

A Wheat Grade Identification and Bad Grain Detection System Based on EasyDL

Hanrui Liu, Yushuo Bai, Jiaxuan Cao, Yinpeng Li, Jiayue Liu
Hebei University of Science and Technology, Shijiazhuang, Hebei, China

Abstract: At present, there are many problems of insufficient professional testing personnel, low testing equipment and low accurate identification. The wheat grade identification and bad grain testing system based on EasyDL will bring profound changes to the agricultural product quality testing industry. This system is implemented in the AI development platform of Baidu fly pulp EasyDL, and the innovative image recognition technology is used in wheat detection. The image features are extracted through the convolutional neural network, and the supervised learning algorithm is used for model training. Using Requests, OpenCV, Xlsx Writer third-party toolkit for data processing, it can realize the detection of seven types of wheat grains and bud gains, and the detection accuracy can reach 95.94%, and the accuracy of quantitative statistics is 98.95%.

Key words: Quality and Safety Detection of Agricultural Products; Wheat Detection; Easydl; AI Development; Image Recognition Technology; Convolution Neural Network for Image Extraction; Internet+ Crop Quality Detection; Technology and Innovation; Python Programing Language

1. Introduction

Wheat grade identification and bad grain detection system can solve the problems of less advanced grain detection equipment, few detection parameters and manual quality inspection in the detection stage. The system uses image recognition technology through convolutional neural network [1] The image features are extracted, the supervised learning algorithm is used for model training, and the corresponding parameters are added according to the national standard document to realize the effect of intelligent recognition of wheat grains. The detection system is written using

Python. The language is simple and easy to use, powerful, efficient, easy to maintain, widely used in data analysis, and has a large number of powerful third-party toolkits used in every development link, throughout the whole development process. The data packets used in this system are include Requests, PyQt 5 , OpenCV [1], XlsxWriter.

Using PyQt 5 for system graphical user interface development, this software is the combination of Python and Qt, on the basis of the high operating efficiency of Qt, and add the characteristics that Python is easy to develop. It is not only very convenient to develop the graphical user interface, but also beautiful style, excellent cross-platform effect, perfect support for Mac, Windows, UNIX and other systems. Binding to the OpenCV. [1] For graphic processing, OpenCV has powerful image processing function, in this system is mainly used to make sample marking results, using the mark data obtained from EasyDL SDK, on the basis of the original picture, add custom color translucent mark box, the mark results one box, thus can be more intuitive observation, and OpenCV has C++, Python, Java and aMATLA interface, more widely used.

The XlsxWriter toolkit is used to generate and save Excel files, the toolkit is a powerful and convenient Excel file writing storage, a few lines can create Excel files, the corresponding content by line and column, relying on Python, powerful data processing function, you can quickly and conveniently export the test results of the system as an Excel file.

2. Specific Implementation Methods

The core function of the system is to upload the local sample picture to the detection service address and obtain the results, which is derived from the service address opened on the machine after EasyDL SDK operation. As shown in Figure 1. Send the post request to the corresponding address using requests' post

method, the parameter threshold represents the control over confidence, only the detection results with confidence exceeding this value are returned, while the local image file is uploaded in the data. After receiving the response, the detection results are deposited in result, variable, and the results are converted to dictionary form in the JSON method for the following data processing.



Figure 1. Process Flow

The graphical interface of the system was made using the PyQt 5 toolkit. First, the layout of the UI should be determined. The system uses a large number of vertical layouts, horizontal layout and grid layout in PyQt 5, and arranges each component in order according to the appropriate way. After determining the layout mode, adding the components, and arranging the various positions of the components, you can make UI classes, with UI names of "UI _ window name", which is used for the UI generation of the window and the arrangement of the components.

After the UI class is made, create the class corresponding to the window, and inherit the QWidget or QMainWindow class and the corresponding UI class. Call the setupUi method of the parent class to realize the display of initialized UI components in the window. After that, functions can be installed, and sub-Windows, window bearer information can be stored into the current window class using member variables, and the specific implementation of various functions can be written into member methods, and then the clicked events of PushButton can be connected to the corresponding method, so as to realize the system functions. Image was processed using the OpenCV toolkit, Parsing the acquired data, Each piece of the data contains the x1, x2, y1, Of y2, the four parameter values, Coordinates representing the upper left and lower right corner of the marked region, Where the coordinate value is the proportional position of the point in the image as a whole, For example, when the x1, And y1 was 0.3, respectively, At 0.4, Suggesting that the point is located from left to right three-tenth of the whole, From top to bottom, The real pixel

coordinates corresponding to this coordinate were obtained using the get Pos function, The function uses a formula of:

$$Y = x * total \quad (1)$$

Where x represents the obtained coordinate value, y represents the actual pixel coordinate position, and total represents the corresponding side length. As shown in Figure 2. In the specified area draw two equal size rectangle, the first rectangle directly in the copy of the original picture of the frame, style for solid line border, no filling, another rectangular frame painted on the blank mask layer, style for full fill pure color rectangular box, after all the mark box is drawn, the blank mask layer with certain transparency and the original picture combination, then get the final mark effect.



