




Article

GrapeNet: A Lightweight Convolutional Neural Network Model for Identification of Grape Leaf Diseases

Jianwu Lin ¹, Xiaoyulong Chen ², Renyong Pan ¹, Tengbao Cao ¹, Jitong Cai ¹, Yang Chen ¹, Xishun Peng ¹ , Tomislav Cernava ³  and Xin Zhang ^{1,*} 

¹ College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China; ljw971121@163.com (J.L.); panry198@163.com (R.P.); 15283673634@163.com (T.C.); gs.jtcai21@gzu.edu.cn (J.C.); cy52cv2022@163.com (Y.C.); pxs19970921@163.com (X.P.)

² College of Tobacco Science, Guizhou University, Guiyang 550025, China; chenxiaoyulong@sina.cn

³ Institute of Environmental Biotechnology, Graz University of Technology, 8010 Graz, Austria; tomlav.cernava@tugraz.at

* Correspondence: xzhang1@gzu.edu.cn

Abstract: Most convolutional neural network (CNN) models have various difficulties in identifying crop diseases owing to morphological and physiological changes in crop tissues, and cells. Furthermore, a single crop disease can show different symptoms. Usually, the differences in symptoms between early crop disease and late crop disease stages include the area of disease and color of disease. This also poses additional difficulties for CNN models. Here, we propose a lightweight CNN model called GrapeNet for the identification of different symptom stages for specific grape diseases. The main components of GrapeNet are residual blocks, residual feature fusion blocks (RFFBs), and convolution block attention modules. The residual blocks are used to deepen the network depth and extract rich features. To alleviate the CNN performance degradation associated with a large number of hidden layers, we designed an RFFB module based on the residual block. It fuses the average pooled feature map before the residual block input and the high-dimensional feature maps after the residual block output by a concatenation operation, thereby achieving feature fusion at different depths. In addition, the convolutional block attention module (CBAM) is introduced after each RFFB module to extract valid disease information. The obtained results show that the identification accuracy was determined as 82.99%, 84.01%, 82.74%, 84.77%, 80.96%, 82.74%, 80.96%, 83.76%, and 86.29% for GoogLeNet, Vgg16, ResNet34, DenseNet121, MobileNetV2, MobileNetV3_large, ShuffleNetV2_×1.0, EfficientNetV2_s, and GrapeNet. The GrapeNet model achieved the best classification performance when compared with other classical models. The total number of parameters of the GrapeNet model only included 2.15 million. Compared with DenseNet121, which has the highest accuracy among classical network models, the number of parameters of GrapeNet was reduced by 4.81 million, thereby reducing the training time of GrapeNet by about two times compared with that of DenseNet121. Moreover, the visualization results of Grad-cam indicate that the introduction of CBAM can emphasize disease information and suppress irrelevant information. The overall results suggest that the GrapeNet model is useful for the automatic identification of grape leaf diseases.

Keywords: convolutional neural network; residual block; attention mechanism; grape leaf disease



Citation: Lin, J.; Chen, X.; Pan, R.; Cao, T.; Cai, J.; Chen, Y.; Peng, X.; Cernava, T.; Zhang, X. GrapeNet: A Lightweight Convolutional Neural Network Model for Identification of Grape Leaf Diseases. *Agriculture* **2022**, *12*, 887. <https://doi.org/10.3390/agriculture12060887>

Academic Editor: Dimitre Dimitrov

Received: 10 May 2022

Accepted: 18 June 2022

Published: 20 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Grapes are one of the most popular fruits in the world and also the main raw material for the production of wine, thus the yield and quality of grapes are of substantial economic value [1]. However, grape leaves are susceptible to various diseases that are influenced by the weather as well as the environment, and mainly caused by fungi, viruses, and bacteria. If the diseased leaves of grapes are not effectively controlled, the disease spreads to the whole plant, thereby affecting the quality and yield of grapes. In the early days, grape

leaf disease identification was mainly conducted by means of classic phytopathology [2]; however, manual identification is time-consuming and labor-intensive. The growing area of land used for grape production makes manual identification methods unreliable. Therefore, automatic identification of grape leaf disease is of great significance for the future development of grape production [3].

With the rapid development of computer technology, a new visual recognition method based on machine learning [4], has been employed for disease recognition. Using machine learning methods to identify crop diseases generally involves three steps: spot segmentation, feature extraction, and classifier recognition [5]. Majumdar et al. extracted wheat disease characteristics and used artificial neural networks (ANNs) to classify diseases, achieving an accuracy of 85% [6]. Guru et al. presented a novel algorithm for extracting lesion areas and applying the probabilistic neural network (PNN) to classify seedling diseases such as anthracnose and frog-eye spots on tobacco leaves, achieving an accuracy of 88.59% [7]. Rumpf et al. accomplished early disease identification of sugar beets using support vector machine (SVM) and hyperspectral techniques [8]. Their experiments showed an accuracy of 97% for healthy and diseased leaves of sugar beets. Moreover, Padol et al. used the SVM classification technique to detect and classify grape leaf diseases [9]. First, the diseased region is identified using segmentation by K-means clustering, and then useful features are extracted. Finally, SVM classification is used to classify the categories of grape leaf diseases, achieving an accuracy of 88.89%. The abovementioned results indicate that it is feasible to use machine learning to identify crop diseases. However, its cumbersome steps lead to low recognition efficiency, and the artificially extracted features are subject to a certain degree of subjectivity, resulting in low recognition accuracy.

Convolutional neural network (CNN) models have been widely used in various application fields, such as facial recognition [10] and license plate detection [11]. The CNN models use sliding window extraction to automatically extract image features and then use fully connected layers for classification to implement an end-to-end disease detection model. Recently, CNN models were used to detect and identify crop diseases instead of traditional machine learning methods [12]. Liu et al. proposed a novel recognition approach based on an improved CNN model for the diagnosis of grape leaf diseases [13]. In this approach, a dense connectivity strategy was introduced to encourage feature reuse and strengthen feature propagation. Finally, a new CNN model named DICNN was built and trained from scratch and achieved an accuracy of 97.22%. Tang et al. proposed a novel method based on a lightweight CNN applying the channel-wise attention mechanism. ShuffleNetV1 and ShuffleNetV2 were chosen as the backbones [14]. The results showed that the proposed model achieved a best trained accuracy of 99.14%, and the model size was only 4.2 MB. Mohanty et al. used GoogLeNet to identify plant disease images from PlantVillage. After GoogLeNet was trained with two methods of training from scratch and isomorphic transfer learning, the accuracy rates were 98.36% and 99.35%, respectively [15]. Pandian et al. proposed a CNN model for image-based plant leaf disease identification using data augmentation and hyperparameter optimization techniques [16]. The results show that the model achieved an accuracy of 98.41% and illustrate the importance of data augmentation techniques and hyperparameter optimization techniques. Chan et al. proposed an early diagnosis method for apple tree leaf diseases based on a deep CNN [17]. The CNN combines DenseNet and Xception, using global average pooling to replace fully connected layers. It achieved an overall accuracy of 98.82% in identifying apple tree leaf diseases. Gao et al. proposed a dual-branch, efficient, channel attention (DECA)-based crop disease recognition model, and the recognition accuracy of the model was 86.65%, 99.74%, and 98.54% on the datasets of PlantVillage, AI Challenger 2018, and Cucumber disease, respectively [18]. Chen et al. introduced the Location-wise Soft Attention mechanism to the pre-trained MobileNetV2 [19]. Furthermore, a two-phase progressive strategy was executed for model training. The experimental results showed that the average accuracy of the model was 99.71% on the open-source dataset. Zeng et al. proposed a lightweight dense-scale network (LDSNet) for corn leaf disease identification under field conditions [20]. The

accuracy of the optimized model on the test data was 95.4%. Kamal et al. proposed a novel deep-separable convolution block, through which MobileNet was able to construct just a few parameters with an accuracy of 98.34% on the PlantVillage dataset [21].

