Leaf Disease Detection in Blueberry Using Efficient Semi- Supervised Learning Approach

Vinh Dinh Nguyen, Ngoc Phuong Ngo, and Narayan C. Debnath

Biochemistry and Plant Physiology Faculty, College of Agriculture, Can Tho University, Can Tho City, Vietnam
Department of Software Engineering School of Computing and Information Technology, Eastern International University, Vietnam

Outline

- Introduction
- Proposed Method
- Experimental Results
- Conclusion and Future Work

Introduction

- Blueberry leaf disease detection is really important to help farmers to early detect leaf disease and find a suitable method to cure the disease. Therefore, this research introduces an approach to detect and classify blueberry leaf disease by using an unsupervised method (auto- encoder), and a supervised method (support vector machine).

- The accuracy of our proposed method was evaluated by conducting the experiments on blueberry dataset captured at Can Tho City, Vietnam. The existing augmentation techniques was also applied to increase the data size of training and testing.

- For the first experiment on normal capturing conditions, the F1 scores of the proposed method and SVM are 89.28% and 81.48%, respectively. For the second experiment with noisy conditions, the F1 scores of the proposed method and SVM are 81.5% and 66.7%, respectively.

- **Nowadays, leaf disease is one of the typical issues that affect farmers. Therefore, several searchers did a study to find a way to help farmers by developing a method that automatically detects and classifies leaf disease, automatically.**

- **The researcher often uses a supervised learning approach to detect and classify blueberry leaf diseases such as SVM, Random Forest, Neural Networks, and Deep learning as discussed [1]. Most of the existing blueberry leaf disease detection is designed by allowing the raw input pixel to the proposed method for training and testing [1]. However, in this research, the benefit of unsupervised learning was used to extract stable groups of features before inputting them into the SVM for train and testing.**

- **To increase the accuracy of the existing support vector machine (SVM)-based method for object detection and classification, the authors study the advantages of the autoencoder [2].**
- **First, the first proposed auto-encoder is used to extract and reduce the features dimension of the input image size from 160x160x3 to 100 features.**

- **Second, the second auto-encoder is used to extract and reduce the feature dimension of the input image resolution from 160x160x3 to 200 features.**
- **Third, the third auto-encoder is used to extract and reduce the feature dimension of the input image resolution from 160x160x3 to 300 features. Fourth, the results of 100, 200, and 300 features are concatenated to establish 600 features for input to the SVM to detect and classify whether the input blueberry leaf has a disease or not.**

- **To detect blueberry leaf disease, various machine learning and deep learning approaches have been investigated. Barbedo et al developed a method to detect and classify various blueberry disease, such as citrus caker, bacterial blight, and black mouth, by using the convolutional neural network [3][4].**

- **Fuentes et al. introduce a deep learn- ing approach to detect various diseases of tomato plants by using SSD and Faster RCNN [5]. Lee et al. developed a new CNN along with a pre-train VGG model to accurately detect and classify leaf disease by using an open dataset [6]. Another re- search was conducted by Too et al. [7] on black rot and early blight by using DenseNet and RestNet. Recently, in the AI-challenge plant disease contest, Zhong et al. use DenseNet-121 to recognize apple leaf disease [8]. More recently, Shrivastava et al. de- veloped an improved version of CNN and SVM to detect and classify Rice plants [9][10].**

