Chapter 5
Results and Conclusion

The study of numerous classification techniques to detect different types of mung leaf diseases is presented in this thesis. Dataset of images are created and processed by image processing methods. In this chapter, the level of accuracies achieved by the classifiers are described via chapter 4. The performance of the classification algorithms are measured by the parameters such as accuracy, precision, recall, and f1 score. Amongst these output performance metrics of classification, the primary focus is on the achieved accuracy level of various classifiers.

5.1 Overall Performance of Accuracy Classifier wise

In Chapter 4 the research work is held with the consideration of three diseases of the mung bean plant leaf and a healthy category in three different dataset environments namely controlled, uncontrolled and combined (controlled + uncontrolled) environments. Seven different classifiers namely SVM, KNN, AdaBoost, GaussianNB, DTC, LogisticRegression and Custom CNN are enabled for classifying the diseases and the healthy images of leaves. For controlled environment in first round total three model were classified and compared with each other based on their accuracies and loses. They are VGG16, MobileNetV2 and Custom CNN. This basic test shows that MobileNetV2 is not suitable for this problem as it overfits the training data despite being less complex than VGG16. VGG16 and Custom CNN both performs well. So for other two environments uncontrolled environment and combined environment Custom CNN is selected.

Below Table 5.1dispalys the overall accuracy achieved by respected classifier in different dataset environments.

	Environment					
	Controlled		Uncontrolled		Combined	
Classifiers	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**
SVM	100%	86.9%	100%	89.00%	99.90%	89.00%
KNN	68.38%	64.41%	43.03%	37.03%	99.59%	73.03%
DTC	85.32%	50.45%	85.75%	35.18%	67.75%	46.06%
AdaBoost	65.05%	63.15%	49.36%	37.96%	60.18%	50.00%
GaussianNB	70.95%	70.27%	63.29%	49.07%	59.57%	50.30%
LogisticRegression	100%	79.72%	100%	59.25%	100%	69.00%
VGG16	93.65%	95.5%	NA	NA	NA	NA
MobileNetV2	99.8%	35.14%	NA	NA	NA	NA
Custom CNN	99.24%	95.05%	99.69%	87.88%	98.81%	90.69%

Table 5. 1: Summary of accuracy achieved by classifiers

From above Table 5.1 we can see that from all the machine learning algorithms SVM gives higher accuracy as hypothesized as in hypothesis H1 but not accurate as deep learning model custom CNN also SVM overfits the dataset.

Custom CNN works well in all the three environments compared to other six classifiers. Custom CNN achieves 99.24% training and 95.05% testing accuracies in controlled environment. MobileNetV2 model performs very poorly and also overfits the dataset with a huge difference. VGG16 also performs well in controlled environment, but we select Custom CNN to train for rest of the two environments uncontrolled and combined. It achieves 99.69% training and 87.88% testing accuracies in uncontrolled environment. Custom CNN again secures 98.81% training and 90.69% testing accuracies in combined environment as hypothesized as in hypothesis H2.

SVM also overfits the dataset. Other five algorithms also does not provide optimum solution. Overall Custom CNN performs best in all the three environments as

hypothesized as in hypothesis H3. The same comparison is displayed using below Figure 5.1.

