Image Automatic Counting Method by Using Deep Transfer Learning

Shasha Huang, Ke Sun, Guoxing Hu, Jia Qiao
Chinese Flight Establishment Test, Xi’an 710089, China

Abstract. In view of the narrow long room, fast speed and roll sport while outside the room, single camera couldn't achieve the measurement of big cargo extraction. This paper puts forward an intersection measurement system of multiple ultra-short focus cameras and high-speed cameras. The system can achieve the parameters of wholly sport process by measuring the marked points on cargo, coordinate conversion and data registration. Measurement data show the trajectory and attitude parameters truly. Measurement precision satisfies the requirement of subject.

1. Introduction

The flight test track test refers to the measurement of the time and space position of the moving target in the flight test of the aviation weapon and equipment. It is to establish the flight test time reference and the spatial reference in a specific real flight environment, and provide the target tracking image information and the track result. Providing qualitative and quantitative real-time monitoring and external parameter information for flight tests is one of the important basic technologies for flight testing [1]. presently, the main methods of track measurement include radar, photoelectric theodolite, measuring camera, and airborne GPS [2].

The measurement accuracy of the radar or airborne GPS device measuring the track is 0.5~1m, and the sampling rate is less than 50frame/s. The advantage is that the flight path of the whole process can be tracked. The disadvantage is that the detailed description of the key track is not clear enough [3]. In order to obtain the flight path details of key segments and the track data with higher precision, flight test often uses high frame rate measurement camera to take photos of high-speed aircraft, and perform data processing such as interpretation and calculation of acquired images. Finally, the sports track data is obtained.

Since such a track measurement method based on the principle of photogrammetry and photographic geometry has a measurement accuracy of up to centimeter level, it has been widely used in flight tests. Therefore, image data processing accounts for most of the workload, and image interpretation is a top priority. In order to complete the flight test task quickly and efficiently and ensure measurement accuracy, the deep migration learning image automatic interpretation method is adopted, and the trained AlexNet volume is adopted. Add multiple adaptive layers in front of the distributor in the neural network, and add adaptive multi-core MMD metrics in the adaptive layer to complete the migration learning of deep neural networks, and then realize automatic recognition, interpretation and image coordinates of moving object images. Extract [4-5].
2. Deep transfer learning principle

2.1. Deep network mobility

As AlexNet won the 2012 ImageNet contest, deep learning began to shine in the field of machine learning research and applications. The neural network has a good hierarchical structure, and the first few layers learn all the general features (general feature). As the network level deepens, the latter network is more focused on learning specific features (specific feature). As shown in Figure 1.

Figure 1. A simple example of feature extraction to classification from a deep neural network

An AlexNet network has a total of 8 layers. Except that the 8th layer is a category-related network that cannot be migrated, the 1st to 7th layers are fine-tuned. It is found that the first three layers of the neural network basically learn the common features, and the effect of migration will be Better, as the number of migrateable layers increases, the performance of the model decreases. If you add fine-tuning in the deep migration network, the effect will be much higher, it may be better than the original network, and it can better overcome the difference between the data. Fine-tuning is to use other people's already trained network to adjust for their own tasks, as shown in Figure 2.
Figure 2. A simple schematic diagram of fine tuning

From Figure 2 we can see that the pre-trained network we use is very complicated. If you use it directly from the beginning, the time cost will be very high. We can transform this network, fix the parameters of the previous layers, and fine-tune the next few layers for our mission. In this way, the speed of network training will be greatly accelerated, and it will also greatly promote the performance of our mission.

2.2. Deep network adaptation

Fine-tuning the deep network can help us save training time and improve learning accuracy. However, it is impossible to deal with the difference between the training data and the test data distribution. Therefore, an adaptive layer needs to be developed for the deep network to complete the adaptation of the source domain and the target domain data. Adaptive can make the data distribution of the source domain and the target domain closer together, thus enhancing the network effect. Typically, the adaptive layer is determined by network loss calculations, and the loss is defined as follows:

$$\ell = \ell_c(D_s, y_s) + \lambda \ell_A(D_s, D_t)$$

Where, $\ell$ represents the final loss of the network, $\ell_c(D_s, y_s)$ represents the general classification loss of the network on the marked data (mostly the source domain). $\ell_A(D_s, D_t)$ represents the adaptive loss of the network. This is unique to migration learning and expresses the distribution differences between the source and target domains. $\lambda$ is a weighting parameter that weighs two parts.

The deep migration learning method studied in this paper adds adaptive multi-core maximum mean difference (MMD) metrics to each of the three layers by adapting the last three layers of the AlexNet network (6th, 7th, and 8th layers). The representation is as shown in Equation 2:

$$\mathcal{K} \triangleq \{k = \sum_{u=1}^{m} \beta_u k_u : \beta_u \geq 0, \forall u\}$$

Therefore, the optimization goal can be changed by Equation 1, as shown in Equation 3:

$$\max_\Theta \frac{1}{n_1} \sum_{i=1}^{n_1} J(\theta(x_i), y_i^f) + \lambda \sum_{i=2}^{n_2} d_k^f(D_s^i, D_t^i)$$

Where, $\Theta$ denotes the weight of the network and the parameters of the bias unit, which is the goal of network learning. $\ell_1, \ell_2$ are 6 and 8, respectively, indicating that the network adaptation is from the sixth layer to the eighth layer, and the previous one is not adapted. $x_a, n_a$ represents a collection of all labeled
data in the source and target domains. \( J(\cdot) \) defines a loss function, which is cross entropy. The deep migration learning network structure added to the adaptation layer and multi-core MMD is shown in Figure 3.