The findings of the abovementioned studies confirm that CNN models have advantages in crop disease identification. However, the objectives of these studies were based on the classification of different disease categories. The classification of different stages of specific diseases was so far neglected. However, accurate identification of different symptom stages of a distinct disease has potential value in modern agriculture. The objective of this study was to examine three defined grape leaf diseases. To that end, the symptoms of the same grape disease were divided into two stages (general symptoms and severe symptoms). We present a well-designed CNN model to provide a novel method for the identification of grape leaf diseases. The main objections of this study were as follows:

- (1) Proposing a lightweight CNN model, named GrapeNet, based on residual feature fusion block (RFFB) modules and convolutional block attention modules (CBAMs) [22], for the identification of different symptom stages for specific grape diseases.
- (2) Implementing ablation experiments and visualization of results of the model to verify the effectiveness of the RFFB modules and the CBAM modules, respectively.
- (3) Comparing GrapeNet with other classical network models to verify the performance advantages of GrapeNet.

2. Materials and Methods

2.1. Image Acquisition

In this study, we obtained seven types of grape leaves in the AI challenger 2018 dataset, for a total of 2850 grape leaf images, including 2456 in the training set and 394 in the test set. Representative images are shown in Figure 1. As the same disease is divided into general and serious symptoms, the inter-class variance in the dataset is small. Therefore, it is challenging for a CNN model to identify the disease accurately.

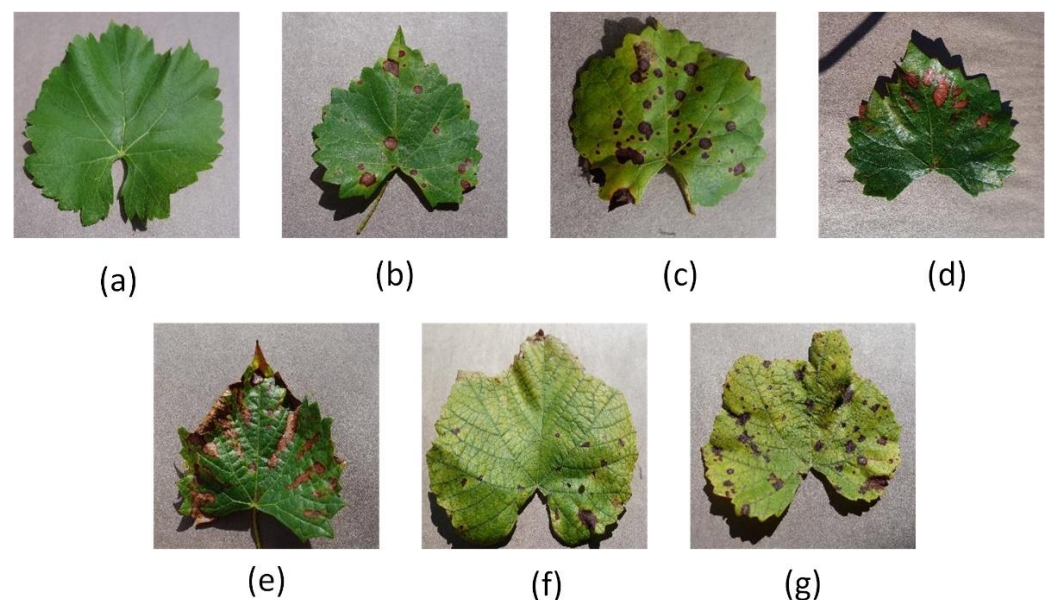


Figure 1. Examples of grape leaves. (a) Grape healthy leaf (GH). (b) Grape black rot fungus with general symptoms (BRF_G). (c) Grape black rot fungus with serious symptoms (BRF_S). (d) Grape black measles fungus with general symptoms (BMF_G). (e) Grape black measles fungus with serious symptoms (BMF_S). (f) Grape leaf blight fungus with general symptoms (LBF_G). (g) Grape leaf blight fungus with serious symptoms (LBF_S).

2.2. Image Preprocessing

The number of grape disease samples was limited, and the number of samples of different categories was not evenly distributed. To reduce overfitting during model training and enhance the generalization ability of the model, the dataset had to be expanded. The following operations were carried out. First, we redivided the training set into a training set and validation set in the ratio of 9:1 and performed data augmentation on the new training set through some operations such as rotation, color enhancement, contrast enhancement, and Gaussian noise. Some of the expanded images are shown in Figure 2. Thereafter, the validation set and test set did not need to be expanded. The validation set was used to verify whether the model training fits, and the test set was used to test the performance of the model. The sample distribution before and after augmentation is shown in Table 1.

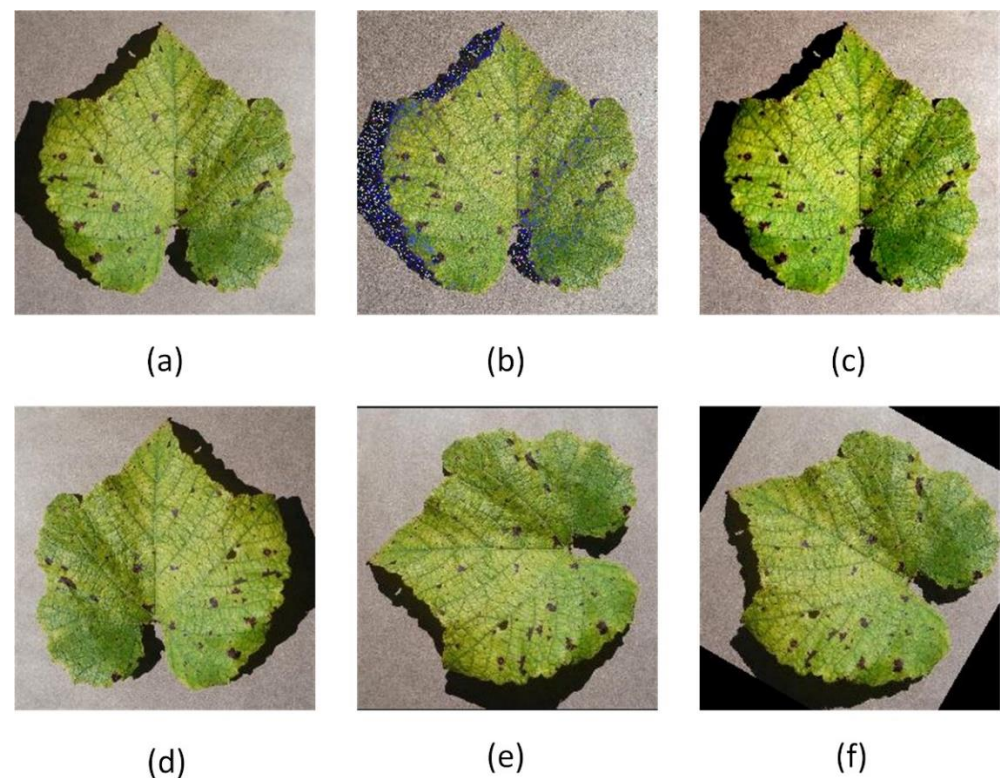


Figure 2. Some of the expanded images. (a) Original image. (b) Image expanded by Gaussian noise. (c) Image expanded by contrast enhancement. (d) Image expanded by horizontal flip. (e) Image expanded by a rotation of 90 degrees counter-clockwise. (f) Image expanded by a rotation of 60 degrees counter-clockwise.

Table 1. Sample distribution before and after augmentation. As the number of gaps between samples was too large in the training set, we iteratively augmented the dataset with fewer samples to ensure a comparable number of samples per class.

Class	Sample Distribution before Augmentation			Sample Distribution after Augmentation		
	Training set	Validation set	Test set	Training set	Validation set	Test set
Grape healthy leaf (GH)	265	29	42	2650	29	42
Grape black rot fungus with general symptoms (BRF_G)	343	38	54	2058	38	54
Grape black rot fungus with serious symptoms (BRF_S)	416	46	66	2496	46	66
Grape black measles fungus with general symptoms (BMF_G)	453	50	74	2718	50	74

Table 1. Cont.