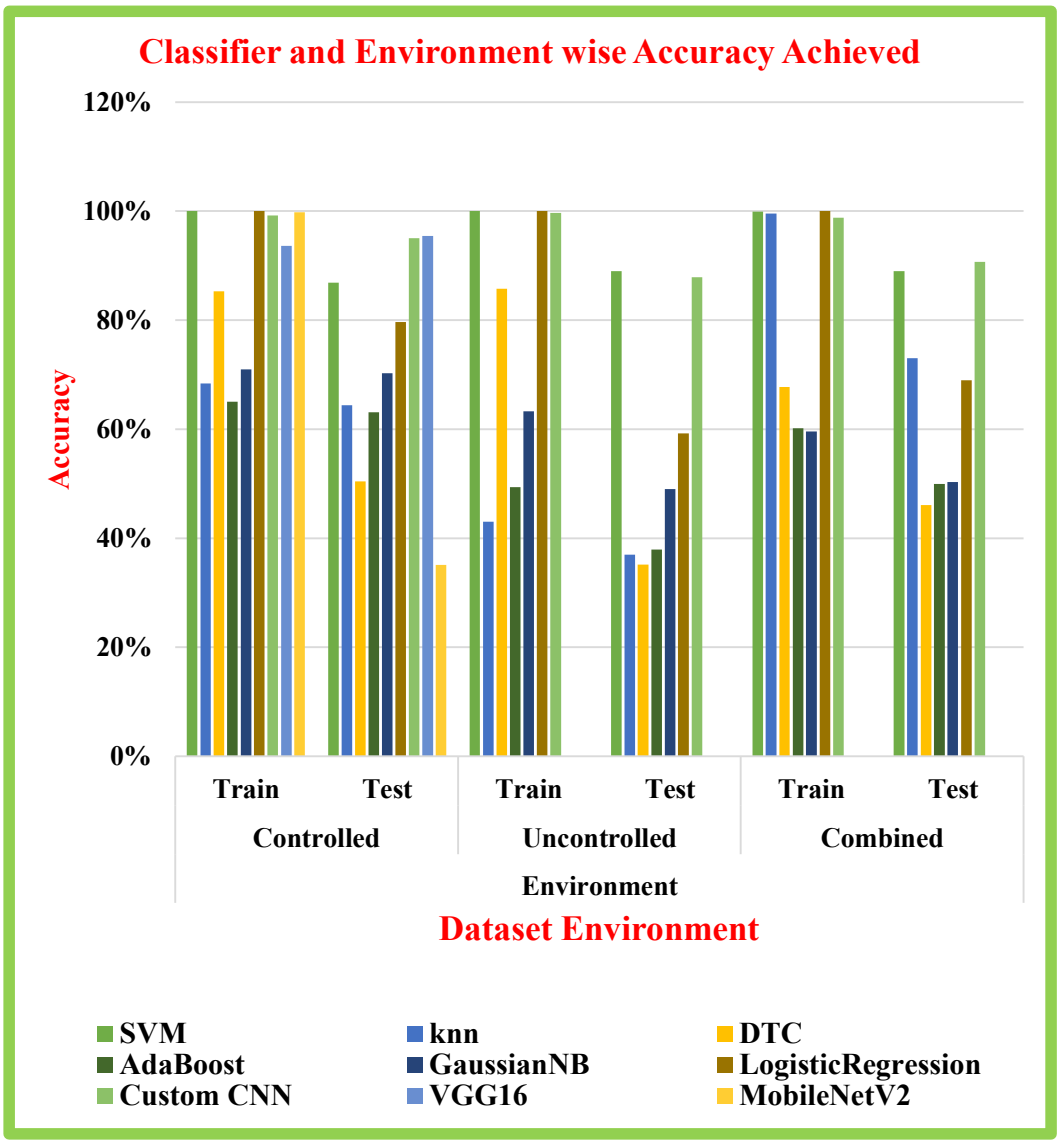

Figure 5.1: Classifier and environment wise accuracy achieved

From the results of the classification process, better performance in terms of accuracy, is gained from the Custom CNN classifier than the other classifiers.

5.2 Dataset Environment wise Comparison

In this study the work has been conducted in three different dataset environments controlled, uncontrolled and combined.

5.2.1 Controlled Environment

In controlled environment an image contains only a single mung leaf with a white background and no noise. Below Table 5.2 shows accuracies achieved by different classifiers in controlled environment. And the comparison of the same is displayed graphically using Figure 5.2.

Controlled		
Classifier	**Train**	**Test**
SVM	100%	86.40%
KNN	68.38%	64.41%
AdaBoost	65.05%	63.15%
GaussianNB	70.95%	70.27%
DTC	85.32%	50.45%
LogisticRegression	100%	79.72%
Custom CNN	99.24%	95.05%

Table 5.2: Accuracy achieved in Controlled Environment

From above table we can see that Custom CNN performs best in training as well as testing phases. SVM overfits the dataset and other algorithms also failed to perform well in controlled environment.

