dimensional space to obtain the resultant kernel matrix as follows:

$$\tilde{K} = (K K^{-1/2} \tilde{W}) (\tilde{W}^T K^{-1/2} K) = K W W^T K. \quad (20)$$

Then, the corresponding kernel matrix value between  $x_i$  and  $x_j$  is

$$\tilde{K}(x_i, x_j) = K_{x_i}^T W W^T K_{x_j}. \quad (21)$$

The transfer learning mapping data  $x_{S_i}$  and  $x_{T_i}$  are obtained by calculating  $X'_S = [K_{S,S} \ K_{S,T}]W$  and  $X'_T = [K_{T,S} \ K_{T,T}]W$ .

The transfer learning for target features is to map the labeled image samples captured from each perspective in the target domain and the labeled samples in the source domain to the common feature subspace at the same time. Through the feature comparison between the tag samples in the source domain and those in the target domain, the target is recognized as a fighter. This method reduces the computation and recognition time of sample recognition in the target domain, and the specific process is shown in Figure 5. To sum up, the specific method of transfer reinforcement learning is shown in Algorithm 1.

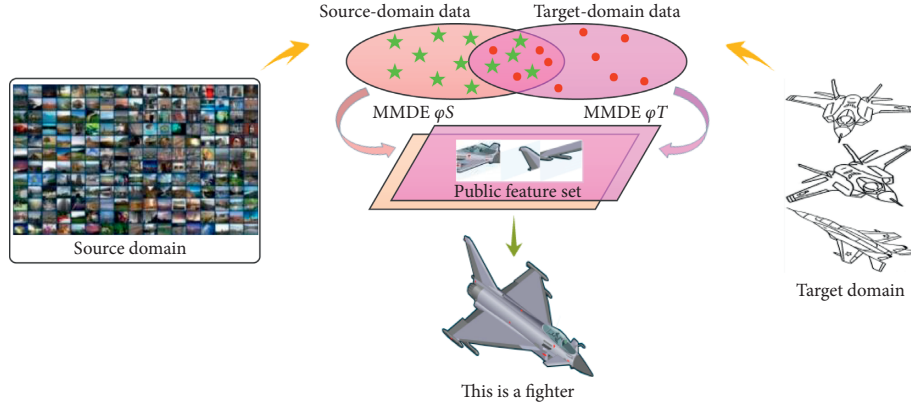


FIGURE 5: Feature transfer learning with tag samples in the target domain. The source domain and the target domain are mapped to the label samples by dimension reduction, and the PCA is used to project the same type labels into the common feature set. The sample tag of the target domain is output by the feature migration learning method.

- (1) **Input:** multiperspective image  $u(x)$  and feature vector  $R$
- (2) The feature clustering analysis is used to quantify  $H$ , and the hash value  $h$  is obtained
- (3) Image feature  $h_n$ , feature environment  $h_0$ , and Hamming distance  $D = h_n - h_0$
- (4) Let  $\tau$  be the similarity threshold
- (5) **for** frame = 1, 2, 3, ... **do**
- (6)   **if**  $S < \tau$  **then**
- (7)     Model judgment features are not similar, so select reinforcement learning
- (8)   **else**
- (8)     Model judgment features are similar, so select transfer learning
- (9) **end**

ALGORITHM 1: Transfer reinforcement learning.

**3.2.4. Multiperspective Light Field Reconstruction Method Based on Transfer Reinforcement Learning.** According to the establishment of the multiperspective representation model, the multiperspective light field is reconstructed using the transfer reinforcement learning algorithm. The original light field image is a  $(x, y)$  ray imaging grid. Each image represents that the light reaches a microlens on the imaging surface, which is from different  $(u, v)$  positions of the main lens, and it is shown in Figure 6.

The original image is composed of a series of pixels, each of which is microlens imaged. Because the aperture is limited, each microlens has a certain field of view, and there is a certain parallax among the different microlenses.

$$E_F(x, y) = \frac{1}{F^2} \iint L_F(x, y, u, v) \cos^4 \theta \, du \, dv, \quad (22)$$

where  $L_F(x, y, u, v)$  is the light field parameter from the target plane  $F$  and  $\cos \theta$  is the attenuation factor due to the optical halo effect. The multiperspective feature information fusion  $u(r)$  and the feature learning  $\varphi$  in transfer reinforcement learning are introduced into the light field reconstruction process. Then, the point imaging function can be obtained in any plane letting  $(x, y, u, v) \rightarrow \varphi(x', y', u(r), v')$ .

$$E_{(\alpha, F)}(x', y') = \frac{1}{\alpha^2 F^2} \iint L_F\left(u\left(1 - \frac{1}{\alpha}\right) + \frac{x'}{\alpha}, v\left(1 - \frac{1}{\alpha}\right) + \frac{y'}{\alpha}, u, v\right) \, du \, dv. \quad (23)$$

According to the similarity threshold of different scenes, the algorithm in this paper autonomously chooses different models of transfer reinforcement learning (i.e., transfer learning or reinforcement learning) to reconstruct the multiperspective light field. Based on the algorithm, the frequency-domain information of images can be obtained by 4D Fourier transform of the multiperspective light field. And then the center slice and the inverse wavelet transform are used to obtain the reconstructed light field image of each perspective, as shown in Figure 7.