	Sample Distribution before Augmentation			Sample Distribution after Augmentation		
Grape black measles fungus with serious symptoms (BMF_S)	378	41	59	2268	41	59
Grape leaf blight fungus with general symptoms (LBF_G)	55	6	9	1980	6	9
Grape leaf blight fungus with serious symptoms (LBF_S)	567	63	90	3402	63	90
Total	2477	273	394	17,572	273	394

2.3. GrapeNet Model Framework

In this study, we propose a lightweight CNN model named GrapeNet for grape leaf disease recognition. GrapeNet extracts rich grape leaf disease features by using the RFFB modules while introducing attention mechanisms to focus on useful disease features and enhance the ability to identify grape leaf diseases. The network structure of the GrapeNet model is shown in Figure 3. It consists of convolutional layers, residual blocks, RFFB modules, CBAM modules, an adaptive average pooling layer, and a classifier. First, the image size is resized to 224×224 when the image is input to the network model before. Then, a convolutional layer with a stride of 2 and a convolutional kernel size of 3 are used to extract shallow feature information such as the contour and color of the grape leaves. Third, alternate structures of residual blocks, RFFB modules, and CBAM modules are used to deepen the network structure while improving the model’s ability to extract disease features, thereby improving the recognition accuracy. Next, the remaining convolutional layers are used to integrate the high-dimensional feature information. Finally, the adaptive average pooling layer integrates the shape of the feature map to $1 \times 1 \times 1280$. The classifier (the fully connected layer) adopts SoftMax for the classification of the extracted features.

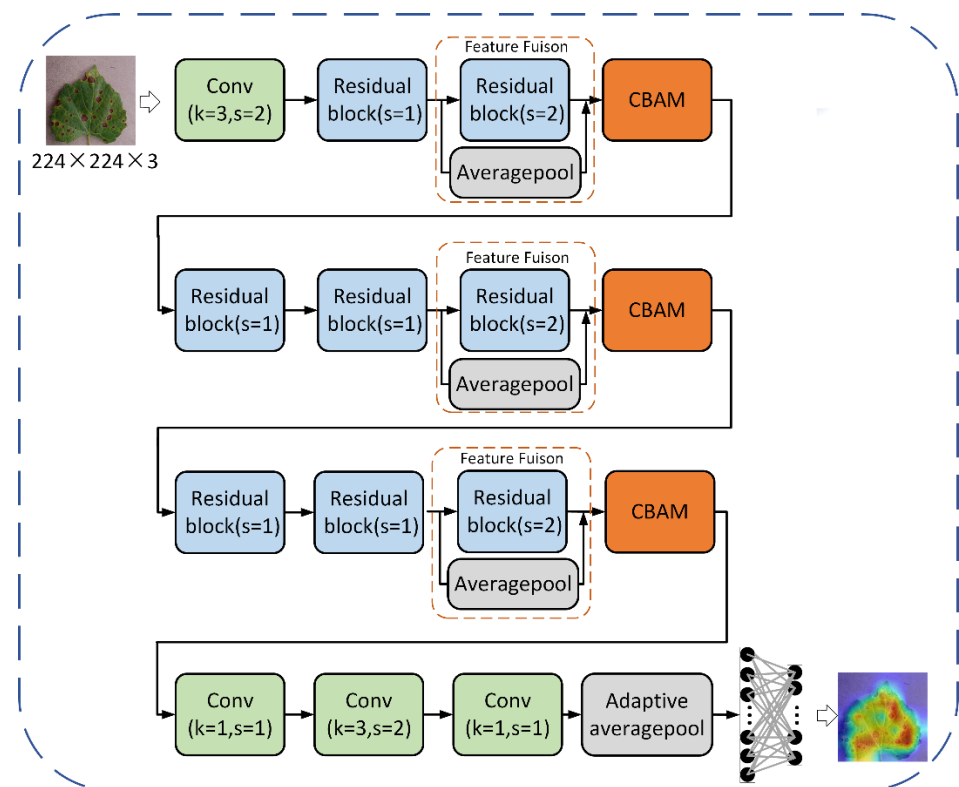


Figure 3. Cont.

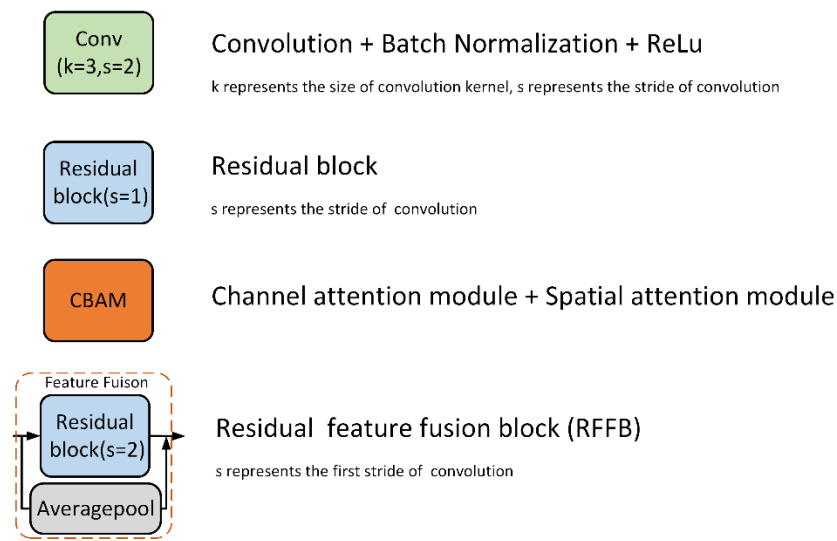


Figure 3. The network structure of the GrapeNet model. It consists of convolutional layers, residual blocks, residual feature fusion block (RFFB) modules, convolutional block attention modules (CBAMs), an adaptive average pooling layer, and a classifier.

2.4. RFFB Module

The residual block is a network structure proposed in the ResNet model. It mainly solves the problem of network degradation caused by the deep structure of the network model through residual learning [23]. He et al. proposed two types of residual blocks in ResNet34. As shown in Figure 4, Figure 4a represents the residual block when the stride is 1. The feature maps of the input and output are added by a skip connection. Figure 4b represents the residual block when the stride is 2. The input feature map is first subjected to a convolution operation with a stride of 2 and a convolution kernel size of 1, and is then added to the output feature map by a skip connection. When designing the GrapeNet model, we found that the residual block when the stride was 2 lost some detailed features, which made the model unable to capture more useful feature information. Therefore, we designed an RFFB module based on this residual block. To preserve the feature information to the greatest extent, we discarded the method of adding by skip connections in the residual module and adopted the method of concatenating by skip connection while using average pooling to replace the convolution operation with a kernel size of 1 on the shortcut branch. In this way, the parameters of the module can be reduced, and more disease characteristics can be preserved. The structure of an RFFB module is shown in Figure 5.

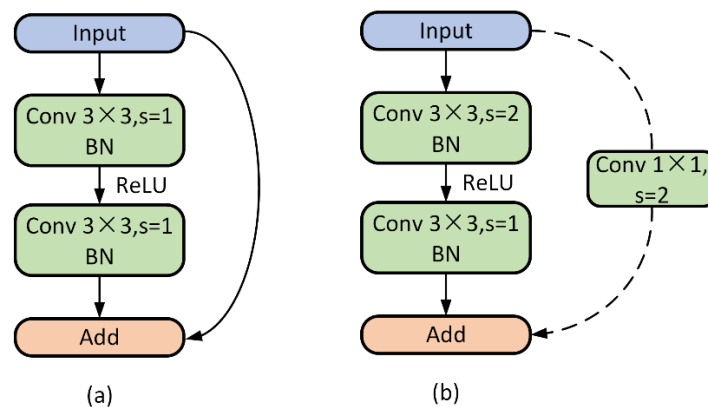


Figure 4. Structures of the residual blocks. (a) The residual block when the stride is 1. (b) The residual block when the stride is 2.