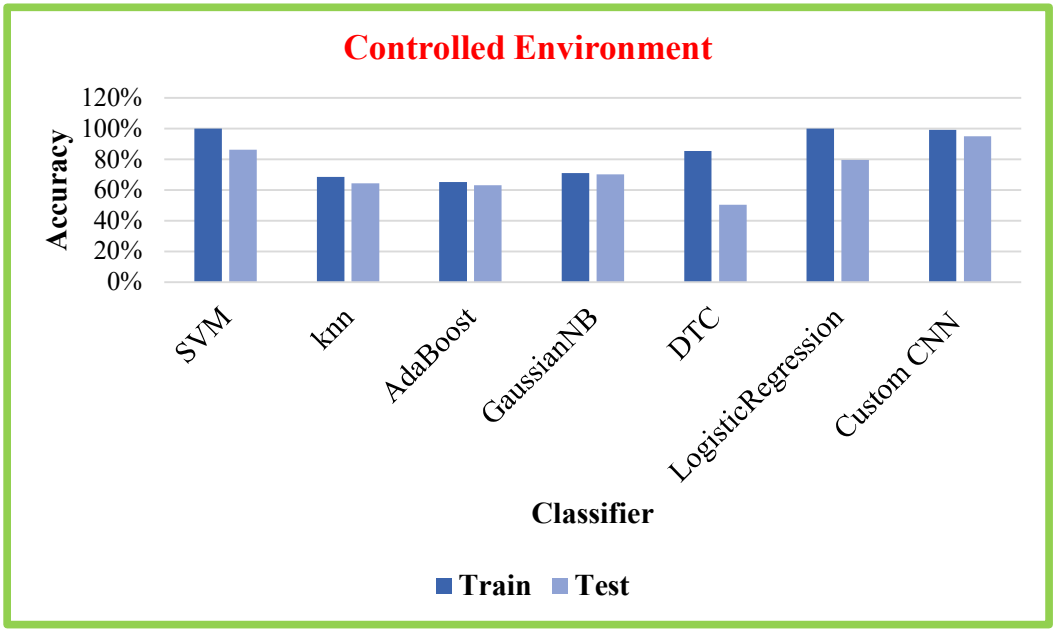

Figure 5.2: Accuracy achieved in Controlled Environment

5.2.2 Uncontrolled Environment

In uncontrolled environment an image contains mung leaf with other details like soil, other leaves with main leaf. Below Table 5.3 shows accuracies achieved by different classifiers in uncontrolled environment. And the comparison of the same is displayed graphically using Figure 5.3.

Uncontrolled		
Classifier	**Train**	**Test**
SVM	100%	89.00%
KNN	43.03%	37.03%
AdaBoost	49.36%	37.96%
GaussianNB	63.29%	49.07%
DTC	85.75%	35.18%
LogisticRegression	100.00%	59.25%
Custom CNN	99.69%	87.88%

Table 5.3: Accuracy achieved in Controlled Environment

From above table we can see that Custom CNN performs best in training as well as testing phases for uncontrolled environment. SVM overfits the dataset and other algorithms also failed to perform well in uncontrolled environment.

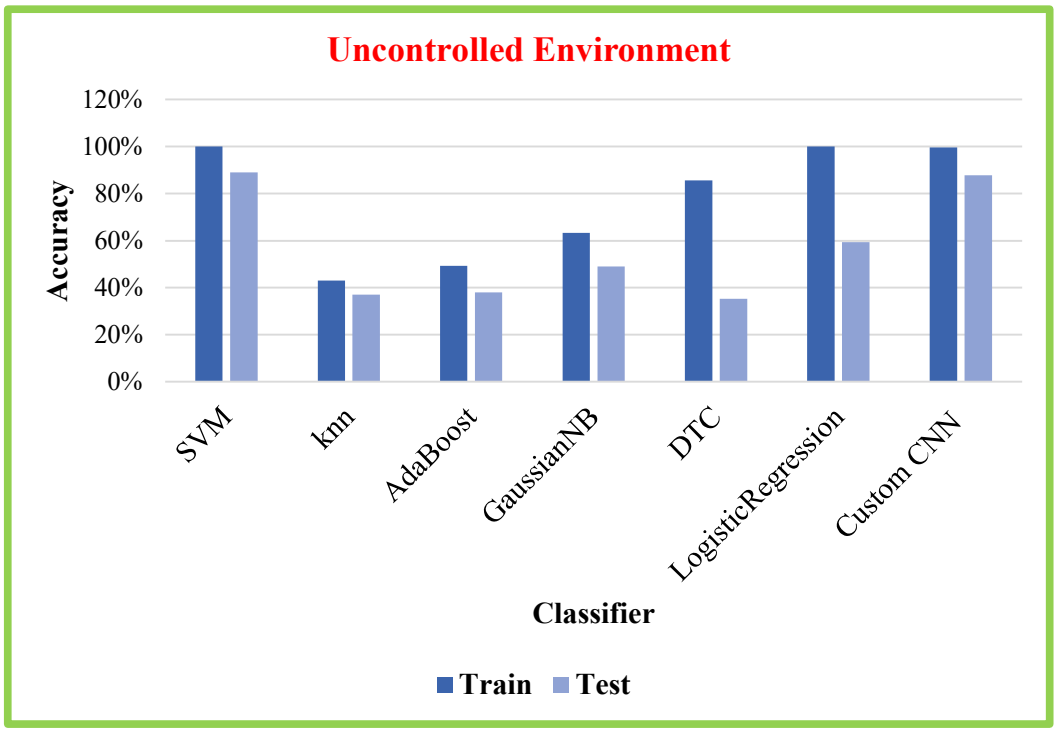

Figure 5.3: Accuracy achieved in Uncontrolled Environment

5.2.3 Combined Environment

In combined environment dataset of controlled and uncontrolled environment are merged together. Below Table 5.4 shows accuracies achieved by different classifiers in combined environment. And the comparison of the same is displayed graphically using Figure 5.4.

Combined		
Classifier	**Train**	**Test**
SVM	100%	89.00%
KNN	99.59%	73.03%

AdaBoost	60.18%	50.00%
GaussianNB	59.57%	50.30%
DTC	67.75%	46.06%
LogisticRegression	100%	69.00%
Custom CNN	98.81%	90.68%

Table 5.4: Accuracy achieved in Combined Environment

From above table we can see that Custom CNN performs best in training as well as testing phases for combined environment also. SVM overfits the dataset and other algorithms also failed to give best in combined environment.