## 4. Experimental Results and Analyses

PASCAL VOC provides a standard set of excellent datasets for image recognition and classification. Among them, the VOC2007 dataset contains 9963 labeled images, consisting of trains, fuel consumption, and tests, with a total of 24640 labeled objects. The VOC2012 dataset is an upgraded version of the VOC2007 dataset, with a total of 11,530 images. The VOC2012 dataset is divided into 20 categories, such as people, birds, dogs, airplanes, and cars. In order to ensure the objectivity and persuasiveness of the experiment, the training set samples used for animal detection, vehicle detection, and human detection are 6800 images with a pixel size of  $500 * 332$ . The test set samples are 1200 images with a pixel size of  $500 * 375$ . In this paper, the detection rate is selected as a performance measure for statistical significance

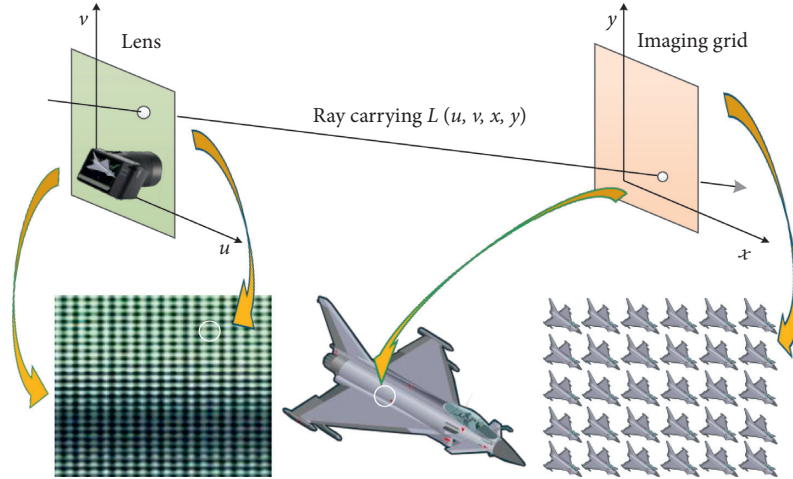


FIGURE 6: Light field imaging model. Taking the aircraft model as an example, the green plane is the camera microlens imaging surface and the orange plane is the imaging grid surface.

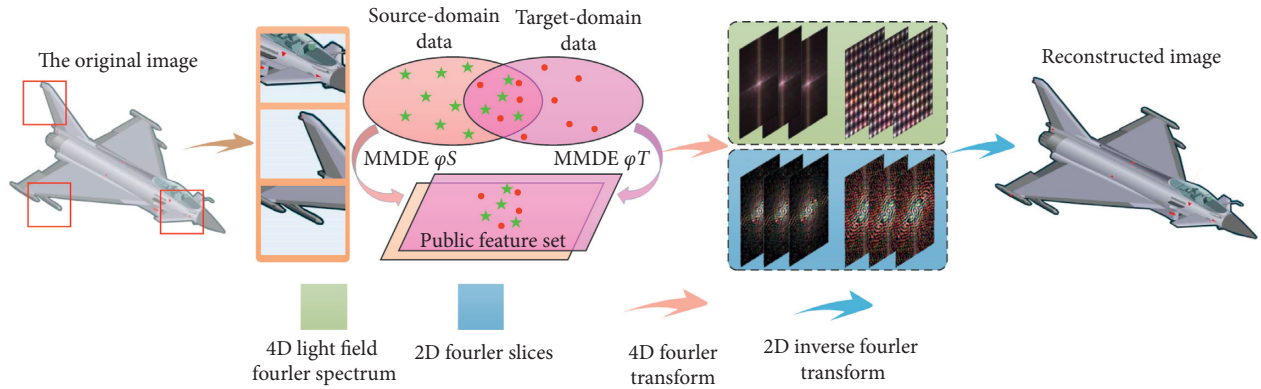


FIGURE 7: Multiperspective light field reconstruction based on transfer reinforcement learning.

test analysis. The detection rate refers to the recognition ratio between the target and the background in the recognition window. We use TensorFlow on the Windows 10 system for experimental simulation. The simulation calculation runs on a small server with an E5-2630 v4 CPU, the main frequency of 2.2 GHz, and a memory of 32 GB. The multiperspective light field reconstruction method training model network structure is shown in Figure 8.

