Implementation of Apple's automatic sorting system based on machine learning

ZhiYong ZOU, Tao LONG, Qi WANG, Li WANG, Jie CHEN, Bing ZOU, Lijia XU

Abstract

In order to reduce post-harvest losses, the classification of fresh apples is crucial. Taking the hierarchical transmission control system as the object, the research was carried out on the verification of the bus network can flexibly expand the motor equipment, the stable and reliable operation of the motor, and the accuracy of Apple's classification. Combine Labview virtual instrument technology to realize the design of Apple's hierarchical transmission control system based on Controller Area Network technology. Fuzzy PID and traditional PID algorithms are used to simulate and realize the operation of brushless DC motor, and compare the advantages of brushless DC motor control based on fuzzy PID to ensure the safe and stable operation of the system. Using the machine learning algorithms model for color detection, the Support Vector Machine algorithm model finally achieved the classification of the three types of apple samples with a recognition rate of 96.7%.

Keywords: brushless DC motor; fuzzy control; machine learning; grading.

Practical Application: Control apple's automated sorting using machine learning.

1 Introduction

Appearance is a very important sensory quality attribute to distinguish fruit quality. It is the external manifestation of the inside of the fruit (sugar content, maturity), which affects the market value of the fruit (Yamamoto et al., 2015). Therefore, ensuring the stable delivery and classification accuracy of apples and other agricultural products in the agricultural production process can increase their value in the market to a certain extent. Recently, many scholars have conducted research on apples (Deng et al., 2022; Lei et al., 2022; León et al., 2022), but few studies have been conducted on the automatic sorting system of apples. The hierarchical transmission control system has gradually combined a variety of advanced technologies in the process of automation development. Such as sensors data collection (Yang et al., 2020), communication network (Arakeri, 2016), automatic control (Siddiq et al., 2019) and intelligent judgment (Blasco et al., 2003; Jahanbakhshi et al., 2021), etc., carry out extensive management of agricultural processes. Continuously improve the complexity, stability and reliability of the system.

On the basis of existing research, the design research is mainly aimed at apple's hierarchical transmission control system. The controller area network (Halder et al., 2021) structure is adopted to realize the application requirements of easy maintenance and flexible expansion of electrical equipment in the process of agricultural automation. And use LabVIEW (Schettino de Souza et al., 2013) to design a control interface to control the work of the motor. The brushless DC motor with long service life, high reliability and good control performance is selected as the power source of the conveyor belt, and the algorithm is simulated and compared and analyzed. The intelligent control algorithm that can make the conveyor belt motor run stably and reliably is selected, and the simulation curve of the conveyor belt operation is obtained. Finally, the machine learning algorithm is used to verify the accuracy of Apple's color grading under the requirements of ensuring that the brushless DC motor has a small operating error, high precision and strong continuity. Machine learning has a wide range of applications in food science (Chen & Yu, 2022; Farah et al., 2021; Hou et al., 2022; Rocha et al., 2020; Wang et al., 2021). It has practical significance for improving the intelligent level of fruit classification.

2 Material and methods

2.1 Mechanical system

The automatic sorting system consists of a control device, a conveying device, multiple sorting devices and multiple sorting crossings. The system can realize the connection and control of multiple Apple grading conveyor belts by one host through the Controller Area Network technology. Each conveyor belt consists of a brushless DC motor (Taibang Group, 5GN-12.5K) and its driver (Yineng Electromechanical, BLD-300B), multiple color recognition sensors (AMS company, TCS3472XFN), multiple load cell (Haixin Technology, HX711), and a stm32 controller (STMicroelectronics, stm32F103VET6). The control device is composed of a host and a hierarchical transmission controller. The virtual instrument technology is used to build a master control panel on the upper computer. The upper computer and the hierarchical transmission controller use controller area network technology to achieve communication. The transmission device is composed of a conveyor belt, a motor, and a motor driver. The controller controls the motor to smoothly transfer the apples. The sorting device is composed of trays, weight sensors, and 14 color recognition sensors. As shown in Figure 1, a transparent tray is fixed on the conveyor belt, and there is a weight sensor and 14 color recognition sensors on the tray. The sensors send the collected information to the upper computer, and the