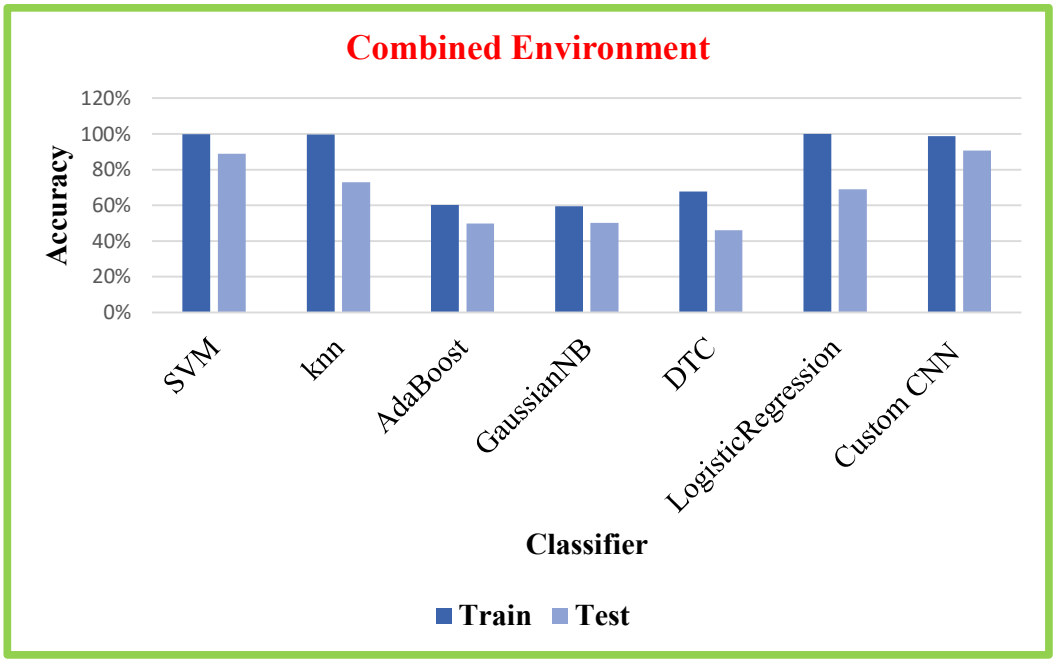

Figure 5.4: Accuracy achieved in Combined Environment

In conclusion the results are as hypothesized as in hypothesis H3 that from all the three environments it is observed that controlled environment gives better results compared to uncontrolled and combined environment. In machine learning, a huge sample of data is needed to decently predict any activity, as well as uncontrolled environment images contains noise which effects on overall accuracy achieved. In

controlled environment the image contains a single leaf image at its centre and a white background i.e. no noise, which makes segmentation process easier.

5.3 Disease wise Comparison of machine learning algorithms

Here performance of each classifier is analyzed based on the leaf category that is three diseases and one healthy category.

5.3.1 Cercospora Leaf Spot Disease

Below Table 5.5 shows accuracy achieved by different classifiers for Cercospora Leaf Spot disease. And Figure 5.5 shows the graphical presentation for the same.

	Controlled	Uncontrolled	Combined
SVM	85	45	72
KNN	73	34	72
AdaBoost	40	7	50
GaussianNB	69	21	50
DTC	54	20	44
LogisticRegression	81	50	69

Table 5.5: Accuracy achieved for Cercospora Leaf Spot disease

From above Table we can see that for Cercospora Leaf Spot disease SVM achieves highest accuracy compared to other classifiers for controlled and uncontrolled environment. For combined environment SVM and KNN gives equal accuracies.

