Identification of Oil Tea (Camellia oleifera C.Abel) Cultivars Using EfficientNet-B4 CNN Model with Attention Mechanism

Xueyan Zhu 1,2, Xinwei Zhang 1,2, Zhao Sun 1,2, Yili Zheng 1, Shuchai Su 3 and Fengjun Chen 1,2,*

1 School of Technology, Beijing Forestry University, Beijing 100083, China; xueyan0111@bjfu.edu.cn (X.Z.); xinwei@bjfu.edu.cn (X.Z.); sunzhao97@bjfu.edu.cn (Z.S.); zhengyili@bjfu.edu.cn (Y.Z.)
2 Beijing Laboratory of Urban and Rural Ecological Environment, Beijing Forestry University, Beijing 100083, China
3 Key Laboratory of Silviculture and Conversation, Ministry of Education, Beijing Forestry University, Beijing 100083, China; sushuchai@bjfu.edu.cn
* Correspondence: chenfengjun1d502@163.com

Abstract: Cultivar identification is a basic task in oil tea (Camellia oleifera C.Abel) breeding, quality analysis, and an adjustment in the industrial structure. However, because the differences in texture, shape, and color under different cultivars of oil tea are usually inconspicuous and subtle, the identification of oil tea cultivars can be a significant challenge. The main goal of this study is to propose an automatic and accurate method for identifying oil tea cultivars. In this study, a new deep learning model is built, called EfficientNet-B4-CBAM, to identify oil tea cultivars. First, 4725 images containing four cultivars were collected to build an oil tea cultivar identification dataset. EfficientNet-B4 was selected as the basic model of oil tea cultivar identification, and the Convolutional Block Attention Module (CBAM) was integrated into EfficientNet-B4 to build EfficientNet-B4-CBAM, thereby improving the focusing ability of the fruit areas and the information expression capability of the fruit areas. Finally, the cultivar identification capability of EfficientNet-B4-CBAM was tested on the testing dataset and compared with InceptionV3, VGG16, ResNet50, EfficientNet-B4, and EfficientNet-B4-SE. The experiment results showed that the EfficientNet-B4-CBAM model achieves an overall accuracy of 97.02% and a kappa coefficient of 0.96, which is higher than that of other methods used in comparative experiments. In addition, gradient-weighted class activation mapping network visualization also showed that EfficientNet-B4-CBAM can pay more attention to the fruit areas that play a key role in cultivar identification. This study provides new effective strategies and a theoretical basis for the application of deep learning technology in the identification of oil tea cultivars and provides technical support for the automatic identification and non-destructive testing of oil tea cultivars.

Keywords: cultivar identification; oil tea; deep learning; CBAM; EfficientNet-B4-CBAM

1. Introduction

Oil tea (Camellia oleifera C.Abel), which is one of four major woody oil plants in the world [1,2], has high nutritional and economic value globally, and its major producing areas include China, Japan, South Korea, and Vietnam [3,4]. Among these numerous producing areas, China, its country of origin, has the largest cultivar group of oil tea. With an increase in the number of new cultivars of oil tea, even professional researchers can become confused when identifying them, which has had a negative impact on the cultivar breeding and industrial structure adjustment of oil tea [5–7]. Owing to the numerous cultivars and their high phenotypic similarity, many difficulties and challenges have been encountered in the cultivar identification of oil tea [8–10]. Therefore, it is extremely challenging to identify oil tea cultivars scientifically and accurately [11–13].

In previous studies, molecular marker [14,15], hyperspectral image [16,17], and traditional computer vision technologies [18,19] have been used to solve the cultivar identifica-
tion problem. Molecular marker technology has the potential to identify oil tea cultivars; however, its cost is relatively high, and it requires long-term laboratory testing [20]. Although hyperspectral image technology has the potential for the non-destructive detection of oil tea cultivars, the high cost of hyperspectral image acquisition limits its application [21]. Traditional computer vision technology is generally used to make decisions based on a various morphological characteristics extracted manually, and the accuracy is generally low [22–24]. It is currently difficult to identify oil tea cultivars scientifically and accurately.

With the development of machine learning technology, a number of researchers have applied image-based deep learning models to the field of plant cultivar identification [25–27]. At present, the cultivar identification application of image-based deep learning has reached numerous plants, ranging from chrysanthemum [28] and wheat [29,30] to maize [31]. The chrysanthemum cultivar identification method proposed by Liu et al. [32] obtained an accuracy of 78% when identifying chrysanthemum images collected from different years. In addition, Zhang et al. [33] designed a 13-layer deep convolutional neural network to identify the category of fruit in factories and supermarkets, which has an overall accuracy of 94.94%. In a study investigating olive cultivars, Ponce et al. [34] established the olive cultivar identification model Inception-ResnetV2, which provides 95.91% precision for the identification of typical olive cultivars. In an investigation into the automatic identification of plum cultivars at early maturity stages, Rodriguez et al. [35] designed a plum maturity identification algorithm based on AlexNet, and the identification accuracy ranged from 91% to 97%. Although the deep learning models in the above research have achieved remarkable results in plant cultivar identification, the majority of cultivar identification objects are images with easily distinguishable color and texture features. Oil tea cultivar identification, however, is more complicated and difficult, whereas the differences in image features (e.g., texture, shape, and color) under different cultivars of oil tea are usually inconspicuous and subtle. Consequently, the oil tea cultivar identification challenge described in this study can be described as a fine-grained image classification problem.

