

PAPER

Unit panel node detection by CNN on FAST reflector

To cite this article: Zhi-Song Zhang *et al* 2019 *Res. Astron. Astrophys.* **19** 011

View the [article online](#) for updates and enhancements.

Unit panel node detection by CNN on FAST reflector

Zhi-Song Zhang^{1,2}, Li-Chun Zhu^{1*}, Wei Tang¹ and Xin-Yi Li¹

¹ National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; songzszs@nao.cas.cn;
lczhu@bao.ac.cn

² University of Chinese Academy of Sciences, Beijing 100049, China

Received 2018 June 19; accepted 2018 August 20

Abstract The Five-hundred-meter Aperture Spherical radio Telescope (FAST) has an active reflector. During observations, the reflector will be deformed into a paraboloid 300 meters in diameter. To improve its surface accuracy, we propose a scheme for photogrammetry to measure the positions of 2226 nodes on the reflector. The way to detect the nodes in the photos is the key problem in this application of photogrammetry. This paper applies a convolutional neural network (CNN) with candidate regions to detect the nodes in the photos. Experimental results show a high recognition rate of 91.5%, which is much higher than the recognition rate for traditional edge detection.

Key words: telescopes — FAST — photogrammetry — nodes detect — convolutional neural network

1 INTRODUCTION

The Five-hundred-meter Aperture Spherical radio Telescope (FAST) relies on the local characteristic of the karst landscape in Guizhou Province, China. The FAST facility is the world's largest single dish radio telescope and it is one of the major infrastructure projects of science and technology in China.

The FAST active reflector includes a reflective surface cable net consisting of nearly 10 000 steel cables with a diameter of 500 meters, reflecting surface unit, hydraulic actuators, ground anchor and ring beam. The reflector cable net is installed on a lattice type circular ring beam, and it has 2225 connecting nodes, through which 4450 reflector units are installed on the cable for reflecting radio waves, each node of which is connected with down-tied cables. Each node can be adjusted according to the observation requirements, thus forming a 300 m paraboloid for the illuminated aperture to complete observations with the telescope. In order to control the shape of the reflector accurately, the positions of all nodes on the entire reflector must be accurately measured in a 90 minute observing session, and the calibration precision should be below 1.5 mm.

The basic task of photogrammetry is to establish the geometrical relationship between an instantaneous im-

age and the object under study. Once this relationship is properly defined, we can carefully extract information about the target object from the image (Atkinson 2003). Photogrammetry is one of the most common methods in telescope reflective surface measurement, which has the advantages of high accuracy and efficiency. By sticking the target point (or projection structure light from the structure) on the antenna surface to be tested, one or more cameras are used to take multiple photos of the target point to calculate the surface type of the antenna (Xu 2006). The Arecibo 300-meter telescope in Puerto Rico has improved its surface accuracy from 15 mm to about 5 mm by using the photogrammetric method (Edmundson & Baker 2001). The 13.7 m millimeter-wave radio telescope near the city of Delingha, in Qinghai Province, China, uses the photogrammetry method to achieve an accuracy level of 0.083 mm (Fan et al. 2010), satisfying the surface precision requirements of millimeter wave observation, and also provides a precise measurement method (Zuo et al. 2011) for study of the gravity deformation.

The surface measurements mentioned above are all transient, but measurements that maintain the FAST surface require repeatability when tracking targets, a process which is very difficult. Based on the requirement of FAST's reflection surface measurement, this paper proposes a photogrammetric scheme without a target to realize

* Corresponding author.

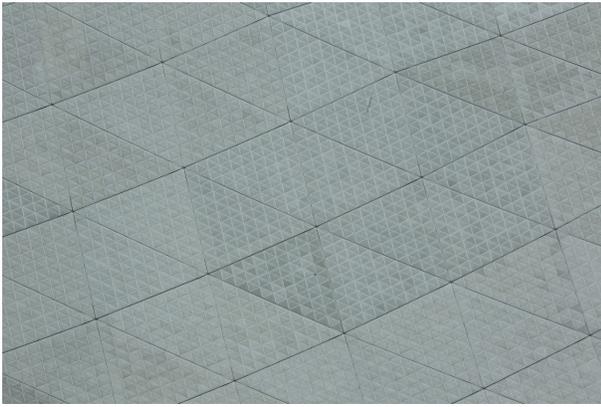


Fig. 1 Image of the FAST reflector.

