Maturity Grading and Identification of *Camellia oleifera* Fruit Based on Unsupervised Image Clustering

Xueyan Zhu 1,2, Deyu Shen 1,2, Ruipeng Wang 1,2, Yili Zheng 1,2, Shuchai Su 3 and Fengjun Chen 1,2,*

1 School of Technology, Beijing Forestry University, Beijing 100083, China
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
* Correspondence: chenfj227@bjfu.edu.cn

**Abstract:** Maturity grading and identification of *Camellia oleifera* are prerequisites to determining proper harvest maturity windows and safeguarding the yield and quality of Camellia oil. One problem in *Camellia oleifera* production and research is the worldwide confusion regarding the grading and identification of *Camellia oleifera* fruit maturity. To solve this problem, a *Camellia oleifera* fruit maturity grading and identification model based on the unsupervised image clustering model DeepCluster has been developed in the current study. The proposed model includes the following two branches: a maturity grading branch and a maturity identification branch. The proposed model jointly learns the parameters of the maturity grading branch and maturity identification branch and used the maturity clustering assigned from the maturity grading branch as pseudo-labels to update the parameters of the maturity identification branch. The maturity grading experiment was conducted using a training set consisting of 160 *Camellia oleifera* fruit samples and 2628 *Camellia oleifera* fruit digital images collected using a smartphone. The proposed model for grading *Camellia oleifera* fruit samples and images in training set into the following three maturity levels: unripe (47 samples and 883 images), ripe (62 samples and 1005 images), and overripe (51 samples and 740 images). Results suggest that there was a significant difference among the maturity stages graded by the proposed method with respect to seed oil content, seed soluble protein content, seed soluble sugar content, seed starch content, dry seed weight, and moisture content. The maturity identification experiment was conducted using a testing set consisting of 160 *Camellia oleifera* fruit digital images (50 unripe, 60 ripe, and 50 overripe) collected using a smartphone. According to the results, the overall accuracy of maturity identification for *Camellia oleifera* fruit was 91.25%. Moreover, a Gradient-weighted Class Activation Mapping (Grad-CAM) visualization analysis reveals that the peel regions, crack regions, and seed regions were the critical regions for *Camellia oleifera* fruit maturity identification. Our results corroborate a maturity grading and identification application of unsupervised image clustering techniques and are supported by additional physical and quality properties of maturity. The current findings may facilitate the harvesting process of *Camellia oleifera* fruits, which is especially critical for the improvement of Camellia oil production and quality.

**Keywords:** *Camellia oleifera*; maturity grading; maturity identification; unsupervised image clustering; harvesting process

**1. Introduction**

As an essential raw material source of edible oil, *Camellia oleifera* is of momentous edible woody oil seed crop in China, India, Japan, and Korea [1,2]. Several studies have reported Camellia oil can not only be used as cooking oil but can also be used as a drug carrier and healthcare product [3,4]. Camellia oil has a high content of unsaturated fatty acids (e.g., oleic acid and omega-6-linoleic acid) as well as natural bioactive components...
(e.g., squalene, polyphenols, and sasanquasaponin) that have promoting effects on health care and anti-aging [5–8]. Therefore, the demand for high-quality Camellia oil has been continuously increasing in recent years [9]. The productivity and quality of Camellia oil are not only solely related to factors such as variety, climate, and squeezing technology but also inextricably related to maturity [10]. As the fruit matures, the oil content of *Camellia oleifera* fruits was increased first and then decreased [11]. However, the acidity increases during the ripening of *Camellia oleifera*, reducing the quality of the produced Camellia oil [12]. Therefore, the yield and quality of Camellia oil are determined by the maturity stage of the harvested fruit. As a rule of thumb, both early and later-harvested *Camellia oleifera* fruits exhibit decreased oil content and quality [13,14]. Hence, accurate identification of *Camellia oleifera* fruit maturity has to be made.

Based on the accumulation of long-term practical experience, the traditional Chinese solar terms (e.g., Autumn Equinox, Cold Dews, Frost’s Descents, and Winter Begins) were commonly used to determine the maturity of different *Camellia oleifera* cultivars [15]. For instance, the traditional Chinese solar terms Frost’s Descents has been considered to be a mature symbol of numerous *Camellia oleifera* cultivars, such as Huaxin, Huajin, Xianglin 1, Xianglin 27, Xianglin 97, and Xianglin 210. Moreover, the traditional Chinese solar term Cold Dews has been used to determine the maturity of different *Camellia oleifera* cultivars, including Xianglin 104, Xianglin 106, Xianglin 107, and Xianglin 131. However, extensive research has found that with global climate change, the phenology of *Camellia oleifera* has altered gradually, which will result in the currently recommended maturity based on traditional Chinese solar terms is not applicable [16,17]. Accordingly, a stable *Camellia oleifera* fruit maturity identification method without being affected by climate change is urgently needed.

For a long time, the oil content has been considered one of the most critical parameters to determine the quality of *Camellia oleifera*, which can be applied as a stable and climate-independent parameter to determine the *Camellia oleifera* fruit maturity stages [18–20]. Based on literature studies, *Camellia oleifera* has an oil content that will significantly increase after entering maturity, and then it will slowly decrease due to being overripe [21]. Conventionally, oil content measurement of *Camellia oleifera* fruit samples is carried out in the laboratory using chemical experiments [22,23]. Despite its stability and efficacy, maturity determine based on oil content suffers from the following several major drawbacks: labor-intensive, time-consuming, sample damage, prone to human error, and not environmentally sustainable.

Based on the facts above, several researchers have attempted to produce a technological breakthrough that can non-invasive identify the maturity of oil fruits such as oil palm [24–26] and olive [27,28]. Tugnolo et al. [29] used visible/near-infrared imaging in a study that automatic identifying the maturity of olive fruits. The sensitivity, specificity, and accuracy of identifying the maturity of olive fruits for the method were 86%, 87%, and 87%, respectively. Raj et al. [30] proposed a method that identification the maturity of oil palm fruitlets based on carotene content from Raman spectra. They reported that the overall accuracy of identification for the best model for oil palm maturity identification was 100%. Jiang et al. [31] conducted a study in which they used a hyperspectral imaging system to access the maturity of *Camellia oleifera* fruit. The *Camellia oleifera* fruit maturity identification results indicated that their model was 81.2% accurate. Although the non-invasive maturity identification techniques used in this research works well, the techniques usually require expensive specialized equipment. However, it is very difficult for the grower to obtain this specialized equipment.

To date, in various fruit maturity grading and identification systems, deep learning-based computer vision has become a substitute for subjective experience, destructive testing, and non-invasive testing, being objective, consistent, efficient, and economical [32,33]. Deep learning-based computer vision has also been used for the identification of the maturity of different oil fruits, including oil palm [34–36], coconut [37], and olive [38]. Khosravi et al. [39] designed a deep convolutional neural network (CNN) in a study that
identified the maturity of on-branch olives for Zard and Roghani cultivars. The results of their study showed that the overall accuracy and prediction time for the model were 91.91% and 12.64 ms, respectively. Parvathi et al. [40] described a method based on Faster R-CNN for the detection of two important maturity stages for coconuts in complex backgrounds. In another study, Mubin et al. [41] described a method based on deep learning for the objective and accurate detection of the maturity of oil palms. The results indicated that the accuracy of detection for young and mature oil palms was 95.11% and 92.96%, respectively.

