Automatic Counting Method of Soybean Seed Based on VGG-T

Ying Wang[a], Yue Li[a], Tingting Wu[b], Shi Sun[b], and Minjuan Wang[a]

[a]Key Laboratory of Agricultural Information Standardization, Ministry of Agricultural and Rural Affairs, College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

[b]Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China

ABSTRACT

In order to count soybean seeds quickly and accurately, improve the speed of seed test and the level of soybean breeding, this dissertation developed a method of soybean seed counting based on VGG-Two (VGG-T). Firstly, in view of the lack of available image dataset in the field of soybean seed counting, a fast target point labeling method of combining pre-annotation based on digital image processing technology with manual correction annotation is proposed to speed up the establishment of publicly available soybean seed image dataset with annotation. This method only takes 197 minutes to mark 37,563 seeds, which saves 1,592 minutes than ordinary manual marking and replaces 96% of manual workload. Secondly, a method that would combine the density estimation-based methods and the convolution neural network (CNN)-based methods is developed to accurately estimate the seed count from an individual threshed seed image with a single perspective. Finally, the model is tested, and verify the effectiveness of the algorithm through three comparative experiments (with and without data enhancement, VGG16 and VGG-T, multiple sets of test set), which respectively provided 0.6 and 0.2 mean absolute error (MAE) in the original image and patch cases, while mean squared error (MSE) is 0.6 and 0.3. Compared with traditional image image morphology operations, ResNet18, ResNet18-T and vgg16, this method improves the accuracy of soybean seed counting.

Keywords: Soybean seed counting, Soybean seed image, density map, point labeling, VGG-T

1. INTRODUCTION

Hundred-seed weight is an important yield trait of soybean, and the premise of measuring weight is to calculate the number of seeds. Fast and accurate counting of soybean seeds can speed up the test, thereby improving the level of soybean breeding, which is of great significance for increasing soybean yield. The traditional method is to manually count the grains, but it's time-consuming and labor-intensive, and errors will inevitably occur after long-term counting.[1] Compared with manual counting, the photoelectric seed counter can easily avoid errors due to chance and subjectivity, but its general shortcomings are expensive and slow of counting speed, which is not conducive to the development of large-scale agricultural production automation;[2–4] With the development of computer technology and the generalization of image information, machine vision is gradually applied to the field of soybean grain counting by scientific researchers, such as corrosion expansion method, watershed algorithm,[1,5] feature point matching,[6] etc. Compared with manual counting and photoelectric seed counting method, the grain counting method based on traditional digital image processing has indeed improved its speed and partly improved its counting accuracy. However, this type of method requires manual extraction of image features and domain expertise, which has complicated parameter tuning process, and each method is specific to specific applications, resulting in poor generalization ability and robustness.[7]

As the convolutional neural network model[8,9] has been successfully applied in many fields,[10–12] deep learning technology is also favored and used by more researchers in the agricultural field. Although convolutional neural networks have relatively few studies and applications in the field of soybean seed counting, there are many studies on other target counting. For example, Pound MP[13] established a new data set called ACID, and proposed a multi-task deep learning method

Further author information: (Send correspondence to Minjuan Wang)
Minjuan Wang: E-mail: minjuan@cau.edu.cn
Ying Wang: E-mail: wyleo@cau.edu.cn

that can accurately locate wheat spikes and spikelets while accurately classifying and counting. Deng et al.[14] established and tested the Faster R-CNN high-precision grain detection model based on the Feature Pyramid Network (FPN), which is used to automatically detect and count the number of grains per panicle. Compared with the results of manual counting of grains, the average accuracy of the model reaches 99.4%, and the detection performance is not affected by varieties and moisture conditions. Wu et al.[15] developed a linear regression model and a deep learning model to calculate the number of grains per ear, and their counting accuracy rates were greater than 96% and 99%, respectively. Compared with traditional digital image processing technology, the advantage of the convolutional neural network model is that it automatically learns and extracts useful features, and realizes automation and intelligent counting.