In recent years, there has been growing academic interest in fine-grained image classification. The positioning–classification method is a vital aspect of fine-grained image classification. Previous positioning–classification methods are more likely to apply to strongly supervised learning, mostly relying on the key areas of manually labeled images. The end-to-end fine-grained image recognition model Mask-CNN proposed by Wei et al. [36] used a partial annotation of the Caltech-UCSD Birds-200-2011 dataset for training a fully convolutional network, which can not only locate the discriminative parts of the bird but can also generate object/part masks. The classification accuracy was 85.5% when Mask-CNN was tested on the Caltech-UCSD Birds-200-2011 dataset. The attention mechanism does not require manual annotation of the key areas of the image and is a method of weak supervision learning. Moreover, the attention mechanism impels the model to make more accurate judgments by distinguishing the importance of the features extracted by the model and extracting key information therein. For instance, the SENet proposed by Hu et al. [37] used channel attention to automatically obtain the importance of each feature channel through network learning and assign different weight coefficients for each feature channel. After different feature channels are processed through channel attention, the importance of the feature channels will be strengthened, and the non-importance feature channels will be suppressed. Compared with traditional Convolutional Neural Network (CNN) models [38,39], the correlation between different channels can be more comprehensively excavated. The two-level attention module proposed by Woo et al. [40] applied both channel attention and spatial attention, where channel attention was used to excavate the correlation between different channels, and spatial attention was used to enhance the specific target area of interest while weakening the irrelevant background area.

At present, the identification of oil tea cultivars has become one of the key issues that limits the efficient development of the oil tea industry. The oil tea industry urgently needs an automatic and accurate method for identifying the cultivars of oil tea. Inspired
by convolutional neural networks and the attention mechanism, this study presents a completely new and more effective convolutional neural network model EfficientNet-B4-CBAM for oil tea cultivar identification. As far as we know, this is the first application of EfficientNet model and CBAM module in the identification of oil tea cultivars. The specific contributions and innovations are summarized as follows: (1) By formulating oil tea cultivar identification as a fine-grained image classification problem, this study presents EfficientNet-B4-CBAM to achieve an adaptive refinement of the feature channels and space. (2) An oil tea cultivar identification dataset was constructed, which contains a total of 4725 images with four cultivars of oil tea. The oil tea cultivar identification dataset was collected using a smartphone and then calibrated by human experts. (3) Based on the oil tea cultivar identification dataset, extensive comparative experiments were conducted on an EfficientNet-B4-CBAM against VGG16 [41], InceptionV3 [42], ResNet50 [43], EfficientNet-B4 [44], and EfficientNet-B4-SE [45]. The results show that the proposed EfficientNet-B4-CBAM is superior to the other methods in comparative experiments, proving the effectiveness of embedding a CBAM module.

2. Materials and Methods

2.1. Study Area

The study area is located in the state-owned Camellia Oleifera Forestry Farm of Huang Lawn, Shaoyang City, Hunan Province, China. The altitude is 470 m above sea level. According to the Köppen climate classification, the study area has a humid subtropical climate, abundant precipitation, simultaneous rain and heat, and four distinct seasons. The average annual temperature is 17.2 °C, and the average rainfall is 1361.6 mm per annual. The State-owned Camellia oleifera Forestry Farm of Huang Lawn is the only state-owned farm named after oil tea in China, with 52.67 ha of a national key oil tea forestry seed base and 133.33 ha of a special oil tea industrial park.

2.2. Data Acquisition and Dataset Construction

The most represented oil tea cultivars, i.e., Xianglin 210, Huashuo, Huaxin, and Huajin, in Hunan Province were selected as the research objects. The images used in this research were taken using an iPhone XR from 8:00 to 17:00 from 4 September to 7 November 2020. To cover the diversity in lighting, shadow, and background, the oil tea images were captured under different weather conditions. The data acquisition process used a minimum resolution of 72 dpi, and the resulting images were saved in the Joint PhotographicExperts Group format with a pixel resolution of 3024 × 4032. After data acquisition, 4725 images were obtained. Sample images of the four oil tea cultivars are shown in Figure 1.