Based on the above-mentioned, it was noted that the majority of published papers target applications of deep learning techniques in the maturity detection of oil fruits, which have maturity grading standards. Up to now, far too little attention has been paid to deep learning for the maturity identification of *Camellia oleifera* fruits that has no maturity grading standard. The review of the previous studies has failed to grade and identification the maturity of *Camellia oleifera* fruits based on images. Therefore, the objective of the present study is to demonstrate a *Camellia oleifera* fruit maturity grading and identification method based on unsupervised image clustering. The contributions of this paper are as follows:

1. Using an unsupervised image clustering method to grade and identification the maturity of *Camellia oleifera* fruits;
2. Conjoint analysis was performed for the physical and quality properties of *Camellia oleifera* fruits with different maturity;
3. The Gradient-weighted Class Activation Mapping (Grad-CAM) visualization analysis reveals the critical regions for *Camellia oleifera* fruit maturity identification.

2. Materials and Methods
2.1. Fruit Sample Preparation

The data used in this study was a set of fruit images and samples of the Huaxin variety obtained from the state-owned *Camellia oleifera* Forestry Farm of Huang Lawn (111°36′8″ N, 27°6′58″ E), Shaoyang City, Hunan Province, China. Color images and samples data of the *Camellia oleifera* fruit were collected 16 times (once every 5 days) starting from early August to mid-November 2020. In image acquisition, the capturing distance between the digital camera built-in smartphone (Xiaomi 10) and the *Camellia oleifera* fruit was approximately 10–30 cm, and the collected images were saved in a 24-bit true-color JPEG format. The images were collected under the condition of sufficient light. Moreover, the digital camera focus mode, white balance, and sensitivity were set to auto, the shutter time was set to 1/2000 s and the exposure compensation was set to 0. Additionally, 20 defect-free fruit samples of *Camellia oleifera* with similar sizes were randomly selected, imaged, and picked during each data collection. After picking fruit samples, physical and quality properties were evaluated. In this study, 320 *Camellia oleifera* fruit samples (16 groups) with different maturity stages were selected. A total of 2788 *Camellia oleifera* fruit images covering the whole mature period were collected, of which 320 images (one image per sample) were corresponding to the sample fruits. The other 2468 *Camellia oleifera* fruit images were randomly collected images of *Camellia oleifera* fruits with phenotypic characteristics roughly similar to the fruit samples. Moreover, the 2468 images were taken from the same *Camellia oleifera* trees of the same variety in the same experimental area as the 320 *Camellia oleifera* fruit sample images mentioned above. The image acquisition equipment, acquisition parameters, and other experimental conditions were identical to the acquisition condition of 320 *Camellia oleifera* fruit sample images. Examples of *Camellia oleifera* images are presented in Figure 1.
2.2. Dataset Details

The *Camellia oleifera* fruit images and samples were divided into the following two sets: a training set (TN) and a testing set (TS), as presented in Table 1. The training set consists of 160 *Camellia oleifera* fruit samples (obtained by randomly selecting 10 samples from each group) and 2628 *Camellia oleifera* fruit images. Among the 2628 *Camellia oleifera* fruit images, there were 160 *Camellia oleifera* fruit images that were corresponding to each of the 160 *Camellia oleifera* fruit samples in the training set. A total of 160 *Camellia oleifera* fruit images and samples were included in the testing set. The *Camellia oleifera* fruits in the testing set were divided into the unripe testing set (Unripe_TS), the ripe testing set (Ripe_TS), and the overripe testing set (Overripe_TS) according to seed oil content [42]. Moreover, the *Camellia oleifera* fruit samples and images of the training set were not manually divided into different maturity according to seed oil content. This is because the unsupervised image clustering technique was used in this study to grade and identified the maturity of *Camellia oleifera* fruits. The typical characteristic of unsupervised image clustering is that the training set is without manual intervention.

Table 1. Dataset overview.

| Dataset Category | Camellia oleifera Fruit Images | Camellia oleifera Fruit Samples |
|------------------|-------------------------------|-------------------------------|
| TN               | 2628                          | 160                           |
| Unripe_TS        | 50                            | 50                            |
| Ripe_TS          | 60                            | 60                            |
| Overripe_TS      | 50                            | 50                            |

2.3. Proposed Method of Fruit Maturity Grading and Identification

The proposed method aims to propose a *Camellia oleifera* fruit maturity grading and identification method based on unsupervised learning. It was divided into two stages, namely, maturity grading and identification, and conjoint analysis of physical and quality properties, where the input for maturity grading was acquired from the *Camellia oleifera* fruit images, and the input for physical and quality property conjoint analysis were acquired from the *Camellia oleifera* fruit samples. In the maturity grading stage, the DeepCluster method was applied to grading *Camellia oleifera* fruit samples and images in training set into different maturity. In the conjoint analysis of the physical and quality properties stage, the ANOVA coupled with Duncan’s test was applied to analyze the differences in physical and quality properties of *Camellia oleifera* fruit with different maturity stages clustered by DeepCluster. An overview of all processes in the proposed method of *Camellia oleifera* fruit maturity grading and identification is depicted in Figure 2.
2.3.1. Grading and Identification of Fruit Maturity

*Camellia oleifera* fruit maturity was graded and identified by DeepCluster [43], an end-to-end unsupervised image clustering model that jointly learns the parameters of neural network (VGG16-D<sub>FC</sub>) and the cluster assignments (K-Means) of the resulting features. As shown in Figure 3, the DeepCluster iteratively groups the features with the K-Means clustering algorithm and updates the parameters of the network using the subsequent assignments as supervision. The DeepCluster model mainly includes the following two main branches: the maturity grading branch (VGG16-D<sub>FC</sub> and PL2-K-Means) and the maturity identification branch (VGG16-D<sub>FC</sub> and L_FC). The maturity grading branch clustering the maturity of *Camellia oleifera* fruits into different stages based on the features extracted by VGG16-D<sub>FC</sub>. Moreover, the maturity identification branch identifies the maturity of *Camellia oleifera* fruits based on the features extracted by VGG16-D<sub>FC</sub>.
As shown in Figure 3, the DeepCluster model receives Camellia oleifera fruit images with unlabeled maturity stages as input data. In the feature extraction stages, the features of Camellia oleifera fruit images in the dataset are extracted. Feature extraction was carried out based on the features extracted by the VGG16-D\_FC model because these features have been used successfully in previous works. The initial weights of the VGG16-D\_FC model were generated by random initialization. Because previous studies have shown that these random features perform much better on standard transfer tasks than on random levels [44]. After the feature extraction, the extracted features (4096 dimensional) of Camellia oleifera fruit images are output through the penultimate fully connected layer of the VGG16-D\_FC model [45].