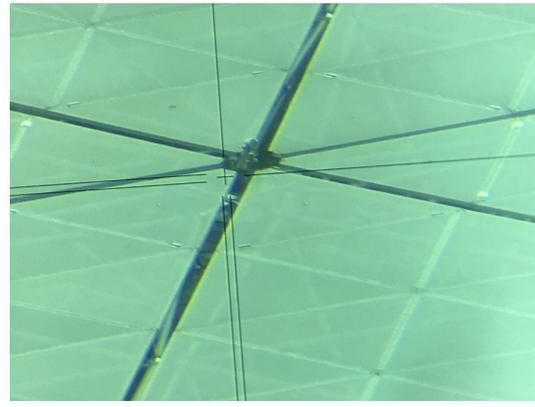


Fig. 2 Six gaps around each node, on which are the targets for all the station measurements.

accurate positioning of the nodes and complete the measurement of reflective surface type precision, through extracting natural characteristics of the nodes in photos. The pre-developed photogrammetry equipment called a digital positioning unit (DPU) (Hu & Zhu 2014) can be used to take photos of the surface on the existing stable foundation pier. The main objective of this paper is to improve the recognition rate for nodes and provide the necessary foundation for implementation of the photogrammetric scheme.

2 DESIGN SCHEME

The FAST reflection plane has unique geometrical features, as shown in Figures 1 and 2. The reflector is mounted on top of the cable net, which has a height difference of 40 cm. The shape of a unit panel is an 11 m triangle, and there is a splicing gap between panels. So, this gives rise to an obvious grey scale contrast between panel and gap when we examine the reflective surface. The cable network node is located at the intersection point of six gaps, which are formed by six panels nearby. This makes it possible to study the measurement method without actively using a target, which is the basis and core technology that can realize reflecting surface photography measurement without a target, using the gap between panels as a foundation for feature recognition to find the nodes in a photograph.

At first, we adopt the traditional Canny edge detection algorithm to identify the nodes (Canny 1986). The experimental result is shown in Figures 3 and 4. In the case of good lighting conditions, the node recognition rate can reach 60%, but the weather changes often in Guizhou Province, and where it is often cloudy and rainy. In the case of uneven or non-ideal illumination conditions due to variable weather, the node recognition rate is bad. Considering the changeable weather conditions at the FAST site, this method cannot achieve a reliably high recognition rate.

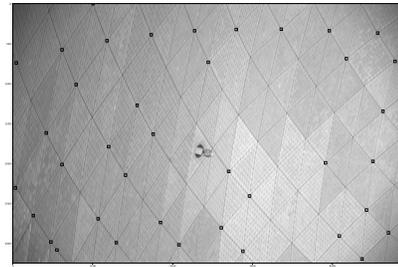


Fig. 3 The recognition of good lighting conditions.

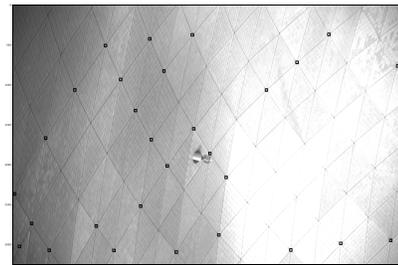


Fig. 4 The recognition of uneven or non-ideal lighting conditions.

To solve this problem, we propose a method to identify nodes using a convolutional neural network (CNN). At present, node recognition for the FAST reflector and the characteristics of CNN perform well:

- (1) The characteristics of the nodes are highly consistent, which makes it easier for the CNN to be trained.
- (2) The target for recognition is always the 2225 nodes. This enables us to collect enough sample data to train the network.

The process of modeling, training and testing a CNN is more complex and time-consuming than that of traditional algorithms. However, the measurement equipment and target image features are fixed and use the same camera to measure the 2225 nodes. After the model training is complete, it is no longer necessary to frequently update the data, which greatly improves the practical feasibility of the method.

Thanks to the CNN and the region proposal algorithm (Gu et al. 2009), object detection has made great advancements since 2014. R-CNN (Girshick et al. 2014), SPP-NET (He et al. 2014), Fast R-CNN (Girshick 2015), Faster R-CNN (Ren et al. 2015), SSD (Liu et al. 2016) and other methods have appeared successively, and these methods have achieved very good results when applied to online training sets. However, due to different tasks and goals, these methods cannot fully meet the specific requirements of recognition for FAST nodes. According to the characteristics of this project and ensuring accurate recognition rate, a method for object detection in a candidate area is proposed in this paper. In the early stages, we developed a set of special photogrammetric DPUs for FAST, an example of which is shown in Figure 5. Combined with the high precision rotating platform for photogrammetry and the related instrumentation, this equipment can not only achieve high precision and fast measurement, but also meets the requirements for a large number of points to be measured and large distribution range. When measuring, the servo motor drives two axes to rotate accurately, so that two or more stereoscopic digital measuring devices are aligned to the measured area at the same time. The focal length of the lens is changed according to the distance between the measured target and the measuring equipment, so as to record data from the target at different distances and keep the object surface of the shooting area the same size. Because there is a big difference in the image of nodes taken in different regions, the model training is carried out in fixed regions in order to eliminate interference caused by image differences in the recognition results. In this paper, we conducted experiments on the fixed area near foundation pier 9 of FAST, and the selection of the region was controlled by the pitch angle and azimuth angle of the DPU. Our proposed approach is illustrated in Figure 6.