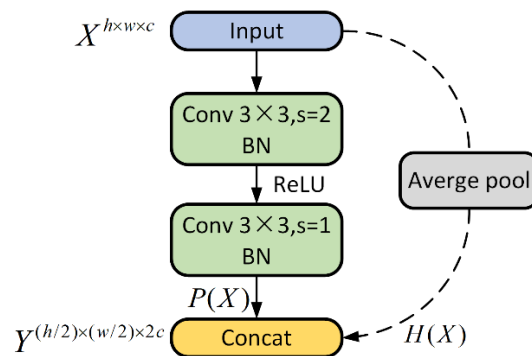


Figure 5. The structure of the RFFB module.

Given an input $X^{h \times w \times c}$, where h , w , and c denotes the height, width, and number of channels of the feature map, respectively. $P(X)$ is the feature map obtained by the convolution operation, which can be expressed as:

$$P(X) = Conv(X) \tag{1}$$

$H(X)$ is the feature map obtained by the average pooling operation, which can be written as:

$$H(X) = Averagepool(X) \tag{2}$$

The final output is obtained by concatenating the high-dimensional feature map $P(X)$ and the high-resolution feature map $H(X)$, thereby yielding:

$$Y^{(h/2) \times (w/2) \times 2c} = Concat\{P(X); H(X)\} \tag{3}$$

The RFFB module prevents the loss of feature information during down sampling by concatenating different forms of feature maps, which retains rich disease feature information, and increases the feature dimension so that the model can more accurately identify grape disease leaves.

2.5. CBAM Module

CNNs can obtain a large amount of useless information when extracting features, including background information and noise. This useless information greatly affects the effect of disease identification. The attention mechanism can ensure that the network focuses on useful feature information, suppresses the background and noise, and improves the recognition accuracy. In this study, the CBAM module was introduced into the network model so that the network can highlight the disease information of grape leaves. The structure of the CBAM module in the GrapeNet model is shown in Figure 6. After the RFFB module extracts a large amount of feature information, the CBAM module assigns different weights to different feature information; for example, it assigns more weight to disease information and assigns less weight to the background and noise. Finally, the residual block integrates the obtained information.

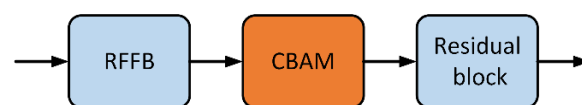


Figure 6. The structure of the CBAM module in the GrapeNet model.

Figure 7 shows the network structure of the CBAM module. This module first extracts the channel information of the feature map through spatial channel attention and then extracts the spatial information through spatial attention. Therefore, the CBAM module can be divided into two sub-modules: the channel attention module and spatial attention module.

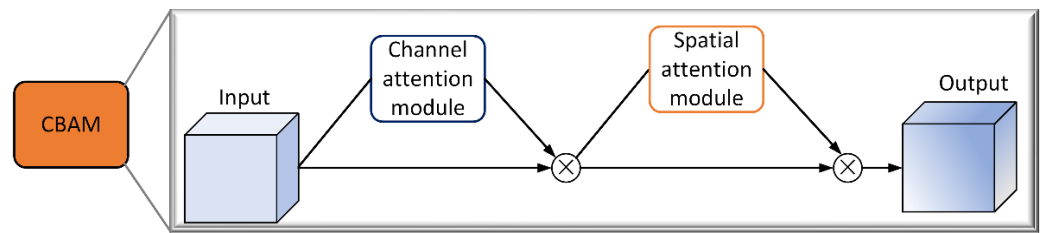


Figure 7. The network structure of the CBAM module.

The network structure of the channel attention module is shown in Figure 8. It mainly weights the channel information of the input feature map and highlights the channels with disease information.

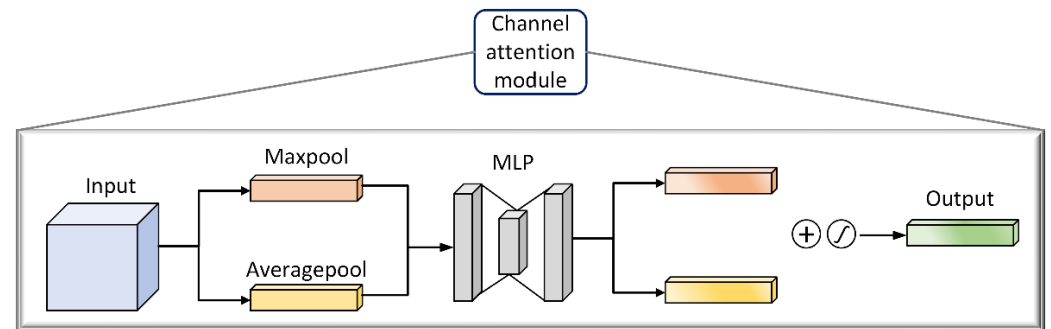


Figure 8. The network structure of the channel attention module.

The input feature map is F , and the output feature map can be expressed as:

$$M_c(F) = \sigma(W_1(W_0(F_{avg}^c) + W_1(W_0(F_{max}^c))) \tag{4}$$

where σ denotes the sigmoid function. The MPL weights (W_1 and W_0) are shared for both inputs and the ReLU activation function.

The network structure of the spatial attention module is shown in Figure 9. It highlights the diseased area of interest in the feature map by weighting the spatial information of the input feature map. The input feature map is F , and the formula can be written as:

$$M_c(F) = \sigma(f^{7 \times 7}(Concat[F_{avg}^s, F_{max}^s])) \tag{5}$$

where $f^{7 \times 7}$ represents a convolution operation with the kernel size of 7×7 , and σ denotes the sigmoid function.

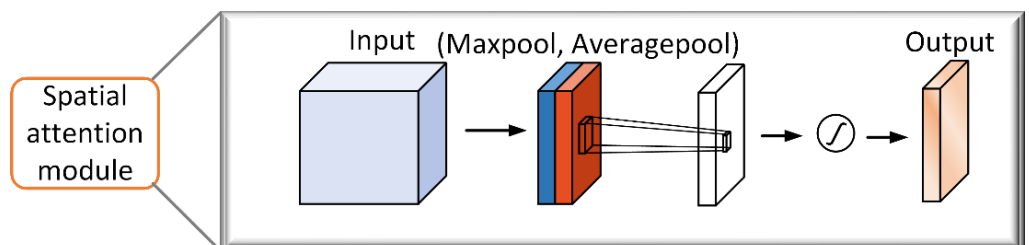


Figure 9. The network structure of the spatial attention module.

2.6. Evaluation Indexes

In this study, we used accuracy, precision, recall, and F1-score as evaluation indicators. The formulas are as follows:

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

$$precision = \frac{TP}{TP + FP} \quad (7)$$

$$recall = \frac{TP}{TP + FN} \quad (8)$$

$$F1 - score = \frac{2TP}{2TP + FP + FN} \quad (9)$$

where TP is the number of true-positive samples, FP is the number of false-positive samples, FN is the number of false-negative samples, and TN is the number of true-negative samples.

2.7. Experimental Environment and Hyperparameter Setting

The experimental environment is shown in Table 2. The hyperparameters were set as follows. The cross-entropy loss function (CE) was used as the loss function, and the Adam optimizer [24] was used to optimize the model. The initial learning rate and batch size during training were set to 0.0001 and 64, respectively. The number of iterations was 120.

Table 2. Experimental environment.

Name	Parameter
CPU	Intel(R) Xeon(R) W-2235
GPU	NVIDIA GeForce RTX 2080Ti
System	Windows 10
Programming language	Python 3.8.8
Deep learning framework	Pytorch 1.6.0

3. Results

3.1. The Impact of Data Augmentation on the Model

Figure 10 shows each epoch of the GrapeNet model with data augmentation and without it. We found that the training loss of the GrapeNet model with data augmentation dropped faster than that of the GrapeNet model without data augmentation, and the average accuracy of the model with data augmentation on the validation set was higher than that of the GrapeNet model without data augmentation on the validation set. This indicates that data augmentation can increase the diversity of data, reduce model overfitting, and enable the model to have better recognition ability. Moreover, the accuracy of the GrapeNet model with data augmentation was 86.25% in the test set, which is 4% higher than that of the GrapeNet model without data augmentation. This also indicates that the data augmentation method enables the model to have a higher generalization ability.