In the feature selection stages, the features extracted were selected. A total of 4096 features were extracted in the feature extraction stage. This leads to complicated clustering processes that also require time-consuming computations. Therefore, feature selection was needed to obtain the critical features with an important contribution to clustering. In this study, the feature selection model—principal component analysis (PCA)—was performed to significantly reduce the number of features. Moreover, the PCA method has been widely applied in many objects related to this work [46–48]. Afterward, normalization was implemented using L2 normalization to minimize the effect of singular features [49]. Because L2 normalization has been successfully used in previous works related to normalization [50,51].

The normalized features were generated by applying the L2 normalization and subsequently used as input in the feature clustering stage. In this study, the K-Means clustering algorithm was implemented to cluster the normalized features into \( k \) distinct groups based on a geometric criterion. Moreover, the K-Means clustering algorithm has been successfully applied in studies related to this work [52–54]. After clustering, each Camellia oleifera image \( i \) in the dataset has a uniquely determined clustering assignment \( y_i \). The optimal cluster assignments \( (y_i^*)_{i=1}^N \) and \( d \times k \) centroid matrix \( C^* \) can be jointly learned using the objective function in Equation (1).

\[
\min_{C \in \mathbb{R}^{d \times k}} \frac{1}{N} \sum_{i=1}^{N} \min_{y_i \in \{0,1\}^k} \|f_\theta(x_i) - C y_i\|_2^2
\]  

(1)

where \( y_i^T 1_k = 1 \) and \( N \) is the total number of Camellia oleifera images in the dataset, \( x_i \) denotes the \( i \)th Camellia oleifera image in the dataset, \( k \) denotes the clustering number, \( C \) denotes the \( d \times k \) centroid matrix, \( f_\theta \) denotes the convnet mapping, \( \theta \) is the set of corresponding parameters of convnet mapping, \( y_i \) is the clustering assignment of \( i \)th Camellia oleifera fruit image.
Then, these assignments were used as pseudo-labels to jointly learn the classifier parameters $W$ and the mapping parameters $\theta$ using the objective function in Equation (2). Moreover, this objective function is optimized by mini-batch stochastic gradient descent and backpropagation to calculate the gradient.

$$\min_{\theta, W} \frac{1}{N} \sum_{i=1}^{N} l(g_w(f_\theta(x_i)), y_i)$$

where $l$ denotes the negative log-softmax function, and $g_w$ denotes the parametrized classifier.

2.3.2. Determination of Physical Properties

After *Camellia oleifera* fruit sample picking, the physical properties of each fruit sample were determined individually. Physical properties were transverse diameter, vertical diameter, dry seed weight, and moisture content of *Camellia oleifera* fruit samples. The transverse and vertical diameters of each *Camellia oleifera* fruit sample were determined on a digital caliper (Deguqmnt, MNT-150T, Hanover, Germany). To acquire the dry seeds, the fresh seeds were deactivated at 105 °C for 15 min and baked at 70 °C for 72 h [13]. Similarly, the fresh seed weight ($W_f$) and dry seed weight ($W_d$) of each *Camellia oleifera* fruit sample were determined with an electronic balance (Shimadzu, UW2200H, Kyoto, Japan). Next, the moisture content ($M_c$) was determined using the following equation:

$$M_c = \frac{W_f - W_d}{W_f} \times 100\%$$

2.3.3. Determination of Quality Properties

Once *Camellia oleifera* fruit samples were taken to the laboratory, the quality property included seed oil content, seed soluble protein content, seed soluble sugar content, and seed starch content were measured for 10 samples in each group. Because the quality properties of individual *Camellia oleifera* fruit samples were difficult to be accurately measured. The oil content in *Camellia oleifera* seed was determined following exhaustive chemical extraction using a Soxtec 2050 extraction system (Foss Analytical, Hillerød, Denmark), as described by Yang et al. [7]. Next, the soluble protein content of the *Camellia oleifera* seed was determined by Coomassie brilliant-blue G-250 staining method on a UV-vis spectrophotometer (Agilent Technologies, Santa Clara, CA, USA) [55]. Subsequently, the soluble sugar content and starch content of the *Camellia oleifera* seed were determined by the anthrone colorimetric method on a UV-vis spectrophotometer (Agilent Technologies, Santa Clara, CA, USA), as described in detail by Zhang et al. [13]. Moreover, the quality properties of each sample in the group were represented approximately by the quality properties within the group.

2.3.4. Conjoint Analysis of Maturity with Physical and Quality Properties

In the fruit maturity clustering stage, the *Camellia oleifera* fruit images are clustered into different maturity stages based on the phenotypic properties extracted by the neural network. Different maturity stages of *Camellia oleifera* fruits had differences in image properties, but it remains to be verified whether differences existed in physical and quality properties. Therefore, analysis of variance (ANOVA) was performed using SPSS software (IBM SPSS 20, Chicago, IL, USA) for the physical and quality properties of *Camellia oleifera* fruit samples in different maturity. Duncan’s multiple comparison tests were used to compare the significance ($p \leq 0.05$) in physical and quality properties among different clusters of *Camellia oleifera* fruits.

2.4. Maturity Identification Performance Evaluation Index

To comprehensively evaluate the effectiveness of the proposed method, four evaluation indexes of image identification, including the precision ($\text{Prec}$), recall ($\text{Rec}$), F1 score ($\text{F1 score}$),
and overall accuracy (OAcc), were used. According to [56], the Prec, Rec, F1score, and OAcc can be calculated through the following equations:

\[
\text{Prec} = \frac{TP}{TP + FP} \times 100\% \tag{4}
\]

\[
\text{Rec} = \frac{TP}{TP + FN} \times 100\% \tag{5}
\]

\[
F1\text{score} = 2 \times \frac{\text{Prec} \times \text{Rec}}{\text{Prec} + \text{Rec}} \times 100\% \tag{6}
\]

\[
\text{OA}_{\text{cc}} = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \tag{7}
\]

where TP represents the positive samples that are correctly identified as a positive class, FP denotes the negative samples predicted as a positive class by the proposed method, FN means the positive samples predicted as a negative class, and TN denotes the negative samples predicted as a negative class by the proposed method.

2.5. Model Visualization

The fruit maturity grading and identification model extract extensive information from the Camellia oleifera fruit images, and these feature data can be visualized and interpreted by Grad-CAM. This method has been successfully applied in the image analysis of lychee and Camellia oleifera cultivar identification [57,58]. The Grad-CAM algorithm obtains a location map that can interpret the key identification regions in the image by weighting and processing the activation feature maps of the last layer according to their weights [59]. The Grad-CAM visualization explains that regions of the image the fruit maturity grading and identification model focused on when determining the Camellia oleifera fruit to be a specific maturity.

3. Results and Discussion

3.1. Implementation Details

The Camellia oleifera fruit maturity grading and identification model was implemented on the Ubuntu 16.04 system and the following hardware configurations: an NVIDIA GeForce GTX 1080 Ti GPU and the memory of 11 G. To construct the Camellia oleifera fruit maturity grading and identification model, we used the Pytorch 1.5.1 platform, which is based on Python ver. 3.6.13. During the model training process, the learning rate was set to 0.1 and the momentum to 0.9.