Proposed Method

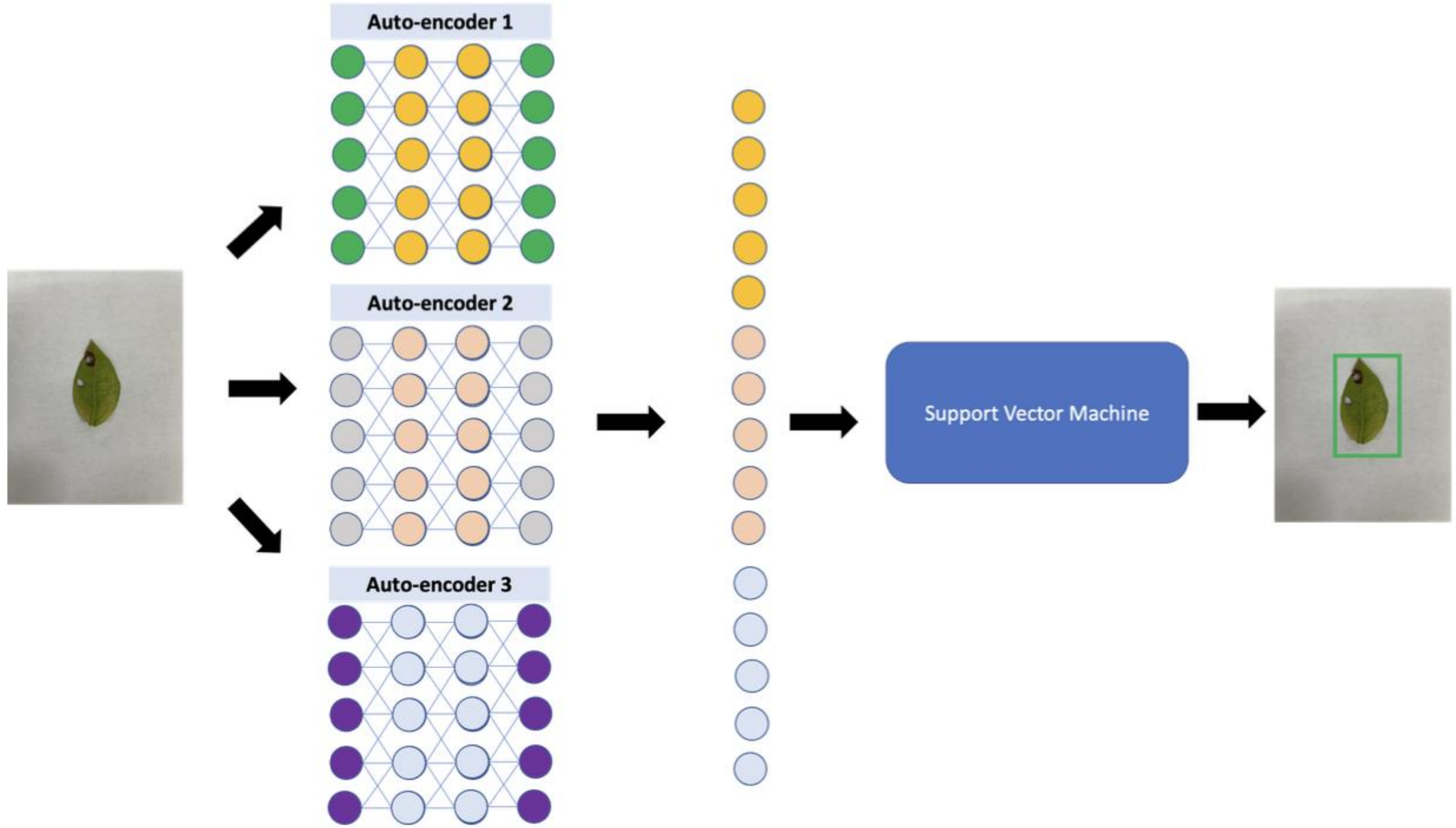

Figure 1 The auto-encoder is used to generate three robust groups of features for training support vector machine [2]

