Computer Life
ISSN: 1819-4818 | Vol. 12, No. 2, 2024

Research on the Recognition Method of Wheat Ears Based on Image Color and Classical Model

Shugang Liu, Yihui Chen

School Of Computer Science, NORTH CHINA ELECTRIC POWER University, Baoding 071003, China.

Abstract: In this paper, we study wheat images during the grouting period, and construct a dataset by extracting wheat spike features through image preprocessing, color space conversion and edge detection. CNN and AlexNet models are used to train and optimize the parameters and structure to improve the recognition accuracy. By adjusting the batch size, learning rate and training rounds, CNN performs optimally with Batch size=64, Epoch=5, and learning rate=0.0001; AlexNet also performs well with similar settings, but the training cost is larger. This study verifies the effectiveness of image processing combined with the classical model CNN for recognizing wheat sheaves, providing data and theoretical support for wheat yield prediction.

Keywords: Digital image processing, edge contour detection, model structure analysis, model training method, constructive dataset.

1. Introduction

As a major grain crop in China, the stability and improvement of wheat yield is directly related to national food security and farmers' economic interests. However, the traditional wheat yield assessment method has the disadvantages of time-consuming, laborious and high cost, which is difficult to meet the demand of modern agriculture for accurate and rapid decision-making. In this paper, a wheat yield estimation [1] method based on image processing [2,3] and neural network modeling is proposed to address this problem. The method focuses on the key period of wheat growth - the grouting period, and extracts wheat spike features through OpenCV image processing technology [4,5] to realize the accurate cutting and recognition of wheat spike samples. Subsequently, a large amount of enhanced wheat spike sample data is used to train and optimize deep learning models [6] such as CNN [7,8] and AlexNet [9], with a view to advance and accurately predict the yield during wheat growth. This study is not only expected to shorten the yield assessment cycle, reduce the assessment cost, and improve the assessment accuracy, but also provide a scientific basis for the management of agricultural production, promote the development of agricultural intelligence and precision, and contribute to the protection of national food security.

2. Research on model structure and image processing process

In this paper, the original picture is a group image of wheat ears during the grouting period taken in the natural environment, due to the shooting angle, shooting distance with variability, the pixel points of the picture obtained from the shooting are inconsistent with the possibility of the picture size unified in advance, according to the resize function provided by the Image library, the expected size data of the picture as a parameter to be passed to the function to be processed, and the picture of the ears of the wheat will be unified to the 700*700 size. Before and after the wheat ears picture size unification, as shown in Fig. 1:

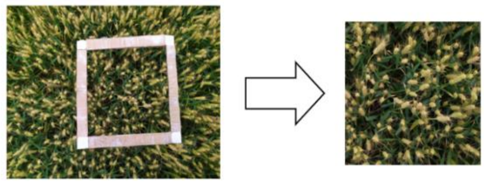


Fig. 1 Before and after harmonization of wheat images

2.1. Convolutional Neural Network Structure.

As described below, the structure of the improved network model is based on the typical CNN and AlexNet structures, combined with the different activation functions, loss functions, optimizers and normalization methods used.

Instead of processing each pixel, the convolutional neural network regionalizes each small piece of data on the image. The filter of the neural network continuously translates and selects the same image region as the kernel size according to a set step size, scales the feature [10,11] information in the image, organizes the collected pixel region information, and then, after several times of the same operation, drops this information into the fully-connected layer for classification. The typical structure of the network model is shown in Fig. 2:

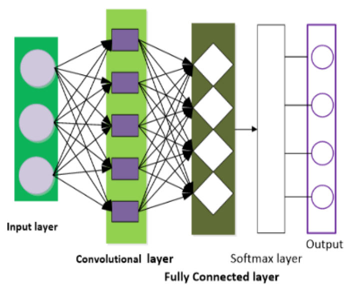


Fig. 2 Typical structure of a network model

2.2. HSV Enhanced Sobel Edge Detection method.

As the grouting period is the key period of wheat yield formation, at this time, the wheat sheaves gradually begin to be full, the shape characteristics of the sheaves are obvious, and the external color of the wheat grain changes from grey-green to bright green to green-yellow, with a glossy surface, and there is an obvious color difference with the rootstock. Therefore, in the wheat population image [12], the color characteristics and shape characteristics of wheat ears can be used to isolate the wheat ears from other backgrounds in the image, so as to extract the wheat ears samples. This is described in stages below:

