ISSN: 2321-2152

IJMECE

International Journal of modern electronics and communication engineering

E-Mail
editor.ijmece@gmail.com
editor@ijmece.com

www.ijmece.com

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

Using Deep Learning Ensemble Models for Grapevine Leaf Image Classification

[1]Mr. K. Lakshmi Narayana, [2] Nalli Sai Pandu,

[1]Assistant Professor, Dept.of Master of Computer Applications, Rajamahendri Institute of Engineering & Technology, Bhoopalapatnam, Near Pidimgoyyi,Rajahmundry,E.G.Dist.A.P. 533107.

[2] Students, Dept.of Master of Computer Applications, Rajamahendri Institute of Engineering & Technology, Bhoopalapatnam, Near Pidimgoyyi,Rajahmundry,E.G.Dist.A.P. 533107.

Abstract—
The nutritional value of grapes is significant, and they have many practical applications beyond just eating and fermenting. The ability to distinguish between different grape varieties by observing their leaf morphology is crucial for grape breeding and variety development since there are so many different grape varieties. The mature leaves of many grape varietals are the subject of this article's study. Collecting and preprocessing leaf pictures is the first step in training and fine-tuning five pre-trained deep learning models: VGG19, VIT, Inception ResnetV2, DenseNet201, and ResneXt. The five models' predictions are then combined using two voting ensemble ML models. The ensemble classifier, which uses soft voting to make decisions, achieves the maximum accuracy of 98.1%. Keywords: grapevine variety identification, hard voting, soft voting, deep learning ensemble model

INTRODUCTION

Identifying grape varieties is crucial for spreading awareness of grape research and promoting this cash crop as the grape market economy grows. The grape variety identification study often uses the leaves as the object of identification. Preprocessing the leaf image is the first step in grape variety recognition using grape leaves. Then, features from deep learning or artificial design are extracted and used as parameter inputs to build a recognition model with a classifier. It relies on human intervention to extract design features, which is tedious, time-consuming, and susceptible to human error. Many domains have found uses for deep learning-based feature extraction, including machine vision, NLU, etc. In machine learning, a method known as ensemble learning teaches several learners and then uses them together. In actual use, this algorithm type often outperforms a single learner when it comes to making predictions. We classified grapevine leaf photos using the ensemble learning approach in this research. We must first argue the training sample set in order to accurately train our deep learning ensemble model. Following the use of augmentation techniques, the training set was enlarged to 2800. Second, to categorize grapevine leaves, we used five different classification models: VGG19[1], VIT[2], Inception ResnetV2[3], DenseNet201[4], and ResneXt [5]. The third point is that hard voting and soft voting were both used. Here is how the remainder of this paper is structured. The relevant literature is examined in Section II. Part III details the study's methodology, including the picture preprocessing approach, deep feature extraction, voting mechanisms, and outcomes comparison. The article is concluded and future work is addressed in Section IV.

RELATED WORKS

The use of machine learning techniques for grapevine leaf image classification has been the subject of much study. Various Bayesian Belief Network, Support Vector Machine, Logistic Regression, and other machine learning approaches have been used. The grapevine leaf classification system was adjusted by Hunar A. Ahmed using DenseNet201. The maximum accuracy that DenseNet 201 was able to attain was 98% [6]. M. Koklu suggested a CNN-SVM investigation into grapevine leaf categorization using chosen deep characteristics. To identify grapevine leaves, they employed support vector machines (SVMs), and to extract features, they used a pre-trained MobileNetv2 Logits layer. Their system's categorization success rate was found to be 97.6% [7]. Photos of grapevine leaves were semantically segmented for phenotyping purposes by Tamvakis Petros using U-Net architecture[8]. Pay close attention to the veins in the leaves and the features of the blades. They used three distinct supervised learning methods: design-and-train, transfer learning, and parameterization of a well-known architecture. A three-stage approach based on deep learning was suggested by Chen Yiping for the identification of illnesses affecting grape leaves. To find lesions on grape leaves, they used ResNet, generative adversarial network for data augmentation, and Faster R-CNN. Their suggested model had strong generalizability, as shown by their experimental findings [9]. Bingpiao Liu put up the YOLOX-RA grape detection model to solve the issue of grape identification in unstructured situations. This model can reliably and swiftly recognize clusters of densely growing grapes, even when they are partially obscured. According to their model, the recognition speed was 84.88 FPS and the mAP was 88.75% [10]. To recognize grape bunches in real time, Sozzi, M. employed six different YOLO object identification methods. They found that YOLOv5x achieved an F1-score of 0.76 and YOLOv4 achieved an F1-score of 0.77 [11].