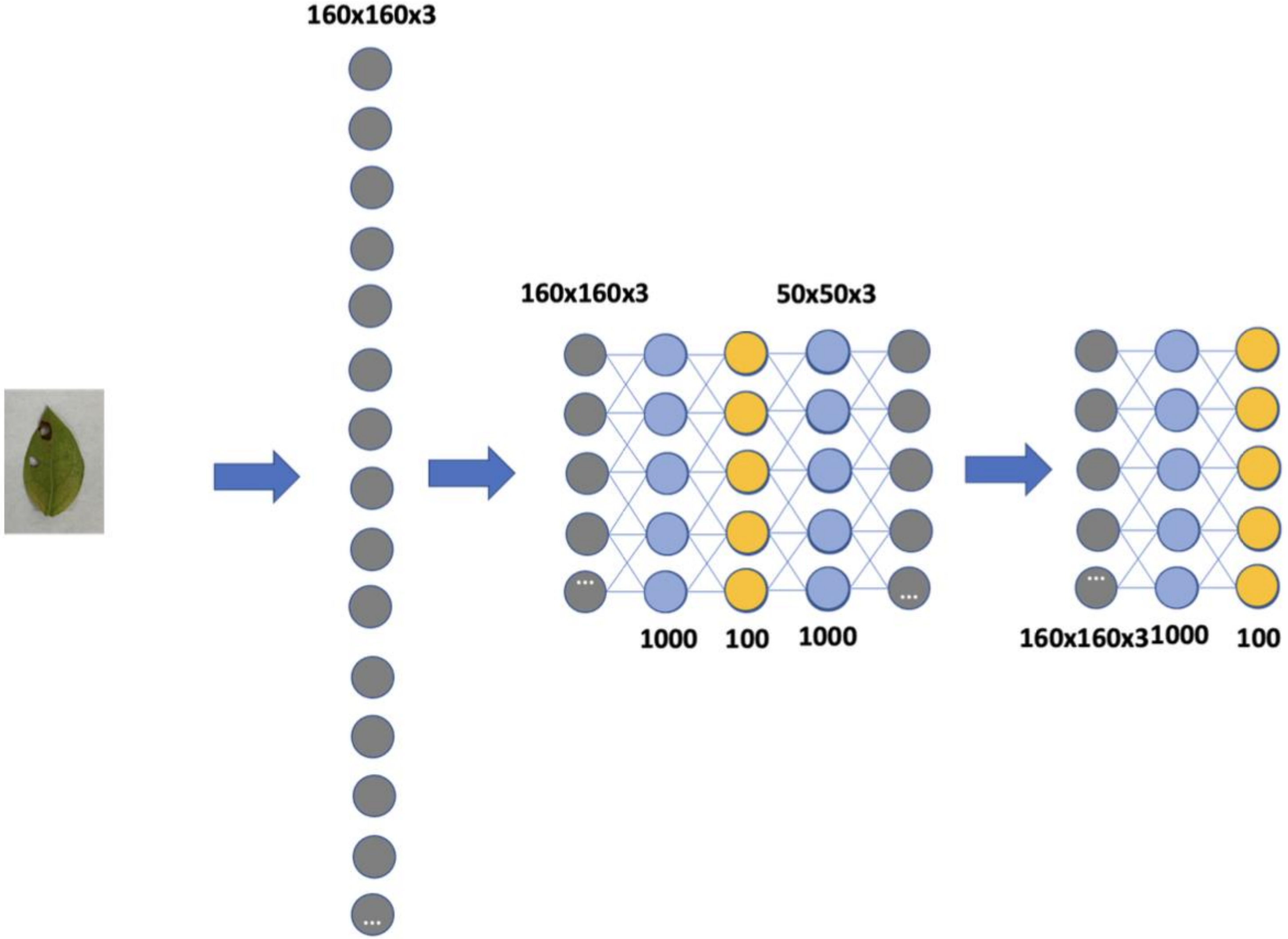

Figure 2 The first proposed auto-encoder to learn and reduce input features (160x160x3) to 100 features.