Received 02 Feb., 2022
Accepted 05 Apr., 2022
1 College of Mechanical and Electronic Engineering, Sichuan Agricultural University, Ya’an, China
*Corresponding author: xulijia@sicau.edu.cn
and retains the characteristics of large DC motor torque, high speed regulation performance, safety and easy maintenance, replace the original mechanical commutation method, which has relatively small, and the anti-interference performance of the brushless DC motor is stronger. However, the armature current adjustment speed of the traditional PID control system and the field-oriented control, and it is easily affected by the environment, and the response curve effect is not ideal. However, the armature current adjustment speed of fuzzy PID control is faster, the response curve fluctuation is relatively small, and the anti-interference performance of the motor system is stronger.

During the operation of the brushless DC motor, the fuzzy PID control algorithm is more stable, and it can be adjusted quickly under interference conditions, and has better dynamic performance of the transmission system. During the simulation operation, the anti-interference ability of the motor also has an important influence on the performance of the transmission system. During the simulation operation, the anti-interference ability of the motor also has an important influence on the performance of the transmission system. When the load is added after the motor is running stably, the load is suddenly added at t=1.5s. It can be seen from Figure 2 that after the load is added, the PID oscillates for 0.06s and then returns to normal speed. The field oriented control oscillates and maintains the current speed. Fuzzy PID hardly needs adjustment.

The torque simulation result is shown in Figure 3. When the load is added, the torque will fluctuate to a certain extent in the three control methods, but the fuzzy PID control reaches a stable value in the shortest time, and the motor runs smoothly.

The response curve of the stator current is shown in Figure 4. From the comparative analysis, it can be seen that the current amplitude fluctuates greatly under the action of the traditional PID control system and the field-oriented control, and it is easily affected by the environment, and the response curve effect is not ideal. However, the armature current adjustment speed of fuzzy PID control is faster, the response curve fluctuation is relatively small, and the anti-interference performance of the motor system is stronger.

During the operation of the brushless DC motor, the fuzzy PID control algorithm is more stable, and it can be adjusted quickly under interference conditions, and has better dynamic performance.

upper computer classifies the apples according to the data. When the apples are sent to the designated sorting crossing, the controller controls the tray to tilt to make the apples roll to the corresponding classification area to complete the apple grading.

### 2.2 Sample and data collection

This study used 400 defect-free and pollution-free Red Fuji apples, purchased in December 2020 and harvested from Yantai (Shandong Province, China). The apple samples are individually packaged in a fruit net made of polyethylene foamed cotton. The weight of the apple samples is in the range of 160 g to 320 g, and the fruit diameter is in the range of 60 mm to 100 mm. 6 and 8 color recognition sensors are evenly installed around the edge of the tray at a height of 25 mm and 45 mm respectively to detect the fruit color near the equator of the apple. When the apple is on the tray, the weight sensor detects the weight of the apple and starts to collect the color of the apple near the equator. And each tray only holds one apple to avoid mutual interference between apples. Each color sensor detects the color three times and obtains the values of R (red), G (green), and B (blue). Take the average of the three detected RGB values as the final RGB value at this point. The RGB values of the 14 color sensors are averaged to obtain the final RGB value of the apple in the dish. And after transformation, the easy-to-observe H (Hue), S (Saturation), and L (Lightness) values are obtained, and then classified.

### 3 Simulation and realization of conveyor control algorithm

The brushless DC motor is a typical mechatronic component. And it has excellent speed and torque characteristics, good static and dynamic characteristics (Vanchinathan & Selvaganesan, 2021). The special electronic commutation method is used to replace the original mechanical commutation method, which has good speed regulation performance, safety and easy maintenance, and retains the characteristics of large DC motor torque, high speed regulation accuracy and no loss of step.