METHODGOLOGY

An ensemble technique for species identification in grape leaves is proposed in this paper. We begin by doing data preparation and getting the data ready. Afterwards, the dataset was trained using five classifiers: VGG19, VIT, Perception ResnetV2, DenseNet201, and ResneXt. For learning, the outputs from all the classifiers are combined. We use both hard voting and soft voting as integration tactics in our strategy. Figure 1 displays the ensemble method's process. Dataset description, picture preparation methods, classification algorithms, and voting procedures are all part of this section.

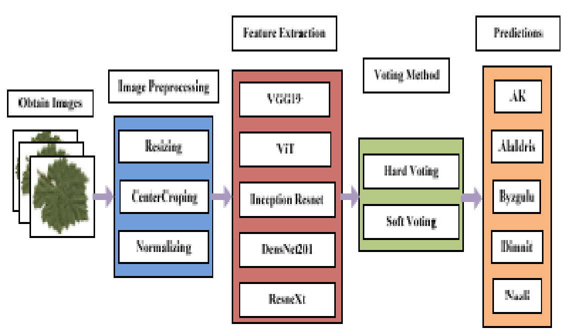

Fig. 1. Workflow of the methodology

There are a total of 500 leaf samples in the grapevine leaf dataset, which includes 5 species with 100 samples per class [7]. Azerbaijan, Ala Idris, Buzgulu, Dimnit, and Nazli are the five groups discussed here. Feature information from one category does not automatically transfer to another. Shape, texture, and size of the leaf are all pieces of feature information. We will train and identify five grape leaves, as shown in Fig. 2.

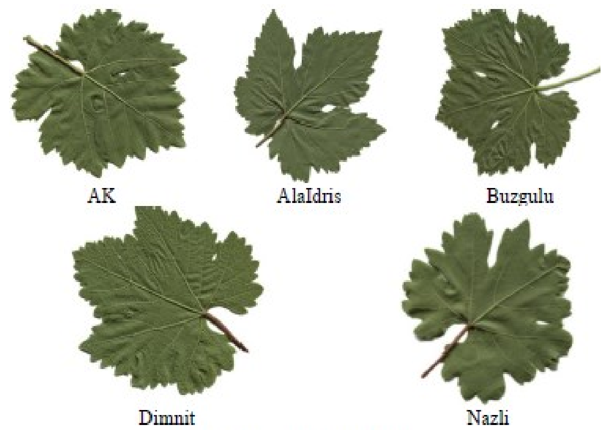

Fig. 2. Five grapevine leaf categories

The original picture has dimensions of 512 by 512 pixels; to save computing expense, we reduce the image size to 256 x 256 pixels. Our next step is to reduce the image's central dimension to 224 × 224. To make the picture more consistent, we choose a standard deviation of 0.2290 and a mean of 0.4850, 0.4560, and 0.4060. Random erasing with parameters p=0.5, scale=(0.02,0.32), and ratio=(0.3, 3.2) is also used. The train set is additionally randomly supplemented with Gaussian noise. The flow diagram of image processing is shown in Figure 3.