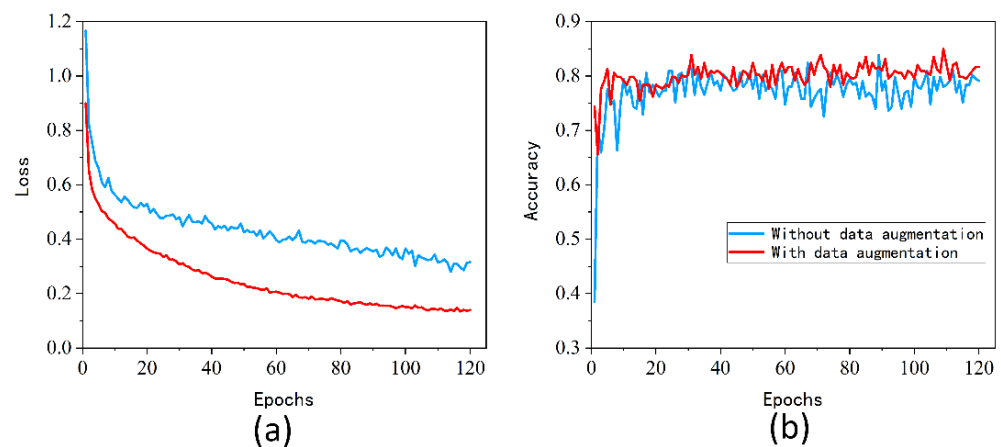


Figure 10. Each epoch of the GrapeNet model. (a) Training loss. (b) Validation accuracy.

3.2. Ablation Experiment

To verify the effectiveness of the RFFB module and CBAM module in the GrapeNet model, we performed ablation experiments on the test set. The obtained results are shown in Table 3. We found that the network model with the RFFB module achieved an accuracy of 82.49%. The accuracy was improved by 3.05% compared to the accuracy of the network model with no module introduced. We also found that the accuracy of the network model with the CBAM module was improved by 2.79% compared to that of the network model that does not introduce any module; the accuracy was 82.23%. The introduction of the RFFB module and the CBAM module did not add too many parameters to the network model. Finally, the network model GrapeNet, in which both modules were introduced simultaneously, achieved the best performance in terms of accuracy, precision, recall, and F1-score values (86.29%, 77.76%, 88.43%, and 79.05%, respectively). This indicates that the RFFB module and CBAM module can effectively enhance the identification of grape leaf disease.

Table 3. Results of ablation experiments.

RFFB	CBAM	Accuracy	Recall	Precision	F1-Score	Param (M)
-	-	0.7944	0.7569	0.7372	0.7413	2.05
✓	-	0.8249	0.7738	0.7878	0.7756	2.14
-	✓	0.8223	0.7658	0.7884	0.7689	2.05
✓	✓	0.8629	0.7776	0.8843	0.7905	2.15

3.3. Visual Comparison of Output Feature Maps

To demonstrate the effect of the RFFB module on the network model, we visualized the output feature maps of GrapeNet without the RFFB module, and GrapeNet with the RFFB module. As shown in Figure 11, the network model extracted the texture, color, and edge of grape leaf diseases in the first several convolutional layers. With the deepening of the network structure, the extracted feature information gradually became abstract feature information. We found that the abstract information of GrapeNet with RFFB module was richer than that of GrapeNet without the RFFB module; this is because the output of the average pooling operation and residual down sampling is concatenated in the RFFB module, and avoids the loss of a large number of detailed features, thus improving the model's capability to identify grape leaf diseases.

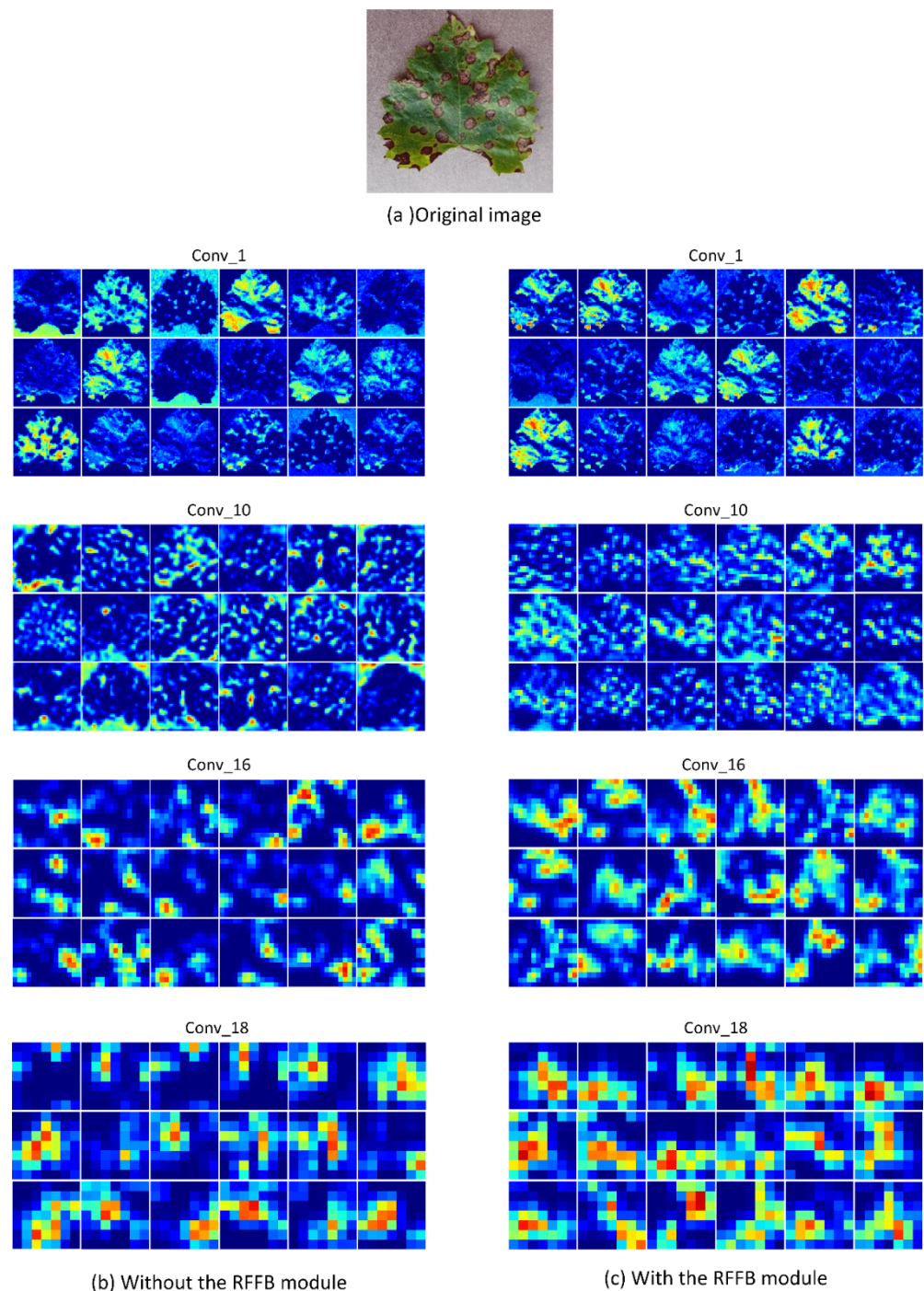


Figure 11. Visualization results of the output feature maps of the convolution layer. (a) The original image. (b) The visualization results of the output feature maps of GrapeNet without the RFFB module. (c) The visualization results of the output feature maps of GrapeNet with the RFFB module.

