# IMPROVED YOLOV8-ALGORITHM FOR SORTING FRESH WHITE TEA: COMBINING FEATURE ENHANCEMENT AND ATTENTION MECHANISM
/

# 改进的 YOLOv8 新鲜白茶分选算法：将特征增强与注意力相结合

Xuedong YU[1,2], Yadong NIU[1], Zhongyou ZHOU[1,2], Bo LU[1,2], Kaiqiang JIN[1,2], Rongyang WANG*[1,2]
[1] College of Intelligent Manufacturing and Elevator Technology, Huzhou Vocational and Technical College, Huzhou, China
[2] The Huzhou Key Laboratory of Robot System Integration and Intelligent Equipment, Huzhou, China
*Tel: +8613957255852; E-mail: 2022024@huvtc.edu.cn*
*DOI: https://doi.org/10.35633/inmateh-75-80*

***Keywords:*** *tea sorting; Dynamic Snake Convolution module; Multi-Head Self-Attention mechanism; YOLOv8*

## ABSTRACT
*In this paper, an improved intelligent sorting algorithm for YOLOv8 white tea fresh leaves is proposed to solve the problems of unclear tea grades and uneven product levels caused by mechanical picking. The algorithm introduces the Dynamic Snake Convolution module (DSConv) for feature enhancement and adds an attention mechanism module, the Multi-Head Self-Attention mechanism (MHSA). Experiments show that the YOLOv8-DsConv-MHSA algorithm has an average accuracy of 96.4% and an average detection rate of 126.6 FPS per second, which is the best algorithm for white tea fresh leaf sorting in the comprehensive comparison. After deploying the proposed YOLOv8-DsConv-MHSA algorithm onto the developed tea sorting machine and conducting experimental comparisons with existing tea sorting machines, it is evident that the screening rate has been enhanced by 10.7%, and the operational efficiency has increased by 20%.*

## 摘要
*本文提出了一种改进的 YOLOv8 白茶鲜叶智能分拣算法，以解决机械采摘造成的茶叶等级划分不清和产品层次不齐的问题。该算法引入了动态蛇形卷积模块（DSConv）进行特征增强，并且添加了注意力机制模块--多头自我注意力（MHSA）。实验表明：YOLOv8-DsConv-MHSA 算法的平均准确率为 96.4%，平均检测率为每秒 126.6 FPS，是综合比较中白茶鲜叶分选效果最好的算法。将所提出的 YOLOv8-DsConv-MHSA 算法部署到研发的茶叶分选机上，与现有茶叶分选机进行实验对比可知，筛分率提升了 10.7%，作业效率提升了 20%。*

## INTRODUCTION
Tea is an important economic crop in China, and the area under tea cultivation is one of the largest in the world, with a total production value of more than 600 billion yuan ($86.6 billion). Scientific studies have shown that tea has great benefits for human health, and the demand for high-quality tea is gradually increasing *(Ruxton et al., 2013)*. Due to the huge production of tea, the picking method is gradually transitioning from labor to mechanization *(Zheng et al., 2011)*. The cost of tea picking through machinery has been greatly reduced, and the efficiency has been improved, but through the mechanical way of tea picking, the tea leaves picked will be mixed with leaf stems and debris, which need to be further sorted out to realize the tea grading of the high-quality fresh tea leaves *(Yuan et al., 2016)*.

Research on grading of machine-picked tea can improve the quality of machine-picked tea and is an effective way to solve the problem of high-quality tea picking. At present, most of the common tea grading devices are mechanized grading, mainly roller *(Wang et al., 2016)*, vibrating screen *(Lv et al., 2022)*, air flow sorting - mesh belt sieving *(Wang et al., 2019)*. These methods separate the normal tea leaves, broken buds and leaves, leaf stalks and other debris of fresh tea, but the grading effect is still a certain distance from the high-quality fresh tea. Achieving the standard of high-quality tea and improving the quality and economic value of tea is an urgent problem for the tea industry.

In recent years, with the continuous emergence of machine vision and artificial intelligence algorithms, it provides many new possibilities in the field of agricultural production and processing, and of course there are some scholars who have achieved research results in the intelligent sorting of fresh tea leaves.

*Chen et al. (2010)* from Nanjing Institute of Mechanical and Agricultural Research (NIMAR) applied BP neural network to classify fresh leaves of tea and achieved good results by extracting a variety of geometric features and texture parameters.

*Liu et al. (2016)* used artificial neural network to categorize and analyze the geometrical features and image characteristics of tea leaves, and identified them with various neural networks such as radial basis, BP neural network, Hopfield neural network, and so on. *Gao et al. (2017)* built a 7-layer convolutional neural network recognition model to realize automatic identification and sorting of fresh tea leaves by sharing weights and adjusting the learning rate. *Song et al.(2018)* took Keemun black tea as the research object, constructed an image acquisition system, extracted six absolute shape features and two relative shape features to construct a feature histogram, and identified it with various classification models such as BP neural network, Extreme Learning Machine (ELM), Support Vector Machine (SVM), Least Squares Support Vector Machine (LS-SVM), which provided experimental data and reference methods for the realization of the digital grade appraisal of tea leaves.

*Zhu et al. (2019)* designed a convolutional neural network (CNN) model with three convolutional layers, two pooling layers and one fully connected layer, the model recognized Huoshanhuangya tea correctly by 95.3% through the training and testing of real-time captured Huoshanhuangya tea images. *Yang et al. (2019)* proposed an improved (You Only Look Once) YOLO-v3 deep convolutional neural network algorithm for recognizing the picking point of young tea buds, which achieves end-to-end target detection and recognition of different poses of high-quality buds, taking into account both efficiency and accuracy. *Chen et al. (2020)* extracted the features of tea by determining the topology of the tea and used SVM to sort and recognize the tea with 94% recognition accuracy. *Gao et al. (2021)* constructed a tea sample dataset and completed the image recognition of tea by building a convolutional neural network, and the correct rate of the trained image recognition model was 96%.

*Chen et al. (2021)* designed a lightweight convolutional neural network model (MobileNetV2-Tea), and by improving the MobileNetV2 network, the MobileNetV2-Tea model obtained has 99% accuracy in fresh tea image recognition, with a model size of only 28.86M and an average recognition time of 45ms.

*Gan et al. (2022)* investigated feature classification methods using improved genetic algorithms to screen features with the best combination of three different classifiers with optimal dimensionality, and finally Support Vector Machines (SVMs) achieved 97% recognition accuracy on a 28-dimensional feature set.