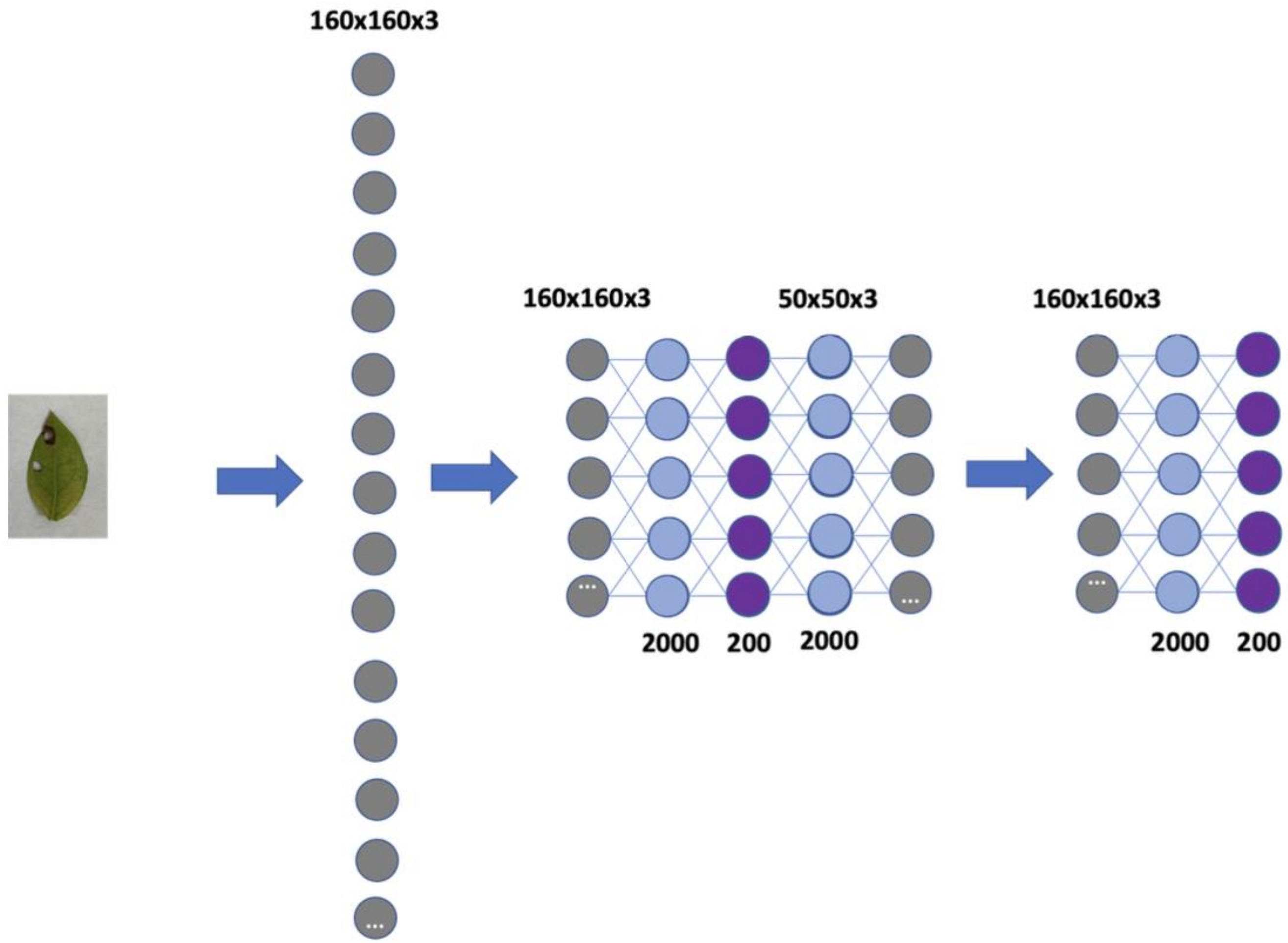

Figure 3 The second proposed auto-encoder to learn and reduce input features (160x160x3) to 200 features.