This section selects several representative data to describe and analyze the target recognition. The experimental data of target recognition are divided into three categories: animals, vehicles, and humans. The test data are sorted according to the number of targets and the complexity of scenarios. The algorithm in this paper identifies targets simultaneously with existing RPN [36], Fast R-CNN [37], YOLO 9000 [38], SSD [39], and R-CNN [40] algorithms.

Figure 9 shows the animal samples for the test dataset, which include sheep, bird, horse, and dog. In the sheep recognition results, we can see that the algorithm has good results in the second, third, and fifth tests. It includes target occlusion and small targets at a distance. In the bird recognition results, we can see from the fifth test data that the algorithm in this paper has accurately identified each target

in the graph. This shows that the algorithm is suitable for multitarget recognition. In the horse recognition results, we can conclude from the third test data that the algorithm can effectively avoid the impact of the light problem on target recognition. In the dog recognition results, we can see the effectiveness of the algorithm in multitarget recognition from the results of the fourth and fifth test data. From Table 1, it can be seen that the average recognition accuracy of the algorithm in the detection of the animal dataset is 78%, which is higher than that of other algorithms.

Figure 10 shows samples of vehicles for the test set, which include daily vehicles such as airplanes, ships, and cars. In the fourth test data of airplanes, it is reflected that all algorithms are less robust to target local information recognition under cluttered background. However, the algorithm identifies the greatest number of targets compared to other algorithms. In the second test data of the airplanes, the algorithm effectively solves the problem of insufficient light. However, in the identification result of the ship, the fifth test data are large due to the fog at sea, and all algorithms are not recognized for distant targets. It can be seen from the recognition results of motorcycles and cars that when the targets are seriously overlapped, other algorithms



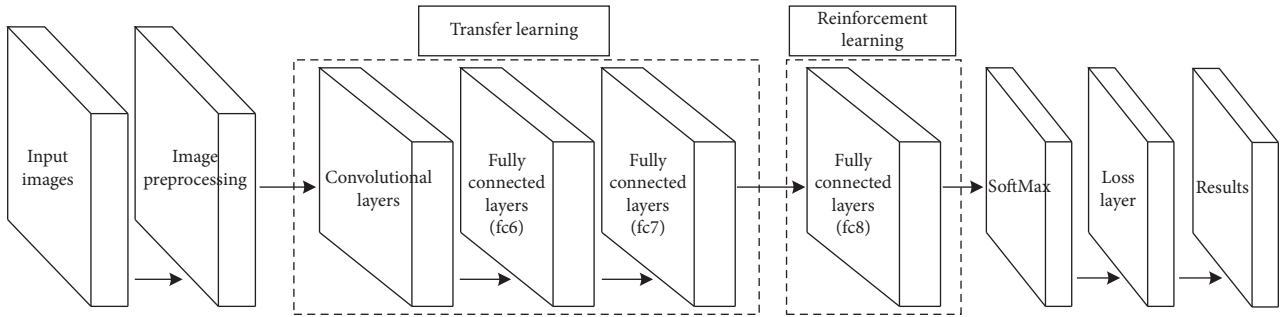


FIGURE 8: Multiperspective light field reconstruction method training model network.