3.2. Fruit Maturity Grading Results

In the model training stages, the 2628 Camellia oleifera fruit images in the training set were used as input to the DeepCluster. After training, the 2628 Camellia oleifera fruit images in the training set were automatically graded into unripe (883 images), ripe (1005 images), and overripe (740 images), as shown in Figure 4. Then, from 2628 Camellia oleifera fruit images after clustering, we located the images that corresponded to each of the 160 Camellia oleifera fruit samples in the training set. The 160 Camellia oleifera fruit samples in the training set were graded into unripe, ripe, and overripe with the numbers 47, 62, and 51. As shown in Figure 4, the features of the unripe Camellia oleifera are mainly distributed in the right half of the cluster image, and the features of the ripe Camellia oleifera are mainly distributed in the middle of the cluster image. Meanwhile, the features of overripe Camellia oleifera are mainly distributed in the left half of the cluster image.
Figure 4. The maturity grading results of the training set. Here, the black pentagram represents the center of each maturity stage, and the *Camellia oleifera* images represent the fruit samples of each maturity stage.

By observing the maturity grading results along with fruit samples of each maturity stage in Figure 4, we can find that the phenotypic properties of *Camellia oleifera* fruit with the same maturity were approximately similar, while differences existed in the phenotypic properties of *Camellia oleifera* fruit of different maturity. It can be seen that feature extraction by VGG16-D$_{FC}$ and feature clustering by K-Means plays an important role in *Camellia oleifera* maturity clustering. Among the three different maturity stages of *Camellia oleifera*, the *Camellia oleifera* fruits that are unripe have a sleek, and indehiscent peel. The *Camellia oleifera* fruits that are ripe have a sleek, and slightly cracked peel, and the *Camellia oleifera* seeds are slightly exposed. The *Camellia oleifera* fruits with overripe have a rough and obviously cracked peel, and the *Camellia oleifera* seeds are obviously exposed. These results indicated that the phenotypic properties of different clusters of *Camellia oleifera* fruits that are significantly different after clustering using the DeepCluster model. It also confirms the effectiveness of using the DeepCluster model to automatically grade the maturity of *Camellia oleifera*.

According to Figure 4, there exists overlap regions between unripe and ripe as well as between ripe and overripe. However, between unripe and overripe there hardly exists overlap regions. To further present the details of the overlap regions, several examples of the overlap regions are depicted in Figure 5. Figure 5a shows the distribution of *Camellia oleifera* fruit images in unripe and ripe overlap regions and the example images located in overlapping regions. Figure 5b shows the distribution of *Camellia oleifera* fruit images in ripe and overripe overlap regions and the example images located in overlapping regions. It is evident from Figure 5a that the *Camellia oleifera* of different maturity stages in the overlapping regions (unripe and ripe) have slight cracks. Moreover, as shown in Figure 5b, the *Camellia oleifera* of different maturity stages in the overlapping regions (ripe and overripe) were obviously cracked, and the *Camellia oleifera* seeds were exposed. Therefore, we speculate that when grading the maturity of *Camellia oleifera*, the DeepCluster model not only considered the crack of the fruit, but also the appearance factors such as peel color and texture.
According to the above, the physical properties of *Camellia oleifera* fruit samples with three different maturity stages graded by DeepCluster are presented in Table 3. As the table gives information about the physical properties of *Camellia oleifera* fruit samples within three different maturity stages graded by DeepCluster. As the table illustrates, no significant differences were found in transverse and vertical diameter values with *Camellia oleifera* maturity stages. As other authors have observed, there were no significant differences in the transverse diameter and vertical diameter of the *Camellia oleifera* with different maturity stages [60,61]. For dry seed weight and moisture content, there were significant differences among the maturity stages. This conclusion was in agreement with the previous findings of Jiang et al. [31]. According to the above, the physical properties of *Camellia oleifera* fruit samples with different maturity stages automatically graded by DeepCluster were basically consistent with previous studies [62]. This result demonstrates the validity of the maturity grading method from the perspective of physical properties.

### Table 2. Physical properties of *Camellia oleifera* with different maturity stages.

| Physical Property         | Maturity Stage       |
|---------------------------|----------------------|
|                           | Unripe               | Ripe                  | Overripe              |
| Transverse diameter (mm)  | 46.05 ± 1.54 a       | 47.31 ± 1.65 a        | 49.32 ± 1.57 a        |
| Vertical diameter (mm)    | 38.78 ± 1.33 a       | 40.10 ± 1.95 a        | 40.64 ± 1.71 a        |
| Dry seed weight (g)       | 8.06 ± 0.41 c        | 12.44 ± 1.02 a        | 11.05 ± 0.38 b        |
| Moisture content (%)      | 69.55 ± 3.74 a       | 58.79 ± 2.13 b        | 48.15 ± 0.62 c        |

Values are the mean of three independent determinations ± standard deviation. Different letters in the same row indicate significant differences ($p < 0.05$).

The quality properties of *Camellia oleifera* fruit samples with three different maturity stages graded by DeepCluster are presented in Table 3. As the table gives information about quality properties, there is a remarkable difference between maturity stages and quality properties. The seed oil content is of particular significance for *Camellia oleifera*. As other authors have observed, the seed oil content of *Camellia oleifera* increased significantly and then decreased slightly with the maturation of the fruits [11]. The highest seed oil content was 46.05%, which was obtained from the ripe stage. The seed oil content decreased amounted to 1.63% from the ripe to the overripe stage. This conclusion is consistent with the findings of Jiang et al. [31]. In addition, the seed-soluble protein content gradually increased as the fruit matured, which was consistent with the results of previous studies [11]. Moreover, as the increases of fruit maturity stage, the seed-soluble sugar content and seed starch content gradually decrease [63]. These results indicate the validity of *Camellia oleifera* maturity as graded by DeepCluster in terms of quality properties.
Table 3. Quality properties of *Camellia oleifera* with different maturity stages.

| Quality Property          | Maturity Stage |          |          |          |
|---------------------------|----------------|----------|----------|----------|
|                           | Unripe         | Ripe     | Overripe |          |
| Seed oil content (%)      | 39.12 ± 0.99 c | 46.05 ± 0.49 a | 44.42 ± 0.53 b |          |
| Seed-soluble protein content (%) | 5.21 ± 0.22 c | 5.61 ± 0.16 b | 6.11 ± 0.07 a |          |
| Seed-soluble sugar content (%) | 25.69 ± 1.16 a | 19.28 ± 3.85 b | 13.75 ± 0.93 c |          |
| Seed starch content (%)   | 3.31 ± 0.55 a  | 2.30 ± 0.24 b | 1.54 ± 0.82 c |          |

Values are the mean of three independent determinations ± standard deviation. Different letters in the same row indicate significant differences ($p < 0.05$).

