LeafNST: An Improved Data Augmentation Method for Classification of Plant Disease using Object-Based Neural Style Transfer

Om Khare[1], Dr. Sunil Mane[1], Harshmohan Kulkarni[1], and Ninad Barve[1]

College of Engineering Pune Technological University
{khareom20.comp, sunilbmane.comp, barvenm20.comp, kulkarnihs20.comp}@coeptech.ac.in

Abstract. Plant diseases significantly threaten global agriculture, impacting crop yield and food security. Nearly 30% of the crop yield is lost due to plant diseases. Efficient identification and classification of plant diseases through computer vision techniques have become imperative for timely intervention. However, popular plant disease datasets often suffer from data imbalance, with certain classes underrepresented, hindering the performance of machine learning models. Traditional data augmentation methods, such as rotation and flipping, are limited in their effectiveness, especially when faced with imbalanced datasets. To address this limitation, we explore advanced data augmentation techniques, including Generative Adversarial Networks (GANs) such as CycleGAN and LeafGAN, which have shown promise in generating synthetic images. However, we propose an innovative approach of Object-based single Style Transfer on a single neural network for augmenting the plant disease dataset. This technique focuses on mitigating data imbalance issues within datasets, which can adversely affect the model's ability to generalize across diverse classes. The proposed method is compared with state-of-the-art data augmentation techniques, highlighting its superiority in addressing data imbalance issues. Our approach aims to produce more realistic and diverse synthetic images, leading to improved model generalization and accuracy in plant disease classification tasks validated using different classifiers. The efficiency of our approach is validated through extensive experimentation and benchmarking against existing methods.

Keywords: Data augmentation · Neural Style Transfer · Plant Disease · Convolutional Neural Networks · Object Detection · Image Processing

1 Introduction

The rapid evolution of agricultural technologies has highlighted the critical importance of plant health management in ensuring global food security and sustainability. As such, the ability to accurately identify and classify plant diseases using technological advancements is not only beneficial but necessary for modern agriculture. This introduction provides a comprehensive overview of the challenges and innovative approaches in the field of plant disease classification using machine learning.

1.1 Background and context

In the realm of agricultural technology, the efficient and accurate identification and classification of plant diseases play a pivotal role in safeguarding crop yield and ensuring food security globally. With plant diseases causing nearly 40% of crop yield losses [8] [17], leveraging computer vision techniques for timely and effective disease detection has become increasingly critical. This backdrop sets the stage for exploring innovative solutions that enhance the performance and generalization capabilities of machine learning models in identifying plant diseases.

1.2 Problem Statement

One of the significant challenges encountered in the classification of plant diseases is the issue of data imbalance within existing datasets. As shown in Figure 1 certain classes within these datasets are underrepresented, which can severely impede the learning process and generalization of machine learning models. This issue can be resolved using data augmentation of imbalanced classes [20]. Traditional data augmentation methods, such as rotation and flipping, while useful, have shown limited efficacy in overcoming the challenges posed by imbalanced datasets, particularly in the context of plant disease classification.

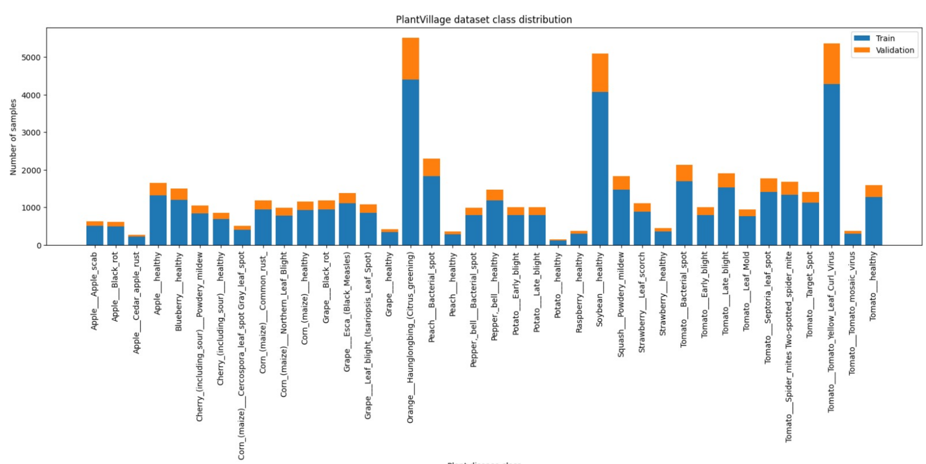


Fig. 1: PlantVillage Dataset Class Distribution

1.3 Previous Work

Our preceding research, "Improved object-based style transfer with a single deep network" [16], laid a foundational framework by introducing an innovative methodology that integrates the YOLOv8 segmentation model for object-specific style transfer using a single neural network. This methodology not only enhanced the visual aesthetics of images but also demonstrated potential applications in various domains by maintaining object characteristics during the style transfer process. This prior work established a basis for further exploration into advanced data augmentation techniques.