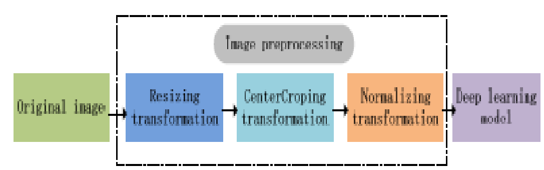

Fig. 3. Schematic diagram of image processing

In this work, grapevine leaf features are extracted using five pre-trained models: VGG19, VIT, Inception ResnetV2, DenseNet201, and ResneXt. At first, all of the chosen models undergo fine-tuning. A convolutional neural network of nineteen hidden layers, including sixteen convolutional layers and three fully connected layers, is known as VGG19 [1]. There are a thousand classes in the original VGG19 model output, but our research only needs five. For this reason, we've modified the layer's out channel number from 1000 to 5. The vit_base_patch16_224 (ViTB/ 16 model) was selected for testing and training purposes in the ViT model [2]. Modules MLP Head, Linear Projection of

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

Flattened Patches, and Transformer Encoder make up the ViT-B/16 model. The dimensions of the input picture used by the ViT-B/16 model are 224 × 224 Ó 3. This patch is 16 x 16 × 3. Every patch embed has a 768-by-12-head dimensionality, and Multi Head Attention makes use of 12 transformer encoder blocks.

$$\hat{y} = \arg\max_{i} \sum_{j=1}^{m} w_j \mathcal{X}_A(C_j(x) = i) \quad (1)$$

Two voting methods are offered in this research. You have two options: hard voting and soft voting. Hard voting relies on the minority caving in to the majority in order to get a final decision. The combined likelihood of all classifiers is used for soft voting. Two voting methods are compared. The suggested methodology's algorithm was shown in Fig. 4 [12].

```
Algorithm 1
procedure Preprocess(grapevine_leaf_data)
return grapevine_leaf_data['AK','AlaIdris','Buzgulu','Dimnit','Nazli']

procedure split_data(grapevine_leaf_data)
    Training_data, Testing_data=split(grapevine_leaf_data)
    returnTraining_data, Testing_data

D1=VGG(Training_data, Testing_data)
D2= ViT(Training_data, Testing_data)
D3=Inception Resnet(Training_data, Testing_data)
D4= DenseNet (Training_data, Testing_data)
D5= ResneXt (Training_data, Testing_data)

procedureensemble_model(Training_data, Testing_data))
    soft_voting_classifier=concatenate(D1,D2,D3,D4,D5)
    soft_voting_classifier.fit(Training_data)
    soft_predictions= hard_voting_classifier.predict(Testing_data)
    hard_voting_classifier=concatenate(D1,D2,D3,D4,D5)
    hard_voting_classifier.fit(Training_data)
    hard_predictions= hard_voting_classifier.predict(Testing_data)
```

Fig. 4. Algorithm for proposed ensemble soft voting Classifier

Result

We split the dataset in three parts: 80% for training, 10% for testing, and 10% for validation. Our method's training hyperparameters are shown in Table 1. We set EPOCHS to 9 since the machine's performance was restricted. Raising the value of EPOCHS really does enhance recognition accuracy.

Table 1. TRAINING HYPERPARAMETER SETTING

Hyperparameter	Value
Batch_Size	64
EPOCHS	9
Learning Rate	0.001

We use F1-score, accuracy, precision, and recall as performance measures to assess the efficacy of our algorithms. These metrics are calculated and shown in Table 2. This table categorizes predictions according to their accuracy: TP for true positive, FP for false positive, TN for true negative, and FN for false negative. Each of these predictions indicates a different outcome: either the prediction is correct or the value is negative.

Table 2. CALCULATION FORMULAS OF PERFORMANCE METRICS

Measure	Formula
Accuracy	(TP+TN)/(TP+TN+FN+FP)
Precision	TP/(TP+FP)
Recall	TP/(TP+FN)
F1-score	2TP/(2TP+FP+FN)

The results of the model-based classifications are shown in Table 3 along with the Accuracy, Precision, Recall, and F1-score values. The soft voting ensemble had better results than both models across the board. Compared to ViT and ResneXt, the soft voting ensemble model achieves an accuracy that is 4% greater. Among all models, its F1-Score is 97.99%, its recall is 98%, and its precision is 98.18%. The comparison of performance measures is shown in Fig. 5.