3 EXPERIMENTAL VALIDATION AND DATA ANALYSIS

3.1 Composing the Training Sets

During about seven days at the field site of the FAST telescope, 460 photos were taken of the designated area around

foundation pier 9 by the DPU on foundation pier 6. Each photo shows about 60 nodes. The photos include images in cloudy, sunny, rainy and other lighting conditions. During the whole shooting process, the FAST reflector was in its initial shape. So, we calculated the coordinates of the nodes in the photos, which will be introduced in Section 3.2. We use the node coordinates to extract about 22 200 node images from the photos, examples of which are shown in Figure 7.

Six representative photos were selected, from which about 48 000 background images were extracted by the sliding window method. The final training set consists of about 22 000 node images and 48 000 background images.

3.2 Node Calibration Method

The relationship between nodes on the reflection surface and in a photo is shown in Figure 8. There are three different coordinates. The object coordinate system is based on the center of the FAST reflector, in which the initial coordinates of the nodes are known. The image plane coordinate system is a two-dimensional coordinate system with the photo center as the origin. The projection center coordinate system is obtained by shifting the image coordinate system by a focal distance. The following relationship can be obtained as follows

$$OP = OS + SP . \quad (1)$$

Let point P be \bar{p} in the space coordinate system $S - xyz$, so SP can be written as

$$SP = M \cdot S\bar{p} , \quad (2)$$

where M is the rotation matrix of the projection center coordinate system to the object coordinate system. $S\bar{p}$ is in line with $S\bar{p}$ and in the opposite direction, so $S\bar{p}$ can be written as

$$S\bar{p} = -kS\bar{p} \quad (k \text{ is the constant of proportionality}) . \quad (3)$$

We can put Equations (2) and (3) in formula (1). In this case, we can find the relationship as follows

$$OP = OS - kM \cdot S\bar{p} , \quad (4)$$

and introduce coordinates defined by formula (4), which can be written as

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} - k \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ -f \end{bmatrix} , \quad (5)$$



Fig. 5 DPU including the body, controller and control computer.

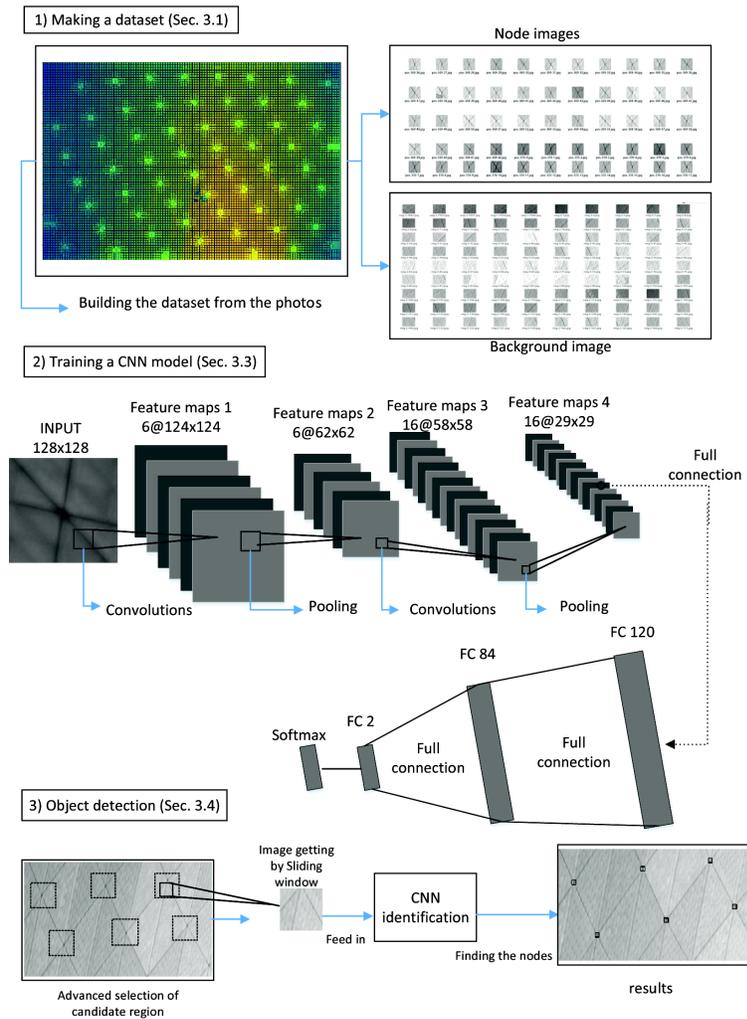


Fig. 6 Our proposed approach.