3.4. Comparison of Results of Different Attention Mechanisms

Table 4 shows the comparison of results of different attention mechanisms (the Squeeze-and-Excitation (SE) module, the Coordinate Attention (CA) module, and the CBAM module) on the network. None of the three attention models were found to produce an excessive number of parameters. However, the GrapeNet model with the introduction of the CBAM module had the highest accuracy of 86.29%, making it 0.76% more accurate than the GrapeNet model with the introduction of the SE module and 0.76% more accurate than the GrapeNet model with the introduction of the CA module. This result indicates that

introducing the CBAM module can make the GrapeNet model better focus on the disease region and reduce the influence of useless features, thereby improving the accuracy of grape leaf disease identification.

Table 4. Comparison of results of different attention mechanisms.

Attention Mechanism	Accuracy	Recall	Precision	F1-Score	Param (M)
SE [25]	0.8553	0.8012	0.8111	0.8053	2.15
CA [26]	0.8553	0.8206	0.8267	0.8217	2.15
CBAM	0.8629	0.7776	0.8843	0.7905	2.15

To show the regions of interest of the network model, Grad-cam [27] was used to visualize the class activation maps of the model using different attention mechanisms. As shown in Figure 12, the first row included the image of grape black rot fungus with serious symptoms (BRF_S). We found that the region of disease captured using the SE module and using the CA module was incomplete; however, the region of disease was captured intact using the CBAM module. For the image of grape black measles fungus with general symptoms (BMF_G) in the second row, the disease area was accurately located using the CBAM module, the disease area was not accurately captured using the SE module, and only a part of the disease area was captured using the CA module. This is because the CBAM module highlights the region of interest along the spatial and channel directions, thereby capturing more complete information about the grape leaf disease. In summary, the obtained results show that the introduction of the CBAM module can focus on the disease information and filter the background information.

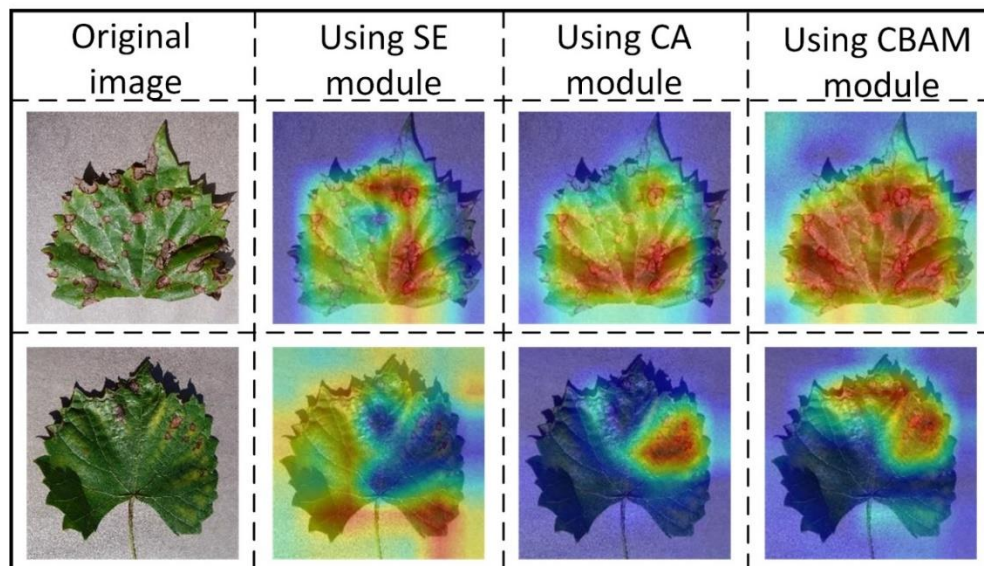


Figure 12. Visualization of results using different attention mechanisms.

3.5. Comparison of Identification Results with Classical CNN Models

Next, we aimed to verify that our proposed GrapeNet model has some advantages compared with classical CNN models such as heavyweight network models, VGG16, EfficientNet, ResNet, DenseNet, GoogLeNet and lightweight networks, MobileNetV2, MobileNetV3, and ShuffleNet. The results of the grape leaf disease test set are shown in Table 5. Furthermore, Table 5 also shows the training time of the nine models. The GrapeNet model achieved good performance on the test set for all evaluation metrics. It achieved a maximum accuracy of 86.29%, the number of parameters was 2.15 million, and the training time was 101 min. Compared to the DenseNet121 model, which was the most accurate of

the classical networks, the GrapeNet model yielded a 1.52% improvement in accuracy, a decrease of 4.81 million parameters, and a two times shorter training time. Compared to the ShuffleNetV2_x1.0 model, which had the fewest parameters and minimum training time among the classical networks, the GrapeNet model had a 5.33% improvement in accuracy, although the number of parameters increased by 0.89 million and the training time increased. Therefore, the GrapeNet model is a lightweight CNN model. It can achieve a balance between accuracy, the number of parameters, and training time, demonstrating the potential for grape leaf disease recognition.

Table 5. Identification results of the nine CNN models.

Model	Accuracy	Recall	Precision	F1-Score	Param (M)	Training Time (mins)
GoogLeNet [28]	0.8299	0.7521	0.8069	0.7601	5.98	107
Vgg16 [29]	0.8401	0.7761	0.7817	0.7777	134.29	254
ResNet34 [23]	0.8274	0.7617	0.77	0.762	21.29	108
DenseNet121 [30]	0.8477	0.7845	0.8357	0.7972	6.96	206
MobileNetV2 [31]	0.8096	0.7327	0.7572	0.74	2.23	98
MobileNetV3_large [32]	0.8274	0.7479	0.7818	0.7569	4.21	84
ShuffleNetV2_x1.0 [33]	0.8096	0.7455	0.7472	0.7424	1.26	64
EfficientNetV2_s [34]	0.8376	0.7738	0.8241	0.7865	20.19	290
GrapeNet	0.8629	0.7776	0.8843	0.7905	2.15	101

The confusion matrix for the nine models is shown in Figure 13. It was evident that the different periods of manifestation of the same grape leaf disease were the biggest factors affecting the recognition effectiveness of the network models. For grape black rot fungus (BRF), the DensNet121 model accurately classified the highest number of samples, 92, and the GrapeNet model accurately classified the next highest number of samples, 90. For both grape black measles fungus (BMF) and grape leaf blight fungus (LBF), the GrapeNet model achieved the highest number of true positive samples, 116 and 92, respectively. This indicates that our proposed GrapeNet model has better identification performance in different periods of the same grape leaf disease.

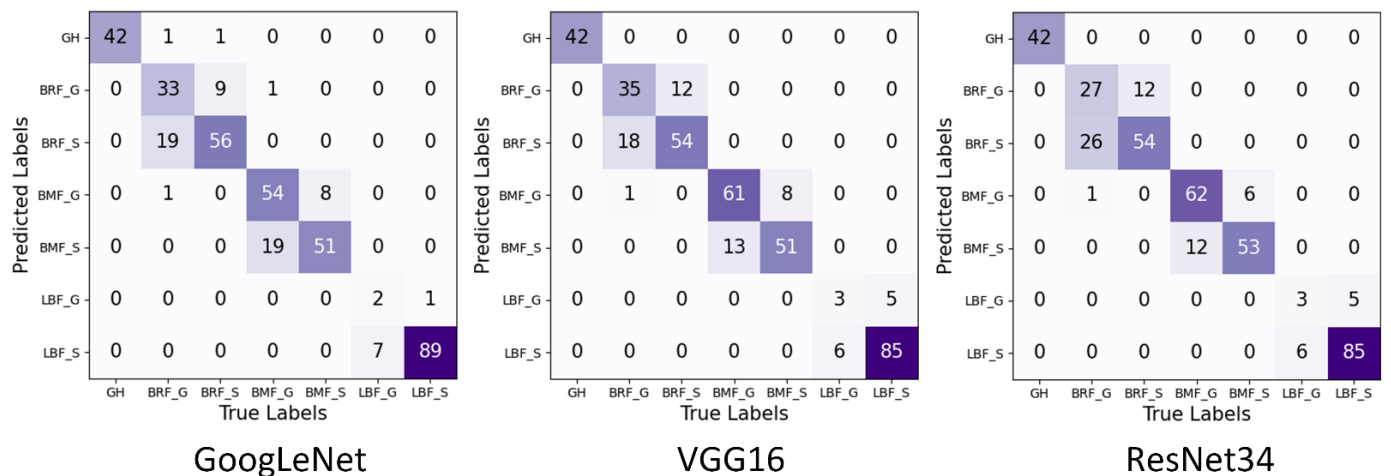


Figure 13. Cont.