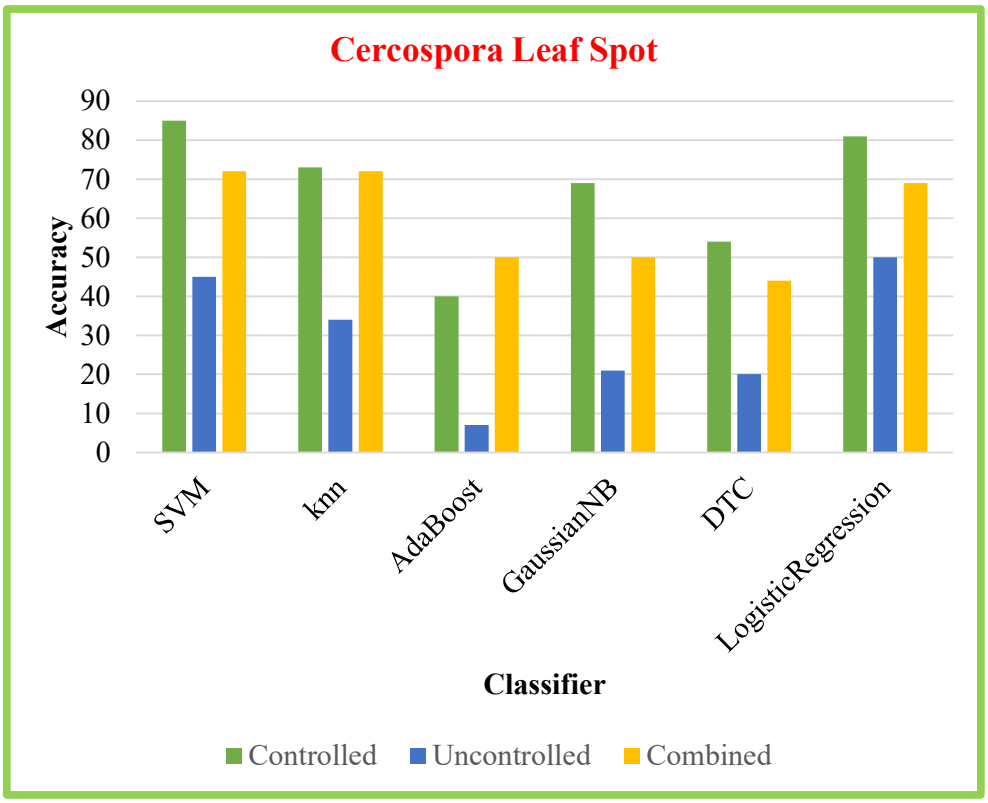

Figure 5.5: Accuracy achieved for Cercospora Leaf Spot disease

5.3.2 Healthy Leaf

Below Table 5.6 shows accuracy achieved by different classifiers for Healthy Leaf category. And Figure 5.6 shows the graphical presentation for the same.

	Controlled	Uncontrolled	Combined
SVM	90	74	77
KNN	53	55	77
AdaBoost	54	54	52
GaussianNB	76	59	41
DTC	37	44	48
LogisticRegression	76	70	70

Table 5.6: Accuracy achieved for Healthy Leaf category

From above Table we can see that for Healthy Leaf category SVM achieves highest accuracy of 90% which is good compared to other classifiers for controlled environment. For uncontrolled environment also SVM gives best accuracy of 74%. For combined environment SVM and KNN gives equal accuracies of 77%.

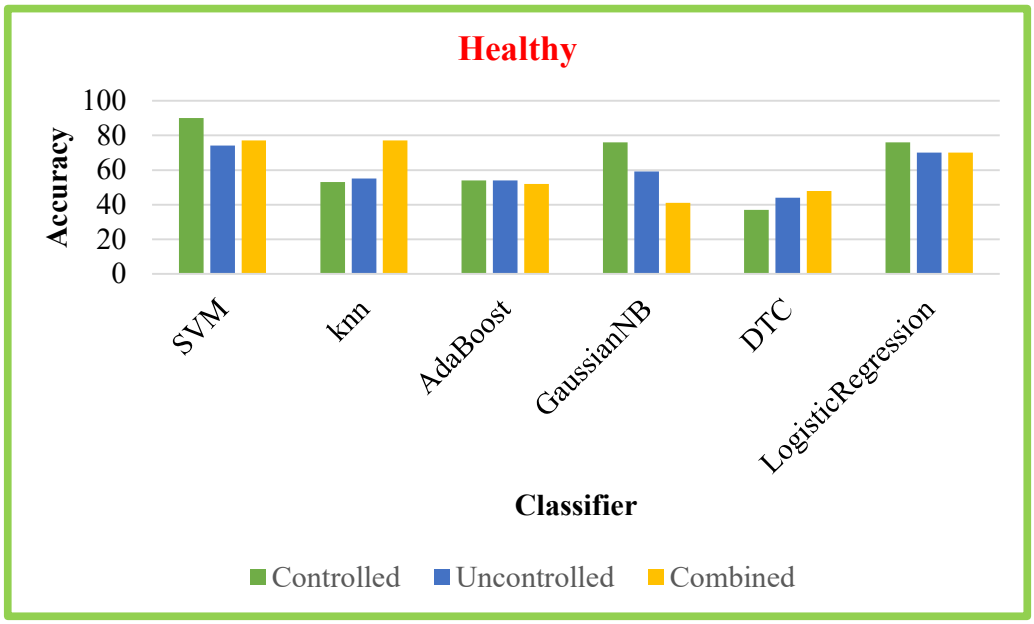

Figure 5.6: Accuracy achieved for Healthy Leaf Category

5.3.3 Powdery Mildew disease

Below Table 5.7 shows accuracy achieved by different classifiers for Powdery Mildew Disease. And Figure 5.7 shows the graphical presentation for the same.

	Controlled	Uncontrolled	Combined
SVM	90	44	77
KNN	65	35	77
AdaBoost	71	17	56
GaussianNB	77	34	58
DTC	53	11	45
LogisticRegression	87	0	76

Table 5.7: Accuracy achieved for Powdery Mildew

From above Table we can see that for Powdery Mildew SVM achieves highest accuracy of 90% which is good compared to other classifiers for controlled environment. For uncontrolled environment also SVM gives best accuracy of 44%. For combined environment SVM and KNN gives equal accuracies of 77%.