Figure 2. Mark Effect

3. Test and Test the Analysis

3.1 Data Source and Scale

Wheat grade identification, bad particle detection system, the data set adopts the data set provided by the enterprise, a total of 87 pictures of wheat grain. After pre-processing, the system marks, trains and verifies the data pictures. Through repeated training and verification, the identification accuracy of the system is 95% and the quantity accuracy is 98%, which has exceeded the national standard, and the training results are super Expected out, to reach the ideal state [2].

3.2 The Model was Built Successfully

The team trained and verified the model for a month and built the entire system. The team runs the system to detect the wheat grain pictures, and we can see each one clearly through the content displayed on the page the type of wheat grain, integrity rate and other information results[3],[4]. It can also be passed through the Click on the tab bar on the right, and then click on the wheat grain for accurate positioning. Users can view the relevant information of each wheat grain according to

their needs, which plays a vital role both for the future applications and for the production development. As shown in Figure 3.

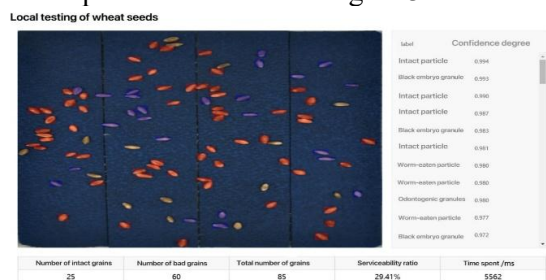


Figure 3. System Operation

3.3 Analysis and Conclusion

Based on the training, calibration and rectification of the system mode, the whole system tends to be mature in the training. By calling the OpenCV library [5]. Mark the original pictures according to the marking results, improve the effectiveness of the system verification results, improve the transparent marking layer, visual expression to enhance the readability of the training results and facilitate users to read the data better.

4. System Implementation

4.1 Data Collection

Data collection uses the way of photography, detection needs to use the uploaded wheat picture provided by the enterprise.

4.2 Data Processing and Management

By using Baidu EasyDL for the marking processing of wheat grains, after uploading pictures, different types of wheat grains are marked with labels, and the intelligent annotation function is used for preliminary large-scale labeling. The total number of grains marked, and the number of different grains will be recorded separately for overall testing. Use the marked picture for model training, set as the local deployment -small general device. After the training, download the SDK to the local, get the detection trial serial number and run the service locally.

4.3 Analysis Results

After the initial training of the model, the results of the image recognition were analyzed. The analysis results of the single picture are shown in the following figure, as shown in Figure 4.



Figure 4. The Analysis Results of the Single Picture

In the case of scattered wheat grains, small numbers, and single species, the detection probability of scab reached 100%.

For the problem that misidentification of black grains and individual wheat grains cannot be recognized., We tried to increase the capacity of the data set (adding dozens of intact images for training) to obtain clearer data to train the model[6]. This method can solve these two problems more satisfactorily. The following figure shows the detection results of the mixed wheat grain after the improved model, which initially achieved the expected target. As shown in Figure 5.

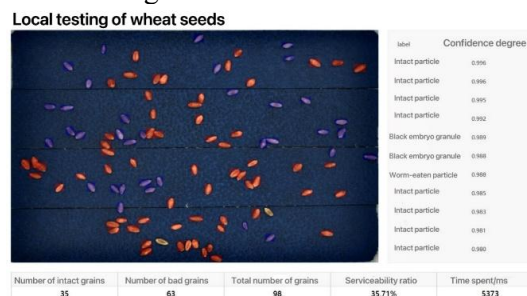


Figure 5. The Detection Results of the Mixed Wheat Grain after the Improved Model

4.4 Use of the Analytical Tools

The 3 Python implementation use Pycharm, build engineering, download PyQt 5 and design tools, make interface style in the design tool, add resource picture path selection, result storage path selection, status display and results display section. Export the ui file to the project. Send please to the local service address deployed via the request module. Get the results of the mark. All the pictures in the path of the resource picture were uploaded and integrated in turn, the total number of wheat grains and the number of intact grains were counted, and the perfect rate of the tested samples and the total perfect rate of the samples were calculated. The results of the

detected samples are stored in the txt file of the corresponding name, and the total results are stored in the txt file named "Results of All _ detection time" and finally stored under the results storage folder. Check the results of intact grain, bad grain, and perfect rate in the picture, and then the results are successfully rendered into the page and presented intuitively in the browser.