![Sample image examples of the four oil tea cultivars: (a) Xianglin 210; (b) Huajin; (c) Huashuo; (d) Huaxin.](image-url)
The 4725 raw experimental images collected from four oil tea cultivars were used to build a high-quality oil tea cultivar identification dataset. As shown in Table 1, the oil tea cultivar identification dataset was divided into three parts: one for training (60% of the dataset, 2845 samples), one for validation (20% of the dataset, 940 samples), and one for testing (20% of the dataset, 940 samples). The training dataset was used to fit the model. The validation dataset was used to adjust the hyperparameters of the model and to select the best model. The testing dataset was used to evaluate the performance of the best model.

Table 1. Number of oil tea cultivar identification dataset.

| Dataset      | Xinglin 210 | Huashuo | Huajin | Huaxin |
|--------------|-------------|---------|--------|--------|
| Training dataset | 795         | 861     | 592    | 597    |
| Validation dataset | 264         | 285     | 195    | 196    |
| Testing dataset | 264         | 285     | 195    | 196    |

2.3. EfficientNet-B4-CBAM Model

EfficientNet is a highly accurate network obtained through a machine search [46]. It uses a simple and effective compound coefficient to uniformly scale the width, depth, and resolution of the networks [47]. In addition, compared to other CNN models that achieve similar accuracy on the ImageNet dataset, EfficientNet is much smaller. To accurately identify different oil tea cultivars, a new deep learning model, called EfficientNet-B4-CBAM, is built combining the EfficientNet-B4 model and CBAM module. Figure 2 illustrates the network structure of the EfficientNet-B4-CBAM model.

Figure 2. Framework of the proposed EfficientNet-B4-CBAM for oil tea cultivar identification. Input denotes input image of oil tea; \( N \times N \times C \) represents the size of feature map, where \( N \times N \) is the 2D map size and \( C \) is the number of channels; Conv denotes pointwise convolution; \( k \times n \times n \) represents the convolution kernel of \( n \times n \); MBConv represents mobile inverted bottleneck convolution; MBConv1 \( \times n \) indicates \( n \) MBConv1 modules; MBConv6 \( \times n \) indicates \( n \) MBConv6 modules; CBAM denotes convolutional block attention module; AdaptiveAvgPool2d represents adaptive average pooling; Fully connected represents fully connected layer; Output represents output layer.

As shown in Figure 2, the EfficientNet-B4-CBAM model is mainly composed of an EfficientNet-B4 model and a CBAM module. In the EfficientNet-B4-CBAM model, the EfficientNet-B4 model was responsible for extracting oil tea features, whereas the CBAM module was responsible for realizing the refinement of the extracted oil tea features [48]. The EfficientNet-B4 model comprises mostly a mobile inverted bottleneck convolution, with
a three-channel image with a pixel resolution of 380 × 380 as input and an identification result as the output. Figure 3 shows an illustration of MBConv6.

![Figure 3](image)

**Figure 3.** The MBConv6 (k5 × 5) fundamental structure. MBConv represents mobile inverted bottleneck convolution; Conv denotes pointwise convolution; BN denotes batch normalization; Swish denotes swish activation function; DWConv represents depthwise convolution; SE Module represents squeeze-and-excitation module; FC represents fully connected layer; H × W × F represents tensor shape (height, width, depth).

Pointwise convolution, depthwise convolution, and a squeeze-and-excitation (SE) module are the three primary components of MBConv. When receiving the feature map, MBConv first executes a 1 × 1 pointwise convolution on it and then changes the channel dimension of the input feature map according to the expansion ratio. Next, a 5 × 5 depthwise convolution is applied, followed by the introduction of a SE module to boost the expressiveness of the model. Subsequently, 1 × 1 pointwise convolution is used to return the feature map to its original channel dimension. Finally, a drop connect is executed, and a skip connection of the input is applied.

As can be seen in Figure 4, the CBAM module mainly consists of a channel attention module and a spatial attention module.

![Figure 4](image)

**Figure 4.** The structure of the attention block: (a) architecture of CBAM module; (b) architecture of channel attention module; (c) architecture of spatial attention module. MaxPool denotes max pooling layer; AvgPool denotes average pooling layer; MLP denotes a two-level fully connected layer; Sigmoid represents sigmoid activation function; Concatenate represents concatenate layer.

As shown in Figure 4a, when the feature map F of the input oil tea image is input to the CBAM module, the CBAM module first sends F to the channel attention module for processing, resulting in the channel attention feature map Mc of the input oil tea image. Then, refined oil tea image feature map F' needed by the spatial attention module is obtained by multiplying Mc with F. Then, F' is fed into the spatial attention module,
yielding the spatial attention feature map $M_s$ of oil tea image, and $F'$ is multiplied by $M_s$ to produce the final feature map $F''$ of oil tea image.

As shown in Figure 4b, the channel attention mechanism compresses the feature map $F$ of the oil tea images in the spatial dimension to produce a one-dimensional vector $M_c$. Average pooling and max pooling are used to aggregate the spatial information of the feature maps, and a shared multilayer perceptron is used to compress the spatial dimensions of the input oil tea image feature maps.