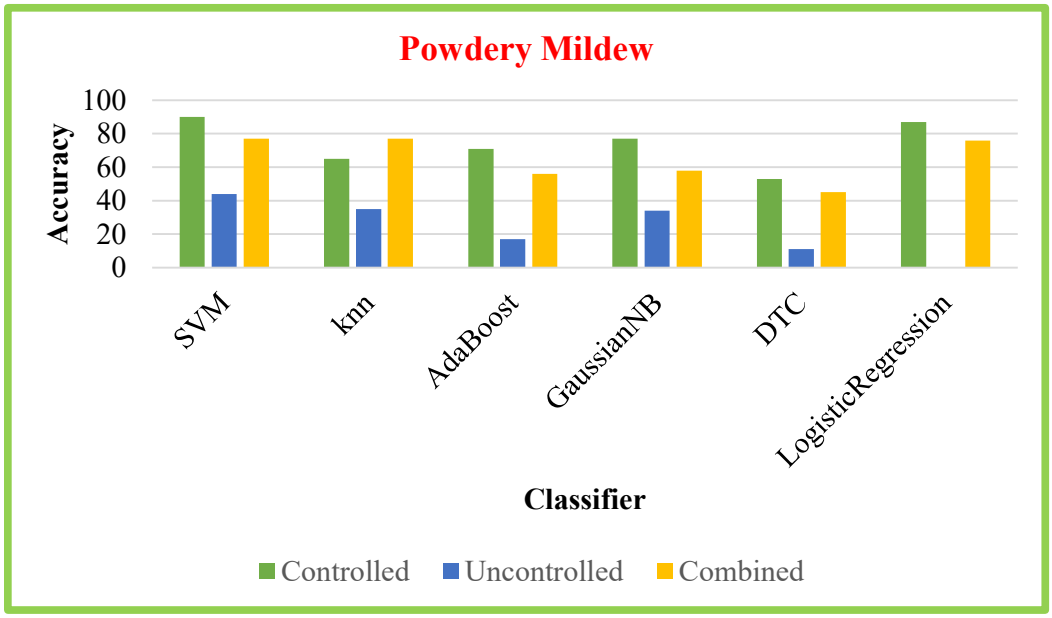

Figure 5.7: Accuracy achieved for Powdery Mildew

5.3.4 Yellow Mosaic Virus disease

Below Table 5.8 shows accuracy achieved by different classifiers for Powdery Mildew Disease. And Figure 5.8 shows the graphical presentation for the same.

	Controlled	Uncontrolled	Combined
SVM	82	68	68
KNN	64	0	68
AdaBoost	47	6	44
GaussianNB	61	58	52
DTC	56	43	47
LogisticRegression	74	63	66

Table 5.8: Accuracy achieved for Yellow Mosaic Virus

From above Table we can see that again for Yellow Mosaic Virus SVM achieves highest accuracy of 82% for controlled environment. For uncontrolled environment also SVM gives best accuracy of 68%. For combined environment SVM and KNN gives equal accuracies of 68%.

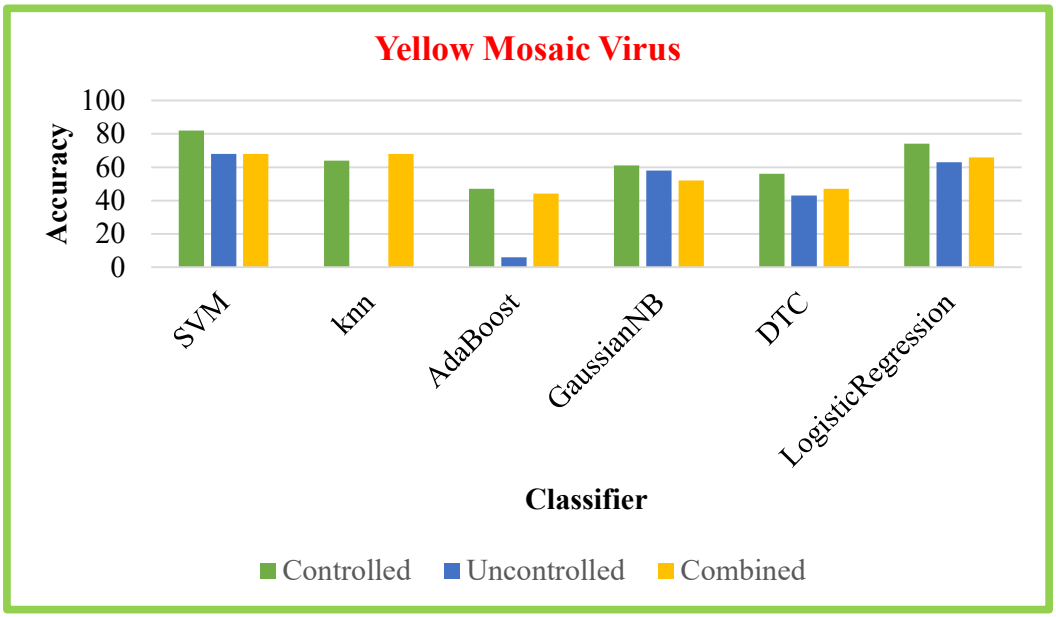

Figure 5.8: Accuracy achieved for Yellow Mosaic Virus

From above sections 5.2 to 5.3 we can see that SVM gives better accuracy compared to other algorithms. But as we have seen in Chapter 4 SVM overfits in all the three environments. And all other machine learning algorithms are also failed to give optimum solution.