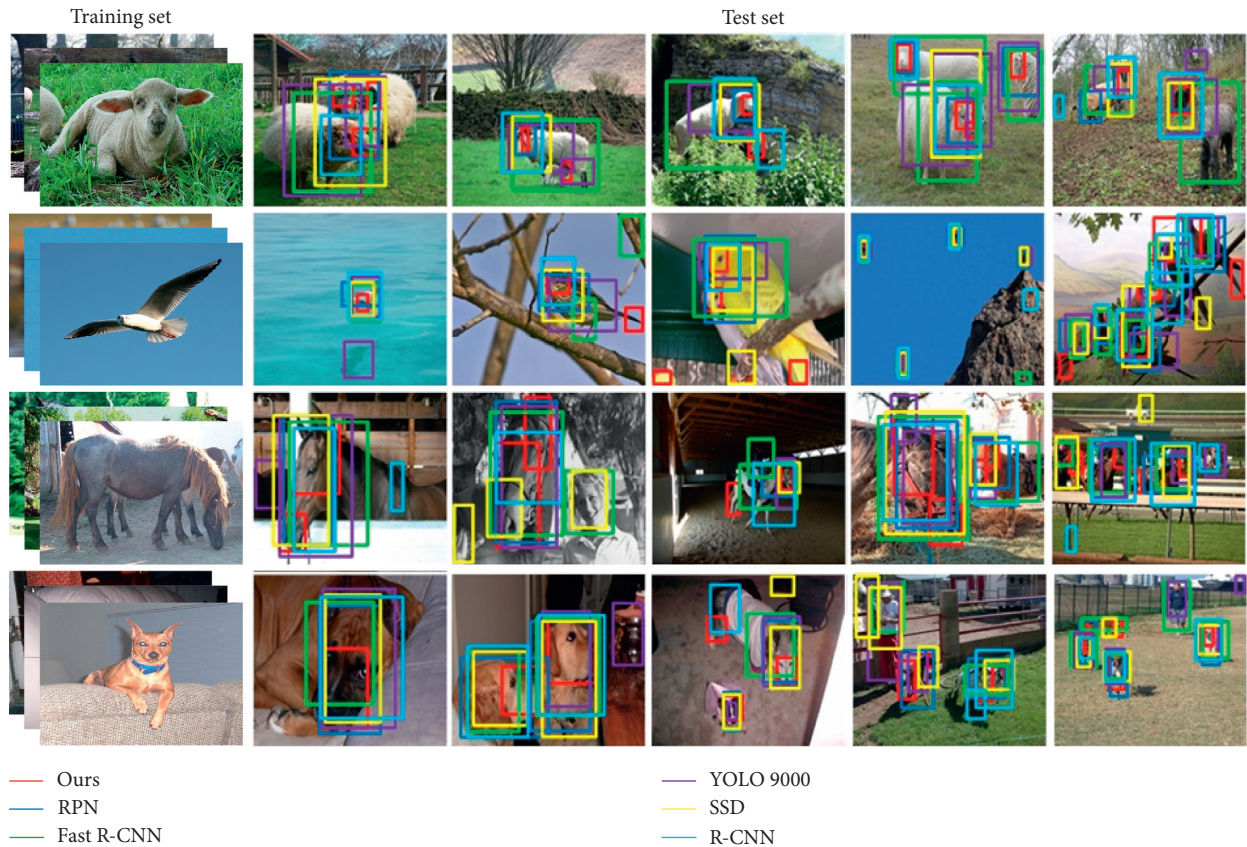


FIGURE 9: Recognition results of animal samples.

mentioned in this paper cannot accurately identify them. From Table 2, it can be seen that the average recognition accuracy of the algorithm in the detection of the vehicle dataset is 76.9%, which is higher than that of other algorithms.

This paper assumes  $H_0: \mu_{Others} = \mu_{Ours}$ , which proves that there is no difference between the two methods for the same target recognition.  $H_A: \mu_{Others} \neq \mu_{Ours}$ , which proves that there is a difference between the two methods when identifying the same target. In these relations, Others represent the existing algorithms and Ours represents the algorithm in this paper. In order to verify the effectiveness of the algorithm in this paper, we set the significance level to the international universal level  $\alpha = 0.05$ . That is, the confidence interval is 95%. The sample is  $n$ , and  $df = 2(n - 1)$ . Therefore, the standard deviation of the

mean is  $S_{Others-Ours} = \sqrt{(S_{Others}^2/n) + (S_{Ours}^2/n)}$ , and the statistic  $t = (|Others| - Ours) / (S_{Others-Ours})$ .

We set the confidence interval to 95% and the sample category to  $n = 4$ . Therefore, the benchmark for the statistical significance test is 2.447. As can be seen from Figure 11(a) and 11(b), the statistical significance test indicators of other algorithms and the algorithm in this paper are higher than 2.447. Then, we can judge the hypothesis. There are statistically significant differences between our algorithm and other algorithms in animal and vehicle datasets.

Figure 12 shows a test dataset of human samples that include men, women, and baby. In the male identification results, both the second and fifth test data showed interference items that were not physical. Therefore, our

TABLE 1: Detection rate evaluation report form for different algorithms in the animal dataset.

	Sheep, average	Bird, average	Horse, average	Dog, average	Animal dataset, average	Standard deviation
RPN	71.7	75.2	74.5	77.4	74.7	2.35
Fast R-CNN	71.6	73.6	73	76.4	73.7	2.02
YOLO 9000	71.7	72.4	73.3	73.1	72.6	0.73
SSD	69.3	71.3	71.6	73.6	71.5	1.76
R-CNN	71.6	71.7	72.3	76.4	73	2.29
<b>Ours</b>	<b>77.2</b>	<b>78.2</b>	<b>78.7</b>	<b>77.8</b>	<b>78</b>	<b>0.63</b>

VOC2012 animal dataset detected accuracy (%). Average: average recognition rate of five samples in each dataset.