*Pi et al. (2023)* applied YOLO-v5 to the task of intelligent detection of tea buds, and added the structure of bidirectional feature pyramid network (BiFPN) in Neck to optimize the structure of the network, which provided an effective detection model for the sorting of high-quality tea leaves.

*Cao et al. (2023)* proposed a novel fresh tea sorting system, which uses a tea recognition model based on YOLOv5 deep learning model for fast, high-precision multi-channel sorting of four grades of tea, with a model recognition accuracy of 88.8%. At the same time, the YOLO deep learning algorithm has demonstrated remarkable effectiveness in the detection of numerous crops, covering various types of crops such as cucurbits' fruits *(Zhao et al., 2022)*, young apple fruits *(Du et al., 2024)*, potato *(Pan et al., 2024)*, grapes *(Tao et al., 2024)*, wheat *(Bi et al., 2024)*.

Although the above research has made some achievements in machine-picked fresh tea leaf sorting, there is still space for improvement in the precision and speed of sorting, especially considering the large-volume and larger-scale fresh tea leaf sorting scenarios. Therefore, this paper takes Anji white tea as the research object, and establishes a database of fresh tea samples containing one bud and one leaf, one bud and two leaves, one bud and three leaves, one bud and four leaves, broken leaves, and single bud in six forms. The state-of-the-art YOLOv8 deep learning algorithm is applied to the fresh tea leaf classification and recognition task to perform deep feature extraction and fresh tea leaf grade classification. On this basis, the Dynamic Snake Convolutional Kernel (DSConv) *(Qi et al.,2023)* is used to carry out the improved replacement of the convolutional kernel in the C2f module, and Multi-Head Self-Attention mechanism (MHSA) *(Vaswani et al., 2017)* is added to enhance the feature representation capability and extract more representative semantic information.

## MATERIALS AND METHODS

### Dataset collection

The fresh tea leaves dataset collected in this study was obtained from Anji County, Huzhou City, Zhejiang Province. The collected Anji white tea was brought back to the laboratory, video shooting was performed by Canon 6D II camera, and then the pictures of fresh tea leaves were serialized and extracted by frames, obtaining 2537 usable pictures of fresh tea leaves of different grades, which was used to obtain to the tea leaf sorting dataset. The details of the dataset are shown in Fig. 1.

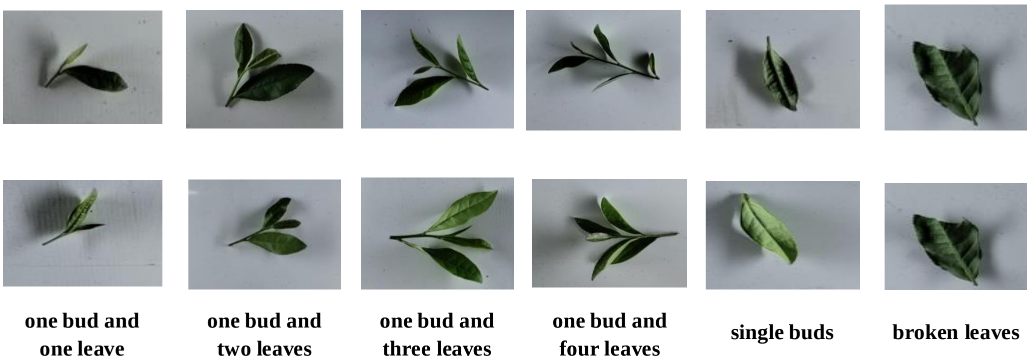


**Fig. 1 - Specific types of fresh white tea leaves**

### Dataset expansion

The amount of data in the realistic fresh white tea leaf sorting scenario is huge. In order to adapt to the realistic demand as well as to meet the training requirements of the YOLOv8 deep learning algorithm, the fresh white tea leaves dataset is expanded to 18 times of the original dataset by rotating the angle, mirroring, sizing, changing the brightness, changing the chromaticity, changing the contrast, changing the sharpness, and so on. Fig. 2 shows a detailed illustration of the expansion of the fresh tea leaf dataset. The expanded dataset was obtained as 45,666 sheets, including 17,388 sheets of one bud and one leaf, 10,674 sheets of one bud and two leaves, 5,742 sheets of one bud and three leaves, 1,836 sheets of one bud and four leaves, 7,236 sheets of single buds, and 2,790 sheets of broken leaves. The fresh white tea leaf dataset of each category was divided into train and test sets in the ratio of 8:2, and the detailed dataset information is shown in Table 1.

**Table 1**

**Number of specific images in the fresh white tea leaf dataset**

| category | original | expansion | train | test |
|---|---|---|---|---|
| one bud and one leaf | 966 | 17388 | 13910 | 3478 |
| one bud and two leaves | 593 | 10674 | 8539 | 2135 |
| one bud and three leaves | 319 | 5742 | 4594 | 1148 |
| one bud and four leaves | 102 | 1836 | 1469 | 367 |
| single buds | 402 | 7236 | 5789 | 1447 |
| broken leaves | 155 | 2790 | 2232 | 558 |