5.4 Machine learning (ML) algorithm Vs. Deep Learning (DL) Algorithm

Since SVM gives highest accuracy amongst all the machine learning (ML) algorithms, in this section we are going to compare the results of SVM with Custom CNN for all the three environments. The overall comparison is represented in below Table 5.9. And graphical presentation for the same is shown in Figure 5.9.

	Controlled		Uncontrolled		Combined	
	Train	**Test**	**Train**	**Test**	**Train**	**Test**
SVM	100%	86.40%	100%	89.00%	100%	89.00%
Custom CNN	99.24%	95.05%	99.69%	87.88%	98.81%	90.68%

Table 5.9: Comparison of ML Vs. DL Algorithm

The deep learning algorithm Custom CNN gives decent results in all the three environments without overfitting. So Custom CNN proved to be best algorithm amongst all the seven algorithms used for classification of mung leaf plant diseases.

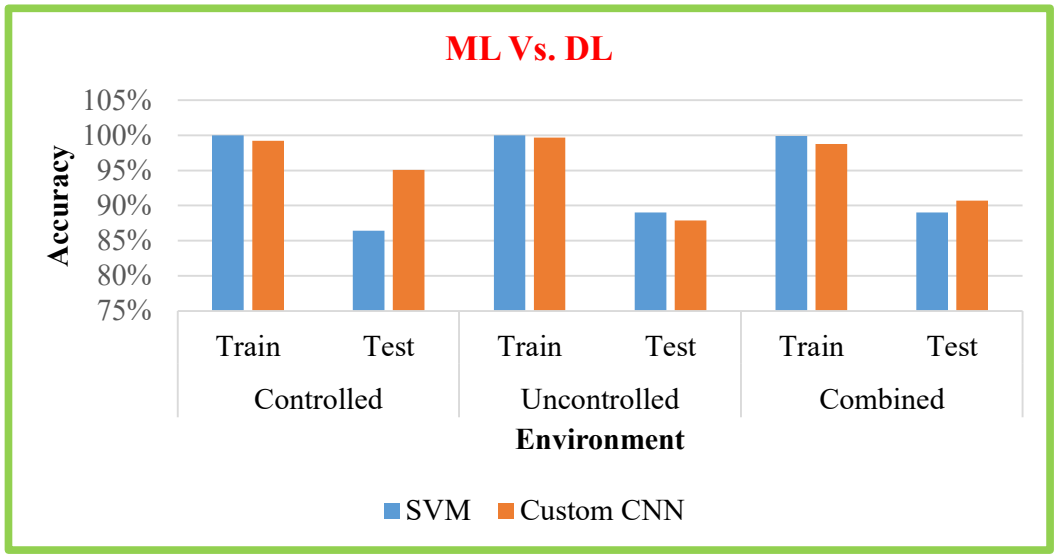

Figure 5.9: Comparison of ML Vs. DL Algorithm

5.5 Comparison with previously done work

Below Table 5.10 presents summary of previously done work in the area of leaf disease identification. Various researcher has tried different techniques for disease identification. Present study achieves better accuracy compared to other work done in the area.

Author	Core Analysis Tool	Plant	Diseases	Environment	Accuracy	Comparison with present study
Akbar Hidayatuloh et al. (2018)	CNN	Tomato	6 Diseases	Uncontrolled, Single Leaf	86.92%	< present study
Sivasakthi (2020)	SVM & ANN	Green house Crop	1 Disease	Controlled	92% (SVM) 87% (ANN)	< present study < present study
Le et et al. (2019)	SVM	NA	Growth Monitoring	Images of plant in testing facility under simulated field	91.85%	< present study
Suresha et al. (2017)	KNN	Padddy Plant	2 Diseases	Controlled	76.59%	< present study
Saleem et al. (2019)	KNN	N/A	Plant Classification	Controlled	96%	< present study
Kawcher Ahmed et al. (2019)	KNN	N/A	3 Diseases	Controlled	97%	< present study
Sridhathan et al. (2018)	K-Mean	N/A	3 Diseases	Uncontrolled, Single Leaf	98.27%	≲ present study
Dhingra et al. (2019)	SVM	Basil	NA	Controlled	98%	≲ present study
Sun et al. (2019)	SVM	Tea Plant	NA	Uncontrolled, Single Leaf	98.5%	≲ present study

Kulkarni et al. (2021)	Random forest classifier	Apple, Corn, Grapes, Potato, Tomato	Multiple	Controlled	93%	< present study
Deepalakshmi et al. (2021)	CNN	N/A	NA	Controlled	94.5%	< present study

Table 5.10: Summary of previous work

As shown in above Table many researchers have done work in controlled environment or uncontrolled environment. Present work is carried out using three different environments controlled, uncontrolled and combined environments. Also previous research work in uncontrolled environment consider only single leaf and soil as background, whereas present study works on multiple leaves with soil, stems and other leaves as background.