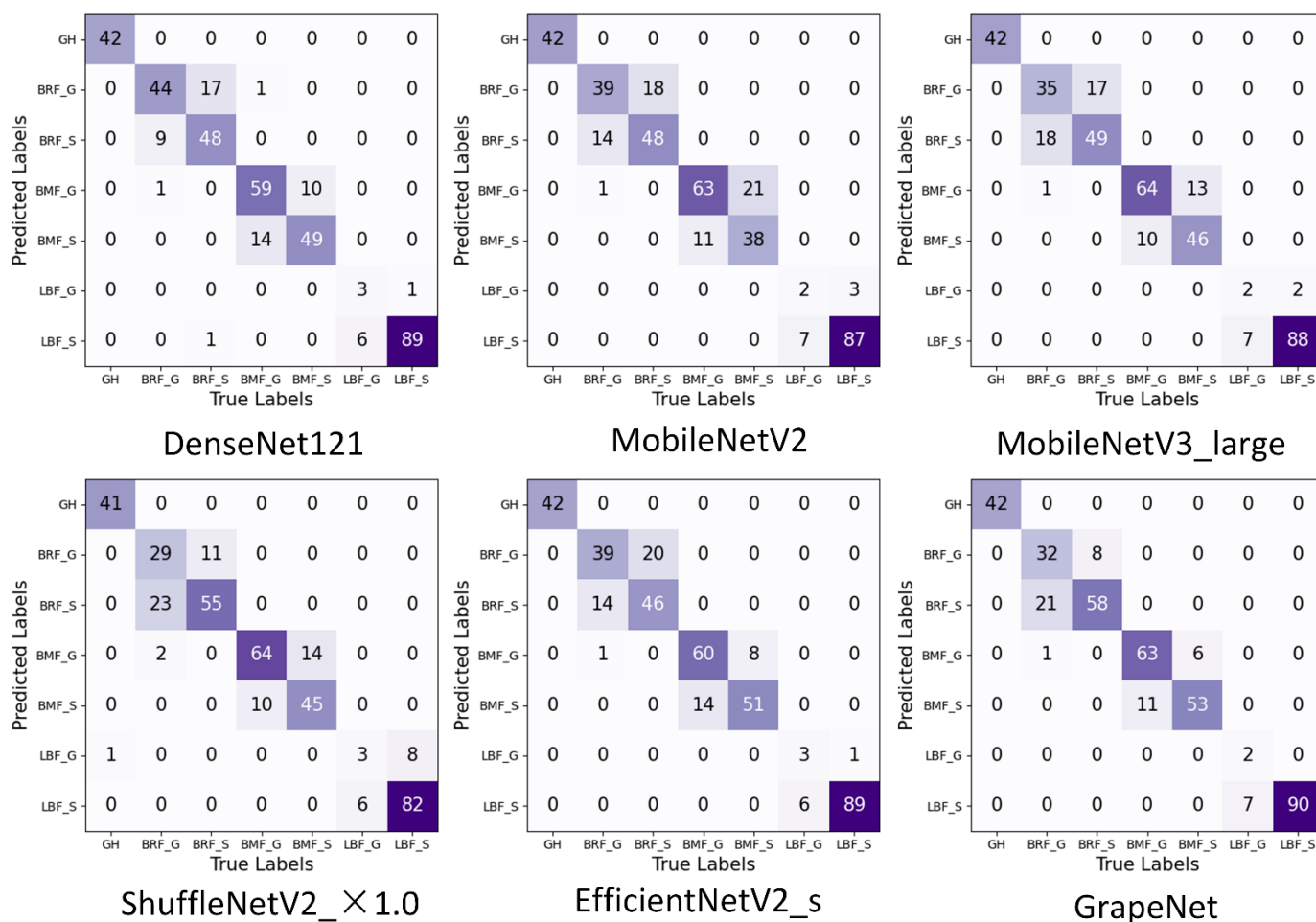


Figure 13. The confusion matrix for the nine models.

4. Discussion

Crop disease is a major threat affecting the safety of global agricultural production. Therefore, it is necessary to utilize new technologies for crop disease identification. Among deep learning methods, the CNN model has the advantages of high speed and high accuracy. It has been widely applied to identify crop diseases [14,16].

In this study, we compared our experimental results with some of the existing literature on crop disease identification. Zhao et al. proposed a deep CNN that combines the SE modules for the identification of tomato diseases [35]. The results suggested that the introduction of the SE module improved the accuracy of ResNet50 by 4.26%. Bao et al. proposed a CNN model called SimpleNet for the identification of wheat diseases [36]. The experiment results showed that the introduction of the CBAM modules improved the accuracy of the benchmark model by 3%. These results demonstrated that the attention mechanism is an effective module for improving the accuracy of crop disease identification. In our approaches, the accuracy of the network model with the CBAM module was improved by 2.79% compared to the accuracy of the network model with no module introduced. Perhaps, it is due to our dataset having a small intra-class variance, which requires fine-grained features for classification for the model. Subsequently, with the increasing difficulty of the classification task, the improved accuracy of the attention mechanism was reduced. Furthermore, the existing models ignored the loss of feature information during feature extraction [18,19]. Thus, we designed a module to reduce the loss of feature information, namely the RFFB module. With the introduction of the RFFB module on the benchmark, the accuracy of the model was improved by 6.85%. The model can retain more feature information while emphasizing them by the combination between the RFFB module and

the attention mechanism. The obtained result suggested that a single attention module is not enough to obtain the best performance of the network model. Hence, we should pay more attention to the loss of feature information to design new models.

Here, GrapeNet was designed for the specific task of grape leaf disease identification based on the above findings. It has a simplified architecture that deepens the network and extracts features with only a few residual blocks. It can reduce the number of parameters and provide an architecture of suitable depth and width for grape leaf disease identification. Afterwards, GrapeNet was compared with the classical models. As shown in Table 5, the disease identification accuracy of GrapeNet achieved 86.29%. The accuracy was 3.3%, 2.28%, 3.55%, 1.52%, 5.33%, 3.55%, 5.33%, and 2.53% higher than GoogLeNet, Vgg16, ResNet34, DenseNet121, MobileNetV2, MobileNetV3_large, ShuffleNetV2_×1.0, and Efficient-NetV2_s. Furthermore, the parameters of GrapeNet amounted to only 2.15 million. Although ShuffleNetV2_×1.0 has a lower number of parameters, its accuracy is 5.33% lower compared to GrapeNet. The results have clearly demonstrated the advantages of the proposed GrapeNet for grape leaf disease identification. Our approach can be deployed in smart mechanical devices or mobile devices to facilitate the rapid identification of grape leaf diseases, thus saving time and labor costs. Moreover, GrapeNet can be used not only for grape leaf diseases, but may have potential research value for other crop diseases with different symptom stages.

5. Conclusions

In this study, a lightweight CNN named GrapeNet was developed for the identification of grape leaf disease. To reduce the loss of disease features, an RFFB module was designed by concatenating the average pooled feature map before the input of the residual block and the output feature map of the residual block. In addition, the CBAM module was included between the RFFB module and the residual block to emphasize the disease features of interest and reduce the influence of redundant information. On the seven-class grape leaf disease test set, the GrapeNet model achieved 86.29%, 77.76%, 88.63%, and 79.05% in the accuracy, recall, precision, and F1-score, respectively. Furthermore, the GrapeNet model only included 2.15 million parameters. The GrapeNet model achieved an excellent balance between high accuracy and low parameter size. The results demonstrated that GrapeNet has high potential for identification of grape leaf diseases on mobile and embedded devices. However, the study has a specific shortcoming. The overly simple background in the grape disease images may lead to a decrease in the recognition accuracy of the model in relevant environments. In the future, we will continue to optimize our model for grape leaf disease recognition under field conditions.