Based on the data in Tables 2 and 3, and the analysis above, it can be seen that the physical and quality properties of *Camellia oleifera* fruits at different maturity stages are basically consistent with previous findings. This conclusion verifies the effectiveness of the *Camellia oleifera* fruit maturity grading and identification method based on unsupervised learning. At the same time, it can be found that the seed oil content of *Camellia oleifera* fruits was optimal at the ripe stage. Therefore, the highest yield was likely obtained by harvesting *Camellia oleifera* fruits at ripe maturity.

3.4. Maturity Identification Performance Evaluation

An experiment was designed to quantitatively evaluate the maturity identification ability of the DeepCluster model. In total, 160 sample *Camellia oleifera* fruits (50 unripe, 60 ripe, and 50 overripe) in the testing set were used to perform the evaluation maturity identification. In the identification problem, a confusion matrix is often used as a way of evaluating the performance of the proposed method [64,65]. Thus, the confusion matrix was used in this study to quantity evaluate the maturity identification results of *Camellia oleifera*. Figure 6 indicates the maturity identification results of the model in the form of a confusion matrix, and the maturity identification results of some *Camellia oleifera* fruit example images are presented in Figure 7.

![Confusion Matrix](image)

**Figure 6.** Confusion matrix for identification *Camellia oleifera* fruit maturity in the testing set.
Figure 7. Maturity identification results of some *Camellia oleifera* fruit example images in testing set: (a,b) denotes maturity identification results of unripe *Camellia oleifera* fruits, (c,d) denotes maturity identification results of ripe *Camellia oleifera* fruits, (e,f) denotes maturity identification results of overripe *Camellia oleifera* fruits.

Generally, the lower the sum of the off-diagonal elements in the confusion matrix, the higher the accuracy of identification would be. As shown in the confusion matrix, the DeepCluster model trained with the *Camellia oleifera* images without labeled maturity can properly identify the maturity of *Camellia oleifera*. Of the 160 *Camellia oleifera* fruit samples with different maturity, 146 were correctly identified, while only 14 were mistakenly identified. Moreover, the misidentification mainly occurs between the adjacent maturity stages such as unripe and ripe, ripe, and overripe. For instance, of the 50 *Camellia oleifera* sample images with unripe maturity, 48 were correctly identified and 2 were mistakenly identified as ripe. Figure 7a,b show the maturity identification results of some unripe *Camellia oleifera* fruits in the testing set. For 60 ripe *Camellia oleifera* sample images, the proposed method correctly identified 52 of the *Camellia oleifera* sample images, and the remaining 8 *Camellia oleifera* sample images were identified as unripe (3) and overripe (5). Figure 7c,d show the maturity identification results of some ripe *Camellia oleifera* fruits in the testing set. Moreover, the *Camellia oleifera* fruit examples depicted in Figure 7c,d are the most challenging examples for the model identification of ripe *Camellia oleifera* fruits. Actually, the *Camellia oleifera* fruit shown in Figure 7c has features similar to unripe,
while the *Camellia oleifera* fruit in Figure 7d was close to overripe. In addition, of the 50 *Camellia oleifera* sample images with overripe maturity, 46 were correctly identified and 4 were mistakenly identified as ripe. Figure 7e,f show the maturity identification results of some overripe *Camellia oleifera* fruits in the testing set. We conjectured that these misidentifications were caused because the phenotypic properties of adjacent maturity *Camellia oleifera* visually appear more similar to each other.

In this study, the following four evaluation indices were employed to quantitatively evaluate the performance of the proposed *Camellia oleifera* maturity identification method: precision, recall, F1 score, and overall accuracy. Table 4 summarizes the performance evaluation results of *Camellia oleifera* maturity identification by implementing the proposed method. Accordingly, the proposed method could identify *Camellia oleifera* fruit sample images into three maturity stages with an overall accuracy of 91.25%. For each maturity stage identification of *Camellia oleifera*, different results were obtained by the proposed method. For example, in the proposed method, the unripe group obtained the best identification with 94.12% precision, 96.00% recall, and 95.05% F1 score. The overripe group obtained the suboptimal identification and achieved precision, recall, and F1 score of 90.20%, 92.00%, and 91.09%, respectively. However, the ripe group reached the value of precision, recall, and the F1 score, which was lower than both other maturity stages. This was due to the fact that the ripe group of *Camellia oleifera* can be confused with both the unripe and overripe groups of *Camellia oleifera*. Based on the *Camellia oleifera* fruit maturity identification results of Figure 7, Table 4, and the analysis above, it can be seen that the proposed method can be used to effectively identify the maturity of *Camellia oleifera* fruits.

Table 4. Quantitative evaluation results of the identification of *Camellia oleifera* fruit maturity.

| Maturity | Prec (%) | Rec (%) | F1score (%) | OA (%) |
|----------|----------|---------|-------------|--------|
| Unripe   | 94.12    | 96.00   | 95.05       |        |
| Ripe     | 89.66    | 86.67   | 88.14       | 91.25  |
| Overripe | 90.20    | 92.00   | 91.09       |        |

OA (%) represents the overall maturity identification accuracy for all three maturity stages.

### 3.5. Maturity Identification Visualization Analysis

According to the heat maps of some *Camellia oleifera* fruit example images that were generated by the Grad-CAM method (Figure 8), the model was well-trained because it identifies similar regions within the same maturity stages and different regions across the maturity stages, which may improve the accuracy of the *Camellia oleifera* maturity identification model. Interestingly, the regions of focus of the model were similar within the maturity of *Camellia oleifera* but different across maturities. For instance, for unripe *Camellia oleifera*, the model focused on the peel regions (Figure 8a,b). However, for ripe *Camellia oleifera*, the model focused on the crack regions and no longer on the peel regions (Figure 8c,d). Moreover, for overripe *Camellia oleifera*, the model focuses more attention on the camellia seed regions (Figure 8e,f). These behaviors are seemingly reminiscent of how human experts would estimate *Camellia oleifera* maturity, either focusing on the crack and camellia seeds regions if the peel is dehiscent or observing the peel regions when the *Camellia oleifera* shows no dehiscent.
Figure 8. Heat maps of some *Camellia oleifera* fruit example images: (a,b) denotes example heat maps of unripe *Camellia oleifera* fruits, (c,d) denotes example heat maps of ripe *Camellia oleifera* fruits, (e,f) denotes example heat maps of overripe *Camellia oleifera* fruits. Here, warm colors indicate that the region has a strong contribution to the maturity prediction.