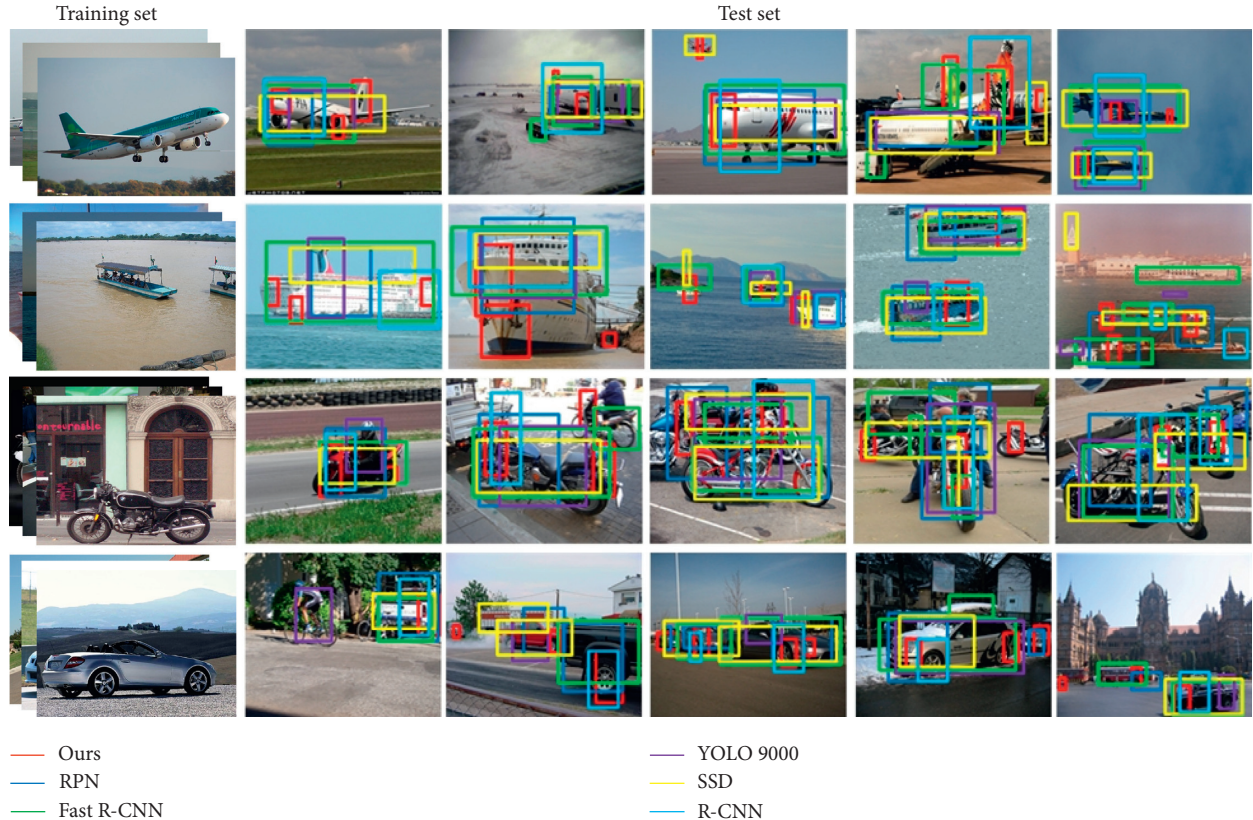


FIGURE 10: Recognition results of vehicle samples.

TABLE 2: Detection rate evaluation report form for different algorithms in the vehicle dataset.

Method	Airplane, average	Boat, average	Motorcycle, average	Car, average	Vehicle dataset, average	Standard deviation
RPN	72.4	72.5	74.3	72.3	72.9	0.95
Fast R-CNN	74.4	74.7	74.8	71.7	73.9	1.48
YOLO 9000	72.8	71.3	70.4	70	71.1	1.24
SSD	74.9	71.1	72.3	71.7	72.5	1.67
R-CNN	70.9	72.7	70.4	71	71.3	1
<b>Ours</b>	<b>78.1</b>	<b>76.8</b>	<b>77.8</b>	<b>75</b>	<b>76.9</b>	<b>1.4</b>

VOC2012 vehicle dataset detected accuracy (%). Average: average recognition rate of five samples in each dataset.

algorithm, RPN algorithm, and SSD algorithm all identify errors in the second test data. However, in the fourth test data, the motion occlusion problem was effectively solved and all targets were correctly identified. In the female recognition results, our algorithm successfully distinguishes the males in the fourth test data. Because the fifth test data have

many tasks, the features of individual characters are not obvious. Therefore, the algorithms mentioned in this paper have different errors in the target recognition process. In the baby's recognition results, all algorithms are accurately identified in both the first and second test data. In the third and fourth test data, the SSD algorithm and the R-CNN