5. Cheng, Preface and Description

5.1 Class Instructions

The class structure of this system is shown in Figure 6.

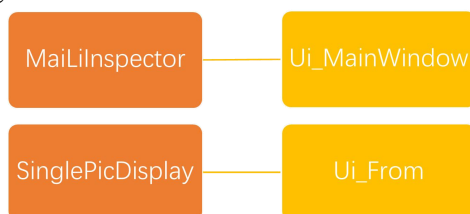


Figure 6. The Class Structure of This System

5.2 UI Class

There are two UI classes: `Ui_MainWindow` and `Ui_Form`, which are the main interface UI class and the single picture viewing window UI class respectively. There are two methods: `setup Ui` and `Retranslate Ui`: the `setup Ui` method is used for setting layout management, creating components and changing styles, and the `retranslate Ui` is used for dynamic translation of component text.

5.3 Window Class

The project has two window categories, `MaiLiInspector` and `SinglePicDisplay`, which are the main interface window class and the window class for single picture view. Inherit the `QtWidgets` separately. `QMainWindow` and `QtWidgets`. `QWidget` Application framework, both construction methods are to call the `in it` method of the parent class and call `setup Ui` to set their own UI layout, and then set the window icon, and then call the respective initialization function. The initialization function shared by both parties is `btnCnt`, which can bind the button events with the corresponding functional implementation method. The unique initialization functions in `MaiLiInspector` are in `itModel`, in `itConfig` and in `itLib`, which are Model of the initial table,

read data and archive from the configuration file, initial file read path and result save file. `Mai Li Inspector` Methods include `start Inspect` to perform detection, `single image detection pic Inspect` and `open single image preview window` to export the data as `open Single Pic Display` and `createXlsx of Excel`.

5.4 Difficulties Encountered and Solutions

The effect of data visualization and image extraction is not intuitive enough, and the validity control of the detection and verification results is insufficient. In order to improve the accuracy of the detection results, the `OpenCV` library is called to mark the original pictures according to the marker results. The `x 1, y 1, y 2, y 2` in the detection data represent the coordinates of two diagonal points marking the rectangular box. Constrained the aspect ratio, because the original data is the proportion of the relative length and width of the value range `0 -1`, using `Eq Actual coordinates = the original data image width` The precise position and size of the excellent block are extracted from the picture, obtaining the actual position of the rectangular box in the picture, and labeling the fluorescein in different colors to distinguish the intact and diseased particles.

Drawing the pie chart to observe that the distribution of diseased particles is affected by the selection of sample data to help confirm the main disease particle types in the current wheat particle samples, the percentage of the wheat particles was counted in the assay, and pie charts were drawn using the `matplotlib` library and displayed on the software panel. All wheat grain samples were initially used as data for drawing the pie charts:

(1) When the proportion of diseased particles is too small, the distribution of diseased particles cannot be seen through the pie chart, so the data of intact particles is not used when drawing the pie chart;

(2) When the integrity rate reaches 100%, a condition minus 0 will occur because of no disease particle sample when drawing the pie chart. Therefore, one grain named "single grain ratio" was added to the data.

(3) The number of sick particles can also be shown by the proportion of individual particles, while preventing the division of 0. The smaller the proportion of single particles, the more the number of diseased particles: Number of

diseased particles = 1 / single grain ratio in addition, the number of disease particles can also be obtained through the formula: the number of a disease particle = a kind of disease particle ratio / single particle ratio

6. Features and Innovation Points

Wheat grade recognition and bad particle detection system innovates the image recognition technology for wheat detection, extracts the image features through the convolutional neural network, and uses the supervised learning algorithm for model training. Image recognition: As an important aspect of artificial intelligence, image recognition has been widely used in various fields. The working principle of image recognition technology is to detect significant area, which contains the most information, image or object, image recognition is the image of pattern recognition, it is mainly divided into the following steps: input image information, extract image features, establish a classifier, according to the extraction image feature classification. (Image acquisition image preprocessing feature extraction image recognition. Can replace the former) using the efficient computing power of the computer, can quickly complete the data characteristics Taking and the calculation of the classification[7],[8]. The three main tasks of image recognition include: target recognition (classifying the picture, determining the target of the picture), target detection (positioning the target, determining the specific position of the target in the picture), target segmentation (classifying the foreground and background of the picture, eliminating the back Scene and describe the shape of the target)[9],[10]. Convolutional neural network (CNN): CNN is essentially a multi-layer perceptron. Multi-layer perceptron (MLP) is a neural network model composed of input layer, hidden layer (one layer or multiple layer) and output layer. The basic structure of CNN is composed of input layer, convolutional layer and pooling layer (Also known as sampling layer), full connection layer and output layer composition[11].