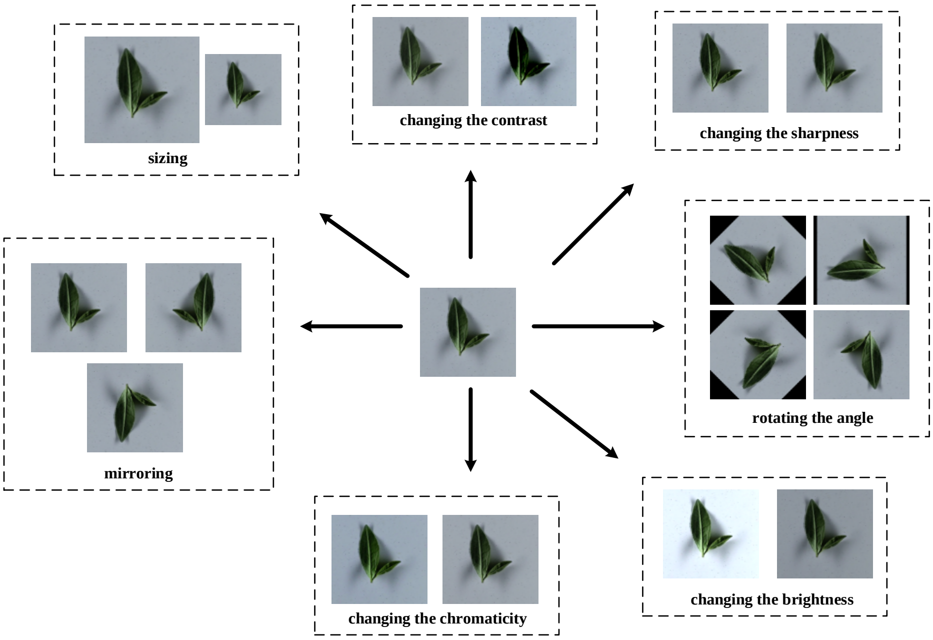


**Fig. 2 - Example diagram of fresh white tea leaf dataset expansion**

### The YOLOv8 deep learning algorithm

The YOLOv8 model is a deep neural network-based target detection algorithm that supports tasks such as image classification, object detection and instance segmentation. Combining the YOLOv8 algorithm with fresh tea leaf sorting can quickly realize end-to-end intelligent grading of fresh tea leaves, which provides a good idea for fresh tea leaf sorting. The specific algorithm implementation process is as follows.

Firstly, from the captured video taken, fresh white tea leaf images are extracted by frame, and the fresh white tea leaf dataset is constructed according to multiple data expansion methods. Then the YOLOv8 deep learning network model is constructed. The recognition ability of the pre-trained YOLOv8 network model is mainly focused on the categories in the ImageNet dataset (the ImageNet dataset covers more than 1,000 image categories that are common in life), so the parameters of the YOLOv8 deep learning network model are dynamically adjusted according to the fresh tea leaves dataset. The model internally undergoes deep feature extraction and sorting, and is finally able to complete sorting of fresh tea leaves.

As shown in Fig.3, the YOLOv8 model applied to the intelligent sorting of fresh tea leaves mainly consists of a backbone network (Backbone) and a predictive output network (Head). The backbone network mainly uses the convolutional kernel to extract multi-scale features from the fresh white tea leaf contour, and operates the fresh tea leaf image layer by layer with multiple Conv, C2f and SPPT modules, so that the proposed features can be more representative. The predicted output network computes probability scores based on the extracted deep feature representations for intelligent sorting of different categories of fresh white tea leaves.

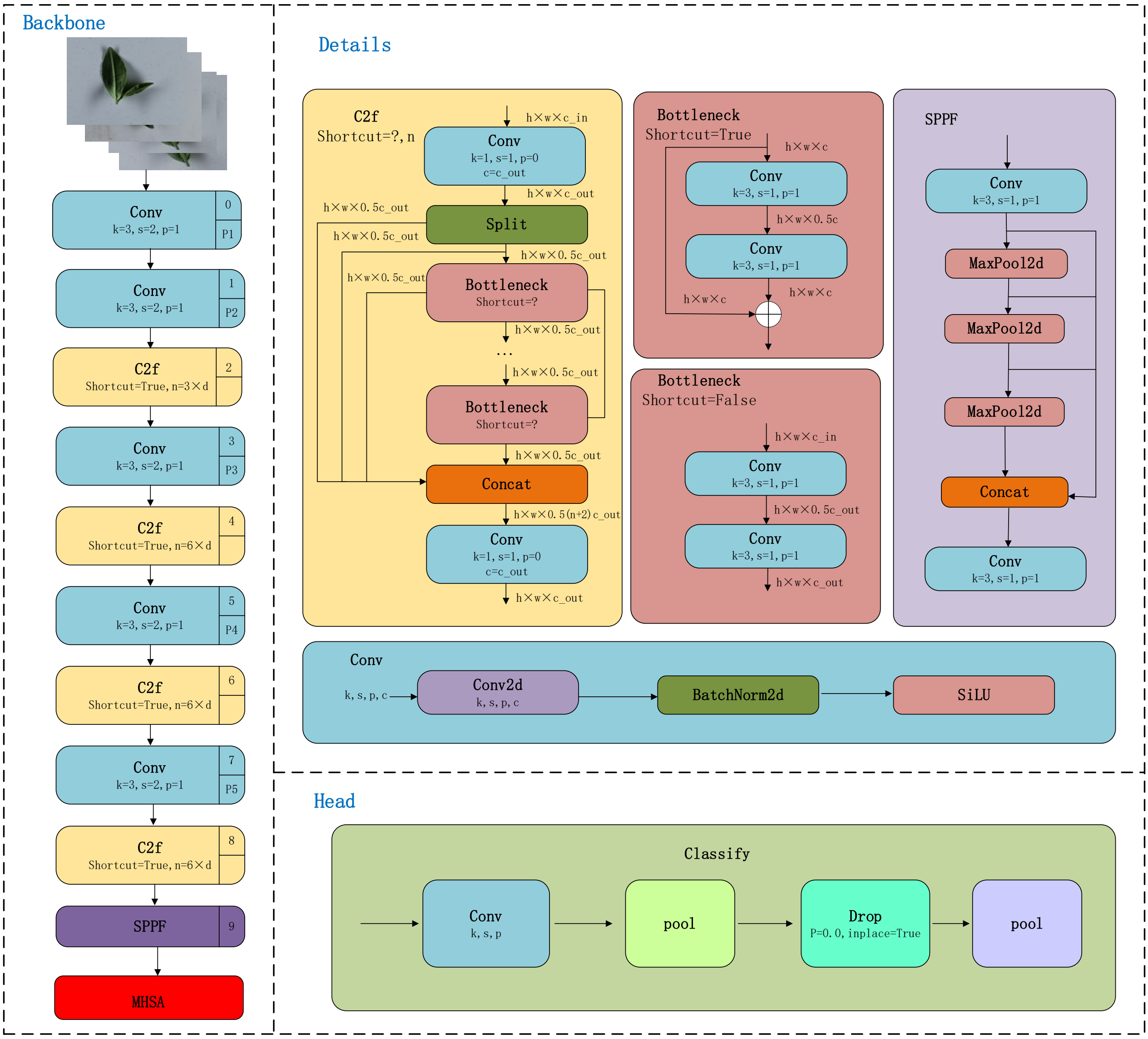


**Fig. 3 - Flowchart of YOLOv8 sorting of fresh white tea leaves**

### Dynamic Snake Convolution Module (DSConv)

The convolution Conv module in the YOLOv8 model performs convolutional computation for the entire region of the input fresh white tea leaf image, in order to be able to focus the convolution kernel on the target features of interest and reduce feature interference in the surrounding region, it is proposed that the dynamic snake convolution kernel (DSConv) be introduced to the fresh white tea leaf feature extraction. The feature extraction process of the Dynamic Serpentine Convolution Kernel (DSConv) is shown in Fig. 4. The next convolution position is calculated by extrapolating from the previous convolution position, starting from the center of the whole image $K_i=(x_i,y_i)$. By iterating continuously, it ensures that the extracted features are concentrated in the region where the fresh tea leaves are located, reduces the bias of the convolutional computation, and improves the accurate representation of the extracted features.