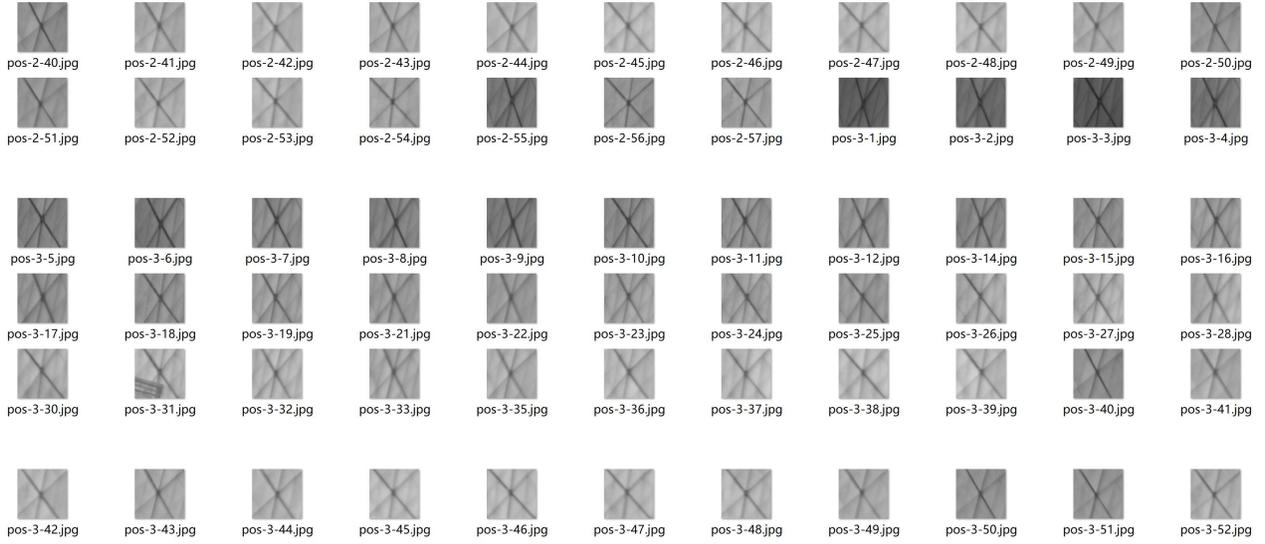


Fig. 7 The node training sets.

where f is the distance from the image plane to the center of photography and

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

is the rotation matrix.

The inverse transformation is as follows

$$\begin{bmatrix} x \\ y \\ -f \end{bmatrix} = -1/k \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} X - X_s \\ Y - Y_s \\ Z - Z_s \end{bmatrix}, \quad (6)$$

and substituting the third subtype of formula (6) into the first two subtypes, we can find the relationship as follows

$$\begin{cases} x = -f \frac{a_1(X - X_s) + b_1(Y - Y_s) + c_1(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)} \\ y = -f \frac{a_2(X - X_s) + b_2(Y - Y_s) + c_2(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)} \end{cases}, \quad (7)$$

The rotation matrix can be written as

$$\begin{aligned} M &= \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\varphi & \sin 2\varphi \\ 0 & -\sin 2\varphi & \cos 2\varphi \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (8) \\ &= \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta \cos 2\varphi & \cos \theta \cos 2\varphi & \sin 2\varphi \\ \sin \theta \sin 2\varphi & -\sin 2\varphi \cos \theta & \cos 2\varphi \end{bmatrix}, \end{aligned}$$

where

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\varphi & \sin 2\varphi \\ 0 & -\sin 2\varphi & \cos 2\varphi \end{bmatrix}$$

is the image plane that rotates 2φ around the x axis, and

$$\begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is the image plane that rotates θ around the z axis. The angle of rotation is determined by the pitch angle and azimuth angle of the DPU in Figure 9. The pitch angle is doubled because the light is reflected from the reflector to enter the lens.

The position coordinates of the nodes on the image plane are calculated by the coordinates of the projection center and the coordinates of nodes in the object coordinate system through Equations (7) and (8).