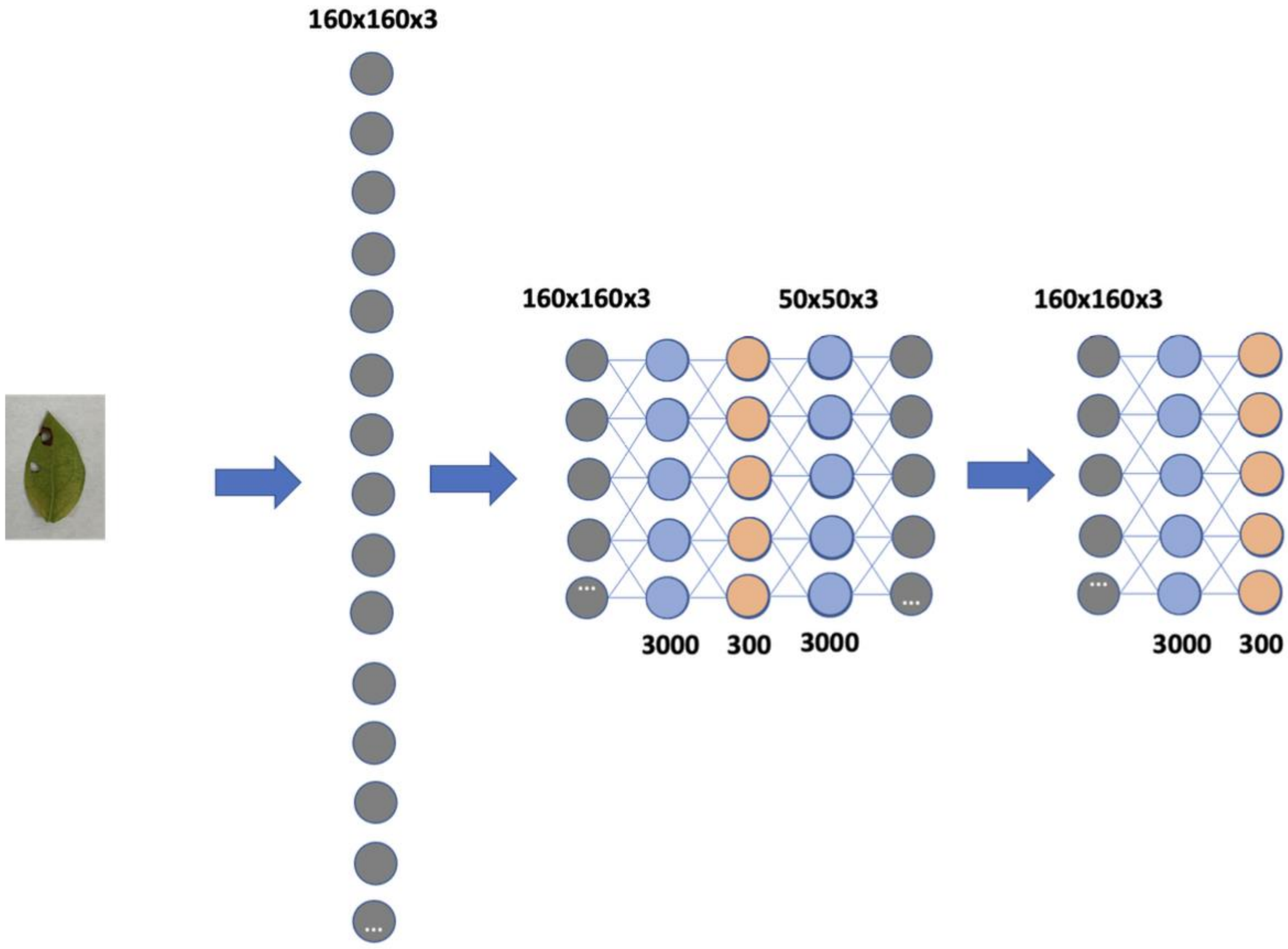

Figure 4 The third proposed auto-encoder to learn and reduce input features (160x160x3) to 300 features.