1.4 Objectives of current work

Building upon our previous research, the current work aims to address the aforementioned data imbalance issue in plant disease datasets. By adopting an advanced approach of CNN-based Style Transfer for data augmentation, we propose to generate synthetic images that more accurately represent underrepresented classes within these datasets. This technique seeks to improve the diversity and realism of the synthetic images, thereby enhancing the machine learning model's ability to generalize across different classes. The project will compare this novel data augmentation approach against traditional methods and state-of-the-art techniques, highlighting its effectiveness in addressing data imbalance issues and its potential to revolutionize plant disease classification tasks.

2 Literature Review

The importance of data augmentation in the field of machine learning cannot be overstated, particularly in domains where data is scarce or imbalanced. This literature review explores various data augmentation techniques, with a focus on their application and effectiveness in different scenarios, including agriculture and healthcare. We also highlight the advancements in style transfer as a means of data augmentation and identify gaps in the current research landscape.

2.1 Data Augmentation

Data Augmentation techniques are crucial in performance enhancement and generalization capabilities of machine learning models across various domains. Data Augmentation is also crucial in the case of small datasets. Multiple fields like agriculture and healthcare suffer from a lack of well-documented large-scale datasets [15]. Smaller datasets lead to under-fitted models and lower performance.

Traditional Augmentation Traditional augmentation methods include rotation, flipping, scaling, cropping, shifting, color jittering(changing the saturation or contrast values), noise addition, and others [22]. These methods are often called unsupervised augmentation methods [22]. They have been widely used to diversify the datasets and offer robust training to the models. However, these traditional methods often possess inherent limitations. These techniques may introduce artifacts or distortions that alter the semantic integrity of the data, potentially leading to overfitting and loss of semantic meaning.

Generative Adversarial Networks for Augmentation The recent development of Generative Adversarial Networks (GANs) has vastly impacted data augmentation. The generative models are designed to maintain the overall distribution of the generated data. The initial work can be seen in Image to Image translation work done by Isola et. al. [13]. The main goal is to learn the mapping between two images. Conditional Adversarial Networks have been used to perform CV tasks like image colorization, boundary filling, and generating content from edge maps in this paper. Notably, the pix2pix software is based on this work. The extension work for this paper was done by Zhu et. al. [32]. They introduced CycleGAN, which can learn the mapping between source and target domains as a whole. This deals with the case when paired images are not available, which is essential to train the pix2pix model of Isola et. al [28]. Intuition tells us that if the source and target are the same domain, CycleGANs can be applied here for data augmentation tasks. Since CycleGANs are only effective for a pair of domains, to handle the case of a multi-domain scenario, Choi et. al. introduced StarGAN [3].

Automated Augmentation Some novel augmentation techniques have been designed to automatically find the best possible augmentation policy for the given source dataset. AutoAugment, introduced by Cubuk et. al. [5] chooses the best possible policy from a pre-defined search space. The target of this technique is based on the Neural Network being trained getting the highest validation accuracy. Fast AutoAugment developed by Lim et. al. has improved the search process. Their results show that their algorithm is faster than standard AutoAugment by 33.33 times on the ImageNet dataset. Similar speedups have been observed on other datasets for image classification tasks. [19] In both the previous methods, there is a separate searching job, where the optimal augmentation policy is located based on a small part of the dataset. This separate search task is time-intensive. Newer and faster algorithms have been shown to produce similar or better results in less time [6]. AutoAugment is usually very computationally heavy and cannot run easily on consumer-grade devices. Also, it is based on selecting a single best schedule for the given task. The schedule selection process, which takes place for the child model, may not produce the best policy for the complete training period. To mitigate this problem and make the process more lightweight, Ho et. al. designed Population-Based Augmentation (PBA) [11]. PBA works by considering augmentation schedule selection as another separate hyperparameter during model training. PBS's search space is also reduced as compared to AutoAugment policies. RandAugment by Cubuk et. al. eliminated the need for a separate searching phase and a "proxy task" [6]. In AutoAugment and Fast AutoAugment, the selection of augmentation policies is performed as a separate step before training. RandAugment does not need a separate step or task.