TABLE 3. CLASSIFICATION RESULTS

Algorithms	Accuracy	Precision	Recall	F1-score
VGG19	0.86	0.8688	0.8600	0.8519
ViT	0.94	0.9436	0.9400	0.9387
DenseNet201	0.86	0.8967	0.8600	0.8600
Inception ResnetV2	0.82	0.8299	0.8200	0.8185
ResneXt	0.88	0.8851	0.8800	0.8752
Hard Voting	0.96	0.9618	0.9600	0.9599
Soft Voting	0.981	0.9818	0.9800	0.9799

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

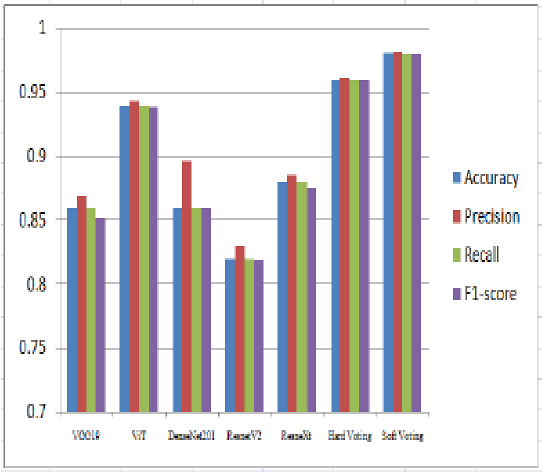

Fig. 5. Comparison of performance metrics

The training set for this research is prepared using preprocessing and image enhance approaches. We examine and train the models using the original data to confirm the effect of data augmentation; table 4 displays the experimental outcomes. Table 3 shows that the algorithm's accuracy, precision, recall, and F1 score may be significantly improved by data preparation and data augment approaches. Tables 3 and 4 show that the accuracy of the soft voting classifier went raised from 92% to 98.1% after picture enhancement and preprocessing. A number of other KPIs have also been enhanced, but to different degrees.

TABLE 4. CLASSIFICATION RESULTS ON ORIGINAL DATA

Algorithms	Accuracy	Precision	Recall	F1-score
VGG19	0.82	0.8392	0.8200	0.8185
ViT	0.88	0.8857	0.8800	0.8802
DenseNet201	0.86	0.8593	0.8600	0.8595
Inception ResnetV2	0.74	0.7453	0.7400	0.7375
ResneXt	0.88	0.8838	0.8800	0.8802
Hard Voting	0.91	0.9120	0.9100	0.9099
Soft Voting	0.92	0.9219	0.9200	0.9202

A confusion matrix is a common tool for visualizing how well a classifier or algorithm is doing. The soft voting ensemble-based model's normalized confusion matrix is shown in Figure 6. For four of the five classes, it successfully identifies all test samples; for the fifth, it makes an error rate of about 2%.

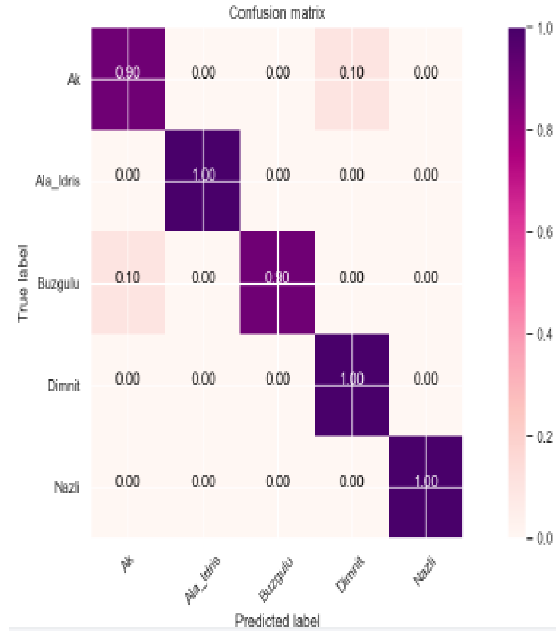

Fig. 6. Normalized confusion matrix of Soft Voting ensemble model

Lastly, we compare our methods to the model in [6] and [7] to assess the suggested model. Table 5 displays the outcomes of the categorization. According to the findings, the suggested model was the most effective. In order to categorize grapevine leaves, one technique [6] modified DenseNet201; in another, [7], they used SVM kernels and a pre-trained MobileNetv2 Logits layer to extract features.