$$K_{i\pm c}=\begin{cases}\left(x_{i+c},y_{i+c}\right)=\left(x_i+c,y_i+\sum_{i}^{i+c}\Delta y\right)\\ \left(x_{i-c},y_{i-c}\right)=\left(x_i-c,y_i+\sum_{i-c}^{i}\Delta y\right)\end{cases} \tag{1}$$

The computational operation of convolution kernel actually starts from the center point and makes convolution computation in x-axis and y-axis directions sequentially, in x-axis direction, as shown in Equation (1), $K_{i\pm c}$ denotes any convolution computational position in the grid, $c=\{0,1,2,3,4\}$ denotes the distance from the center point, and the bias $\Delta=\{\delta|\delta\in[-1,1]\}$ is the offset of the next convolution position $K_{i+1}$ relative to the previous one $K_i$. Since the convolution operation is a non-stop iterative process, $\sum$ needs to be added outside the bias to perform a linear accumulation.

$$K_{j\pm c}=\begin{cases}\left(x_{j+c},y_{j+c}\right)=\left(x_j+\sum_{j}^{j+c}\Delta y,y_j+c\right)\\ \left(x_{j-c},y_{j-c}\right)=\left(x_j+\sum_{j-c}^{j}\Delta y,y_j-c\right)\end{cases} \tag{2}$$

Equation (2) represents the y-axis direction to do the convolutional computation operation.

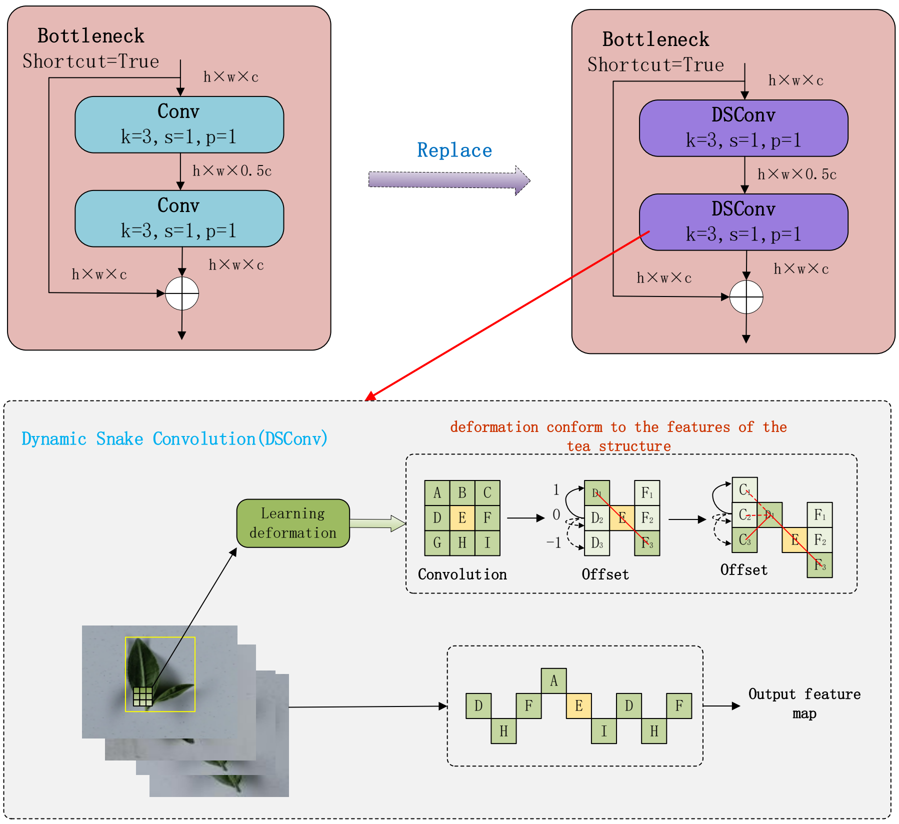


**Fig. 4 - Flowchart of dynamic snake convolution kernel feature extraction**

### The Multi-Head Self-Attention mechanism (MHSA)

In the intelligent sorting of fresh tea leaves, the completeness of the extracted feature information is directly related to the subsequent performance of fresh tea leaves sorting and identification. To further enhance the feature representation capability, the multi-head self-attention mechanism (MHSA) is added to the YOLOv8 network model as a way to ensure the completeness and accuracy of the proposed features. As shown in Fig. 5, in the self-attention layer, all the keys $K$, values $V$, and query operations $Q$ come from the feature parameters of the last fresh tea leaves, and the mapping calculation can make each position in the self-attention layer pay attention to all the positions in the previous layer, which increases the completeness of feature extraction. Each self-attention mechanism focuses on different subspace information of different positions, and the results of multiple attention mechanism calculations are fused and then linearly mapped to obtain the final feature representation.

$$A(Q,K,V)=softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{3}$$

where $A$ denotes the Attention computation operation, $d_k$ denotes the vector dimension computed by the input key and the query. The keys $K$, values $V$, and query operations $Q$ computed from the previous layer are matrices.

$$MH(Q,K,V)=Concat(h_1,...,h_i)w^o \tag{4}$$

$$h_i=A(W_i^Q,KW_i^K,VW_i^V) \tag{5}$$

where $W_i^Q \in R^{d_{model} \times d_k}$, $W_i^K \in R^{d_{model} \times d_k}$, $W_i^V \in R^{d_{model} \times d_v}$, $W_i^O \in R^{hd_v \times d_{model}}$ are composed of parameter matrices of the corresponding dimensions. $d_{model}$ denotes the unity parameter dimensions of the model, $h$ denotes the number of attention heads, which is set to $h=4$ in this model. Multiple self-attention mechanisms are spliced and fused, where *MH* denotes multi-head self-attention mechanism and $h_i$ denotes one of the $head_i$ attention mechanisms.