#### 3.1 Research on control algorithm of Brushless DC motor

According to the process analysis of fuzzy PID and the actual demand of maintaining stable operation of the conveyor belt, the two quantities of speed deviation value $n_e$ and speed deviation rate $\frac{dn_e}{dt}$ are used as inputs for fuzzy processing (Varshney et al., 2017). Output the correction values of kp, ki, kd parameters to complete the speed regulation of the motor, and finally output the set current to the current controller of the inner loop to ensure the stable operation of the motor. Map the actual value to the fuzzy universe, and determine the fuzzy subset as {NB, NM, NS, ZO, PS, PM, PB}, and the universe as [-6, +6]. The membership function of input speed deviation and deviation rate selects Gaussian curve, and the membership function of kp, ki, kd selects triangular curve.

Determine the initial values of Kp, Ki, and Kd as Kp=70, Ki=0.5, Kd=0.08, set the speed to 2000r/min, simulate the brushless DC motor to start under no load, set the simulation time to $t=2.5$ s, the simulation result is shown in Figure 2. In the simulation diagram, the horizontal axis represents time, and the vertical axis represents speed. The blue represents the simulation curve of field-oriented control, the black represents the simulation curve of PID control, and the orange represents the simulation curve of fuzzy PID control.

Fuzzy PID control stably rises after $t=0.007$ s and arrives first and stabilizes near the set value, magnetic orientation control reaches and stabilizes near the set value after $t=0.021$ s, adjustment time $t=0.009$ s, overshoot $\sigma% = 0.19%$. The PID control reaches a stable value after $t=0.059$ s, the adjustment time $t=0.032$ s, and the overshoot $\sigma% = 1.1%$. It can be seen that fuzzy PID can make the motor reach the set speed faster. During the simulation operation, the anti-interference ability of the motor also has an important influence on the performance of the transmission system. During the simulation operation, the anti-interference ability of the motor also has an important influence on the performance of the transmission system. When the load is added after the motor is running stably, the load is suddenly added at $t=1.5$ s. It can be seen from Figure 2 that after the load is added, the PID oscillates for 0.06s and then returns to normal speed. The field oriented control oscillates and maintains the current speed. Fuzzy PID hardly needs adjustment.

The torque simulation result is shown in Figure 3. When the load is added, the torque will fluctuate to a certain extent in the three control methods, but the fuzzy PID control reaches a stable value in the shortest time, and the motor runs smoothly.

The response curve of the stator current is shown in Figure 4. From the comparative analysis, it can be seen that the current amplitude fluctuates greatly under the action of the traditional PID control system and the field-oriented control, and it is easily affected by the environment, and the response curve effect is not ideal. However, the armature current adjustment speed of fuzzy PID control is faster, the response curve fluctuation is relatively small, and the anti-interference performance of the motor system is stronger.
performance. Compared with the traditional PID control, the application of fuzzy PID control in the transmission control system can make the motor run stably.

### 3.2 Motor control

According to apple’s graded transmission control requirements, combined with the characteristics of the conveyor belt drive motor. This article selects the 5GN-12.5K brushless DC motor produced by Taibang Group as the drive motor of the hierarchical control system. The maximum linear speed of the control conveyor belt is 1.5 m/s, and the roller diameter of the conveyor belt is 178 mm. According to the conversion relationship between the rotational speed and the linear speed, the maximum motor speed can be obtained as 160 r/min. Through the configuration of the control program, the duty cycle of the PWM is changed.
Apple automatic sorting system

In the case of no-load operation, the sampling period is set to 1 min, and sampling is performed every 1 s.