As shown in Figure 4c, the spatial attention module compresses the oil tea image feature maps $F$ output by the channel attention module within the spatial dimension to generate a one-dimensional vector and then generates spatial attention feature $M_s$ after a sigmoid. Max pooling is used to extract the maximum value on the channel, and average pooling is used to extract the average value on the channel.

2.4. Evaluation Indicators

In this study, six quantitative criteria are used to evaluate the oil tea cultivar identification results. The accuracy ($ACC$), precision ($P$), recall ($R$), $F_1$-score ($F_1$-score), overall accuracy ($OA$), and kappa coefficient ($K_c$) are used to assess and compare the performance of the cultivar identification, as shown in Equations (1)–(6).

\[
ACC = \frac{TP + TN}{TP + FN + FP + TN} \times 100\% \quad (1)
\]

\[
P = \frac{TP}{TP + FP} \times 100\% \quad (2)
\]

\[
R = \frac{TP}{TP + FN} \times 100\% \quad (3)
\]

\[
F_{1\text{-score}} = 2 \cdot \frac{P \cdot R}{P + R} \quad (4)
\]

\[
OA = \frac{\sum_{i=1}^{r} x_{ii}}{N} \times 100\% \quad (5)
\]

\[
K_c = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} \times x_{+i})} \times 100\% \quad (6)
\]

where $TP$ is a true positive, indicating the positive samples that are correctly classified by the model; $FN$ is a false negative, indicating the positive samples that are misclassified by the model; $FP$ is a false positive, indicating the negative samples misclassified by the model; and $TN$ is a true negative, indicating the negative samples correctly classified by the model. In addition, $N$ is the number of observation points of the confusion matrix, $r$ is the number of oil tea cultivars. The $x_{ij}$ represents the element in the $i^{th}$ row and the $j^{th}$ column of the confusion matrix. The $x_{i+}$ represents the sum of the $i^{th}$ row of the confusion matrix. The $x_{+i}$ represents the sum of the $i^{th}$ column of the confusion matrix.

2.5. Experimental Environment Configuration

All CNN models in this study were modeled using the Python tool in the TensorFlow and Keras framework, and all the confusion matrix and heat maps were exported using Matplotlib. The training and testing of the convolutional neural network were conducted on a PC with a 3.7-GHz Intel i7-8700k CPU and an NVIDIA GeForce GTX 1080 Ti graphics processing unit.

The input image size and training epochs of the EfficientNet-B4-CBAM model were $380 \times 380$ pixels and 30 epochs, respectively. To accelerate the model training and improve
the model performance, transfer learning was exploited, and the initial freeze layer was set to 12. The learning rate and batch size were set to 0.0001 and 8, respectively.

3. Results and Discussion

3.1. Analysis Results of Cultivar Identification Using EfficientNet-B4-CBAM

In this section, the cultivar identification performance of the EfficientNet-B4-CBAM model was assessed using images from the testing dataset. Figure 5 shows the confusion matrix of EfficientNet-B4-CBAM for the identification of oil tea cultivars in the test dataset.

![Confusion matrix of EfficientNet-B4-CBAM](image)

Figure 5. Confusion matrix of EfficientNet-B4-CBAM.

From the confusion matrix, it can be seen that the misidentification of oil tea cultivars mainly occurs between Huashuo and Huajin. After observing the dataset of oil tea, we found that there were two main reasons: (1) there are extremely small and subtle differences between different oil tea cultivars, and (2) leaf-shading and illumination influence the identification of the oil tea cultivars. In Table 2, we report the quantitative evaluation results of EfficientNet-B4-CBAM on the testing dataset.

Table 2. Quantitative evaluation results of EfficientNet-B4-CBAM.

| Cultivars     | TP  | TN  | FP  | FN  | ACC (%) | P (%) | R (%) | F1-score (%) | OA (%) | Kc  |
|--------------|-----|-----|-----|-----|---------|-------|-------|-------------|--------|-----|
| Xianglin 210 | 260 | 671 | 5   | 4   | 99.04   | 98.11 | 98.48 | 98.29       | 97.02  | 0.96|
| Huashuo      | 278 | 642 | 13  | 7   | 97.87   | 95.53 | 97.54 | 96.52       |        |     |
| Huajin       | 184 | 738 | 7   | 11  | 98.09   | 96.34 | 94.36 | 95.34       |        |     |
| Huaxin       | 190 | 741 | 3   | 6   | 99.04   | 98.44 | 96.94 | 97.68       |        |     |