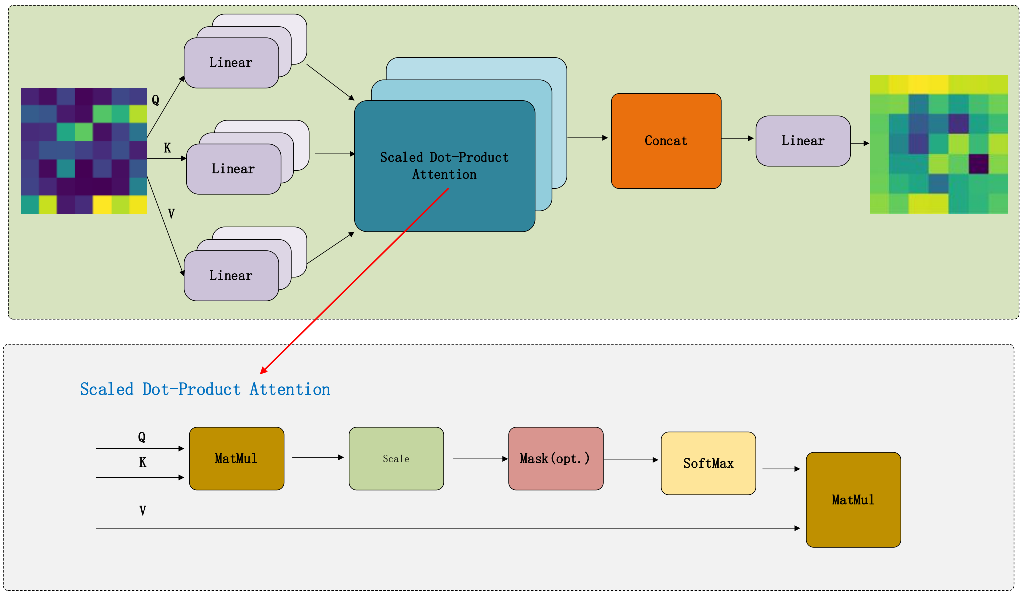


**Fig. 5 - Flowchart of feature extraction for Multi-head Self-Attention Mechanism**

### Improving the network structure parameters of YOLOv8

The Dynamic Snake Convolution Module (DSConv) and Multi-head Self-Attention Mechanism (MHSA) are added into the YOLOv8 deep learning algorithm model to obtain the improved YOLOv8 model. In order to more clearly visualize the specific parameters of the model, the parameters of the improved YOLOv8 network structure are listed in Table 2.

**Table 2**

**Network structure parameters of the improved YOLOv8**

| Layers | From | n | Params | Module | Arguments |
|---|---|---|---|---|---|
| 0 | -1 | 1 | 464 | Conv | [3, 16, 3, 2] |
| 1 | -1 | 1 | 4672 | Conv | [16, 32, 3, 2] |

| Layers | From | n | Params | Module | Arguments |
|---|---|---|---|---|---|
| 2 | -1 | 1 | 18888 | C2f_DySnakeConv | [32, 32, True] |
| 3 | -1 | 1 | 18560 | Conv | [32, 64, 3, 2] |
| 4 | -1 | 2 | 134800 | C2f_DySnakeConv | [64, 64, True] |
| 5 | -1 | 1 | 73984 | Conv | [64, 128, 3, 2] |
| 6 | -1 | 2 | 507024 | C2f_DySnakeConv | [128, 128, True] |
| 7 | -1 | 1 | 295424 | Conv | [128, 256, 3, 2] |
| 8 | -1 | 1 | 982088 | C2f_DySnakeConv | [256, 256, True] |
| 9 | -1 | 1 | 164608 | SPPF | [256, 256, 5] |
| 10 | -1 | 1 | 197376 | MHSA | [256, 14, 14, 4] |
| 11 | -1 | 1 | 337926 | Classify | [256, 6] |
| summary: 370 layers, 2735814 parameters, 2735814 gradients, 5.8 GFLOPs | | | | | |

### Experimental equipment and parameterization

The specific parameters of the computer used for the experimental training in this study are: the operating system is Windows 11, the CPU is i9, the GPU is NVIDIA RTX A4000, the deep learning modeling framework used is PyTorch 1.12.0, and the programming version is Python 3.9, and the specific parameter settings at the time of model training are shown in Table 3.

**Table 3**

**Parameter settings for model training**

| Parameters | Value |
|---|---|
| Image-size | 224 |
| Epochs | 60 |
| Batch | 30 |
| Momentum | 0.937 |
| Workers | 8 |
| Optimizer | Auto |
| Lr | 0.01 |
| Loss | VFL Loss |

The improved YOLOv8 deep learning algorithm model built in this study is a classification problem, and the accuracy rate is mainly utilized as an evaluation metric when performing model evaluation *(Li et al., 2023)*.

$$ACC=\frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

$$mACC=\frac{\sum_{i=1}^{N} ACC_i}{N} \tag{7}$$

where $ACC$ denotes accuracy, $TP$ denotes the number of positive samples correctly identified, $TN$ denotes the number of negative samples correctly identified, $FP$ denotes the number of negative samples misreported, $FN$ denotes the number of positive samples omitted, $\mathrm{ACC_i}$ denotes the accuracy of classification of the ith class of fresh white tea leaves, $N$ denotes the type of fresh white tea sorting, which was six in this study, and $mACC$ denotes the mean accuracy.

## RESULTS

### Comparison of the impact of improved modules

In this study, the Dynamic Snake Convolution module (DSConv) was replaced and the Multi-head Self-Attention Mechanism (MHSA) module was added in order to further improve the sorting of fresh white tea leaves. In order to validate the usefulness of the improvements that had been made, the test images were randomly selected from the database of fresh white tea leaves and experiments were conducted, and the extracted comparative plots are shown in Fig. 6.

As can be seen in Fig. 6, after replacing the dynamic snake convolution module in the C2f module, the features extracted from each channel are clearer and crisper. Taking the 10th channel as an example (as shown in the red box in Fig. 6), after the dynamic serpentine convolution module, the extracted features more closely match the contour map of fresh white tea leaves. This is due to the fact that the dynamic serpentine convolution kernel (DSConv) is able to accurately extract adaptive features based on the overall external shape and structure of the tea leaves, focusing on the shape of the white tea fresh leaves themselves, which effectively improves the accuracy of the features.


**Fig. 6 - Comparison of features before and after replacing the dynamic snake convolution**

In the YOLOv8 deep learning algorithm model, with the depth of the network layers, the deep network will focus more on abstract features, and for the Multi-head Self-Attention Mechanism (MHSA), it is added after the SPPF layer. As it can be seen from Fig.7, the multi-head self-attention mechanism (MHSA) focuses more on the detail region of fresh white tea leaves, which can effectively capture the robust features in the image of fresh tea leaves and realize the accuracy and completeness of the feature extraction, at the same time, it helps to further improve the recognition rate of the intelligent sorting of fresh white tea leaves.