A new method called KeepAugment by Gong et. al. [10] utilizes a relatively old concept of saliency maps. Saliency maps contain information about important regions within an image. In simple terms, they highlight the regions that grab visual attention in an image. Saliency maps can be used as a convolution visualization method as shown in [23]. KeepAugment has the main focus that the important regions in an image are to be "kept" intact in the image. Every step has an importance score calculation substep. Based on the highest score, data is augmented in such a manner that the regions scored highest are untouched.

Table 1: Comparison of the results achieved in other research using style transfer for data augmentation.
† *All metrics are accuracy or increase in accuracy unless specified*
* *Hong et. al. have provided a comparison of error rates.*
NST: Neural Style Transfer
NSGT: Neural Style Transfer + Geometric Transformations

Paper	Dataset	Model	Old Metrics†	New Metrics†
Zheng et. al. [30]	Caltech-101 and Caltech-256	VGG-16 VGG-19	83.34 84.50	85.26 85.81
Jackson et. al. [14]	Office dataset	Inception V3 ResNet 18 ResNet 50 VGG-16	78.9 39.9 48.8 55.8	88.2 49.5 61.4 55.1
Xi et. al. [26]	NA	Faster R-CNN YOLOv5	- -	Inc by 0.044 Inc by 0.024
Cicalese et. al. [4]	NA	NA	85.2	88.5
Hong et. al. [12]	CIFAR-10	PyramidNet-200	3.85*	2.55*
Darma et. al. [7]	Self Collected	MobileNet	75.4	91.6 (NSGT) 87.9 (NST)
Xu et. al. [27]	Car Dataset	Inception V2	80.3	88.8
Zhong et. al. [31]	Market-1501 DukeMTMC-reID	ResNet-50 ResNet-50	38.6 27.4	64.7 51.7

2.2 Data Augmentation using Style Transfer

Style transfer is the process of image transformation where the style of an image is applied to a target image, without affecting its contents. Neural Style Transfer was first introduced by Gatys et. al. [9]. Initially, the whole content image would be affected by the style. To counter this and make the style transfer process more targeted, [25], [21], [18], [2] implemented object detection-based style transfer. Object detection models like SAM, and DeepLab V2 have been used to segment objects in the content image. The style is applied only to those objects and the background is unaltered. Using style transfer for data augmentation has shown promising results on tasks like object detection and image classification. In many cases with small-sized datasets, there is a common style present within all the images. For example, plant diseases display a similar-looking pattern on the leaves. Table 1 lists various improvements over the baseline model as a result of style transfer-based data augmentation.

2.3 Research Gaps and Findings

Summarizing the literature review, it has been proven that data augmentation vastly helps unbalanced datasets and improves final model accuracy. In the literature review, we have encountered research utilizing style transfer-based data augmentation techniques. In all of them, an increase in the model accuracy and a decrease in the error rate after augmentation has been noted. However, the methodology of using object-based style transfer for dataset augmentation has not been explored yet.

Object detection-based style transfer models seen till now have used separate object detection and style transfer modules. The unification of both tasks into a single deep neural model and using it for data augmentation has not been explored. Some approaches have successfully shown improvements in the dataset after augmentation using NST.

Plant diseases usually show up on leaves in the form of a pattern, which can be utilized as a "style" for this method. We propose to demonstrate the innovation of our unified model, object detection-based style transfer, using the plant disease dataset.

3 Proposed Methodology

Figure 2 shows the overall architecture of the proposed methodology. The entire process is divided into three components namely the segmentation module, the style transfer module, and the blending module. In the entire process, YOLOv8 is taken as the backbone architecture.

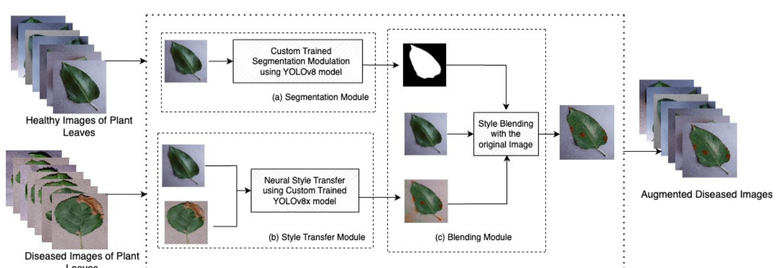


Fig. 2: Data Augmentation Architecture

3.1 Segmentation Module

In this module, the YOLOv8 segmentation model is trained on a custom dataset since it does not support the detection and segmentation of leaves by itself. The custom-trained YOLOv8 segmentation model showed an accuracy of 92%. This model generates a binary mask for each healthy leaf, depicting the leaf in white and the background in black. This mask serves as a guide for the subsequent steps in our augmentation process.