5.6 Conclusion

In agriculture, one of the developing research area is computerization of identification and classification of various mung leaf diseases. The computerization may help to improve quantity and quality of the crop yield. Manual identification method is very time consuming and may also provide lack of accuracy for the farmers. Therefore, it requires an image processing technique for timely and accurately detection and classification of plant diseases. In large farms, it is not possible to monitor the crop using manual or traditional methods. The computerized technique of image processing is adapted for the disease detection. On mung bean plant leaves, which uses color information to detect the disease on mung bean plant leaf. From the stated information from the thesis, an interface is encouraged to the farmers of mung bean for the early recognition of mung bean plant leaf diseases. It has been achieved by considering three different mung bean plant leaf diseases such as, CercosporaLeaf Spot, Powdery Mildew, and Yellow Mosaic Virus. The key contribution of the presented study is the early detection of mung bean plant leaf diseases. For all the chapters, presented in this

thesis the images are captured from the Navsari Agriculture University, Navsari, Gujarat and farms of nearer villages, for the collection of three different disease infected leaf images and healthy leaf images of mung bean plant. The study will help mung bean farmers who generally depend on the agricultural experts or advisors for the disease diagnosis.

In the stage of pre-processing, the original RGB image is transformed into gray scale image. The segmentation is done using HOG. The final classification process of the discussed chapters are done with the application of seven different classifiers such as SVM, KNN, AdaBoost, GaussianNB, DTC, LogisticRegression and Custom CNN.

The performance level of different classifiers in the contribution of detecting disease of mung leaf crop is discussed in the Chapter 4 and 5. Custom CNN classifier delivers superior accuracy outcomes, when discovering the diseases of the plant.

In future quantification or severity of the diseases can be measured and also more diseases can be considered apart from the three diseases covered in the present study.

References

1. Akbar H, Nursalman M, Nugraha E. Identification of Tomato Plant Diseases by Leaf Image Using Squeezenet Model. International Conference on Information Technology Systems and Innovation (ICITSI) Bandung - Padang. October 22–25, 2018.
2. Sivasakthi, S.: Plant leaf disease identification using image processing and svm, ann classifier methods. International Conference on Artificial Intelligence and Machine learning. Journal of Analysis and Computation (JAC). (2020).
3. Le, V.N.T., Apopei, B., Alamesh, K.: Effective plant discrimination based on the combination of local binary pattern operators and multi-class support vector machine methods. Inf. Process. Agricult. 6, 116–131 (2019).
4. Suresha M, Shreekanth K N, Thirumalesh B V (2017). Recognition of diseases in paddy leaves using KNN classifier. 2017 2nd International Conference for Convergence in Technology (I2CT).
5. Saleem, G., Akhtar, M., Ahmed, N., Qureshi, W.S.: Automated analysis of visual leaf shape features for plant classification. Comput. Electron. Agricult. 157, 270–280 (2019).
6. Kawcher A, Tasmia R, Syed M. Irfanul A, Sifat M (2019). Rice leaf disease detection using machine learning techniques. International Conference on Sustainable Technologies for Industry 4.0 (STI), IEEE, (pp. 1–5). Dhaka, Bangladesh: IEEE..
7. Sridhathan, C., Senthil Kumar, M.: Plant Infection Detection Using Image Processing. Int J Mod Eng Res 8(7), 13–16 (2018).
8. Dhingra, G., Kumar, V., Joshi, H.D.: A novel computer vision based neutrosophic approach for leaf disease identification and classification. Measurement 135, 782–794 (2019)
9. Sun, Y., Jiang, Z., Zhang, L., Dong, W., Rao, Y.: SLIC_SVM based leaf diseases saliency map extraction of tea plant. Cpmput. Electron. Agricult 157, 102–109 (2019)
10. Kulkarni, P., Karwande, A., Kolhe, T., Kamble, S., Joshi, A., & Wyawahare, M. (2021). Plant Disease Detection Using Image Processing and Machine Learning. arXiv preprint arXiv:2106.10698.

11. Deepalakshmi P., Prudhvi Krishna T., Siri Chandana S., Lavanya K. and Parvathaneni Naga Srinivasu.: Plant Leaf Disease Detection Using CNN Algorithm. International Journal of Information System Modeling and Design (IJISMD)12(1), 2021