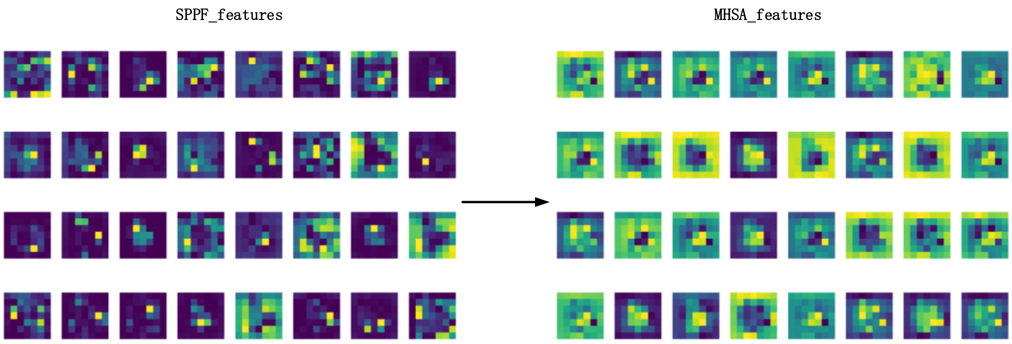


**Fig. 7 - Comparison of features before and after the addition of the Multi-head Self-Attention Mechanism (MHSA) module**

## Comparison of Sorting Performance between Different Models

In this study, in order to verify the effectiveness of the proposed algorithms, YOLOv5, YOLOv7, YOLOv8, YOLOv8-DsConv, YOLOv8-MHSA and our algorithm (YOLOv8-DsConv-MHSA) are used in this study for comparative verification. In white tea fresh leaf sorting, the algorithm speed and sorting accuracy are two indicators that need to be considered in practical applications, so these two indicators are used for comparative analysis.

**Table 4**

**Sorting effect of fresh white tea leaves with different algorithms**

| Model | ACC | | | | | | mACC | FPS |
|---|---|---|---|---|---|---|---|---|
| | OL | TwL | ThL | FL | SiB | BrL | | |
| YOLOv5 | 93.6 | 92.9 | 90.5 | 89.9 | 95.5 | 96.2 | 92.9 | 138.9 |
| YOLOv7 | 94.1 | 93.4 | 90.9 | 90.6 | 95.2 | 96.5 | 93.4 | 147.1 |

| Model | ACC | | | | | | mACC | FPS |
|---|---|---|---|---|---|---|---|---|
| | OL | TwL | ThL | FL | SiB | BrL | | |
| YOLOv8 | 95.2 | 94.6 | 93.5 | 91.8 | 96.2 | 97.1 | 94.6 | 161.3 |
| YOLOv8-DsConv | 96.5 | 95.7 | 94.1 | 93.2 | 97.9 | 97.5 | 95.7 | 137.0 |
| YOLOv8-MHSA | 96.3 | 96.1 | 94.9 | 93.6 | 97.3 | 97.4 | 95.9 | 140.9 |
| Ours | 97.4 | 96.4 | 94.9 | 94.8 | 97.0 | 97.8 | 96.4 | 126.6 |

As can be seen from Table 4, under the same experimental conditions, the YOLOv8-DsConv-MHSA algorithm proposed in this study is far ahead of the other algorithms in terms of accuracy, with an average of 96. 4%, followed by the YOLOv8-MHSA algorithm with an average of 95.9%, the YOLOv8 algorithm, with an average of 94.6%, and the YOLOv8- DsConv algorithm with an average accuracy of 95.7%. From the processing speed of each model for white tea fresh leaves, YOLOv8 algorithm has the best recognition frame rate of 161.3 FPS, while YOLOv8-DsConv-MHSA algorithm has a certain degree of decline in recognition speed due to the increase in the parameters of the model, which is 126. 6 FPS. Combining the results of the above data measured results, it can be seen that the proposed YOLOv8-DsConv-MHSA algorithm has the highest accuracy and the recognition speed of each fresh leaf can meet the requirements of practical applications. Among the various types of white tea fresh leaves, the sorting recognition effect of broken leaves is the best, up to 97.8%. The sorting effect of one bud and one leaf, one bud and two leaves, single bud is also more satisfactory, while the sorting effect of one bud and three leaves, one bud and four leaves is obviously poorer, which indicates that the task of recognizing the complex structure still has some difficulty.

**Practical Application**

Through continuous research and testing, our research team has successfully deployed the proposed algorithm model onto a developed tea sorting machine and conducted application debugging. The components of this tea sorting machine include sensors, a control cabinet, machine vision cameras, a server, robotic arms, and more. The implemented algorithm model is the proposed YOLOv8-DsConv-MHSA. First, the trained model is deployed on the server. When the tea leaves pass through the sensor on the conveyor belt, the sensor sends a signal to the control cabinet. The control cabinet then triggers the machine vision camera to capture an image. The captured image is transmitted to the server, where the fresh tea leaf image is input into the model to complete the sorting and recognition task. Subsequently, the robotic arm controls a vacuum suction cup to pick up the fresh tea leaves that do not meet the requirements and moves them to a parallel secondary conveyor belt for placement, thereby achieving automatic sorting of the fresh tea leaves.