3.2 Style Transfer Module

In the style transfer module, we implement Neural Style Transfer (NST) to transfer the style observed in diseased leaves onto healthy leaves. This NST process is performed using the intermediate layers of the pre-trained YOLOv8 neural network, ensuring a detailed and accurate style transfer.

Style Transfer

For the proposed methodology, our previous work [16] is used as a technique of image-to-image style transfer. The approach leverages the You Only Look Once version 8 (YOLOv8) segmentation model and its backbone neural network for style transfer. It stands out for its innovation in combining segmentation and style transfer within a single deep convolutional neural network. As mentioned in the segmentation module above, custom-trained YOLOv8 is used for the task of Segmentation and Image Style Transfer. This integration of using the same models for two different tasks eliminates the need for multiple stages or models, simplifying the training and helping decrease the time required for augmentation.

Blending Module

In this module, the mask generated by the custom-trained YOLOv8 model delineating the targeted leaf object and the leaf image styled with a particular disease is received. This module involves the pixel-wise multiplication of the styled image with the mask. Through this process, only the pixels corresponding to the styled object remain in the image, while the rest are marked out. Subsequently, the isolated object is reintegrated into the original image by adding the styled portion back, thereby completing the object-based neural style transfer.

Similarity algorithm

To select the style to transfer, we developed and implemented a computational framework to identify the most similar diseased leaf images to a given healthy leaf image, leveraging texture analysis and similarity measurement techniques. Our approach begins with extracting the texture features from the leaf images using the Gray Level Co-occurrence Matrix (GLCM) method. Specifically, for each image, we convert it to grayscale and then calculate the GLCM, from which we derive four key texture features: contrast, correlation, energy, and homogeneity. These features capture the spatial distribution of pixel intensities and are crucial for distinguishing between different leaf conditions based on their texture patterns.

We employ the Euclidean distance metric to quantify the similarity between the texture features of healthy and diseased leaf images. This involves computing the root sum squared difference between corresponding texture feature values of a healthy leaf image and each diseased leaf image within our dataset, which spans multiple disease conditions. The process iterates over every healthy leaf image, comparing it against a collection of diseased leaf images to find the one with the minimum Euclidean distance, hence deemed the most similar in terms of texture characteristics.

3.3 Validation and Performance Analysis of Augmented Dataset

Once the augmentation process is concluded, our next step involves validating the effectiveness of the augmentation by employing a classifier trained on the original dataset. This classifier will identify correctly augmented images from the pool of styled images. By passing the augmented images through this pre-trained classifier, we can determine which ones have been augmented accurately, as shown in Figure 3. Only those images classified correctly to their respective disease categories will be selected for further augmentation.

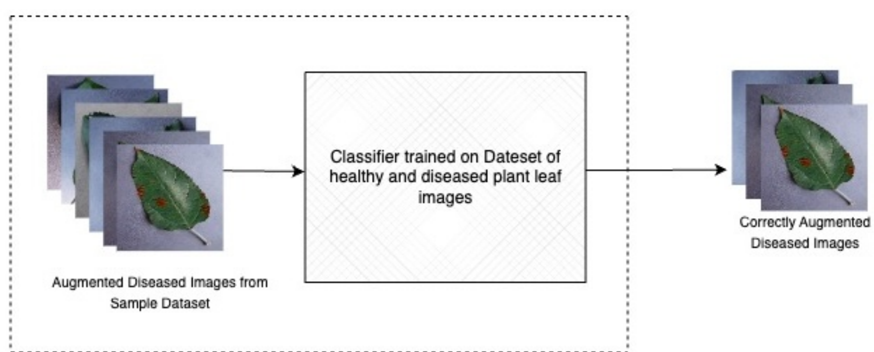


Fig. 3: Classifier used for identifying correctly augmented images

The selected accurately augmented images will then be added to their corresponding disease categories within the dataset, thus expanding and enriching the dataset with augmented data. The classifier will then be retrained on the augmented dataset to accommodate the newly added images. The purpose of this retraining is to ensure that the classifier is updated with the augmented data and can accurately classify images, including the newly augmented ones.

To assess the performance improvements achieved through augmentation, a comparative analysis will be conducted between the new classifier trained on the augmented dataset and the previous classifier trained solely on the original dataset. This comparison will involve evaluating various parameters such as Recall, F1 score, testing accuracy, etc. By analyzing these metrics, we can gauge the effectiveness of the augmentation process in enhancing the classifier's performance and its ability to classify images, particularly those that have undergone augmentation accurately.