Learning is divided into two broad categories of parameters: learning network parameter \( \Theta \) and multi-core MMD \( \beta \). The learning of \( \Theta \) depends on the calculation of the multi-core MMD distance. The multi-core MMD metric can be expanded into the inner product form by the kernel mechanism, but the computation time complexity is high, so the unbiased estimation of the multi-core MMD metric needs to be performed, as shown in Equation 4:

\[
d_{k}^2(p, q) = \frac{2}{n_s} \sum_{i=1}^{n_s/2} g_{k}(Z_i)
\]

Where, \( Z_i \) is a quad group: \( Z_i \triangleq \left( X_{2i-1}, X_{2i}, X_{2i-1}^l, X_{2i}^l \right) \). After the nuclear mechanism is applied to \( Z_i \), Equation 4 changes to Equation 5:

\[
g_{k}(Z_i) \triangleq k(X_{2i-1}^l, X_{2i}^l) + k(X_{2i-1}, X_{2i}^l) - k(X_{2i-1}^l, X_{2i}) - k(X_{2i}, X_{2i-1}^l)
\]

The above changes can reduce the time complexity from \( O(n^2) \) to \( O(n) \)!

Learning \( \beta \) is mainly to determine the weight of multiple nuclear mechanisms, the goal is to ensure that the variance of the MMD distance generated by each nuclear mechanism is the smallest, as shown in Equation 6:

\[
\max_{k \in \mathcal{X}} d_{k}^2(D_{\ell}^l, D_{\ell}^l) \sigma_k^{-2}
\]

\( \sigma_k^{-2} = [E[g_k^2(z)] - [E(g_k(z)))]^2 \) is the estimated variance.

3. Test and data analysis

First of all, the model needs to be trained. The number of iterations in the experiment is nearly 3,000. The loss rate before 300 times is large, but it has a sharp downward trend. The subsequent loss rate is lower, and finally converges to 2.98%, as shown in Figure 4 shows.
After the model training is completed, the coordinate data of the resistance umbrella center coordinate in the flight test is taken as the input source data. First, the image is segmented out of the attention area, and then the gray level equalization processing is performed. The partial image interpretation results are shown in Fig. 5, and the test results prove the paper. The algorithm model studied can be directly migrated to the field of image interpretation applications.

**Figure 5.** partial image interpretation results

In this experiment, a total of 300 pictures were imported, and the interpretation took nearly 60 seconds, which greatly improved the interpretation speed and data processing efficiency.

4. Conclusion

Deep migration learning image automatic interpretation method, through the integration of multiple adaptive layer and multi-core MMD metrics in front of the distributor in the trained AlexNet convolutional neural network, complete the migration learning of deep neural network, and then realize the automatic image of moving objects Recognition, interpretation, and image point coordinate extraction have achieved good results. The experimental data proves that this method is more efficient and accurate than the manual interpretation and traditional migration learning methods, and is more suitable for flight test mission requirements. However, there is not much space left for model innovation in this method. For future research, it is a technical problem, and more in-depth algorithm research is needed to further improve efficiency and accuracy.
references

[1] ZHOU ziquan. Flight Test Engineering [M].Beijing:China Aviation Publishing,2010.

[2] ZUO Yi-hong,HE Hong-li,FENG Qiao-ning,etc. Application of High Speed Video in Load Test for Aircraft Landing Gear.ELECTRO-OPTIC TECHNOLOGY APPLICATION.June,2011,Vol.26,No.3:82-84.

[3] ZHANG Jie,FENG Qiaoning. Flight Path Measurement During Takeoff and Landing Phase Based on High Speed Camera Array. OPTICS & OPTOELECTRONIC TECHNOLOGY. August, 2013.Vol.11,No.4:45-49.

[4] Dan Zhiping,Sang Nong,Chen Yanfei,etc. A new target recognition method in remote sensing image.J.Huazhong Univ.of Sci.& Tech.(Natural Science Edition. Oct.2013.Vol.41 Sup. Ⅰ : 140-143.

[5] Yang Yang,Zhang Wensheng. Image Auto-Annotation Based on Deep Learning. Journal of Data Acquisition and Processing.Vol.30, No.1,Jan.2015:88-98.

[6] Cai Yingfeng, Wang Hai. Vision based vehicle detection transfer learning algorithm. JOURNAL OF SOUTHEAST UNIVERSITY ( Natural Science Edition). Mar 2015, Vol.45,No.2:275-280.

[7] LIU Jie, DU Jun-ping. A Transfer Learning Based Text-Image Feature Mapping Algorithm. Journal of Beijing University of Posts and Telecommunications. Dec.2012, Vol.35,No.6:1-5.

[8] SHAO Song, LIU Hong, WANG Xiangdong, et al. Review of pedestrian detection based on transfer learning. Computer Engineering and Applications, 2014, 50(24):156-163.