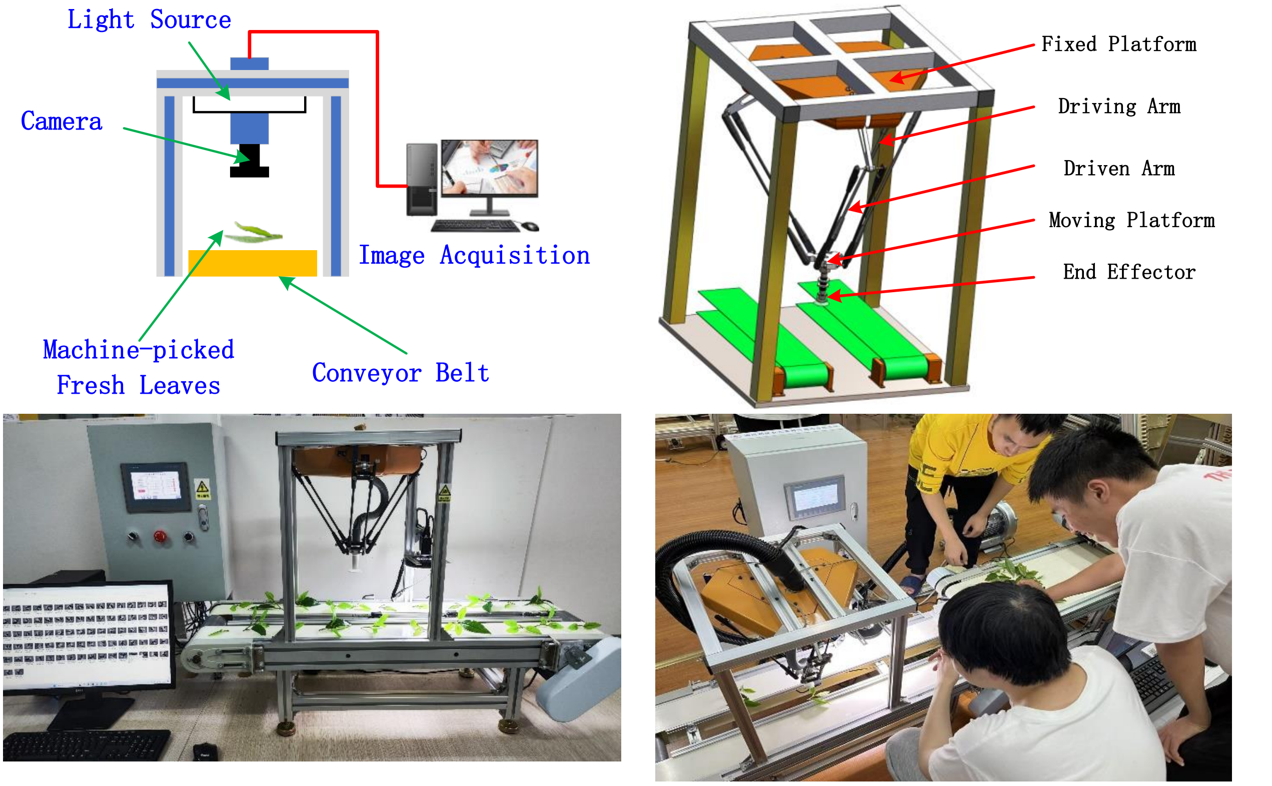


**Fig. 8 - Research and Development Process of the Tea Fresh Leaf Sorting Machine**

After the successful development of our tea sorting machine, a comparative analysis between the tea sorting machine developed by our team and existing tea sorting machines available on the market was conducted. Three tea sorting machines were identified for comparison: the YJY-2 developed by the Tea Research Institute of the Chinese Academy of Agricultural Sciences, the DJFJ developed by Zhejiang Sci-Tech University, and the 6CXF-70 developed by Zhejiang Chunjiang Tea Machinery Co., Ltd. Our analysis primarily focused on two key metrics: screening rate and operational efficiency. The screening rate is defined as the percentage of the weight of sorted tea leaves relative to the total weight of tea leaves, while operational efficiency refers to the amount of tea that can be processed per unit of time. As shown in Table 5, compared to the most advanced tea sorting machines currently available, our equipment achieved a 10.7% improvement in screening rate and a 20% increase in operational efficiency.

**Table 5**

**Performance Comparison Analysis of Different Tea Sorting Machines**

| Sorting Machine | YJY-2 | DJFJ | 6CXF-70 | Ours |
|---|---|---|---|---|
| Sorting Grade | 4 | 4 | 5 | 6 |
| Screening Rate | 62.4% | 81.6% | 76.5% | 92.3% |
| Operational Efficiency | 78kg/h | 115kg/h | 150kg/h | 180kg/h |

## CONCLUSIONS

This paper investigates the problems encountered in sorting fresh white tea leaves and proposes an intelligent sorting algorithm for them using an improved YOLOv8. The algorithm introduces the DSConv and MHSA to enhance the accuracy of feature characterization and enable intelligent sorting of the established fresh white tea leaves dataset. The main results of this study are presented below:

1) A complete dataset of fresh white tea leaves and an improved YOLOv8 deep learning algorithm model were established. The accuracies for one bud and one leaf, one bud and two leaves, one bud and three leaves, one bud and four leaves, single bud, and broken leaves were 97.4%, 96.4%, 94.9%, 94.8%, 97.0%, and 97.8%, respectively, resulting in an average accuracy of 96.4%.

2) The comparison experiments demonstrate that the improved YOLOv8 deep learning model performs much better in intelligent sorting, with an improved average accuracy of 0.7% over YOLOv8-DsConv, 0.5% over YOLOv8-MHSA, 1.8% over YOLOv8, and 3.5% over YOLOv5;

3) In the improved YOLOv8 deep learning algorithm model, the DSConv is able to extract adaptive feature representations based on the shape structure of fresh white tea leaves, and the proposed features are concentrated in the region where the fresh white tea leaves are located, which reduces the bias of the convolutional computation, and improves the accurate representation of the extracted features;

4) In the improved YOLOv8 deep learning algorithm model, each self-attention mechanism in the MHSA focuses on different subspace information at different locations, and fuses the computation results of multiple attention mechanisms to make the captured image features of the fresh white tea leaves more robust, and to realize the accuracy and completeness of feature extraction.

5) The proposed YOLOv8-DsConv-MHSA was deployed on the server for the development of the tea sorting machine. Comparative analysis with existing tea sorting machines revealed that the screening rate improved by 10.7%, and the operational efficiency increased by 20%.

The experimental results show that the algorithm proposed in this study has broad application prospects in the intelligent sorting and recognition of tea leaves, providing a good solution for the subsequent sorting of mechanically harvested fresh white tea leaves. Additionally, the proposed YOLOv8-DsConv-MHSA was deployed on the server and the practical application development of the tea sorting machine was carried out. Of course, there are some limiting factors in the intelligent sorting of fresh tea leaves, such as mutual shading between tea leaves and the influence of light intensity. These issues need to be addressed in subsequent research, and the performance of the tea sorting machine should be further optimized.

## ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (52105283), the Huzhou Public Welfare Application Research Project (2024GZ26,2021GZ27), and the teacher Professional Development Project for Domestic Visiting Scholars in Universities (FX2022113).