3.4 CycleGAN and LeafGAN for comparison

Another comparative analysis with the existing Style transfer method of CycleGAN has been done. CycleGAN uses a cyclical training method. Unlike other GANs, it trains the Generator and Discriminator within the same epoch. If A and B are two domains, this training method trains generators for $A \rightarrow B$ and $B \rightarrow A$. The discriminators for domains A and B are also trained in the same epoch. An important part of cyclical training is the use of Cycle Loss. In the case of domains A and B, cycle loss for A is calculated between an image from A and the image generated after translating the first image from A to B and back to A. Here, considering domain A to be that of healthy leaf images, and domain B of diseased leaf images, the generator which translates from A to B is of interest. Healthy leaf images are passed to this model and translated into diseased images.

LeafGAN is a novel method targeted to leaf disease augmentation developed by Cap et. al. [1]. Using CycleGAN as its backbone, a special Label-Free Leaf Segmentation (LFLSeg) module has been introduced. This causes the complete leaf to be considered as the region of importance, hence generating better-augmented images than CycleGAN. In contrast to CycleGAN, this method examines and transforms only the segmented leaf, resulting in natural images.

The dataset was augmented using the healthy leaf images of the corresponding species as the source. For each disease, both CycleGAN and LeafGAN were trained. After the training process, the healthy images are passed through the generators converting from the healthy domain to diseased domain.

3.5 Traditional Style Transfer Methods for Comparison

Traditional augmentation methods were used on the PlantVillage dataset to establish a baseline model for comparison of the proposed augmentation method. Traditional augmentation includes rotation, resizing, horizontal flipping, and shearing. Rotation involves rotating the image to a certain degree, helping the creation of variations in the orientation of the leaf. Resizing involves changing the dimensions of the image to introduce scale variance. Horizontal flipping mirrors the image along the vertical axis, offering variations in the plant's appearance.

Shearing distorts the image by shifting one side of the image along the horizontal or vertical axis, contributing to shape diversity. These methods collectively enhance the dataset by introducing diverse perspectives and balancing the dataset, aiding in robust model training and generalization.

4 Experimental Setup

For our experiments, we utilized an NVIDIA RTX 3090 graphics card, known for its high computational power and efficiency in handling parallel processing tasks. Each model was run individually to ensure that the GPU's resources were fully dedicated to the task, providing an optimal environment for accurate timing and performance analysis.

The operating system used was Ubuntu 20.04, coupled with Python 3.8 and TensorFlow 2.4, forming a robust platform for executing our experiments. Specific configurations and hyperparameters for each model, such as learning rate and batch size, were set according to the best practices identified in prior studies to ensure consistency and reproducibility of our results.

The PlantVillage dataset was used for training this model. This public dataset contains 54304 images of diseased and healthy plant leaf images collected under a controlled environment. The data spans 14 species with 39 classes: 12 healthy and 27 diseased. Post training, the testing was done against the Plant Pathology competition dataset [24].

5 Results

This section presents the empirical findings and provides a comprehensive view of the efficacy of traditional versus innovative augmentation techniques in improving disease classification accuracy in plant leaves. Initially, the performance of pre-trained classifiers such as VGG16, ResNet, and InceptionV3 was assessed using the original dataset. Subsequent tests were conducted using the same dataset augmented through both traditional and novel methods to investigate the impact of these techniques on the classifiers' ability to generalize. The results are organized into subsections detailing the performance of different classifiers, the effect of augmentation methods, and the computational efficiency of the techniques employed.

5.1 Performance on different classifiers

Initially, the original dataset was utilized to train advanced pre-trained classification models—such as VGG16, ResNet, and InceptionV3—via transfer learning to identify plant leaf diseases. The dataset was divided into training and validation subsets, with 90% and 10% of the data, respectively. The classification accuracy achieved by these models on the original dataset, without any augmentation, is visualized in Figure 4.

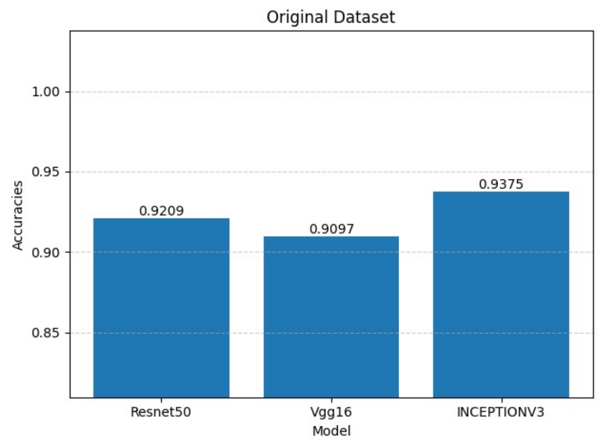


Fig. 4: Performance accuracy on different classifiers on original data

To enhance the dataset, traditional augmentation techniques, including rotation, resizing, horizontal flipping, and shearing, were applied. For this type of augmentation, 1000 images were generated, which included all types of traditional augmentation techniques for each class. The results, shown in Figure 5, indicate that the classifiers performed better on the Plant Village dataset augmented using traditional techniques compared to the original dataset. This suggests that traditional augmentation methods contribute positively to the models' ability to generalize.