As shown in Table 2, the proposed EfficientNet-B4-CBAM has a remarkable effect on the identification of oil tea cultivars in the testing dataset. From the perspective of single cultivar identification, the accuracy of EfficientNet-B4-CBAM for identifying four oil tea cultivars was higher than 97.87%, the precision was higher than 95.53%, the recall was higher than 94.36%, and the F1-score was higher than 95.34%. Taking Xianglin 210 as an example, the accuracy, precision, recall, and F1-score of EfficientNet-B4-CBAM for Xianglin 210 identification were 99.04%, 98.11%, 98.48%, and 98.29%, respectively. The identification results of a single oil tea cultivar showed that EfficientNet-B4-CBAM could accurately identify each cultivar in the testing dataset. From the perspective of the overall cultivar identification, we found that the overall accuracy of EfficientNet-B4-CBAM for the testing dataset identification was 97.02%, and the kappa coefficient was 0.96, indicating that EfficientNet-B4-CBAM has a stable ability to identify all oil tea cultivars.

Moreover, we found that the EfficientNet-B4-CBAM model also made some misjudgments when identifying oil tea cultivars. For example, EfficientNet-B4-CBAM identifies Xianglin210, Huashuo, Huajin, and Huaxin with FP of 5, 13, 7, and 3, respectively, and FN
of 4, 7, 11, and 6, respectively. This was because the differences between some cultivars were extremely small and subtle [49]. For instance, the difference between Xianglin 210 and Huashuo was only the rind color of the fruit. In addition, between Huajin and Huashuo, the only difference was the shape of the fruit.

3.2. Comparison of Cultivar Identification Results with Different Models

To estimate the cultivar identification performance of EfficientNet-B4-CBAM model, the EfficientNet-B4-CBAM model was compared with other CNN models on the testing dataset. The cultivar identification results of these models were presented in the form of a confusion matrix, as shown in Figure 6.

![Confusion Matrix](image)

Figure 6. Confusion matrix of six oil tea cultivar identification models. (a) InceptionV3; (b) VGG16; (c) ResNet50; (d) EfficientNet-B4; (e) EfficientNet-B4-SE; (f) EfficientNet-B4-CBAM.

As Figure 6 indicates, the cultivar misidentification of EfficientNet-B4-CBAM in the testing dataset was far lower than that of the other models in the comparison experiment. Comparing the confusion matrices of VGG16, InceptionV3, ResNet50, and EfficientNet-B4, it can be seen that the cultivar misidentification of EfficientNet-B4 is lower, which proves that it is reasonable to choose EfficientNet-B4 as the base model in this study. In addition, after adding the CBAM module to the EfficientNet-B4 model, the correctly identified Xianglin210, Huashuo, Huajin, and Huaxin increase by 11, 1, 6, and 35. The test results prove that the CBAM module can effectively improve the oil tea cultivar identification ability of the EfficientNet-B4 model [50].

To comprehensively assess the EfficientNet-B4-CBAM, we used the accuracy, precision, recall, F1-score, overall accuracy, and kappa coefficients as evaluation indicators to quantitatively evaluate all methods of the comparative experiment. A comparative evaluation of several state-of-art identification methods on the testing dataset is presented in Table 3.
It can be clearly seen from Table 3 that, compared with InceptionV3, VGG16, and ResNet50, EfficientNet-B4 has a better ability to identify oil tea cultivars. The identification overall accuracy of EfficientNet-B4 for the oil tea cultivar was 91.38%, which was 22.34%, 19.36%, and 5.85% higher than that of InceptionV3, VGG16, and ResNet50, respectively. The kappa coefficient of the EfficientNet-B4 was 0.88, which was 0.29, 0.26, and 0.07 higher than that of InceptionV3, VGG16, and ResNet50, respectively.

By comparing the experimental data of EfficientNet-B4, EfficientNet-B4-SE, and EfficientNet-B4-CBAM listed in Table 3, we found that adding an attention mechanism to EfficientNet-B4 can significantly improve the ability of oil tea cultivar identification. When the attention mechanism was not added, the overall accuracy and kappa coefficients of EfficientNet-B4 for oil tea cultivar identification in the testing dataset were 91.38% and 0.88, respectively. When the SE module was added to EfficientNet-B4, such as EfficientNet-B4-SE shown in Table 3, the overall accuracy and kappa coefficient of oil tea cultivar identification were increased by 1.60% and 0.02, respectively. When the CBAM module was added to EfficientNet-B4, such as EfficientNet-B4-CBAM in Table 3, the overall accuracy and kappa coefficient of oil tea cultivar identification were increased by 5.64% and 0.08, respectively. From the cultivar identification results of EfficientNet-B4, EfficientNet-B4-SE, and EfficientNet-B4-CBAM, we discovered that the performance improvement of EfficientNet-B4 when using the CBAM module is superior to that of the SE module, which may be associated with the fact that the spatial attention module of the CBAM module can locate the key information more accurately.

In summary, after a comprehensive comparison of different models on the testing dataset, we found that the EfficientNet-B4-CBAM proposed in this paper can accurately identify oil tea cultivars under natural conditions. Compared with other models used in comparative experiments, EfficientNet-B4-CBAM has obvious advantages for most evaluation indicators.