3.3 Classifier Training

Since the node image and background image structure are relatively simple, and are 128×128 greyscale images, we select the convolutional network structure LeNet5 (Haykin & Kosko 2001), which is relatively mature based on handwriting experiments. By fine tuning its structure, the input becomes an 128×128 matrix, and the output takes the form of a 1×2 vector.

Table 1 shows two CNN structures. Figures 10 and 11 display the accuracy and loss functions for the training processes of two network structures. It can be seen from

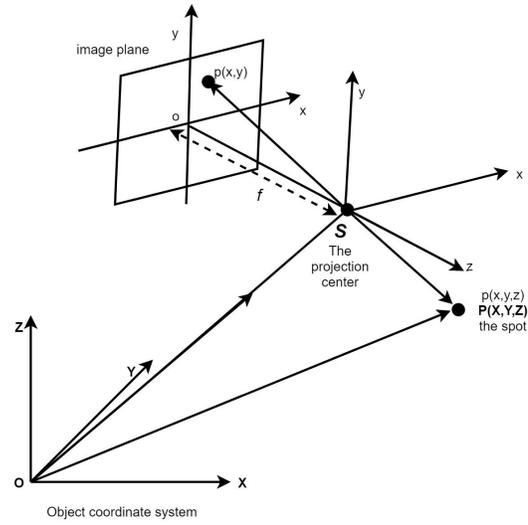


Fig. 8 Central perspective projection.

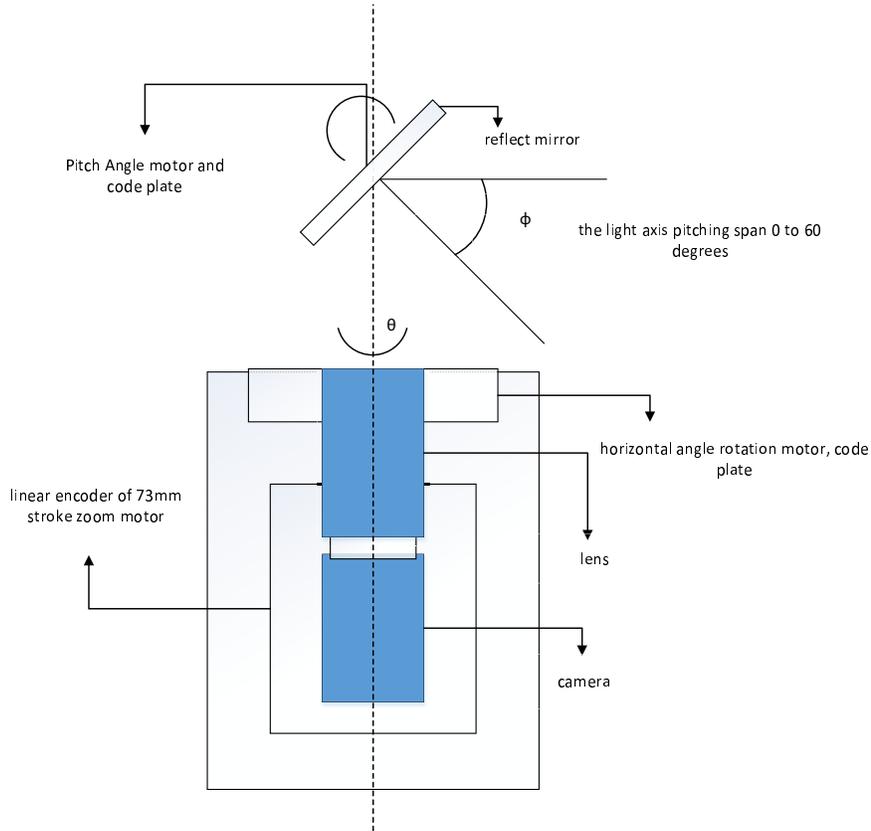


Fig. 9 DPU structure.

Table 1 ConvNet Configurations

Configuration										
ConvNet Configuration1	5 weight layers	input (128 × 128 × 1 image)	Conv5-6	Maxpool	Conv5-16	Maxpool	FC-120	FC-84	FC-2	
ConvNet Configuration2	6 weight layers	input (128 × 128 × 1image)	Conv5-6	Maxpool	Conv5-16	Maxpool	Conv5-32	FC-120	FC-84	FC-2