The realization of target counting based on convolutional neural network(CNN) provides a new idea for soybean grain counting. Image-based target counting methods can be summarized into two categories: detection-based and regression-based, including direct regression and density map regression. Due to the different density of soybean grain images and small grains, the method based on small target detection needs to train the detector to capture information, detect the target through it and calculate its number, but the training of the detector is more complicated and the amount of calculation is large.[16] At the same time, after multiple downsampling in the deep convolutional neural network architecture, the deep feature maps will lose spatial information. And the disadvantage of counting based on direct regression is that there is no precise positioning, but the method based on density regression skips the arduous identification and classification tasks, directly generates density maps, and learns the mapping between the local features of the image and its corresponding density maps, then integrating the density map to get the target count.[17] Therefore, this paper combines density estimation and CNN to construct a VGG-T model based on the grain characteristics, and then realizes the rapid and accurate counting of soybean grains from a single visual image.

2. MATERIALS AND METHODS

2.1 Materials

Soybean materials grown in the Crop Science of the Chinese Academy of Agricultural Sciences were selected. Randomly select soybean plants in 5 planting areas, and then perform operations such as harvesting, picking pods, and manually removing dirt, then performing digital imaging in the image acquisition area. Randomly spread soybean grains on a black light-absorbing background cloth. In the daytime, a camera was used to capture the original soybean seed image in a plant factory under diffuse natural lighting conditions. The camera manufacturer is SONY (ILCE-5000) with the f/4 of the aperture, the focal length is set to 16 mm, and the exposure time is 1/60 s. Place it 30-50 cm directly above the tiled seed. Figure 1 shows the acquisition device of images and the original image after preprocessing steps such as adjusting the image contrast, brightness and size.

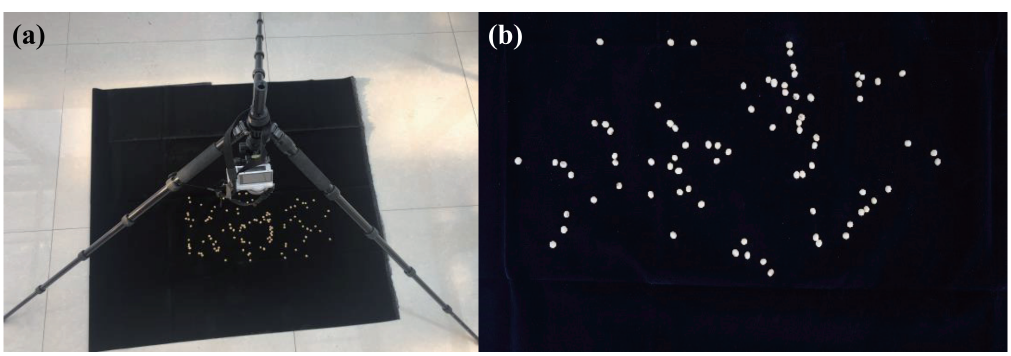

Figure 1. (a) Image acquisition device; (b) The original image after pretreatment.

2.2 Image annotation

In order to quickly establish the public soybean seed image dataset with labeling, we used a series of image analysis algorithms to automatically obtain the centroid coordinates according to the seed characteristics of the image set, and delete the wrong coordinates of adhered soybean, then manually correct them to avoid labeling errors, as shown in Fig.2.

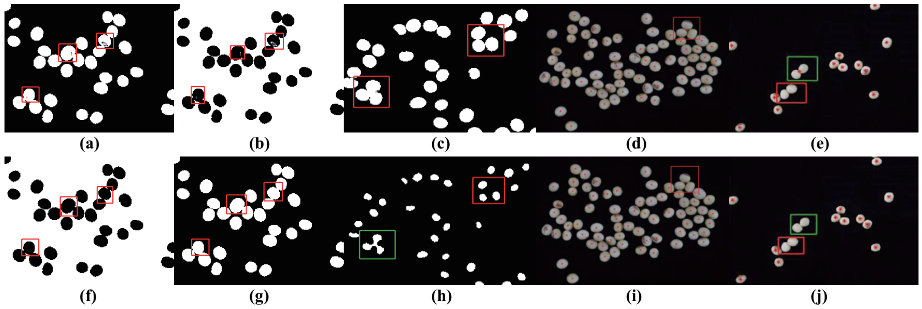

Figure 2. (a)(b)(f)(g) Sketch map of deleting small area; (c)(h) Morphological erosion rendering; (d)(i) Annotation diagram before and after deleting large area; (e)(j) Annotation diagram after deleting different large area.

2.3 VGG-T

In order to realize the network model to estimate seed density map from the input seed image, the VGG-T network needs to be trained in the early stage, and the training network needs high-quality training data. In addition to the grain images described in section 2.1, the data set used in this chapter also includes the ground truth density map corresponding to each image, as shown in Figure 3(b).