### 3.3. Visual Analysis of Cultivar Identification Results

To investigate the reason of why EfficientNet-B4-CBAM outperforms other CNN models, Grad-CAM [51] is adopted to visualize the cultivar identification results of oil tea cultivars. The visualization results are shown in Figure 7.
tea. The attention heat maps of EfficientNet-B4 and EfficientNet-B4-CBAM are displayed in Figure 7.

Figure 7. Attention heat maps of EfficientNet-B4 and EfficientNet-B4-CBAM. (a) Original image of Xianglin 210; (b) Original image of Huashuo; (c) Original image of Huajin; (d) Original image of Hauxin; (e) EfficientNet-B4 identification heat map of Xianglin 210; (f) EfficientNet-B4 identification heat map of Huashuo; (g) EfficientNet-B4 identification heat map of Huajin; (h) EfficientNet-B4 identification heat map of Hauxin; (i) EfficientNet-B4-CBAM identification heat map of Xianglin 210; (j) EfficientNet-B4-CBAM identification heat map of Huashuo; (k) EfficientNet-B4-CBAM identification heat map of Huajin; (l) EfficientNet-B4-CBAM identification heat map of Hauxin.

As shown in Figure 7, it is apparent that when identifying the cultivars of oil tea, EfficientNet-B4 focused on the areas of the fruit and background, whereas EfficientNet-B4-CBAM paid substantial attention to the areas of the fruit, but paid little attention to the background areas. According to the experience of human experts, when identifying oil tea cultivars, the fruit areas of oil tea images can often provide key information for cultivar
identification, whereas the background areas of the oil tea image generally interfere with the cultivar identification [52].

By comparing the heat maps shown in Figure 7, it can be seen that after the CBAM module was added to EfficientNet-B4, the attention of the model focused more on the fruit areas. Based on this phenomenon, we can determine that when identifying oil tea cultivars, the CBAM module can not only determine the location of key information but also improve the expression of information in key areas, thereby improving the identification of oil tea cultivars based on the use of EfficientNet-B4. The heat map analysis experiment proved the capability of the EfficientNet-B4-CBAM model to identify oil tea cultivars from a visual perspective.

4. Conclusions

There are significant differences in the yield, oil content, and camellia oil quality among different oil tea cultivars, and the identification of oil tea cultivars has an important impact on cultivar breeding and an adjustment of the industrial structure. In this study, four typical oil tea cultivars were identified using computer vision technology based on deep learning, and an oil tea cultivar identification model, EfficientNet-B4-CBAM, was proposed.

When identifying the four typical oil tea cultivars in the testing dataset, the overall accuracy and kappa coefficient of EfficientNet-B4-CBAM proposed in this paper were 97.02% and 0.96, respectively. Compared with other CNN models used in the comparative experiment, EfficientNet-B4-CBAM has obvious advantages in all evaluation indexes. The experiment results of the visual analysis show that the EfficientNet-B4-CBAM proposed in this paper can not only accurately locate fruit areas but can also fully express the information of such areas when identifying the cultivars of oil tea.

This research can provide more advanced technical options for the identification of oil tea cultivars and lay a foundation for further research on image-based non-destructive recognition of oil tea cultivars. We suggest that future work can further enrich the oil tea cultivar identification data set, and provide a sufficient data basis for the research on the oil tea cultivar identification algorithm. In future research, we will optimize the speed of the EfficientNet-B4-CBAM model and attempt to deploy it using a mobile phone.

Author Contributions: Conceptualization, X.Z. (Xueyan Zhu); methodology, X.Z. (Xueyan Zhu); software, X.Z. (Xueyan Zhu); validation, X.Z. (Xueyan Zhu) and X.Z. (Xinwei Zhang); formal analysis, X.Z. (Xueyan Zhu); investigation, Z.S.; data curation, Z.S. and S.S.; writing—original draft preparation, X.Z. (Xueyan Zhu); writing—review and editing, F.C. and Y.Z.; visualization, X.Z. (Xinwei Zhang); supervision, F.C.; project administration, F.C.; funding acquisition, F.C. and S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the “National Key Research and Development Program of China: grant number 2019YFD1002401”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Wu, L.; Li, J.; Li, Z.; Zhang, F.; Tan, X. Transcriptomic analyses of Camellia oleifera ‘Huaxin’ leaf reveal candidate genes related to long-term cold stress. Int. J. Mol. Sci. 2020, 21, 846. [CrossRef]
2. Zhang, M.; Wang, A.; Qiu, M.; Qin, X.; Yang, S.; Su, S.; Sun, Y.; Zhang, L. Direct and indirect somatic embryogenesis induction in Camellia oleifera Abel. Front. Plant Sci. 2021, 12, 644389. [CrossRef]
3. Han, J.; Sun, R.; Zeng, X.; Zhang, J.; Xing, R.; Sun, C.; Chen, Y. Rapid classification and quantification of Camellia (Camellia oleifera Abel.) oil blended with rapeseed oil using FTIR-ATR spectroscopy. Molecules 2020, 25, 2036. [CrossRef] [PubMed]
4. Deng, Q.; Li, J.; Gao, C.; Cheng, J.; Deng, X.; Jiang, D.; Li, L.; Yan, P. New perspective for evaluating the main Camellia oleifera cultivars in China. Sci. Rep. 2020, 10, 20676. [CrossRef] [PubMed]
5. Chen, Y.; Wang, B.; Chen, J.; Wang, X.; Wang, R.; Peng, S.; Chen, L.; Ma, L.; Luo, J. Identification of rubisco rbcL and rbcS in *Camellia oleifera* and their potential as molecular markers for selection of high tea oil cultivars. *Front. Plant Sci.* **2015**, *6*, 189. [CrossRef]