However, our maturity identification model also misidentified the maturity of some *Camellia oleifera* fruits. Figure 9 presents some example heat maps that misidentified the maturity stages of *Camellia oleifera*. The Grad-CAM analysis revealed that the incorrect identification of *Camellia oleifera* maturity was caused by failing to locate the typical region of focus for maturity identification. For example, when identifying unripe *Camellia oleifera*, the model incorrectly focused on the non-fruit regions, which resulted in unripe *Camellia oleifera* being mistakenly identified as ripe (Figure 9a). Besides, the model failed to accurately focus on the crack region when identifying the ripe *Camellia oleifera*, resulting in ripe *Camellia oleifera* mistakenly identified as unripe (Figure 9b). Thus, training with a higher number of *Camellia oleifera* images and adopting an attention mechanism may be another suitable strategy to improve the effectiveness of *Camellia oleifera* maturity identification [66]. In addition, the *Camellia oleifera* in the background also influences the ultimate results of *Camellia oleifera* maturity identification (Figure 9c). Therefore, although the model may potentially be capable of identifying *Camellia oleifera* maturities, further model improvements will be needed for practical applications to adapt to the complex environment.
4. Conclusions

In conclusion, to acquire high-quality and high-yield Camellia oil, it is indispensable to grade and identification the maturity of *Camellia oleifera* fruits. This study developed a *Camellia oleifera* fruit maturity grading and identification method based on unsupervised image clustering techniques. Measurements of overall accuracy indicated that the proposed method successfully achieved 91.25%. Moreover, Grad-CAM visualization analysis suggests that peel regions, crack regions, and seed regions were the key positions for *Camellia oleifera* fruit maturity identification. Besides, the *Camellia oleifera* fruits with ripe maturity stages were the most suitable for harvesting. The misidentification occurred due to the phenotypic properties of adjacent maturity *Camellia oleifera* visually appearing more similar to each other. In addition, the complex environment also had an impact on the maturity identification of *Camellia oleifera*. The limitations of this study will be changed by acquiring more *Camellia oleifera* fruit images and suppressing the complex environment. Moreover, the proposed *Camellia oleifera* fruit maturity grading and identification methods are still wide open for further enhancement to simultaneously identify the maturity of multiple *Camellia oleifera* fruits in the image. The current findings contribute towards refining the harvesting process of *Camellia oleifera* fruits and facilitate the *Camellia oleifera* industry to improve quality and efficiency.

**Figure 9.** Some example heat maps which identified the maturity stages of *Camellia oleifera* by the proposed model incorrectly: (a) unripe identified as ripe, (b) ripe identified as unripe, and (c) overripe identified as ripe.
Author Contributions: Conceptualization, X.Z.; methodology, X.Z.; software, X.Z.; validation, X.Z. and D.S.; formal analysis, X.Z. and R.W.; investigation, D.S. and R.W.; data curation, X.Z. and S.S.; writing—original draft preparation, X.Z.; writing—review and editing, F.C. and Y.Z.; visualization, X.Z. and D.S.; 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 (2019YFD1002401) and the Fundamental Research Funds for the Central Universities (2021ZY74).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

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

References

1. Azadnia, R.; Kheirali-pour, K. Evaluation of hawthorns maturity level by developing an automated machine learning-based algorithm. Ecol. Inform. 2022, 71, 101804. [CrossRef]
2. Wang, M.; Wan, Y.; Liu, T.; Zeng, X.; Liang, X.; Wu, X.; Fu, G. Effect of refining degree on the quality changes and lipid oxidation of Camellia (Camellia oleifera) oil during heating. Foods 2022, 11, 2232. [CrossRef] [PubMed]
3. Cheng, X.; Yang, T.; Wang, Y.; Zhou, B.; Yan, L.; Teng, L.; Wang, F.; Chen, L.; He, Y.; Guo, K.; et al. New method for effective identification of adulterated Camellia oil basing on Camellia oleifera-specific DNA. Arab. J. Chem. 2018, 11, 815–826. [CrossRef]
4. Quan, W.; Wang, A.; Gao, C.; Li, C. Applications of Chinese Camellia oleifera and its By-Products: A Review. Front. Chem. 2022, 10, 921246. [CrossRef] [PubMed]
5. Lei, X.; Liu, Q.; Liu, Q.; Cao, Z.; Zhang, J.; Kuang, T.; Fang, Y.; Liu, G.; Qian, K.; Fu, J.; et al. Camellia oil (Camellia oleifera Abel.) attenuates CCl4-induced liver fibrosis via suppressing hepatocyte apoptosis in mice. Food Funct. 2020, 11, 4582–4590. [CrossRef] [PubMed]
6. Weng, M.H.; Chen, S.Y.; Li, Z.Y.; Yen, G.C. Camellia oil alleviates the progression of Alzheimer’s disease in aluminum chloride-treated rats. Free. Radic. Biol. Med. 2020, 152, 411–421. [CrossRef] [PubMed]
7. Yang, C.; Liu, X.; Chen, Z.; Lin, Y.; Wang, S. Comparison of oil content and fatty acid profile of ten new Camellia oleifera cultivars. J. Agric. Food Chem. 2016, 64, 62–650. [CrossRef] [PubMed]
8. Cheng, Y.T.; Wu, S.L.; Ho, C.Y.; Huang, S.M.; Cheng, C.L.; Yen, G.C. Beneficial effects of Camellia Oil (Camellia oleifera Abel.) on ketoprofen-induced gastrointestinal mucosal damage through upregulation of HO-1 and VEGF. J. Agric. Food Chem. 2014, 62, 624–630. [CrossRef]
9. Zhang, F.; Zhu, F.; Chen, R.; Su, E.; Chen, Y.; Cao, F. Composition, bioactive substances, extraction technologies and the influences on characteristics of Camellia oleifera oil: A review. Food Res. Int. 2022, 156, 111159. [CrossRef] [PubMed]
10. Wu, Y.; Ma, L.; Zhang, L.; Zhang, Y.; Zhou, H.; Wang, Y.; Liu, Y. Photosynthetic carbon and nitrogen metabolism of Camellia oleifera Abel during acclimation to low light conditions. J. Plant Physiol. 2022, 153814. [CrossRef]
11. Ye, T.; Liu, X.; Liang, X.; Su, S. Comparison of fruit quality differences between two kinds of Camellia oleifera at different maturity. China Oils Fats 2022, 12, 2158. Available online: https://kns.cnki.net/kcms/detail/61.1099.TS.20220518.1932.022.html (accessed on 19 May 2022).
12. Zhu, G.; Liu, H.; Xie, Y.; Liao, Q.; Lin, Y.; Liu, Y.; Liu, Y.; Xiao, H.; Gao, Z.; Hu, S. Postharvest processing and storage methods for Camellia oleifera seeds. Food Rev. Int. 2019, 36, 319–339. [CrossRef]
13. Zhang, F.; Li, Z.; Zhou, J.; Gu, Y.; Tan, X. Comparative study on fruit development and oil synthesis in two cultivars of Camellia oleifera. BMC Plant Biol. 2021, 21, 348. [CrossRef] [PubMed]
14. Gong, W.; Song, Q.; Ji, K.; Gong, S.; Tan, L.; Chen, L.; Zhang, J.; Yuan, D. Full-Length transcriptome from Camellia oleifera seed provides insight into the transcript variants involved in oil biosynthesis. J. Agric. Food Chem. 2020, 68, 14670–14683. [CrossRef] [PubMed]
15. Wang, J.; Gao, X.; Liu, M.; Li, J.; Yang, H.; Wang, Z.; Yang, L. Camellia oil trait and DIA quantitative proteomics analyses reveal the impact of harvesting time on the oil content and quality of the late-maturing cultivar C. oleifera ‘Huaxin’. Food Qual. Saf. 2022, 6, 54ac055. [CrossRef]
16. 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]
17. Iler, A.M.; CaraDonna, P.J.; Forrest, J.R.K.; Post, E. Demographic consequences of phenological shifts in response to climate change. Annu. Rev. Ecol. Evol. Syst. 2021, 52, 221–245. [CrossRef]
18. Xie, H.; Chen, F.; Yin, H.; Peng, G.; You, C.; Qin, P.; Jiang, S.; Guo, X. Characterization and comparison of lipids in Camellia oleifera kernels of XL210 and XL1 based on LC-MS/MS. Reprod. Breed. 2021, 1, 193–203. [CrossRef]
19. Yang, J.; Chen, B.; Manan, S.; Li, P.; Liu, C.; She, G.; Zhao, S.; Zhao, J. Critical metabolic pathways and SAD/FADs, WR1s, and DGATs cooperate for high-oleic acid oil production in developing oil tea (Camellia oleifera) seeds. *Hortic. Res.* **2022**, *9*, 109887. [CrossRef]