Because density regression is mostly used for population counting, and many works[18,19] have achieved significant results on the VGG16[20] network, all of which use VGG16 as the backbone to obtain good performance on many test datasets. According to the good performance on different datasets, this research also uses VGG16 as the basic network. In order to better detect small objects, shallow and deep features are combined to obtain necessary spatial and semantic information. The VGG-T network in this study has the following characteristics: after Conv4_3, two feature data streams are branched, the first one directly generates the first density map, and the second one passes through Conv5_3 to generate the first data stream. The final estimated density map was obtained by fusing the Two density maps, and the architecture diagram of VGG-T was shown in Figure 3(a).

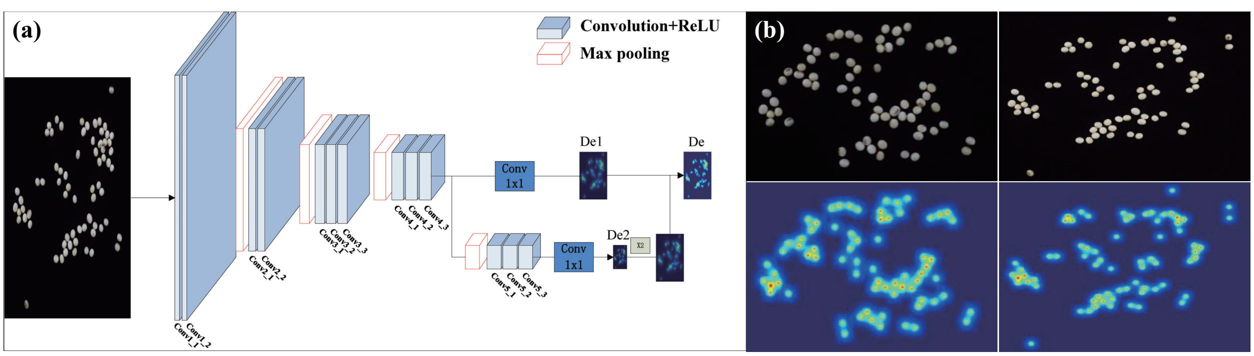

Figure 3. (a) VGG-T network architecture; (b) ground truth density map.

3. RESULTS AND DISCUSSION

3.1 Comparison of different counting methods

Use 239 original soybean seed images and 2151 patches as the training dataset to train the basic VGG16 network and VGG-T network, ResNet18 network and ResNet18-T network, while using traditional morphological operations for grain counting, table 1 is the relevant result data. According to the values of MAE and MSE, it can be seen that when performing traditional morphological operations, the corresponding errors are relatively large; for the original image, VGG-T is significantly better than the basic VGG16, ResNet18 and ResNet18-T in terms of estimation accuracy and stability; for the patch data, the performance of VGG-T is equivalent to that of ResNet18-T, meanwhile better than VGG16 and ResNet18. In summary, the results show that the fusion of data features of the proposed two branches can further improve the performance of the model, and the VGG-T network has the best comprehensive performance in all data.

Table 1. Results of different networks

Method	Patch	MAE	MSE
Traditional morphological operations		4.36	7.17
ResNet18		2.1	2.4
	✓	0.7	0.8
VGG16		2.2	2.9
	✓	0.4	0.6
ResNet18-T		0.7	0.8
	✓	0.2	0.3
VGG-T		**0.6**	**0.6**
	✓	**0.2**	**0.3**

Table 2. Comparison of seed counting time cost

Indicator	Manual counting	Photoelectric seed counting apparatus	VGG-T
Efficiency	100/80 s	1000/3 min	1/0.00857 s
Elapsed Time	2.52 h	0.23 h	0.027 h

3.2 Time cost

At present, manual counting is the most common counting method used by soybean breeders. At the same time, the photoelectric seed counter can easily avoid errors caused by accident and subjectivity. Therefore, the method in this paper is compared with the two methods in terms of counting time, as shown in Table 2. The test set established in this paper contains a total of 103 soybean seed images, including 11,350 seeds. Assuming that the counting efficiency is unchanged, it takes about 2.52 h for uninterrupted manual counting and 0.23 h for the photoelectric seed counter, but the cost is relatively high. The method in this paper takes about 0.027 h with lower cost. This method saves about 2.493 h and 0.203 h for manual counting and particle counter respectively, and the time cost is 1/94 and 1/9 of them.