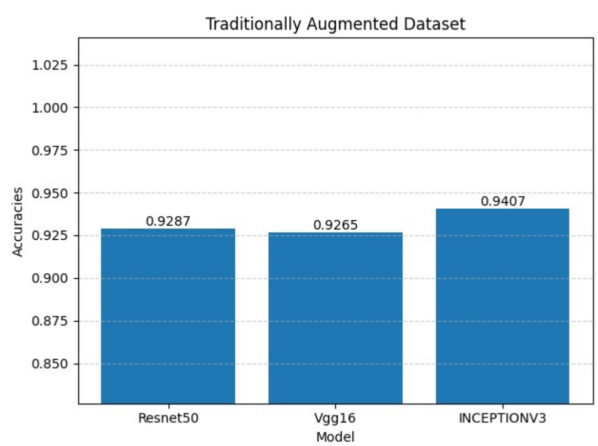


Fig. 5: Performance accuracy on different classifiers on traditionally augmented data

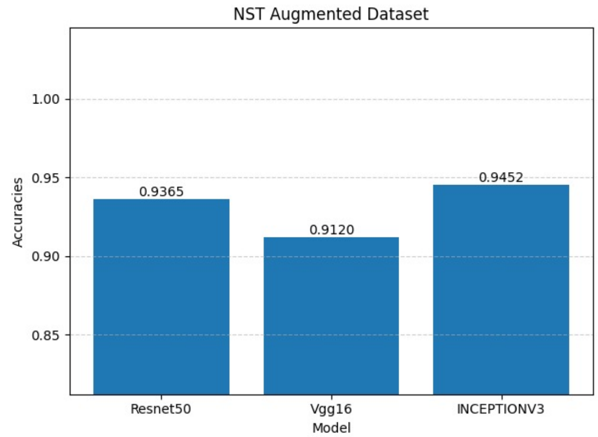
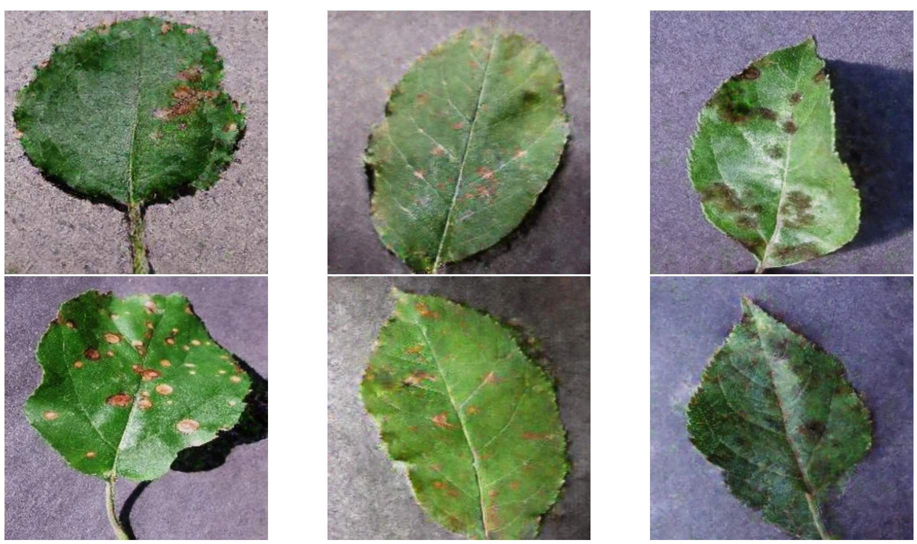


Fig. 6: Performance accuracy on different classifiers on data augmented using proposed object-based NST



(a) Apple Blackrot (b) Apple Rust (c) Apple Scab

Fig. 7: Augmented images for apple disease leaves using proposed approach

In Figure 6, the classification accuracy achieved by the models on the dataset augmented using the proposed object-based NST technique is displayed. The proposed augmentation technique generated a total of 20,723 images, out of which 7,593 were accurately classified by the InceptionV3 model. These correctly classified generated images were subsequently utilized to augment the dataset. InceptionV3 was selected as the classifier for assessing the accuracy of augmented images due to its superior performance on the original Plant Village dataset. Figure 7 shows the images generated by our approach. It is observed that augmentation through object-based NST outperforms the original dataset in generalizing the classifiers. However, when compared with the traditionally augmented dataset, the object-based NST augmentation method exhibits superior performance primarily with the ResNet50 and InceptionV3 models. This highlights the effectiveness of the proposed object-based NST augmentation technique in improving classification accuracy for certain models.