Experimental Results

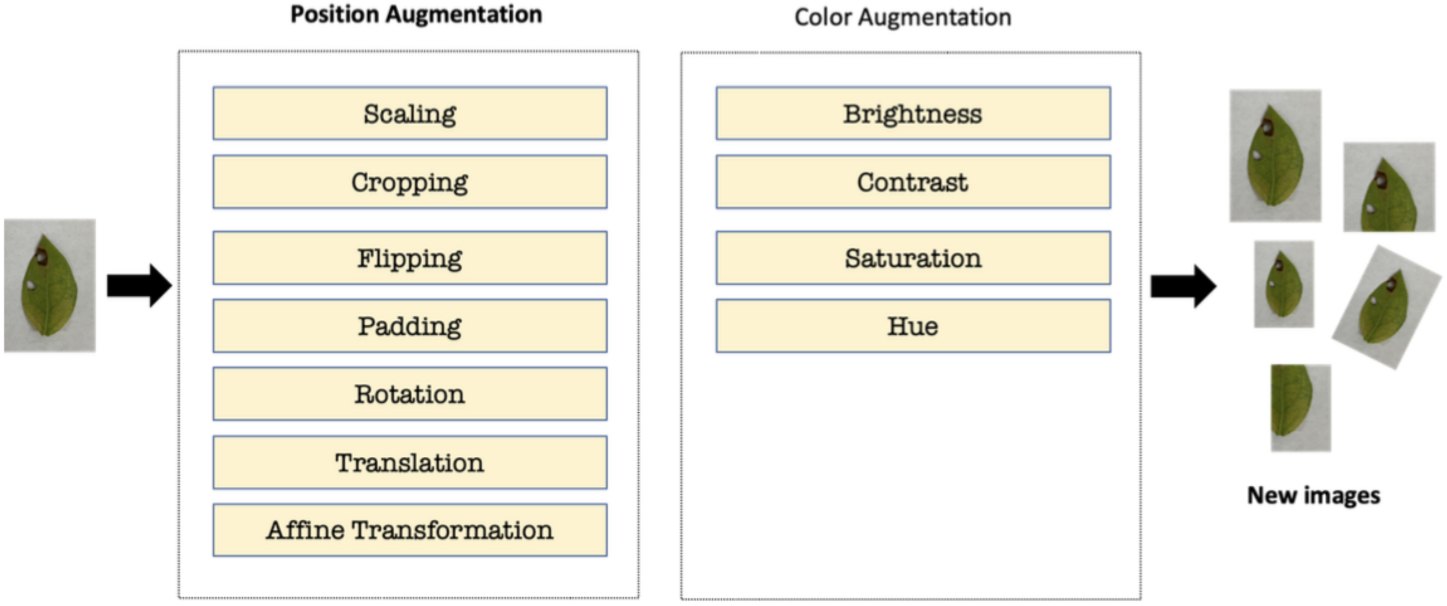

Figure 5 Flow chart of generating images for training and testing using the augmentation techniques[11]

Experimental Results

- A cell phone Samsung Galaxy A33 5G was setup to capture and label blueberry images for training and testing. The dataset was built at Can Tho University, Vietnam.

- The total number of images captured and labeled is 200 images. However, the captured images were still not enough to train and evaluate the accuracy of the proposed method. Various common augmentation techniques [11] were applied to generate 1000 images for training, 200 images for testing, and 200 images for validation as shown in Figure 5.

- The proposed system was trained by using free GPU Google CoLab provided by Google. The accuracy of the proposed method with the original support vector machine (SVM) [1] were evaluated by using three popular metrics: precision., recall, and F1 score.

Figure 6. Blueberry detection results by using our dataset under normal condition. Left image is the results of the proposed method. Right image is the results of the SVM.

Figure 7. Blueberry detection results by using our dataset under noisy conditions. Left image is the results of the proposed method. Right image is the results of the SVM.