Stage 1: After the image size is unified, in order to distinguish the difference between the external color of wheat ears during the grouting period and the background of other objects, the HSV [13] spatial conversion method is used to extract the hue (H), saturation (S), and luminance (V) parameters of the image for quantification, which facilitates the judgement of the color of the objects in the image, and the quantified picture is binarized to enhance the color contrast. The color enhancement results are shown in Fig. 3:

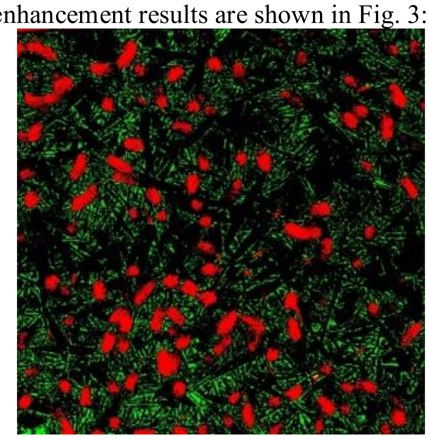


Fig. 3 Wheat ears after color enhancement

Stage 2: After the color has been enhanced, a variety of colors can be observed to exist, as the research object is wheat ears, so only the wheat ears color needs to be retained, in order to exclude interference with the color, the above image to create a mask, set the mask color threshold, the mask and binarized image for the mask operation, alone to retain the characteristics of the wheat ears in the picture. The results of the mask processing are shown in Fig. 4:

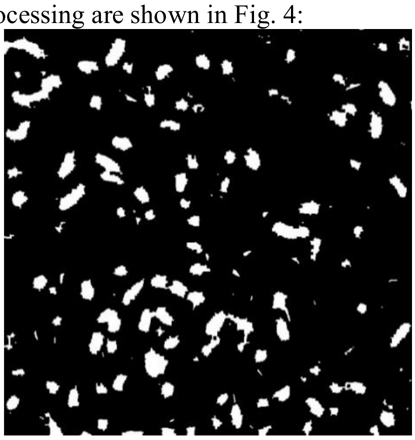


Fig. 4 Wheat ears after image mask processing

Stage 3: After specification and HSV space conversion, the most characteristic data of the wheat spike image is the edge of the wheat spike, the edge contour of the wheat spike is detected, the data of the closed contour is taken, and the Sobel operator is used to calculate the overall edge of the wheat spike, and the minimum approximation matrix of each contour is obtained for cutting, and the wheat spike sample is obtained. The minimum approximation matrix after Sobel operator detection is shown in Fig. 5 below:

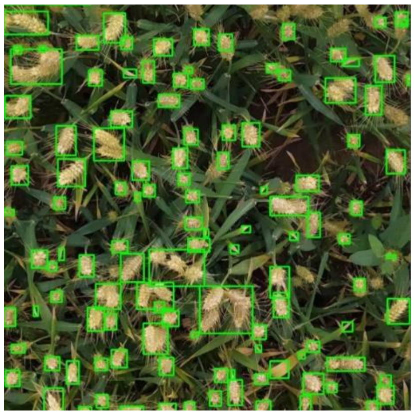
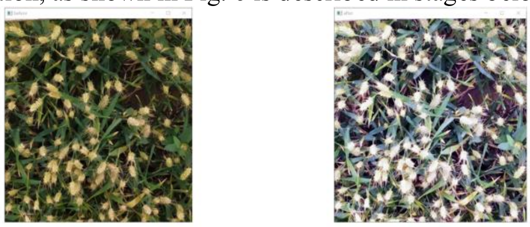


Fig. 5 Minimum approximation matrix plot for wheat ears

2.3. Equalized K-mean clustering image processing method.

After the image size is unified, because the difference between the external color of wheat ears and the background of other objects during the irrigation period is more obvious, this paper uses the histogram [14] to calculate the probability of the color distribution in the image, calculates the transformation relationship between the original image and the target image, and scales the grey level of the pixels under the effect of the transformation relationship to redistribute the colors in the image, so as to make the distribution of the colors more uniform, and get the color differences more concentrated in the Wheat ears image. Before and after image equalization, as shown in Fig. 6 is described in stages below:



Before picture equalization After image equalization

Fig. 6 Comparison of wheat ears before and after equalization

Stage 1: Histogram equalization strengthens the features of each sheaf in the picture, K-means [15] clustering analysis edge detection algorithm is used to cluster the sheaf pictures whose features have been strengthened, and a distance threshold is set to automatically assign similar samples to a category [16], which constitutes a target cluster, i.e. a sheaf head. The clustered image is shown in Fig. 7:

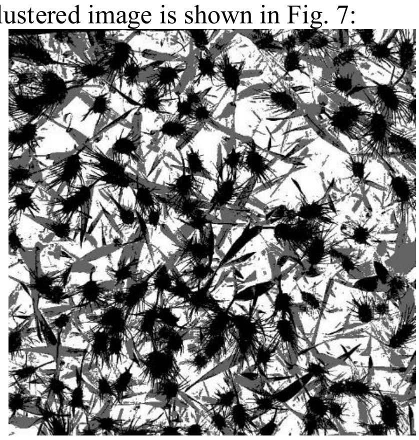


Fig. 7 After clustering of wheat images

Stage 2: According to the difference to set the threshold, to the wheat ears features [17] will be obtained, open and close operation to remove the details [18], identify the contour after the calculation of the area of each closed contour, take the minimum approximation matrix for each wheat ears contour according to the cut to save, to get the initial wheat ears sample dataset [19]. The minimum approximation matrix plot of wheat ears is shown in Fig. 8:

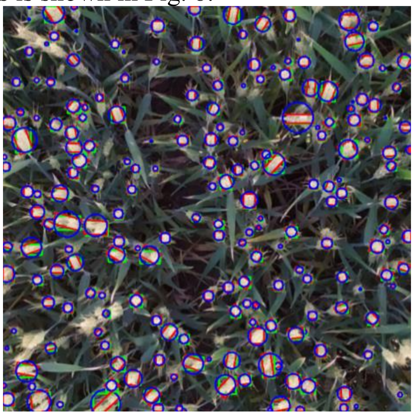


Fig. 8 Plot of the minimum approximation matrix of wheat ears

Based on the improved and newly proposed image processing techniques, a more accurate feature recognition [20,21] was obtained than the image technique in the literature [22], which improved the number of wheat ears recognized with a 30% improvement in effectiveness.

3. Results and analysis

After building the neural network in the computer program, the training of the model was started. The two models have the same dataset, 4 categories, 6000 sample images per category, a kind of 24000, and the training set ratio ration=0.7 is used to train the model.

3.1. Model Training and Corresponding Results.

After building the neural network in the computer program, the training of the models began. Both models have the same dataset, 4 categories, 6000 sample images for each category, 24000 for one kind. The training set ratio of 0.7 was set in the computer program to train the models. For the CNN model, the AlexNet model was trained by constant testing, choosing a different Epoch, Batch size, STEP, or learning rate each time.

Improved CNN model training process and results

The CNN model is convolved with ReLU and uses LRN to prevent overfitting, the pooling layer uses the same kernel size, the step size is set to 1 in the computer program, and the loss function Reduce mean and the Adam optimizer are used. Detailed training improvement records are shown under Table 1:

Table 1 CNN model training improvement records

Epoch	Batch size	Step of 1 epoch	Actual STEP	learning rate	accuracy
0.25	20	1200	300	0.000001	0.24
1	32	750	750	0.0001	0.62
1	64	375	375	0.0001	0.64
1	128	188	188	0.00001	0.25
1	128	188	188	0.0001	0.50
2	256	94	188	0.0001	0.53
3	20	750	3600	0.000001	0.35
3	64	375	1125	0.0001	0.56
3	128	188	564	0.0001	0.70
3	256	94	282	0.0001	0.55
3	512	47	141	0.0001	0.50
5	64	375	1875	0.0001	0.85
5	64	375	1875	0.00001	0.55
5	128	188	940	0.0001	0.80
5	256	94	470	0.0001	0.72