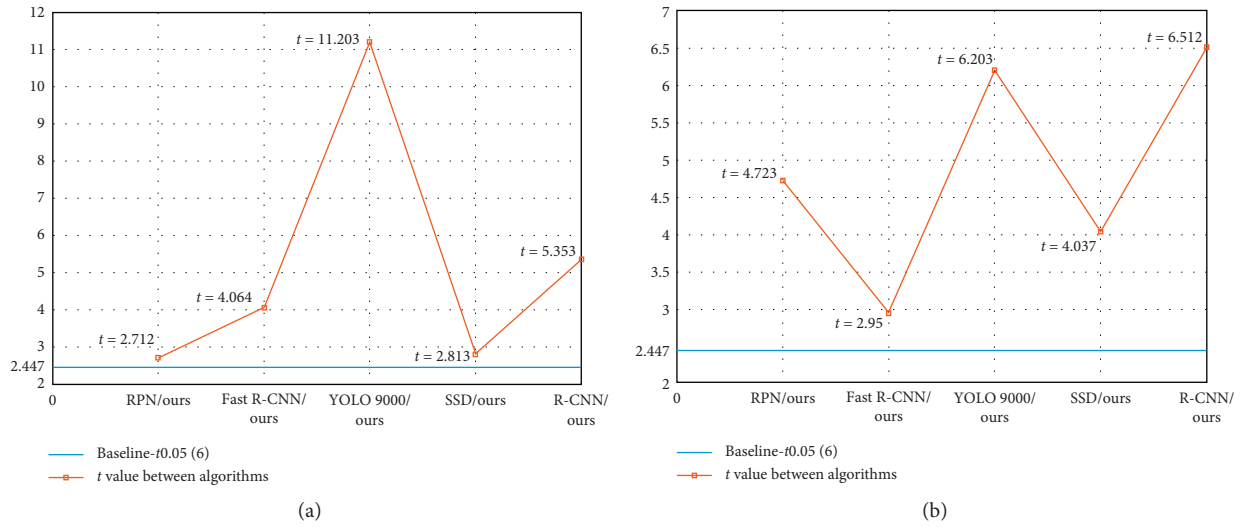


FIGURE 11: Analysis chart of the statistical significance test of each target algorithm in the animal dataset (a) and vehicle dataset (b).



FIGURE 12: Recognition results of human samples.

algorithm identify objects outside the target. However, in the fifth test data, all algorithms cannot be accurately identified because the scene is too complicated and the target is small. As can be seen from Table 3, when detecting the baby dataset, the recognition rate of the RPN algorithm and the algorithm in this paper reached 71.4% at the same time. In the male dataset and female dataset, the recognition accuracy of the algorithm in this paper is 72.5%, which is higher than that of other algorithms.

Since the sample category of the human dataset is  $n = 3$ , the benchmark for the statistical significance test is 2.776. As can be seen from Figure 13 and Table 3, the accuracy of the proposed target recognition algorithm on the human dataset is 72.5%, which is higher than that of RPN and Fast R-CNN. However, the statistical significance test indicators are lower

than the baseline. This shows that the hypothesis test is valid, and the recognition results of the algorithm in this paper are not significantly different from those of the PRN and Fast R-CNN algorithms. Compared with those of YOLO 9000, SSD, and R-CNN algorithms, the statistical significance test indicators of the algorithm in this paper are higher. This shows that there are obvious differences between the algorithm in this paper and them.

Based on the experimental results of target recognition and statistical significance test analysis, the algorithm in this paper fully reflects the advantages of multiple perspectives in target recognition. In target recognition, the larger the boundary box, the less accurate the target recognition result. When detecting any feature of the target, the algorithm uses several small red boxes to label it. Therefore, the above target

TABLE 3: Detection rate evaluation report form for different algorithms in the human dataset.

Method	Male, average	Female, average	Baby, average	Human dataset, average	Standard deviation
RPN	71.8	72.3	<b>71.4</b>	71.8	0.45
Fast R-CNN	71.3	71.3	71.1	71.2	0.12
YOLO 9000	70	71.1	69.9	70.3	0.67
SSD	70	69.8	70.3	70	0.25
R-CNN	69.8	70.7	70.1	70.2	0.46
<b>Ours</b>	<b>73.6</b>	<b>72.5</b>	<b>71.4</b>	<b>72.5</b>	1.1

VOC2012 human dataset detected accuracy (%). Average: average recognition rate of five samples in each dataset.

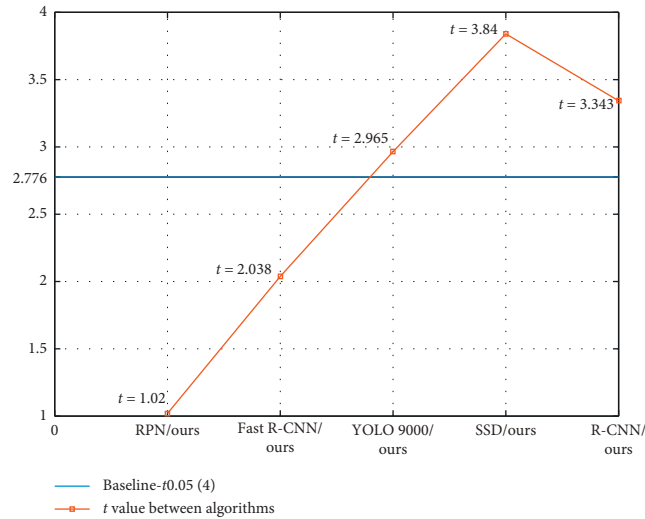


FIGURE 13: Analysis chart of the statistical significance test of each target algorithm in the human dataset.

recognition results verify the effectiveness of the proposed algorithm.