6. Wen, Y.; Su, S.; Ma, L.; Yang, S.; Wang, Y.; Wang, X. Effects of canopy microclimate on fruit yield and quality of *Camellia oleifera*. *Sci. Hortic.* **2018**, *235*, 132–141. [CrossRef]

7. Zeng, W.; Endo, Y. Effects of Cultivars and geography in China on the lipid characteristics of *Camellia oleifera* seeds. *J. Oleo Sci.* **2019**, *68*, 1051–1061. [CrossRef]

8. Sahari, M.; Ataii, D.; Hamedi, M. Characteristics of tea seed oil in comparison with sunflower and olive oils and its effect as a natural antioxidant. *J. Am. Oil Chem. Soc.* **2004**, *81*, 585–588. [CrossRef]

9. Xu, Z.; Cao, Z.; Yao, H.; Li, C.; Zhao, Y.; Yuan, D.; Yang, G. The physicochemical properties and fatty acid composition of two new woody oil resources: *Camellia hainanica* seed oil and *Camellia sinensis* seed oil. *CyTA-J. Food* **2021**, *19*, 208–211. [CrossRef]

10. He, Z.; Liu, C.; Wang, X.; Wang, R.; Tian, Y.; Chen, Y. Leaf transcriptome and weight gene co-expression network analysis uncovers genes associated with photosynthetic efficiency in *Camellia oleifera*. *Biochem. Genet.* **2021**, *59*, 398–421. [CrossRef] [PubMed]

11. Wang, R.; Chen, L.; Chen, J.; Chen, Y.; Zhang, Z.; Wang, X.; Peng, Y.; Peng, S.; Li, A.; Wei, X. Different nitrate and ammonium ratios affect growth and physiological characteristics of *Camellia oleifera* Abel. seedlings. *Forests* **2018**, *9*, 784. [CrossRef]

12. Liu, C.; Chen, L.; Tang, W.; Peng, S.; Li, M.; Deng, N.; Chen, Y. Predicting potential distribution and evaluating suitable soil condition of oil tea Camellia in China. *Forests* **2018**, *9*, 487. [CrossRef]

13. Cheng, J.; Jiang, D.; Cheng, H.; Zhou, X.; Fang, Y.; Zhang, X.; Xiao, X.; Deng, X.; Li, L. Determination of *Camellia oleifera* Abel. germplasm resources of genetic diversity in China using ISSR markers. *Not. Bot. Horti Agrobot. Cluj-Napoca* **2018**, *46*, 501–508. [CrossRef]

14. Kim, M.; Jung, J.; Shim, E.; Chung, S.; Park, Y.; Lee, G.; Sim, S. Genome-wide SNP discovery and core marker sets for DNA barcoding and variety identification in commercial tomato cultivars. *Sci. Hortic.* **2021**, *276*, 109734. [CrossRef]

15. Shamustakimova, A.; Mavllyutov, Y.; Klimenko, I. Application of SRAP markers for DNA identification of Russian alfalfa Cultivars. *Russ. J. Genet.* **2021**, *57*, 540–547. [CrossRef]

16. Wang, X.; Li, Z.; Zheng, D.; Wang, W. Nondestructive identification of millet varieties using hyperspectral imaging technology. *J. Appl. Spectrosc.* **2020**, *87*, 54–61. [CrossRef]

17. Zhao, F.; Yan, Z.; Xue, J.; Xu, B. Identification of wild black and cultivated goji berries by hyperspectral image. *Spectrosc. Spectr. Anal.* **2021**, *41*, 201–205. [CrossRef]

18. Sachar, S.; Kumar, A. Survey of feature extraction and classification techniques to identify plant through leaves. *Expert Syst. Appl.* **2021**, *167*, 114181. [CrossRef]

19. Waldchen, J.; Mader, P. Machine learning for image based species identification. *Methods Ecol. Evol.* **2018**, *9*, 2216–2225. [CrossRef]