20. Yuan, J.; Wang, C.; Chen, H.; Zhou, H.; Ye, J. Prediction of fatty acid composition in *Camellia oleifera* oil by near infrared transmittance spectroscopy (NITS). *Food Chem.* **2013**, *138*, 1657–1662. [CrossRef]

21. Song, Q.; Ji, K.; Mo, W.; Wang, L.; Chen, L.; Gao, L.; Gong, W.; Yuan, D. Dynamics of sugars, endogenous hormones, and oil content during the development of *Camellia oleifera* fruit. *Botany* **2021**, *99*, 515–529. [CrossRef]

22. Liu, X.X.; Luo, X.F.; Luo, K.X.; Liu, Y.L.; Fan, T.; Li, Z.Z.; Duns, G.J.; He, F.L.; Qin, Z.D. Small RNA sequencing reveals dynamic microRNA expression of important nutrient compounds during development of *Camellia oleifera* fruit. *Int. J. Biol. Sci.* **2019**, *15*, 416–429. [CrossRef]

23. Wei, T.; Dong, L.; Zhong, S.; Jing, H.; Deng, Z.; Wen, Q.; Li, J. Chemical composition of *Camellia chekiangoleosa* Hu seeds during ripening and evaluations of seed oils quality. *Ind. Crops Prod.* **2022**, *177*, 114499. [CrossRef]

24. Sinambela, R.; Mandang, T.; Subrata, I.D.M.; Hermawan, W. Application of an inductive sensor system for identifying ripeness and forecasting harvest time of oil palm. *Sci. Hortic.* **2020**, *265*, 109231. [CrossRef]

25. Misron, N.; Alithe, N.A.; Harun, N.H.; Tashiro, K.; Sato, T.; Wakiwaka, H. Relative estimation of water content for flatlay inductive-based oil palm fruit maturity sensor. *Sensors* **2016**, *17*, 52. [CrossRef]

26. Saeed, O.M.B.; Sankaran, S.; Shariff, A.R.M.; Shahri, H.Z.M.; Ehsani, R.; Alfatni, M.S.; Hazir, M.H.M. Classification of oil palm fresh fruit bunches based on their maturity using portable four-band sensor system. *Comput. Electron. Agric.* **2012**, *82*, 55–60. [CrossRef]

27. Alamprese, C.; Grassi, S.; Tognolo, A.; Casiraghi, E. Prediction of olive ripening degree combining image analysis and FT-NIR spectroscopy for virgin olive oil optimisation. *Food Control* **2021**, *123*, 107755. [CrossRef]

28. Greco, M.; Spadafora, N.; Shine, M.; Smith, A.; Muto, A.; Muzzalupo, I.; Chiappetta, A.; Bruno, L.; Muller, C.; Rogers, H.; et al. Identifying volatile and non-volatile organic compounds to discriminate cultivar, growth location, and stage of ripening in olive fruits and oils. *J. Sci. Food Agric.* **2022**, *102*, 4500–4513. [CrossRef]

29. Tognolo, A.; Giovenzana, V.; Beghi, R.; Grassi, S.; Alamprese, C.; Casson, A.; Casiraghi, E.; Guidetti, R. A diagnostic visible/near infrared tool for a fully automated olive ripeness evaluation in a view of a simplified optical system. *Comput. Electron. Agric.* **2021**, *180*, 105887. [CrossRef]

30. Raj, T.; Hashim, F.H.; Huddin, A.B.; Hussain, A.; Ibrahim, M.F.; Abdul, P.M. Classification of oil palm fresh fruit maturity based on carotene content from Raman spectra. *Sci. Rep.* **2021**, *11*, 18315. [CrossRef]

31. Jiang, H.; Hu, Y.; Jiang, X.; Zhou, H. Maturity Stage Discrimination of *Camellia oleifera* fruit using visible and near-infrared hyperspectral imaging. *Molecules* **2022**, *27*, 6318. [CrossRef]

32. Behera, S.K.; Rath, A.K.; Sethy, P.K. Maturity status classification of papaya fruits based on machine learning and transfer learning approach. *Inf. Process. Agric.* **2021**, *8*, 244–250. [CrossRef]

33. Begum, N.; Hazarika, M.K. Maturity detection of tomatoes using transfer learning. *Meas. Food Food* **2022**, *7*, 100308. [CrossRef]

34. Alfatni, M.S.M.; Khairunniza-Bejo, S.; Marhaban, M.H.B.; Saeed, O.M.B.; Mustapha, A.; Shariff, A.R.M. Towards a real-time oil palm fruit maturity system using supervised classifiers based on feature analysis. *Agriculture* **2022**, *12*, 1461. [CrossRef]

35. Khan, N.; Kamaruddin, M.A.; Sheikh, U.U.; Yusup, Y.; Bakht, M.P. Oil palm and machine learning: Reviewing one decade of research. *Ind. Crops Prod.* **2018**, *125*, 515–529. [CrossRef]

36. Tan, X.J.; Cheor, W.L.; Yeo, K.S.; Leow, W.Z. Expert systems in oil palm precision agriculture: A decade systematic review. *J. King Saud Univ.-Comput.* **2021**, *33*, 1569–1594. [CrossRef]

37. Caladacid, J.A.; Cabahug, S.; Catamco, M.R.; Villacencio, P.E.; Cosgafa, L.; Cabizares, K.N.; Hermosilla, M.; Piedad, E., Jr. Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal. *Comput. Electron. Agric.* **2020**, *179*, 105858. [CrossRef]

38. Sola-Guirado, R.R.; Bayano-Tejero, S.; Aragón-Rodríguez, F.; Bernardi, B.; Benalia, S.; Castro-García, S. A smart system for the automatic evaluation of green olives visual quality in the field. *Comput. Electron. Agric.* **2020**, *179*, 105858. [CrossRef]

39. Khosravi, H.; Saedi, S.I.; Rezaei, M. Real-time recognition of on-branch olive ripening stages by a deep convolutional neural network. *Sensors* **2021**, *21*, 832. [CrossRef]