4. CONCLUSIONS

This research proposes a rapid and high-precision counting method for soybean seeds. First, a soybean grain labeling system is designed. The fast point labeling of target was completed by combining pre-labeling based on traditional digital image processing and manual correction labeling, and an open available soybean seed image dataset containing 342 labeled images was constructed. Combining the density estimation method to construct a soybean seed number estimation model based on VGG-T, the MAE in the original image and the patch case are 0.6 and 0.2, and the MSE is 0.6 and 0.3, meanwhile the detection of the testset is completed in 0.027 h, finally realizing Soybean seed high-precision and fast counting task.

REFERENCES

[1] H., Z., *Design and implementation of rice seed test system based on image processing*, Master's thesis, University of Electronic Science and Technology (2019).

[2] S., R. W., J., B. D., and W., M., "A photoelectric seed counting detector," *Journal of Agricultural Engineering Research* **21**(2).

[3] Y., Z., X., Q., and Cheng, W., “A photoelectric seed counting detector,” *Automatic seed counting instrument* (3), 62–64 (2005).

[4] R., S., J., H., and J., L., “Design of photoelectric automatic particle counting instrument with rotary table,” *Transactions of the Chinese Society for Agricultural Machinery* **42**(11).

[5] F., R., “Development of crop seed automatic counting software based on image processing,” *Industrial Design* (7), 126–127 (2011).

[6] L., C., *Crop seed counting detection system based on machine vision*, Master’s thesis, North China University (2016).

[7] R., P., Q., X., and W., Z., “Digital image technology and its application in crop phenotype research,” *Journal of Changjiang University* **13**(21).

[8] Chua, L. O., [*CNN: A paradigm for complexity*], vol. 31, World Scientific (1998).

[9] Yandong, L. I., Hao, Z., and Lei, H., “Survey of convolutional neural network,” *Journal of Computer Applications* **56**(1).

[10] Alsmirat, M. A., Al-Alem, F., Al-Ayyoub, M., Jararweh, Y., and Gupta, B., “Impact of digital fingerprint image quality on the fingerprint recognition accuracy,” *Multimedia Tools and Applications* **78**(3), 3649–3688 (2019).

[11] Meden, B., Mallı, R. C., Fabijan, S., Ekenel, H. K., Štruc, V., and Peer, P., “Face deidentification with generative deep neural networks,” *IET Signal Processing* **11**(9), 1046–1054 (2017).

[12] Yu, H., He, F., and Pan, Y., “A novel segmentation model for medical images with intensity inhomogeneity based on adaptive perturbation,” *Multimedia Tools and Applications* **78**(9), 11779–11798 (2019).

[13] Pound, M. P., Atkinson, J. A., Wells, D. M., Pridmore, T. P., and French, A. P., “Deep learning for multi-task plant phenotyping,” in [*Proceedings of the IEEE International Conference on Computer Vision Workshops*], 2055–2063 (2017).

[14] Deng, R., Tao, M., Huang, X., Bangura, K., Jiang, Q., Jiang, Y., and Qi, L., “Automated counting grains on the rice panicle based on deep learning method,” *Sensors* **21**(1), 281 (2021).

[15] Wu, W., Liu, T., Zhou, P., Yang, T., Li, C., Zhong, X., Sun, C., Liu, S., and Guo, W., “Image analysis-based recognition and quantification of grain number per panicle in rice,” *Plant Methods* **15**(1), 1–14 (2019).

[16] Liu, Y., Sun, P., Wergeles, N., and Shang, Y., “A survey and performance evaluation of deep learning methods for small object detection,” *Expert Systems with Applications* , 114602 (2021).

[17] Babu Sam, D., Surya, S., and Venkatesh Babu, R., “Switching convolutional neural network for crowd counting,” in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 5744–5752 (2017).

[18] Varior, R. R., Shuai, B., Tighe, J., and Modolo, D., “Multi-scale attention network for crowd counting,” *arXiv preprint arXiv:1901.06026* (2019).

[19] Zhu, L., Zhao, Z., Lu, C., Lin, Y., Peng, Y., and Yao, T., “Dual path multi-scale fusion networks with attention for crowd counting,” *arXiv preprint arXiv:1902.01115* (2019).

[20] Simonyan, K. and Zisserman, A., “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556* (2014).