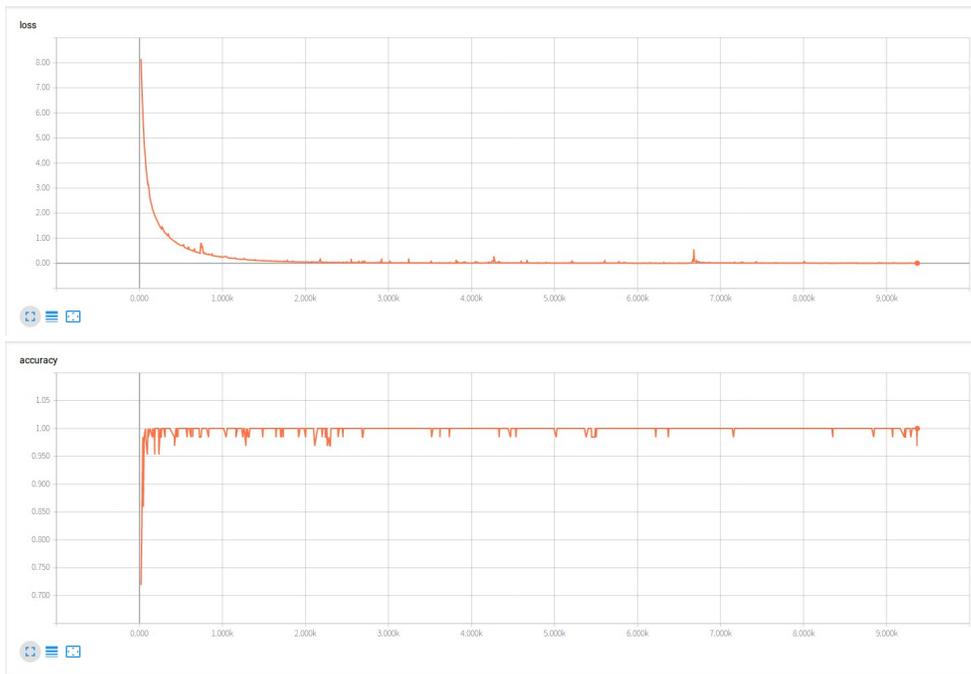


Fig. 10 Training results of configuration 1 by Tensorboard.

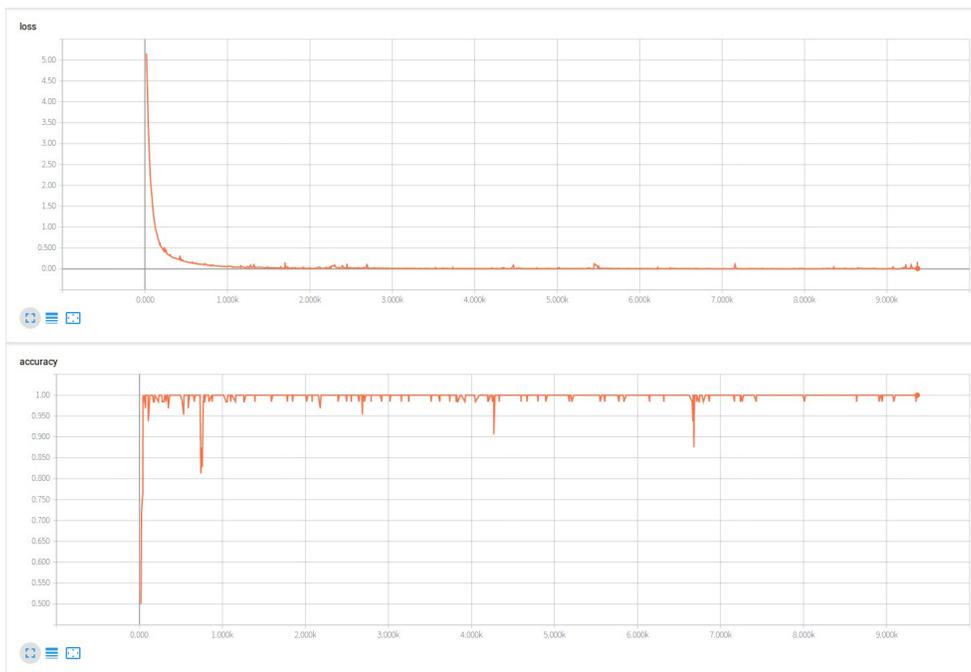


Fig. 11 Same as Fig. 10, but for configuration 2.