5.2 Leaf GAN and Cycle GAN

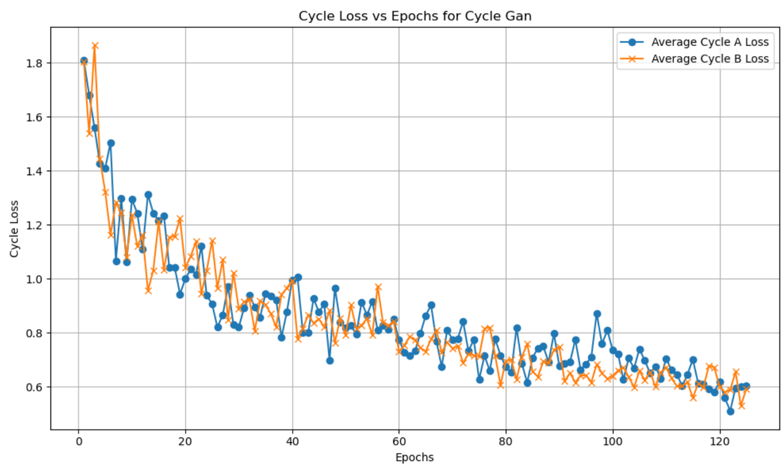


Fig. 8: Cycle Loss vs epochs for Cycle GAN

Figures 8 and 9 illustrate the training progression of a CycleGAN and LeafGAN applied to the plant pathology dataset, as evidenced by the plotted cycle loss for both cycle A (blue) and cycle B (orange) losses over 120 epochs. The y-axis represents the average cycle loss, a measure of how well the GAN maintains the original content after a round-trip transformation (A to B and back to A, or vice versa). The x-axis tracks the number of training epochs.

From the graph, it is apparent that both Average Cycle A Loss and Average Cycle B Loss experience sharp declines within the initial epochs, indicating rapid learning and adaptation by the network. After this initial phase, both losses display a trend of gradual decrease, interspersed with fluctuations that are common during GAN training due to its adversarial and unsupervised nature.

By the end of 120 epochs, the losses have mostly stabilized, with Average Cycle A Loss showing slightly more variability than Average Cycle B Loss. This

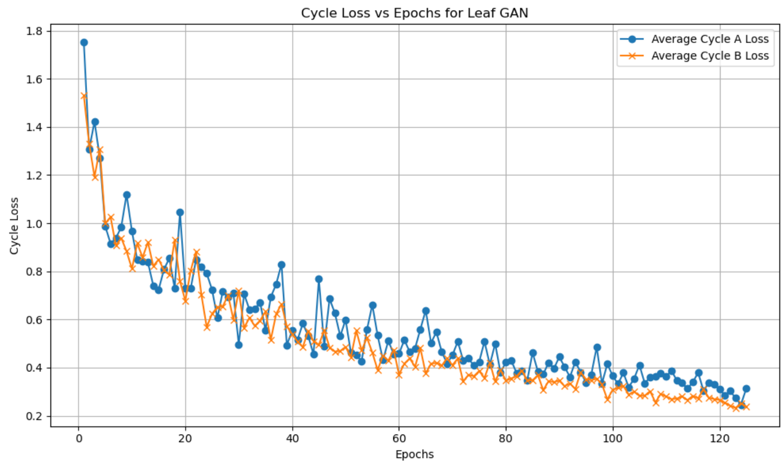


Fig. 9: Cycle Loss vs epochs for Leaf GAN

convergence towards lower loss values suggests that the GAN is achieving a more consistent and accurate mapping between the domains as training progresses.

5.3 Time taken

Table 2 presents the training durations required for the proposed object-based NST, CycleGAN, and LeafGAN techniques to augment apple scab class images comprising 1316 healthy apple images. These durations reflect the total time taken to train each model and subsequently generate the augmented image dataset using the specified number of epochs. The data delineates the differences in computational efficiency and processing time among the techniques. This comparison of training times across different techniques has not been previously addressed by other approaches, highlighting the novel aspect of this analysis in evaluating the efficiency of each method. NST clearly outperforms other techniques.