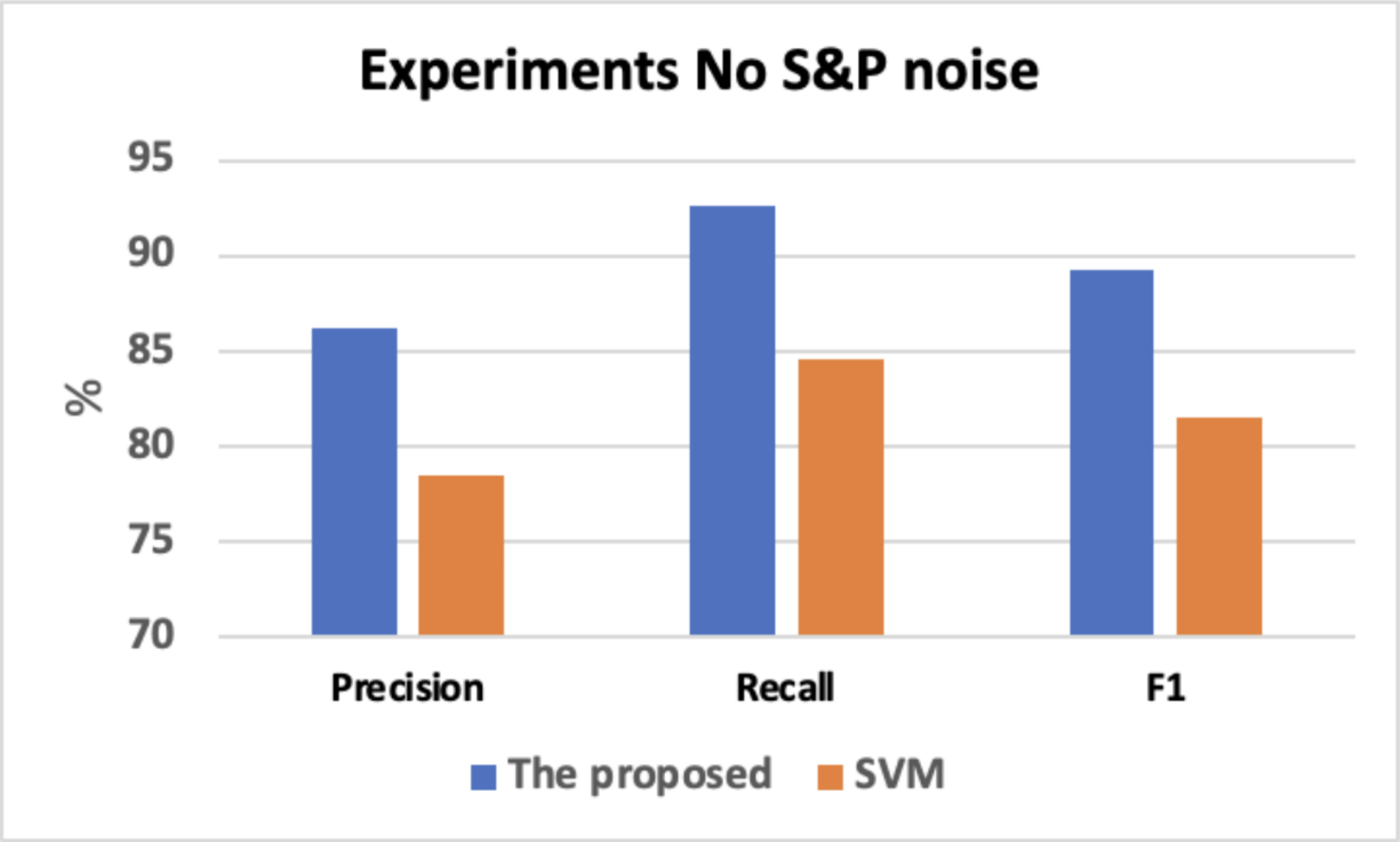

Figure 8. Precision and Recall of the proposed method and support vector machine under the testing condition without adding Salt and Pep noise