Improved AlexNet model training process and results

In the computer program AlexNet model convolutional layer is set with different convolutional kernel sizes and move steps, ReLU is used after convolution, a total of three pooling layers are set up, each using the same kernel sizes, step sizes, Cross Entropy Loss function Cross Entropy Loss and Adam Optimizer, and finally Dropout is added in the Fully Connected Layer to prevent overfitting. The training improvement record is shown in Table 2:

Table 2 AlexNet model training improvement records

Epoch	Batch_size	Step of 1 epoch	Actual STEP	optimizer	learning rate	accuracy
1	128	188	188	SGD	0.0001	0.25
1	128	188	188	SGD	0.000001	0.25
1	128	188	188	Adam	0.0001	0.63
3	64	375	1125	SGD	0.0001	0.25
3	128	188	564	SGD	0.0001	0.25
3	128	188	564	Adam	0.0001	0.85
5	128	188	640	Adam	0.0001	0.25

3.2. Analysis.

The evaluation metric is the accuracy rate, which reflects how good a model is. The higher the accuracy rate, the better the learning ability of model training. In this paper, we evaluate which model is better from four aspects: Batch size, Learning rate, Epoch, and the time used to train the model. The detailed evaluation conclusions are as follows:

(1) Batch size:

Sample batch size has a significant effect on the learning ability during model training. Experiments indicate that CNN model performs best at Batch size=64, while AlexNet performs better at 128. Larger Batch size requires longer processing time, so CNN performs better in this regard.

(2) Learning rate:

The learning rate directly affects the gradient descent rate, which in turn affects the accuracy. Experiments indicate that the training effect is best when the learning rate is set to 0.0001.

(3) Epoch:

Epoch represents the number of training rounds, which affects the training time and effect. Experiments indicate that the accuracy of CNN increases to 85% when Epoch=5, while AlexNet reaches this level at Epoch=3. Considering the training time, CNN learns the features earlier.

In conclusion, this paper experimentally analyzes the effects of Batch size, Learning rate and Epoch on model accuracy, and finds that appropriate parameter settings can significantly improve the model performance, as shown in the following Fig.9:

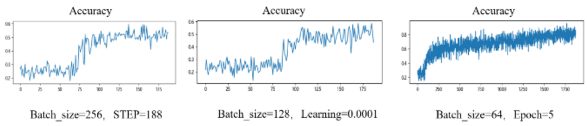


Fig. 9 Effect of Batch size, Learning rate, and Epoch improvement

4. Summary

In this paper, for wheat images during the irrigation period, by optimizing the image processing parameters with the introduction of HSV color space [23,24] and Sobel operator, the number of wheat ears recognition was significantly improved, and the accuracy of the optimized image processing method was verified. Subsequently, using these process sed image data, the CNN and AlexNet models were trained and optimized, and by fine-tuning the model structure and parameters, higher wheat spike recognition accuracy was achieved. This research not only enhances the accuracy of wheat growth monitoring, but also provides technical support for agricultural intelligence. In the future, it is planned to integrate the research results into the WEB platform and build a visualization system to make the management of agricultural production more intuitive and convenient, and help the efficient development of modern agriculture.

References

[1] Jahui Zhang. Research on the model of jujube yield estimation based on image processing [D]. Hebei Agricultural University, 2019.

[2] MorenoLucio Mireyaetal. Extraction of Pest Insect Characteristics Present in a Mirasol Pepper (Capsicum annuum L.) Crop by Digital Image Processing[J]. Applied Sciences, 2021, 11(23):11166-11166.

[3] Fangzheng Tianetal. Automatic coronary artery segmentation algorithm based on deep learning and digital image processing[J]. Applied Intelligence, 2021:51(12):1-15.

[4] Hasan Aetal. Quantitative assessment of mosaic disease severity based on digital image processing[J]. IOP Conference Series: Earth and Environmental Science, 2021, 694(1):012043-.

[5] Samadzadegan Farhadetal. Detection and Recognition of Drones Based on a Deep Convolutional Neural Network Using Visible Imagery[J]. Aerospace, 2022, 9(1) : 31-31.