## 5. Conclusions

Duplicate processing of target data and lack of required data affect the accuracy and real-time performance of reconstruction of the complete light field. In view of the problems, this paper proposes the transfer reinforcement learning method for multiperspective light field reconstruction. According to the similarity threshold, reinforcement learning or transfer learning can be selected independently. Our algorithm effectively solves the repeated calculation of the same data and shortens the time of multiperspective light field reconstruction. The experimental result shows that the transfer reinforcement learning algorithm is better than other algorithms in target recognition. In the experiment, the recognition efficiency of the algorithm in this paper makes it not ideal for recognition in the foggy environment. In the future work, we will introduce the GAN to transfer learning. We will use the characteristics of the GAN to generate data and enhance data to process foggy images and increase their target recognition accuracy.

## Data Availability

The dataset contains confidential information such as the performance parameters of the aircraft model and cannot be released.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (61703143), Science and Technology Project of Henan Province (192102310260), Young Backbone Teacher Training Project of Henan University (2017GGJS123), and Science and Technology Major Special Project of Xinxiang City (ZD18006).

## References

- [1] A. Gershun, "The light field," *Studies in Applied Mathematics*, vol. 18, no. 1–4, pp. 51–151, 1939.
- [2] M. Levoy and P. Hanrahan, "Light field rendering," *Conference on Computer Graphics and Interactive Techniques*, vol. 2, no. 3, pp. 31–42, 1996.
- [3] S. Pathak, L. Pulina, and A. Tacchella, "Verification and repair of control policies for safe reinforcement learning," *Applied Intelligence*, vol. 48, no. 4, pp. 886–908, 2017.
- [4] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [5] L. Tran, X. Yin, and X. Liu, "Disentangled representation learning GAN for pose-invariant face recognition," in *Proceedings of the IEEE Conference on Computer Vision and*