The local connection and shared weights adopted by CNN, on the one hand, reduces the number of weights and makes the network easy to optimize, and on the other hand, reduces the risk of over fitting. Supervised

learning algorithm: the process of adjusting a set of known categories and adjusting the parameters of the classifier to achieve the required performance. Supervised learning is to infer a functional machine learning task from labeled training data. In supervised learning, each instance is composed of an input object and a desired output value. The supervised learning algorithm is developed by analyzing the training data to produce an inferred function that can be used to map new instances. An optimal scheme would allow the algorithm to correctly determine the class labels of those invisible instances, requiring that the learning algorithm is formed from a training data to an invisible condition in a "reasonable" way.

convolutional neural network[12]. Using the BP framework for learning in supervised learning is one of the first deep algorithms for learning in the BP framework. BP in the convolutional neural network is divided into three parts, namely the back propagation of the fully connected layer and the convolution core and the reverse pathway (backward pass) of the pooling layer. The BP calculation of the fully connected layer is the same as the traditional feedforward neural network, and the backpropagation of the convolutional layer is a cross-correlation calculation similar to the forward propagation.

7. Conclusions

The wheat grade recognition and bad particle detection system innovatively applies image recognition technology to wheat detection, extracts image features through convolutional neural networks, and uses supervised learning algorithms for model training. By analyzing existing pain points and collecting sufficient data, accurately building models, and improving the industrialization process, it is of great significance to improve food productivity and ensure food security. Effectively solved the problems of insufficient detection parameters, insufficient detection equipment and instruments, and outdated crop detection stages in crop detection. Where E is the error calculated by the cost function, f' is the derivative of the excitation function, and α is the learning rate (learning rate). If the forward propagation of the convolution kernel is calculated using convolution, the back propagation also lips the convolution kernel to

perform the convolution operation. The error function of the convolutional neural network can be.

There are a variety of choices, common including Softmax, loss function, hinge loss function, triplet loss function, etc.

Reference

- [1] Zhou Feiyan, J in Linpeng, Dong Jun. Review of convolutional neural network research. *Journal of Computer Science*, 2017,40 (6) :23.
- [2] Zhang Yurong, Wang Qiangqiang, Wu Qiong, et al. Based on the Python-OpenCV Wheat imperfect grain recognition study by image processing technology. *Journal of Henan University of Technology: Natural Science Edition*, 2021,42 (6): 8-10.
- [3] Deng Qiang, Jing Yuan, CAI Yanhong, et al. Agricultural product quality and safety testing Analysis of the input and output risks in the organization workflow. *Journal of Food Safety and Quality Testing*, 2016,7 (11): 6-9.
- [4] Hou Guohu. Agricultural product quality and safety supervision and testing work of the difficulties and Strategy for inquiry. *Seed Technology*, 2020,38 (5): 12-15.
- [5] Green rice and my wood. Agricultural product quality safety detection exists in the work and its countermeasures. *Southern Agriculture*, 2020,14 (5): 2.
- [6] Feng Xiaoxia. Research on the image recognition algorithm based on deep learning. North China University of Technology,2021.
- [7] Gao Junyu, Yang Xiaoshan, Zhang Tianzhu, Xu Chang sheng. Based on deep learning Robustness of the visual tracking method. *Journal of Computer Science*.2016 (07):56-72.
- [8] Gao Junyu, Yang Xiaoshan, ZhangTianzhu, Xu Chang sheng. Based on deep learning Robustness of the visual tracking method. *Journal of Computer Science*.2016 (07)
- [9] Wang Haichao, Kaohsiung, Chen Tiej ing, et al. Crop disease detection Status and prospects of spectroscopic and image processing technologies. *Agricultural Research*, 2015,37 (10): 69-72.
- [10] Li Yandong. Key technique of computer vision based on convolutional neural network Surgical study. University of Electronic Science and Technology of China,2019.
- [11] Li Yandong, Hao Zongbo, Lei Hang. Review of convolutional neural network research. *Computer App lications*, 2016,36 (009) :2508-2515.
- [12] Bradsky, K ler, Yu Shiq i, et al.study OpenCV: Chinese version. Tsinghua University Press, 2020.