As shown in Figure 5, first adjust the BLD-300B built-in potentiometer RV knob of the brushless DC motor driver to set it to other speed modes. Connect the Hall signal terminals HW, HV, and HU with the motor terminal. Connect the BRK end of the motor and the common port COM end with the ground wire of the single-chip microcomputer to realize normal operation. The speed control signal SV is connected with the microcontroller to input the PWM signal. During the test, sampling is performed through the drive speed signal output port SPEED, and the pulse frequency corresponding to the speed of the brushless DC motor at this time is collected.

The speed curves of the brushless DC motor obtained by changing the duty cycle are shown in Figure 6. Compared with the traditional PID, the brushless DC motor controlled by fuzzy PID has greatly reduced speed fluctuation, which makes the operation more stable. Applying it to apple's hierarchical transmission control system can better ensure the stability of the system.

**Figure 4.** Stator current response curve. Fuzzy PID control motor system has stronger anti-interference performance.

**Figure 5.** Speed sampling of brushless DC motor. Collect the pulse frequency corresponding to the speed of the brushless DC motor.
4 Apple color grading

The RGB values collected by the color sensor are converted into HSL values, and 400 data are processed by machine learning algorithms. Use support vector machine (SVM), k-nearest neighbor (KNN), extreme gradient boosting (XGBoost), CatBoost four machine learning algorithms. Randomly select 70% of the Red Fuji apple data as the training set, and the remaining 30% of the Red Fuji apple data from the three types of samples as the test set. Construct a discriminant model to verify the color attributes of red, light red, and light yellow apples. Accuracy, F-Score, Logloss and Hamming Loss are used to evaluate the effect of model training and prediction.

4.1 Support vector machine algorithm

Support vector machine is mainly used to solve the data classification problem in the field of pattern recognition (Jiang et al., 2021). It is a kind of supervised learning algorithm and a novel small-sample learning method with a solid theoretical foundation. Support vector is the training result of SVM, and it is the support vector that plays a decisive role in SVM classification decision. A small number of support vectors determine the final result, which not only helps us to grasp the key samples, but also destined that the method is not only simple in algorithm, but also has better robustness (Arias Velasquez, 2021).

4.2 K-nearest neighbor algorithm

K-nearest neighbor is a well-known statistical method for pattern recognition, and it occupies a considerable position in machine learning classification algorithms (Jhantani et al., 2021). It belongs to the category of non-parametric learning algorithms (Ho & Yu, 2021) and is a theoretically mature method (Dong et al., 2021). It can be used for classification and regression, and is a supervised learning algorithm. KNN algorithm has been widely used in pattern recognition (Ni et al., 2009), data mining (Adeniyi et al., 2016), intrusion detection and other fields.

4.3 Extreme gradient boosting algorithm

Extreme gradient boosting was proposed by Chen Tianqi et al. in 2016 (Chen & Guestrin, 2016). It is a scalable machine learning algorithm for tree boosting (Friedman et al., 2000; Friedman, 2001; Kotsiantis, 2013; Ruiyi et al., 2021), which is efficient, flexible and portable. In the training process, by changing the weight of the training sample, learning multiple classifiers, and finally obtain the optimal classifier. After each round of training is over, reduce the weight of the correctly classified training sample and increase the weight of the incorrectly classified sample. After multiple training sessions, some of the training samples that have been misclassified will get more attention, and the weight of the correct training samples will approach zero. Multiple simple classifiers are obtained, and a final model is obtained by combining these classifiers.

4.4 CatBoost algorithm

CatBoost is a GBDT framework with fewer parameters, support for categorical variables and high accuracy based on the decision tree as the base learner (Albaqami et al., 2021). CatBoost is composed of Categorical and Boosting, which can handle categorical features efficiently and reasonably. In addition, CatBoost also solves the problems of gradient bias and prediction shift. So as to reduce the occurrence of over-fitting, and then improve the accuracy and generalization ability of the algorithm (Ben Jabeur et al., 2021).