- Pattern Recognition (CVPR)*, pp. 1283–1292, Honolulu, HI, USA, July 2017.
- [6] S. Pang, J. J. del Coz, Z. Yu, O. Luaces, and J. Díez, “Deep learning and preference learning for object tracking: a combined approach,” *Neural Processing Letters*, vol. 47, no. 3, pp. 859–876, 2017.
  - [7] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
  - [8] Z. Liu, P. Luo, X. Wang et al., “Deep learning face attributes in the wild,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3730–3738, Las Condes, Chile, December 2015.
  - [9] A. Kangasrääsio and S. Kaski, “Inverse reinforcement learning from summary data,” *Machine Learning*, vol. 107, no. 8–10, pp. 1517–1535, 2018.
  - [10] Y. Zou, T. Liu, D. Liu, and F. Sun, “Reinforcement learning-based real-time energy management for a hybrid tracked vehicle,” *Applied Energy*, vol. 171, pp. 372–382, 2016.
  - [11] S. J. Gershman and N. D. Daw, “Reinforcement learning and episodic memory in humans and animals: an integrative framework,” *Annual Review of Psychology*, vol. 68, no. 1, pp. 101–128, 2017.
  - [12] B. Zoph, V. Vasudevan, J. Shlens et al., “Learning transferable architectures for scalable image recognition,” in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8697–8710, Salt Lake City, UT, USA, June 2018.
  - [13] N. A. Goussies, S. Ubalde, and M. Mejail, “Transfer learning decision forests for gesture recognition,” *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3667–3690, 2017.
  - [14] A. Van Opbroek, M. A. Ikram, M. W. Vernooij, and M. de Bruijne, “Transfer learning improves supervised image segmentation across imaging protocols,” *IEEE Transactions on Medical Imaging*, vol. 34, no. 5, pp. 1018–1030, 2015.
  - [15] L. Cai, P. Luo, G. Zhou et al., “Maneuvering target recognition method based on multi-perspective light field reconstruction,” *International Journal of Distributed Sensor Networks*, vol. 15, no. 8, pp. 1–12, 2019.
  - [16] G. Zhou, C. Wen, and J. Gao, “Object recognition based on reconstruction of light field,” in *Proceedings of the 2015 International Conference on Estimation, Detection and Information Fusion (ICEDIF)*, pp. 82–87, Harbin, China, January 2015.
  - [17] Y. Hong, R. Lu, R. Zhu et al., “Spatial multiplexing reconstruction for Fourier-transform ghost imaging via sparsity constraints,” *Optics Express*, vol. 26, no. 3, pp. 2181–2190, 2018.
  - [18] C. Sun, B. Wang, Y. Fang, K. Yang, and Z. Song, “High-resolution ISAR imaging of maneuvering targets based on sparse reconstruction,” *Signal Processing*, vol. 108, pp. 535–548, 2015.
  - [19] H. Lu, Y. Li, T. Uemura, H. Kim, and S. Serikawa, “Low illumination underwater light field images reconstruction using deep convolutional neural networks,” *Future Generation Computer Systems*, vol. 82, pp. 142–148, 2018.
  - [20] Z. Cai, X. Liu, X. Peng, and B. Z. Gao, “Ray calibration and phase mapping for structured-light-field 3D reconstruction,” *Optics Express*, vol. 26, no. 6, pp. 7598–7613, 2018.
  - [21] J. Zhang, L. Ye, Q. Zhang, and J. L. Wang, “Three-dimensional object surface reconstruction based on camera calibration and SIFT,” *Applied Mechanics and Materials*, vol. 719–720, pp. 1191–1197, 2015.
  - [22] J. Chen, J. Hou, and L.-P. Chau, “Light field compression with disparity-guided sparse coding based on structural key views,” *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 314–324, 2018.
  - [23] Z. Cai, X. Liu, X. Peng et al., “Universal phase-depth mapping in a structured light field,” *Applied Optics*, vol. 57, no. 1, pp. 26–32, 2018.
  - [24] D. C. Schedl, C. Birkbauer, and O. Bimber, “Optimized sampling for view interpolation in light fields using local dictionaries,” *Computer Vision and Image Understanding*, vol. 168, pp. 93–103, 2018.
  - [25] Y. Wang, F. Liu, K. Zhang, G. Hou, Z. Sun, and T. Tan, “LFNet: a novel bidirectional recurrent convolutional neural network for light-field image super-resolution,” *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4274–4286, 2018.
  - [26] S. Zhou, Y. Yuan, L. Su, X. Ding, and J. Wang, “Multiframe super resolution reconstruction method based on light field angular images,” *Optics Communications*, vol. 404, pp. 189–195, 2017.
  - [27] M. Cúth, O. F. K. Kalenda, and P. Kaplický, “Isometric representation of lipschitz-free spaces over convex domains in finite-dimensional spaces,” *Mathematika*, vol. 63, no. 2, pp. 538–552, 2017.
  - [28] J. Lieman-Sifry, M. Le, F. Lau, S. Sall, and D. Golden, “FastVentricle: cardiac segmentation with ENet,” *Functional Imaging and Modelling of the Heart*, Springer, Berlin, Germany, pp. 127–138, 2017.
  - [29] E. Shelhamer, J. Long, and T. Darrell, “Fully convolutional networks for semantic segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2017.
  - [30] H. Noh, S. Hong, and B. Han, “Learning deconvolution network for semantic segmentation,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1520–1528, Las Condes, Chile, December 2015.
  - [31] V. Badrinarayanan, A. Kendall, and R. Cipolla, “Segnet: a deep convolutional encoder-decoder architecture for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
  - [32] L. Chen, X. Hu, T. Xu, H. Kuang, and Q. Li, “Turn signal detection during nighttime by CNN detector and perceptual hashing tracking,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3303–3314, 2017.
  - [33] C. Distanto, A. Anglani, and F. Taurisano, “Target reaching by using visual information and Q-learning controllers,” *Autonomous Robots*, vol. 9, no. 1, pp. 41–50, 2000.
  - [34] B. Saha, S. Gupta, D. Phung, and S. Venkatesh, “Multiple task transfer learning with small sample sizes,” *Knowledge and Information Systems*, vol. 46, no. 2, pp. 315–342, 2016.
  - [35] M. Mehdipour Ghazi, B. Yanikoglu, and E. Aptoula, “Plant identification using deep neural networks via optimization of transfer learning parameters,” *Neurocomputing*, vol. 235, pp. 228–235, 2017.
  - [36] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
  - [37] R. Girshick, “Fast R-CNN,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1440–1448, Las Condes, Chile, December 2015.
  - [38] J. Redmon and A. Farhadi, “Yolo 9000: better, faster, stronger,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6517–6525, Honolulu, HI, USA, July 2017.
  - [39] W. Liu, D. Anguelov, D. Erhan et al., “SSD: single shot MultiBox detector,” in *Proceedings of the European Conference on Computer Vision*, pp. 21–37, Las Vegas, NV, USA, 2015.

- [40] R. Girshick, J. Donahue, T. Darrelland et al., “Rich feature hierarchies for object detection and semantic segmentation,” in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, June 2014.