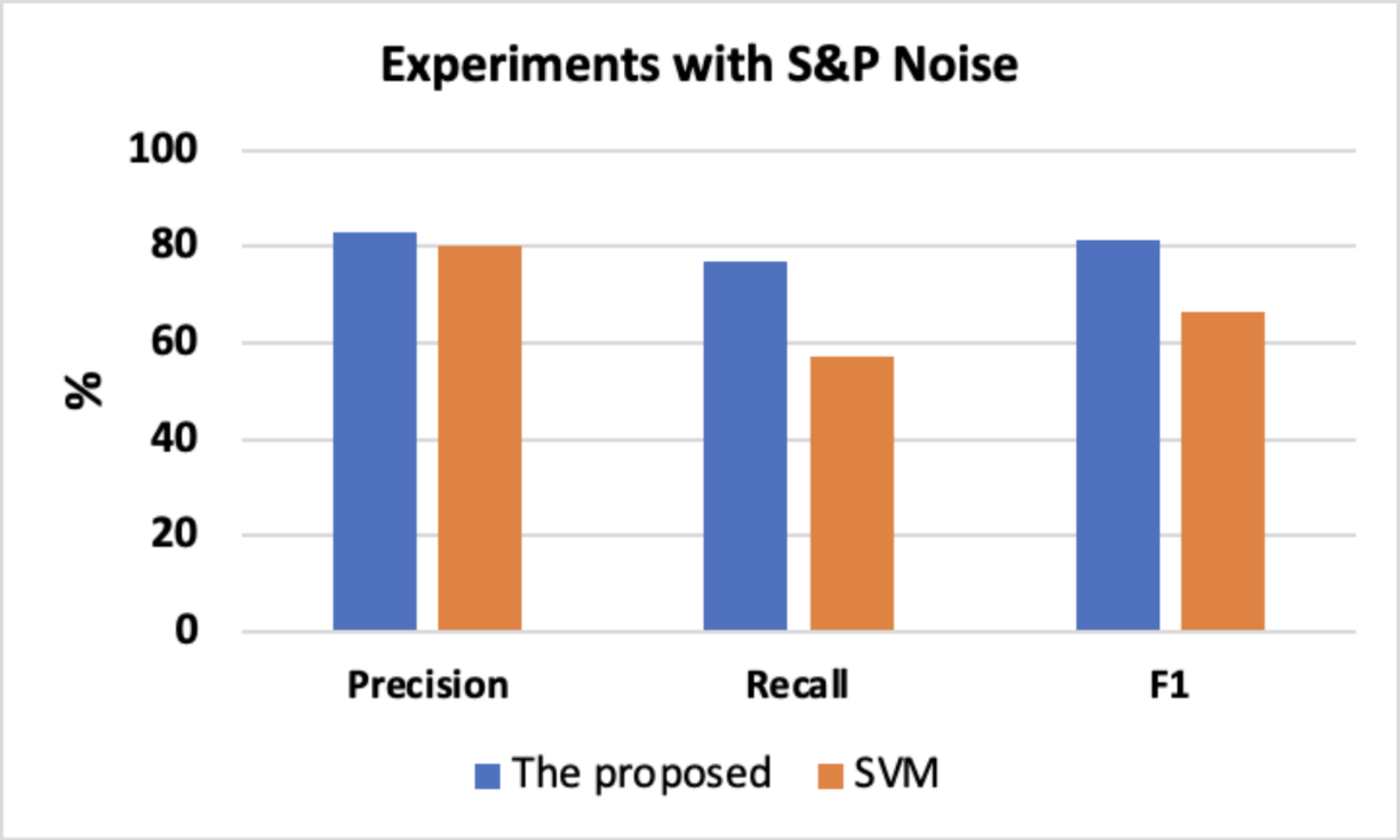

Figure 9. Precision and Recall of the proposed method and support vector machine under the testing condition by adding salt and pepper noise

Conclusions

- Blueberry leaf disease detection is really important to help farmers to early detect and find a way to reduce the cost. The proposed method, by combining the unsupervised method (auto-encoder) and supervised method (SVM), can detect and classify blue- berry leaf disease under various testing environments. However, the processing time of our system is still slow due to the proposed auto-encoder for learning and extracting three groups of features. Therefore, in the future, The effect of various auto-encoder variants are investigated in order to improve the performance of our system.

Thank You!