40. Parvathi, S.; Tamil Selvi, S. Detection of maturity stages of coconuts in complex background using Faster R-CNN model. *Biosyst. Eng.* **2021**, *202*, 119–132. [CrossRef]

41. Mubin, N.A.; Nadarajoo, E.; Sharifi, H.Z.M.; Hamedifar, A. Young and mature oil palm tree detection and counting using convolutional neural network deep learning method. *Int. J. Remote Sens.* **2019**, *40*, 7500–7515. [CrossRef]

42. Wu, B.; Ruan, C.; Han, P.; Ruan, D.; Xiong, C.; Ding, J.; Liu, S. Comparative transcriptomic analysis of high- and low-oil *Camellia oleifera* reveals a coordinated mechanism for the regulation of upstream and downstream multigenes for high oleic acid accumulation. *3 Biotech* **2019**, *9*, 257. [CrossRef]

43. Niu, C.; Zhang, J.; Wang, G.; Liang, J. GATCluster: Self-supervised gaussian-attention network for image clustering. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 735–751. [CrossRef]

44. Mehdi, N.; Paolo, F. Unsupervised learning of visual representations by solving jigsaw puzzles. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; pp. 69–84. [CrossRef]

45. Li, Y.; Xue, J.; Wang, K.; Zhang, M.; Li, Z. Surface defect detection of fresh-cut cauliflowers based on convolutional neural network with transfer learning. *Foods* **2022**, *11*, 2915. [CrossRef]
46. Zhang, X.; Zhu, Y.; Su, Y.; Xie, B.; Gu, Q.; Zheng, K. Quantitative extraction and analysis of pear fruit spot phenotypes based on image recognition. *Comput. Electron. Agric.* 2021, 190, 106474. [CrossRef]

47. Al-Dakheel, A.J.; Hussain, M.I.; Abdulrahman, A.; Abdullah, A. Long term assessment of salinity impact on fruit yield in eighteen date palm varieties. *Agric. Water Manag.* 2022, 269, 107683. [CrossRef]

48. Hu, Y.; Wu, Y.; Sun, J.; Geng, J.; Fan, R.; Kang, Z. Distinguishing different varieties of oolong tea by fluorescence hyperspectral technology combined with chemometrics. *Foods* 2022, 11, 2344. [CrossRef]

49. Aytekin, C.; Ni, X.; Cricri, F.; Aksu, E. Clustering and Unsupervised Anomaly Detection with L2 Normalized Deep Auto-Encoder Representations. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–6. [CrossRef]

50. Jegou, H.; Perronnin, F.; Douze, M.; Sanchez, J.; Perez, P.; Schmid, C. Aggregating local image descriptors into compact codes. *IEEE Trans. Pattern Anal. Mach. Intell.* 2012, 34, 1704–1716. [CrossRef]

51. Wang, C.; Xu, R.; Xu, S.; Meng, W.; Zhang, X. CNDesc: Cross normalization for local descriptors learning. *IEEE Trans. Multimed.* 2022, in press. [CrossRef]

52. Jain, M.; Kaur, G.; Saxena, V. A K-Means clustering and SVM based hybrid concept drift detection technique for network anomaly detection. *Expert Syst. Appl.* 2022, 193, 116510. [CrossRef]

53. Zhang, H.; Peng, Q. PSO and K-means-based semantic segmentation toward agricultural products. *Future Gener. Comput. Syst.* 2022, 126, 82–87. [CrossRef]

54. Fredes, C.; Valero, C.; Diezma, B.; Mora, M.; Naranjo-Torres, J.; Wilson, M.; Delgadillo, G. A model based on clusters of similar color and NIR to estimate oil content of single olives. *Foods* 2021, 10, 609. [CrossRef]

55. Dhindsa, R.S.; Plumb-Dhindsa, P.; Thorpe, T.A. Leaf senescence: Correlation with increased levels of membrane permeability and lipid peroxidation and increased levels of superoxide dismutase and catalase. *J. Exp. Bot.* 1981, 32, 93–101. [CrossRef]

56. Septiarini, A.; Sunyoto, A.; Hamdani, H.; Kasim, A.-A.; Utaminingrum, F.; Hatta, H.-R. Machine vision for the maturity classification of oil palm fresh fruit bunches based on color and texture features. *Sci. Hortic.* 2021, 286, 110245. [CrossRef]

57. Osako, Y.; Yamane, H.; Lin, S.-Y.; Chen, P.-A.; Tao, R. Cultivar discrimination of litchi fruit images using deep learning. *Sci. Hortic.* 2020, 269, 109360. [CrossRef]

58. Zhu, X.; Zhang, X.; Sun, Z.; Zheng, Y.; Su, S.; Chen, F. Identification of oil tea (*Camellia oleifera* C.Abel) cultivars using EfficientNet-B4 CNN model with attention mechanism. *Foods* 2021, 10, 609. [CrossRef]

59. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Int. J. Comput. Vis.* 2019, 128, 336–359. [CrossRef]

60. Hu, Y.; Jiang, H.; Zhou, H.; Xu, L.; Ju, H.; Wang, Y. Maturity detection of *Camellia* oleifera by hyperspectral imaging. *Food Sci.* 2022, 43, 324–331. [CrossRef]

61. Xu, H.; Zeng, H.; Liu, Z.; Liu, L. Changes of phenotypic characters and oil content of *Camellia* oleifera during fruit development. *China Oils Fats* 2021, 46, 102–107. [CrossRef]

62. Hongzhe, J.; Wei, W.; Bo, W.; Xuan, C.; Xin, Z. Design and experimental research of shelling and sorting machine for fresh *Camellia* oleifera fruit based on physical properties. In Proceedings of the 2018 ASABE Annual International Meeting, St. Joseph, MI, USA, 19 July–1 August 2018; p. 1800815. [CrossRef]

63. Ma, L.; Zhong, H.; Chen, Y.; Peng, S.; Zhu, N.; Chen, L.; Wang, R. The effect of post—Harvest treatment of Camellia Fruit on quality of Camellia Seed. *J. Chin. Cereals Oils Assoc.* 2014, 12, 73–76. Available online: http://www.cnki.net/kcms/detail/11.2864.TS.20141210.1100.001.html (accessed on 10 December 2014).

64. Yue, X.-Q.; Shang, Z.-Y.; Yang, J.-Y.; Huang, L.; Wang, Y.-Q. A smart data-driven rapid method to recognize the strawberry maturity. *Inf. Process. Agric.* 2020, 7, 575–584. [CrossRef]

65. Mazen, F.M.A.; Nashat, A.A. Ripeness classification of bananas using an artificial neural network. *Arab. J. Sci. Eng.* 2019, 44, 6901–6910. [CrossRef]

66. Zhu, X.; Yu, Y.; Zheng, Y.; Su, S.; Chen, F. Bilinear attention network for image-based fine-grained recognition of oil tea (*Camellia oleifera* Abel.) cultivars. *Agronomy* 2022, 12, 1846. [CrossRef]