[6] JU Wenhui. Research on visual counting method based on deep learning [D]. Shandong Normal University, 2021.

[7] Cheng-Cheng Huang,Hua-Mei Huang. Architecture and application development of convolutional neural networks[J]. Information and Computer (Theoretical Edition), 2019(07):143-144.

[8] Zhang Chen. Face recognition method and application based on AlexNet convolutional neural network model[J]. Journal of Ezhou University, 2022, 29(01):102-104.

[9] Kong Yunchen and Ma Xue and Wen Chenglin. A New Method of Deep Convolutional Neural Network Image Classification Based on Knowledge Transfer in Small Label Sample Environment[J]. Sensors, 2022, 22(3):898-898.

[10] WANG Xin, PAN He, ZHAO Ying. Research on corn kernel shape feature detection technology - based on image processing[J]. Agricultural Mechanization Research, 2015, 37(06):46-48+52.

[11] MA Xiaodan, ZHU Kexin, GUAN Haiou, FENG Jiarui, LIU Meng, ZHENG Ming. Crop image feature extraction technology and its application in disease diagnosis[J]. Journal of Heilongjiang Bayi Agricultural Reclamation University, 2019, 31(02):93-99.

[12] LI Yi-Nian, DU Shi-Wei, YAO Min, YI Ying-Wu, YANG Jian-Feng, DING Qi-Shuo, HE Rui-Yin. A method for counting wheat ears and yield prediction in the field based on wheat population images[J]. Journal of Agricultural Engineering, 2018, 34(21):185-194.

[13] ZHAO Xiaoyu, CHEN Jianjun, ZHANG Kaiqi, QIN Qiaoting, YOU Huotian, HAN Xiaowen. Automatic extraction of vegetation coverage from UAV images based on HSV color space and Otsu algorithm[J]. Science, Technology and Engineering, 2021, 21(35):15160-15166.

[14] Haiqiong Wen,Jiancheng Li. Adaptive thresholding image enhancement algorithm based on histogram equalization[J]. China Integrated Circuit, 2022, 31(03):38-42.

[15] LIU Jun,YU Qing.Research on K-Means clustering algorithm and its performance optimization[J]. Specialized Equipment for Electronic Industry, 2020, 49(05):46-49.

[16] BAO Zhiqiang, ZHAO Yuanyuan, HU Xiaotian, ZHAO Yan. A new K-Means clustering algorithm insensitive to isolated points[J]. Modern Electronic Technology, 2020, 43(05):109-112.

[17] Zhu Xiaoqian. Research on bubble feature extraction based on image processing [D]. Central China Normal University, 2021.

[18] LI Zongnan, CHEN Zhongxin, WANG Limin, LIU Jia, ZHOU Qingbo. Extraction of corn downed area based on small UAV remote sensing[J]. Journal of Agricultural Engineering, 2014, 30(19):207-213.

[19] Hao Wu. Research on deep learning method for image processing based on synthetic dataset [D]. Lanzhou University, 2021.

[20] ZANG Jing, GUO Qianqian, LI Chenghua, FU Lidong. Research on identification and statistical methods of multiple corn kernels under complex conditions[J]. Journal of Shenyang Agricultural University, 2014, 45(5):633-636.

[21] DU Shi-Wei, LI Yi-Nian, YAO Min, LI Ling, DING Qi-Shuo, HE Rui-Yin. Seed counting method based on spikelet image segmentation in wheat spike[J]. Journal of Nanjing Agricultural University, 2018, 41(4):742-751.

[22] Li Haiyang. Recognition and counting of wheat ears based on image processing and deep learning [D]. Zhengzhou: Henan Agricultural University, 2021.

[23] LIU Changhua, HU Linlong, YANG Kang, WU Xiaoming. Segmentation of aerial rape flower images based on color space[J]. Journal of Wuhan Light Industry University, 2020, 39(01):13-21.

[24] JI Sweet, LI Zebin, ZHAO Jiangdong, YAO Youfeng, HUANG Ji. Segmentation method of plant leaf disease image based on color variability[J]. Hubei Agricultural Science, 2018, 57(18):94-97.