**REFERENCES**


[1] Bi, Z., Li, Y., Guan, J., Zhang, X., (2024). Real-time wheat detection based on lightweight deep learning network RepYOLO model [J]. *INMATEH-Agricultural Engineering*, Vol. 72(1), pp-601-610, Romania. DOI: https://doi.org/10.35633/inmateh-72-53

[2] Cao, J., Wu, Z., Zhang, X., Luo, K., Zhao, B., Sun, C., (2023). Sorting of fresh tea leaf using deep learning and air blowing. *Applied Sciences*, Vol. 13(6), 3551, Switzerland.

[3] Chen, Y., Chang, C., Xiao, H., Song, W., Zhang, P., (2010). Application of artificial neural network technology in fresh tea leaf sorting. *Agricultural Network Information*, Vol. 7, pp. 37-40+43, China.

[4] Chen, Z., He, L., Ye, Y., Chen, J., Sun, L., Wu, C., Chen, L., Wang, R., (2020). Automatic sorting of fresh tea leaves using vision-based recognition method. *Journal of Food Process Engineering*, Vol. 43(9), e13474, USA.

[5] Chen, Z., Wang, D., Fang, H., Wang, G., Xie, B., (2021). Rapid identification method of fresh tea leaves based on lightweight model. *In: 2021 IEEE International Conference on Real-time Computing and Robotics (RCAR). IEEE*, pp. 1158-1163, USA.

[6] Du, Y., Gao, A., Song, Y., Guo, J., Ma, W., Ren, L., (2024). Young apple fruits detection method based on improved YOLOv5 [J]. *INMATEH-Agricultural Engineering*, Vol. 73(2), Romania.

[7] Gan, N., Sun, M., Lu, C., Li, M., Wang, Y., Song, Y., Ning, J., Zhang, Z., (2022). High-speed identification system for fresh tea leaves based on phenotypic characteristics utilizing an improved genetic algorithm. *Journal of the Science of Food and Agriculture*, Vol. 102(15), pp. 6858-6867, England.

[8] Gao, M., Shi, M., Li, C., (2021). Research and implementation of image recognition of tea based on deep learning. *In: 2021 21st ACIS International Winter Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD-Winter)*. IEEE, pp. 63-68, USA.

[9] Gao, Z., Wang, A., Liu, Y., Zhang, L., Xia, Y., (2017). Intelligent fresh-tea-leaves sorting system research based on convolution neural network. *Transactions of the Chinese Society of Agricultural Machinery*, Vol. 48(7), China.

[10] Ruxton, C., (2013). Emerging evidence for tea benefits. *Nutrition Bulletin*, Vol. 38(3), pp. 287-301, England.

[11] Li, P., Zheng, J., Li, P., Long, H., Li, M., Gao, L., (2023). Tomato maturity detection and counting model based on MHSA-YOLOv8. *Sensors*, Vol. 23(15), 6701, Switzerland.

[12] Liu, Y., Li, H., (2016). Neural network image technology design of intelligent fresh tea leaf sorter principle research. *Fujian Tea*, Vol. 38(3), pp. 196-197, China.

[13] Lv, H., Wu, C., Tu, Z., Chen, J., Jia, J., Chen, Z., Ye, Y., (2022). Optimization of grading parameters of vibratory classifier for machine-picked tea fresh leaves based on EDEM. *Journal of Tea Science*, Vol. 42(1), pp. 120-130, China.

[14] Pan, Z., Zhang, H., Liu, Z., Yang, R., Su, Z., Li, X., Liu, Z., Shi, C., Wang, S., Wu, H., (2024). Potato appearance detection algorithm based on improved YOLOv8[J]. *INMATEH-Agricultural Engineering*, Vol. 74(3), pp.864-874, Romania. DOI: https://doi.org/10.35633/inmateh-74-76

[15] Pi, W., Wang, R., Sun, Q., Wang, Y., Lu, B., Liu, G., Jin, K., (2023). White tea bud detection based on deep learning research. *INMATEH-Agricultural Engineering*, Vol. 70(2), Romania.

[16] Qi, Y., He, Y., Qi, X., Zhang, Y., Yang, G., (2023). Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. *In: Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6070-6079, USA.

[17] Song, Y., Xie, H., Ning, J., Zhang, Z., (2018). Grade recognition of Keemun black tea based on machine vision shape feature parameters. *Transactions of the CSAE*, Vol. 34(23), pp. 279-286, China.

[18] Tao, S., Wen, S., Hu, G., Ge, Y., Wen, J., Cao, X., Chen, J., (2024). Improved YOLOv8N-based detection of grapes in orchards [J]. *INMATEH-Agricultural Engineering*, Vol. 74(3), pp.473-484, Romania. DOI: https://doi.org/10.35633/inmateh-74-42

[19] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., Polosukhin, I., (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, Vol. 30, USA.

[20] Wang, C., Zhang, T., Zheng, Y., Chen, J., Hu, S., (2016). Simulation study of tea sieving by two degrees of freedom reciprocating tumbler screen. *Journal of Tea Science*, Vol. 36(6), pp. 613-620, China.

[21] Wang, R., Wei, Y., Wu, C., Sun, L., Zheng, W., (2019). Numerical simulation of motion characteristics of flexible fresh tea leaf in Poiseuille shear flow via combined immersed boundary-lattice Boltzmann method. *International Journal of Modern Physics C*, Vol. 30(5), 1950038, USA.

[22] Yang, H., Chen, L., Chen, M., Ma, Z., Deng, F., Li, M., Li, X., (2019). Tender tea shoots recognition and positioning for picking robot using improved YOLO-V3 model. *IEEE Access*, Vol.7, pp. 180998-181011, USA.

[23] Yuan, H., Hua, J., Deng, Y., Jiang, Y., Wang, Y., Chen, G., Zhang, L., Yin, J., (2016). Optimization of grading and classification technology for machine-picking leaves based on YJY-2 type classifier. *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 32(6), pp. 276-282, China.

[24] Zheng, N., Wang, Z., (2011). Mechanization of tea picking: Break out the bottleneck of tea industry. *Chinese Country Technology and Science*, Vol. 12(9), pp. 14-17, China.

[25] Zhu, Y., Fu, M., Shi, Y., (2019). Study on sorting technology of fresh tea leaves based on convolutional neural network model. *In: 2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI). IEEE*, pp. 614-617, USA.

[26] Zhao, F., Zhang, J., Zhang, N., Tan, Z., Xie, Y., Zhang, S., Han, Z., Li, M., (2022). Detection of cucurbits' fruits based on deep learning [J]. *INMATEH-Agricultural Engineering*, Vol. 66(1), pp.321-330, Romania. DOI: https://doi.org/10.35633/inmateh-66-32