Table 2 Recognition Rate for the Two Methods

Method	Sphere for all	Sphere for good condition	Sphere for bad condition	Paraboloid for all	Paraboloid for good condition	Paraboloid for bad condition	Total
Traditional Edge Detection	52.1%	56.2%	47.6%	51.3%	56.9%	44.9%	51.5%
CNN with Candidate Region	95.5%	96.8%	93.6%	89.4%	90.7%	86.8%	91.1%

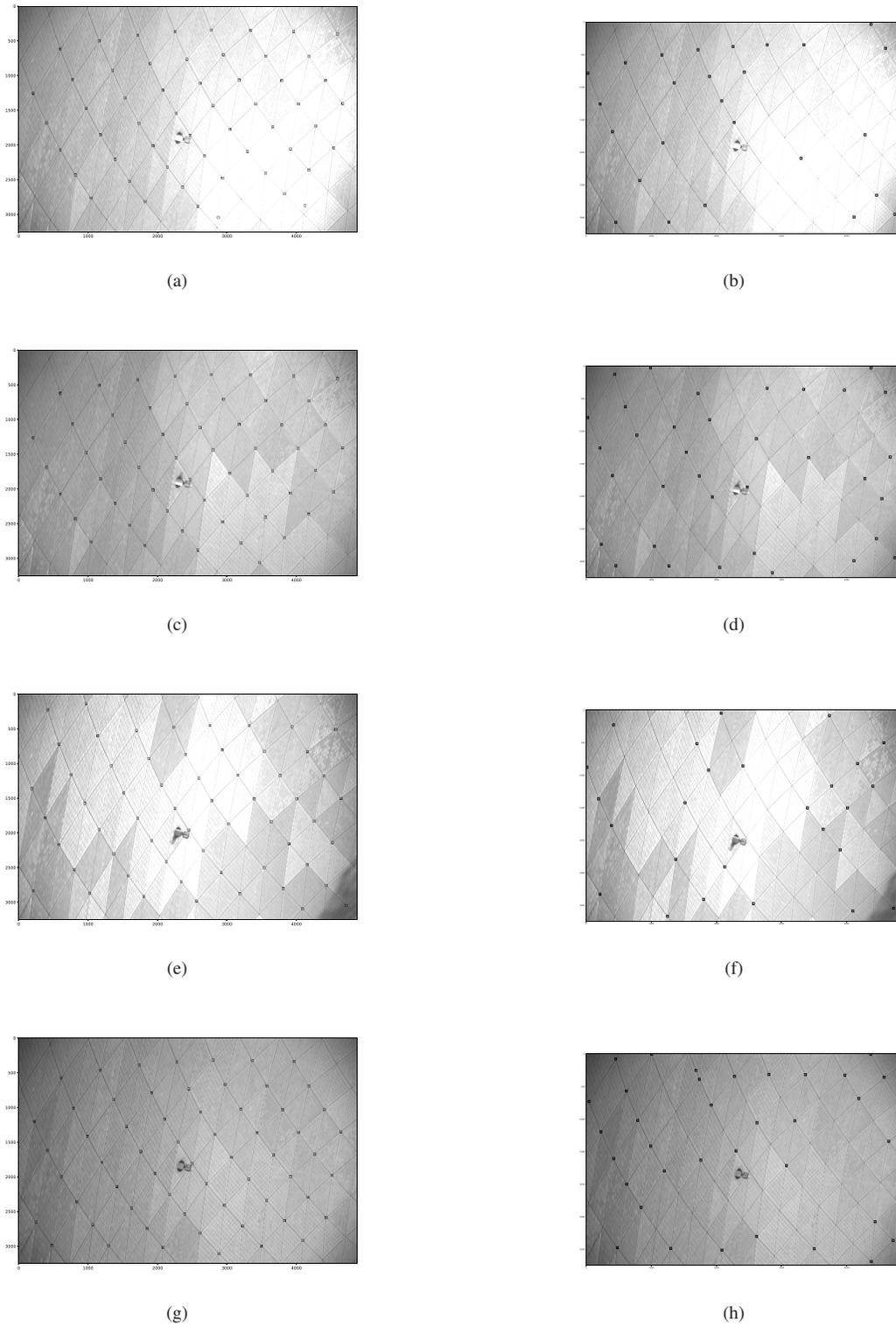


Fig. 12 Left column is detected by CNN with a candidate region, while right column by traditional edge detection. Panels (a) and (b) are paraboloid for bad lighting conditions. Panels (c) and (d) are paraboloid for good lighting conditions. Panels (e) and (f) are sphere for bad lighting conditions. Panels (g) and (h) are sphere for good lighting conditions

the figures that the loss value quickly converges to a small value, but comparing the accuracy, structure 1 converges better, so we finally choose structure 1.

Eight-thousand images are randomly selected from the more than 60 000 training images to verify the training of the model, and the other 50 000 images are employed as training sets. The final accuracy of the model in the training set is 99.89% and the final accuracy in the verification set is 99.97%. Analysis results show that the model has been able to accurately classify nodes and backgrounds in the training set and validation set.