5 Results and discussion

The prediction results of different algorithm models are shown in Table 1. The SVM model has the highest recognition rate of 97.6%. The recognition rate of the XGBoost model is second, with a prediction accuracy of 95.8%. The prediction accuracy of the CATBoost model ranks third at 93.3%. The KNN model has the worst prediction effect, at 86.7%. And the average time consumed for building each model is sorted from high to low into CatBoost, XGBoost, SVM, and KNN. It can be seen that the recognition rate of the SVM algorithm model is the highest, reaching 96.7%. At the same time, the Log Loss and Hamming Loss of this model are the...
Food Sci. Technol, Campinas, 42, e24922, 2022

Apple automatic sorting system

The confusion matrix of the classification results of the four models, the observation results of the three types of apple samples of red, light red, and light yellow are shown in Figure 7. The four algorithm models all predict red apples as light red apples to varying degrees; or predict light red apples as red apples; or predict light yellow apples as light red apples. And XGBoost and CatBoost algorithm models also predict light red apples to light yellow apples. The SVM model distinguishes three types of apple samples with a prediction rate of 96.7%, which shows that the established prediction model has a good recognition effect and can be used to effectively monitor the quality of apples.

6 Conclusion

This research has developed an automatic inspection and classification system for apple quality based on fuzzy PID motor control and machine learning technology. In order to reduce the cumbersome processing steps after apple harvesting, a processing mode of detection and classification is proposed. Fuzzy PID and traditional PID algorithm are used to simulate and realize the operation of brushless DC motor. By comparison, the advantages of brushless DC motor control based on fuzzy PID are obtained to ensure the safe and stable operation of the system. Four algorithm models were used to classify and predict the HSL data of 400 apple samples. In the end, the SVM algorithm model achieved the classification of the three types of apple samples with a recognition rate of 96.7%. Using Apple’s multi-point HSL data, under the condition of ensuring accurate classification, compared with image processing, the amount of

Table 1. The prediction results of different algorithm models.

<table>
<thead>
<tr>
<th>Models</th>
<th>ACC</th>
<th>Fscore Red</th>
<th>Fscore Light red</th>
<th>Fscore Light yellow</th>
<th>Log Loss</th>
<th>Hamming Loss</th>
<th>Fit time</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>0.967</td>
<td>0.919</td>
<td>0.978</td>
<td>0.957</td>
<td>712.654</td>
<td>0.033</td>
<td>0.052</td>
</tr>
<tr>
<td>KNN</td>
<td>0.867</td>
<td>0.778</td>
<td>0.915</td>
<td>0.500</td>
<td>1127.119</td>
<td>0.133</td>
<td>0.002</td>
</tr>
<tr>
<td>XGBoot</td>
<td>0.958</td>
<td>0.919</td>
<td>0.972</td>
<td>0.917</td>
<td>747.193</td>
<td>0.042</td>
<td>0.173</td>
</tr>
<tr>
<td>CatBoost</td>
<td>0.933</td>
<td>0.889</td>
<td>0.956</td>
<td>0.818</td>
<td>850.809</td>
<td>0.067</td>
<td>0.277</td>
</tr>
</tbody>
</table>

Four models, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Extreme Gradient Boosting (XGBoost) and CatBoost, are used, respectively. ACC: The accuracy of classification.

Figure 7. Confusion matrix of prediction results of four models. (a) Support vector machine (SVM). (b) K-nearest neighbor (KNN). (c) Extreme gradient boosting (XGBoost). (d) CatBoost. Support vector machine has the highest recognition rate.
data calculation is reduced. In addition, the SVM algorithm model took a short time to build, and the performance requirements of the host computer were lower. More Apple classification data can be processed in an effective time to meet the requirements of using controller area network technology to expand multiple sorting conveyor belts. In conclusion, machine learning is a very promising tool for rapid apple detection and should be valued and encouraged in food sorting systems.

References