Author Contributions: Conceptualization, X.Z. and J.L.; methodology, X.Z. and X.C.; software, R.P.; validation, X.Z., J.L. and R.P.; formal analysis, J.L. and T.C. (Tengbao Cao); investigation, J.C.; data curation, Y.C. and T.C. (Tengbao Cao); writing—original draft preparation, J.L., R.P. and T.C. (Tomislav Cernava); writing—review and editing, X.Z., T.C. (Tomislav Cernava), and X.C.; visualization, X.P.; supervision, X.Z. and X.C.; project administration, X.Z. and X.C.; funding acquisition, X.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key Research and Development Plan Key Special Projects, grant number 2021YFE0107700, National Nature Science Foundation of China, grant numbers 61865002 and 31960555, Guizhou Science and Technology Program, grant number 2019-1410, and Outstanding Young Scientist Program of Guizhou Province, grant number KY2021-026. In addition, the study received support by the Program for Introducing Talents to Chinese Universities, 111 Program, grant number D20023.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Peng, Y.; Zhao, S.Y.; Liu, J.Z. Fused Deep Features-Based Grape Varieties Identification Using Support Vector Machine. *Agriculture* **2021**, *11*, 869. [[CrossRef](#)]
2. Ji, M.; Zhang, L.; Wu, Q. Automatic grape leaf diseases identification via UnitedModel based on multiple convolutional neural networks. *Inf. Process. Agric.* **2020**, *7*, 418–426. [[CrossRef](#)]
3. Singh, V.; Misra, A. Detection of plant leaf diseases using image segmentation and soft computing techniques. *Inf. Process. Agric.* **2017**, *4*, 41–49. [[CrossRef](#)]
4. Zhang, Z.; Khanal, S.; Raudenbush, A.; Tilmon, K.; Stewart, C. Assessing the efficacy of machine learning techniques to characterize soybean defoliation from unmanned aerial vehicles. *Comput. Electron. Agric.* **2022**, *193*, 106682. [[CrossRef](#)]
5. Jaisakthi, S.; Mirunalini, P.; Thenmozhi, D. Grape leaf disease identification using machine learning techniques. In Proceedings of the 2019 International Conference on Computational Intelligence in Data Science (ICCIDIS), Vatsala, Australia, 21–23 February 2019; pp. 1–6.
6. Majumdar, D.; Kole, D.K.; Chakraborty, A.; Majumder, D.D. An integrated digital image analysis system for detection, recognition and diagnosis of disease in wheat leaves. In Proceedings of the Third International Symposium on Women in Computing and Informatics, Kerala, India, 10–13 August 2015; pp. 400–405.
7. Guru, D.; Mallikarjuna, P.; Manjunath, S. Segmentation and classification of tobacco seedling diseases. In Proceedings of the Fourth Annual ACM Bangalore Conference, Bangalore, India, 25–26 March 2011; pp. 1–5.
8. Rumpf, T.; Mahlein, A.K.; Steiner, U.; Oerke, E.C.; Dehne, H.W.; Plümer, L. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Comput. Electron. Agric.* **2010**, *74*, 91–99. [[CrossRef](#)]
9. Padol, P.B.; Yadav, A.A. SVM classifier based grape leaf disease detection. In Proceedings of the 2016 Conference on Advances in Signal Processing (CASP), Pune, India, 9–11 June 2016; pp. 175–179.
10. Martins, P.; Silva, J.S.; Bernardino, A. Multispectral Facial Recognition in the Wild. *Sensors* **2022**, *22*, 4219. [[CrossRef](#)]
11. Khan, I.R.; Ali, S.T.A.; Siddiq, A.; Khan, M.M.; Ilyas, M.U.; Alshomrani, S.; Rahardja, S. Automatic License Plate Recognition in Real-World Traffic Videos Captured in Unconstrained Environment by a Mobile Camera. *Electronics* **2022**, *11*, 1408. [[CrossRef](#)]
12. Orchi, H.; Sadik, M.; Khaldoun, M. On Using Artificial Intelligence and the Internet of Things for Crop Disease Detection: A Contemporary Survey. *Agriculture* **2022**, *12*, 9. [[CrossRef](#)]
13. Liu, B.; Ding, Z.; Tian, L.; He, D.; Li, S.; Wang, H. Grape leaf disease identification using improved deep convolutional neural networks. *Front. Plant Sci.* **2020**, *11*, 1082. [[CrossRef](#)]
14. Tang, Z.; Yang, J.; Li, Z.; Qi, F. Grape disease image classification based on lightweight convolution neural networks and channelwise attention. *Comput. Electron. Agric.* **2020**, *178*, 105735. [[CrossRef](#)]
15. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Front. Plant Sci.* **2016**, *7*, 1419. [[CrossRef](#)] [[PubMed](#)]
16. Pandian, J.A.; Kanchanadevi, K.; Kumar, V.D.; Jasińska, E.; Goño, R.; Leonowicz, Z.; Jasiński, M. A Five Convolutional Layer Deep Convolutional Neural Network for Plant Leaf Disease Detection. *Electronics* **2022**, *11*, 1266. [[CrossRef](#)]
17. Chao, X.; Sun, G.; Zhao, H.; Li, M.; He, D. Identification of Apple Tree Leaf Diseases Based on Deep Learning Models. *Symmetry* **2020**, *12*, 1065. [[CrossRef](#)]
18. Gao, R.; Wang, R.; Feng, L.; Li, Q.; Wu, H. Dual-branch, efficient, channel attention-based crop disease identification. *Comput. Electron. Agric.* **2021**, *190*, 106410. [[CrossRef](#)]
19. Chen, J.; Zhang, D.; Suzauddola, M.; Zeb, A. Identifying crop diseases using attention embedded MobileNet-V2 model. *Appl. Soft Comput.* **2021**, *113*, 107901. [[CrossRef](#)]
20. Zeng, W.; Li, H.; Hu, G.; Liang, D. Lightweight dense-scale network (LDSNet) for corn leaf disease identification. *Comput. Electron. Agric.* **2022**, *197*, 106943. [[CrossRef](#)]
21. Kamal, K.; Yin, Z.; Wu, M.; Wu, Z. Depthwise separable convolution architectures for plant disease classification. *Comput. Electron. Agric.* **2019**, *165*, 104948.
22. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
23. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
24. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
25. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.
26. Hou, Q.; Zhou, D.; Feng, J. Coordinate attention for efficient mobile network design. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 13708–13717.
27. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 618–626.
28. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.

29. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
30. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
31. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
32. Howard, A.; Sandler, M.; Chu, G.; Chen, L.-C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for MobileNetV3. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV 2019), Seoul, Korea, 27 October–2 November 2019; pp. 1314–1324.
33. Ma, N.; Zhang, X.; Zheng, H.-T.; Sun, J. Shufflenet v2: Practical guidelines for efficient cnn architecture design. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 116–131.
34. Tan, M.; Le, Q.V. Efficientnetv2: Smaller models and faster training. *arXiv* **2021**, arXiv:2104.00298.
35. Zhao, S.; Peng, Y.; Liu, J.; Wu, S. Tomato Leaf Disease Diagnosis Based on Improved Convolution Neural Network by Attention Module. *Agriculture* **2021**, *11*, 651. [[CrossRef](#)]
36. Bao, W.; Yang, X.; Liang, D.; Hu, G.; Yang, X. Lightweight convolutional neural network model for field wheat ear disease identification. *Comput. Electron. Agric.* **2021**, *189*, 106367. [[CrossRef](#)]