3.4 Method for Advanced Selection of Candidate Areas

At first, we attempt to scan the whole picture acquired by the sliding window method. The result is not ideal, the node recognition rate is not high and there are a lot of errors. There are two reasons for this:

- (1) The background interference encountered in the scanning process is so much that it makes the error rate increase.
- (2) The sliding window includes 20 pixels, missing many nodes.

Because a DPU's pitch and azimuth are controlled by providing the initial position of the node, small candidate areas can be decided around the initial position of the nodes. From Li & Zhu (2012), we know that the radial displacement of nodes is less than 0.8 meters. The magnitude of the longitude and latitude displacements is even smaller, so we do not need to consider this aspect. The candidate areas can be determined as a box with 256×256 pixels around the initial location of the nodes. Then we use the sliding window and non-maximum suppression methods to find objects as a result. In order to further correct the position accuracy of a node, the result of the position is shifted up and down with an interval of 1 pixel, then we select the optimal location.

3.5 Experimental Results and Analysis

In May 2018, we collected data on FAST with a spherical surface and paraboloidal surface consisting of a total of 92 photos. The photos were divided into four categories according to the surface type and light condition, and were detected by the traditional edge detection and the method described in this paper. The recognition rate is listed in Table 2.

Compared with the traditional edge detection, the node recognition rate with this method exhibits great improve-

ment, and this method has good robustness to a change in illumination. The recognition has no significant decline under good or bad light conditions. The overall recognition rate can reach 91.5%, which satisfies the requirements of photogrammetry for recognition rate and can be continuously photographed and identified under different lighting conditions.

Figure 12 demonstrates the contrast between the traditional edge detection method and the effect of this method. We can find that under the condition of varying illumination, the recognition rate of this method is better than that of the edge detection method. In addition, the effect of our method is very stable.

4 DISCUSSION AND CONCLUSIONS

In recent years, object detection has developed rapidly, and many innovative methods have been proposed but these methods are mostly based on online training sets. According to the actual work requirements, we apply the object detection method to identifying FAST reflecting surface nodes. Considering the actual conditions, this paper proposes a method for composing the training sets and a method for choosing the candidate regions. The classification accuracy of the network successfully improves to 99.9% and the node recognition rate increases to 91.5%. The experiments show that this method is better than the traditional edge detection method and lays a foundation for feasibility of the whole non-target photogrammetry scheme. Node identification is a key part of photogrammetry. Next, we will calibrate the DPU and camera, and finally get the position of the node in the object coordinate system through the joint calculation of two DPUs.

Acknowledgements This work was supported by study on the fusion of total station dynamic tracking measuring and IMU inertial measuring for the feed support measurement in FAST (Grant No. 11503048), the Open Project Program of the Key Laboratory of FAST, NAOC, Chinese Academy of Sciences and the Key Laboratory of Radio Astronomy, Chinese Academy of Sciences.

References

- Atkinson, K. B. 2003, *The Photogrammetric Record*, 18, 329
 Canny, J. 1986, *IEEE Trans. Pattern Anal. Mach. Intell.*, 8, 679
 Edmundson, K., & Baker, L. 2001, in *17th Annual Coordinate Measurement Systems Committee Conference*, Albuquerque, NM, August 13–17, 2001, id.1, 1
 Fan, Q. H., Fang, S. H., Zuo, Y. X., et al. 2010, *Acta Astronomica Sinica*, 51, 210

- Girshick, R. 2015, arXiv:1504.08083
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. 2014, in Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR '14 (Washington, DC, USA: IEEE Computer Society), 580
- Gu, C., Lim, J. J., Arbelaez, P., & Malik, J. 2009, in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 1030
- Haykin, S., & Kosko, B. 2001, GradientBased Learning Applied to Document Recognition, in Intelligent Signal Processing (Wiley-IEEE Press), 608
- He, K., Zhang, X., Ren, S., & Sun, J. 2014, in Computer Vision – ECCV 2014, eds. D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Cham: Springer International Publishing), 346
- Hu, J., & Zhu, L. 2014, Science Technology and Engineering, 14, 190
- Li, M.-H., & Zhu, L.-C. 2012, Journal of Guizhou University (Natural Science), 24
- Liu, W., Anguelov, D., Erhan, D., et al. 2016, in Computer Vision – ECCV 2016, eds. B. Leibe, J. Matas, N. Sebe, & M. Welling (Cham: Springer International Publishing), 21
- Ren, S., He, K., Girshick, R., & Sun, J. 2015, in Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1, NIPS'15 (Cambridge, MA, USA: MIT Press), 91
- Xu, W.-X. 2006, Research and Application of Large Antenna Measurement Method., PhD thesis, The PLA Information Engineering University
- Zuo, Y. X., Li, Y., Sun, J. X., et al. 2011, Acta Astronomica Sinica, 52, 152