20. Kumar, S.; Singh, L. Traditional system versus DNA barcoding in identification of bamboo species: A systematic review. *Mol. Biotechnol.* **2021**, *63*, 651–675. [CrossRef]

21. Li, H.; Zhang, L.; Sun, H.; Rao, Z.; Ji, H. Identification of soybean varieties based on hyperspectral imaging technology and one-dimensional convolutional neural network. *J. Food Process Eng.* **2021**, *44*, e13767. [CrossRef]

22. Yang, C. Plant leaf recognition by integrating shape and texture features. *Pattern Recognit.* **2021**, *112*, 107809. [CrossRef]

23. Alexandru Borz, S.; Păun, M. Integrating offline object tracking, signal processing, and artificial intelligence to classify relevant events in sawmilling operations. *Forests* **2020**, *11*, 1333. [CrossRef]

24. Roshani, M.; Sattari, M.A.; Ali, P.J.M.; Roshani, G.H.; Nazemi, B.; Corniani, E.; Nazemi, E. Application of GMDH neural network technique to improve measuring precision of a simplified photon attenuation based two-phase flowmeter. *Flow Meas. Instrum.* **2020**, *75*, 101804. [CrossRef]

25. Javanmardi, S.; Ashian, S.; Verbeek, F.; Martynenko, A. Computer-vision classification of corn seed varieties using deep convolutional neural network. *J. Stored Prod. Res.* **2021**, *92*, 101800. [CrossRef]

26. Jin, X.; Che, J.; Chen, Y. Weed identification using deep learning and image processing in vegetable plantation. *IEEE Access* **2021**, *9*, 10940–10950. [CrossRef]

27. Cheţa, M.; Marcu, M.V.; Iordache, E.; Borz, S.A. Testing the capability of low-cost tools and artificial intelligence techniques to automatically detect operations done by a small-sized manually driven bandsaw. *Forests* **2020**, *11*, 739. [CrossRef]

28. Zhang, R.; Tian, Y.; Zhang, J.; Dai, S.; Hou, X.; Wang, J.; Guo, Q. Metric learning for image-based flower cultivars identification. *Plant Methods* **2021**, *17*, 65. [CrossRef]

29. Yang, B.; Gao, Z.; Gao, Y.; Zhu, Y. Rapid detection and counting of wheat ears in the field using YOLOv4 with attention module. *Agronomy* **2021**, *11*, 1202. [CrossRef]

30. Kayaa, A.; Keceli, A.; Catal, C.; Yalic, H.; Temucin, H.; Tekinerdogan, B. Analysis of transfer learning for deep neural network based plant classification models. *Comput. Electron. Agric.* **2019**, *158*, 20–29. [CrossRef]

31. Zhang, C.; Zhao, Y.; Yan, T.; Bai, X.; Xiao, Q.; Gao, P.; Li, M.; Huang, W.; Bao, Y.; He, Y.; et al. Application of near-infrared hyperspectral imaging for variety identification of coated maize kernels with deep learning. *Infrared Phys. Technol.* **2020**, *111*, 103550. [CrossRef]

32. Liu, Z.; Wang, J.; Tian, Y.; Dai, S. Deep learning for image-based large-flowered chrysanthemum cultivar recognition. *Plant Methods* **2019**, *9*, 146. [CrossRef] [PubMed]
33. Zhang, Y.; Dong, Z.; Chen, X.; Jia, W.; Du, S.; Muhammad, K.; Wang, S. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. Multimed. Tools Appl. 2019, 78, 3613–3632. [CrossRef]
34. Ponce, J.; Aquino, A.; Andújar, J. Olive-fruit variety classification by means of image processing and convolutional neural networks. IEEE Access 2019, 7, 147629–147641. [CrossRef]
35. Rodríguez, F.; García, A.; Pardo, P.; Chávez, F.; Luque-Baena, R. Study and classification of plum varieties using image analysis and deep learning techniques. Prog. Artif. Intell. 2018, 7, 119–127. [CrossRef]
36. Wei, X.; Xie, C.; Wu, J.; Shen, C. Mask-CNN: Localizing parts and selecting descriptors for fine-grained bird species categorization. Pattern Recogn. 2018, 76, 704–714. [CrossRef]
37. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018. [CrossRef]
38. Nazemi, B.; Rafiean, M. Forecasting house prices in Iran using GMDH. Int. J. Hous. Mark. Anal. 2021, 14, 555–568. [CrossRef]
39. Borz, S.A. Development of a modality-invariant multi-layer perceptron to predict operational events in motor-manual willow felling operations. Forests 2021, 12, 406. [CrossRef]
40. Wei, X.; Xie, C.; Wu, J.; Shen, C. Mask-CNN: Localizing parts and selecting descriptors for fine-grained bird species categorization. Pattern Recogn. 2018, 76, 704–714. [CrossRef]