TABLE 5. CLASSIFICATION RESULTS

`	Accuracy	Precision	Recall	F1-score
M.Koklu[2]	0.976	0.9762	0.760	0.9760
Hunar[1]	0.9802	0.9800	0.9818	0.9800
Hard Voting	0.96	0.9618	0.9600	0.9599
Soft Voting	0.981	0.9818	0.9800	0.9799

IV. CONCLUSION

This work aimed to categorize five species of grapevine leaves using ensembling techniques and transfer learning models like VGG19, VIT, Inception ResnetV2, DenseNet201, ResneXt, and others. The research found that using VGG19, VIT, Inception ResnetV2, DenseNet201, and ResneXt via a soft voting classifier outperformed both models. much if the models perform at the SOTA level, they may be much better. Here are some things to think about: Given the short size of the sample, it would be beneficial to collect more data for this research in order to enhance identification accuracy and lower the likelihood of misclassification. For the purpose of this work, we used five different deep learning models to identify grape leaves. Actually, EfficientNet and Swin Transformer are only two of several learning models that are capable of good object recognition. We may attempt to use these models for grape leaf identification in the next stage of our study.Other ensemble approaches that are used: We employed both hard voting and soft voting, two ensemble approaches, for classification in this research. A stacked ensemble learning classifier may be trained to recognize grapevine leaves using information learnt from several classification models.

References

[1] K. Simonyan, A. Zisserman. (2014)Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556

[2] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner. (2021)An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In Proceedings of the 2021 International Conference on Learning Representations (ICLP), Colombo, Sri Lanka. pp:20–27.

[3] C. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi. (2016)Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, ArXiv, abs/1602.07261

[4] G. Huang, Z. Liu, Van Der Maaten, L. Weinberger, K.Q.(2017)Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI. pp:4700-4708.

[5] Xie, R. Girshick, P. Dollár, Z. Tu , K. He.(2017)Aggregated Residual Transformations for Deep Neural Networks, In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI. pp:5987-5995

[6] Ahmed Hunar, Hama Hersh, Jalal, Shayan, Ahmed Mohammed.(2023) Deep Learning in Grapevine Leaves Varieties Classification Based on Dense Convolutional Network. Journal of Image and Graphics, 11:98- 103.

[7] M. Koklu, M. F. Unlersen, I. A. Ozkan. (2021)A CNN-SVM study based on selected deep features for grapevine leaves classification. Measurement ,188(110425):1-10

[8] Tamvakis Petros, Kiourt Chairi, Solomou Alexandra, Ioannakis George Alexis, Tsirliganis, Nestoras. (2022)Semantic Image Segmentation with Deep Learning for Vine Leaf Phenotyping. IFACPapersOnLine, 55: 83-88.

[9] Chen Yiping , Wu Qiufeng.(2022)Grape leaf disease identification with sparse data via generative adversarial networks and convolutional neural networks. Precision Agriculture. 24: 235–253.

[10] Bingpiao Liu, Lufeng Luo, Jinhai Wang, Qinghua Lu, Huiling Wei, Yunzhi Zhang, Wenbo Zhu. (2023)An improved lightweight network based on deep learning for grape recognition in unstructured environments. Information Processing in Agriculture(In Press).

[11] Sozzi, M., Cantalamessa, S., Cogato, A., Kayad, A., F. Marinello. (2022)Automatic Bunch Detection in White Grape Varieties Using YOLOv3, YOLOv4, and YOLOv5 Deep Learning Algorithms. Agronomy,12(2): 319

[12] Saloni Kumari, Deepika Kumar, Mamta Mittal. (2021)An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. International Journal of Cognitive Computing in Engineering, 2:40-46