Technique	**Time Required**
LeafGAN (125 epochs)	12 hours
CycleGAN (125 epochs)	7 hours
NST	3 hours

Table 2: Time required by different techniques to train and generate apple scab class images

5.4 Performance comparison on Plant Pathology Dataset

Multiple classifiers were initially trained on the PlantVillage dataset with and without augmentation to assess the effectiveness of a proposed augmentation technique until convergence was reached. Following this, the performance of the proposed augmentation technique was evaluated on the PlantPathology dataset, as detailed in [29]. The comparison involved training the ResNet152 classifier on various augmentation techniques, including CycleGAN, traditional augmentation methods, and the proposed NST augmentation. The classifiers were then tested on the PlantPathology dataset, which comprises plant leaf data categorized into four classes: healthy and diseased.

In addition to the techniques mentioned in [29], recent advancements in Generative Adversarial Networks (GANs) have introduced newer approaches for data augmentation. Expanding upon the previous study, ResNet152 was trained on the PlantVillage dataset augmented with LeafGAN [1], representing one of the latest advancements in GAN-based augmentation methods. By including LeafGAN alongside CycleGAN and traditional augmentation techniques, a comprehensive evaluation of the proposed augmentation strategy was conducted to assess its impact on the classification performance of the ResNet152 classifier on the Plant Pathology dataset.

Evaluation Metrics				
Method	Precision	Recall	F1-score	Accuracy
None	0.749	0.664	0.665	0.664
Traditional	0.794	0.721	0.722	0.721
Balancing	0.713	0.639	0.642	0.639
CycleGAN	0.702	0.691	0.683	0.691
LeafGAN	0.727	0.706	0.698	0.701
NST	0.799	0.782	0.777	0.782
Proposed Object-based NST	0.807	0.786	0.783	0.795

Table 3: Evaluation of proposed Augmentation Technique on Plant Pathology Dataset

Table 3 illustrates that our proposed NST augmentation technique delivers results on par with the NST-based augmentation method presented in [29]. Demonstrating superiority over other augmentation techniques, our approach outperforms all others by employing a single model for object-based NST augmentation. Notably, LeafGAN exhibits superior performance compared to CycleGAN. Our proposed object-based NST shows slightly better results than previous NST-based augmentation methods. Figure 10 provides better visualization of the accuracy of ResNet152 Classifier on augmented Plant Pathology Dataset for all the augmentation techniques.

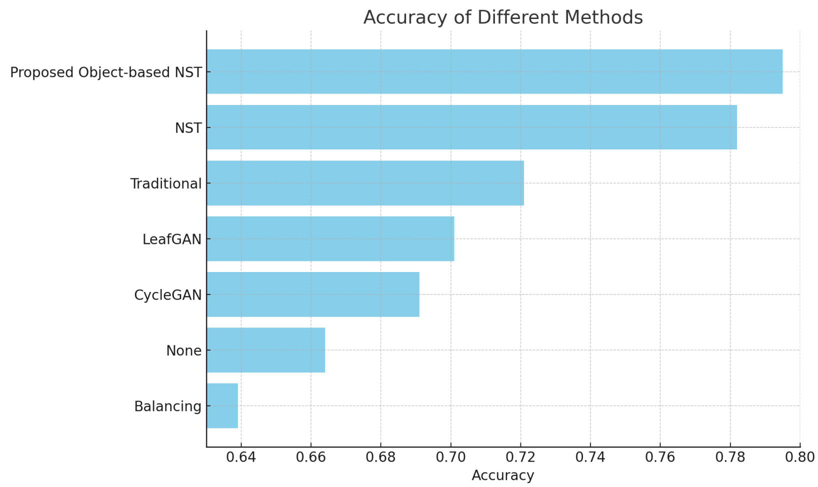


Fig. 10: Accuracy of ResNet152 Classifier on augmented Plant Pathology Dataset for all the augmentation techniques

6 Conclusion

This study effectively demonstrates the significant role of data augmentation in enhancing the performance of machine learning models, specifically in the domain of plant disease classification. Our investigation revealed that the application of advanced pre-trained models, including VGG16, ResNet, and InceptionV3, substantially benefited from augmentation techniques in identifying plant leaf diseases. Traditional augmentation methods, such as rotation and flipping, showed improved model generalization over the original dataset. However, our proposed object-based Neural Style Transfer (NST) technique further amplified this effect, yielding superior results, particularly with the ResNet50 and InceptionV3 models.

Performance metrics across different augmentation methods on the Plant Pathology dataset further underscored the efficacy of our proposed object-based NST augmentation. It not only achieved the highest accuracy but also showed robust precision and recall metrics, outperforming existing methods. These findings advocate for the integration of NST-based augmentation in practical applications where model accuracy is critical to success.

In conclusion, the results of this study support the integration of the proposed object-based neural style transfer method to augment the dataset in order to improve the machine learning model. The superior performance of object-based NST with respect to time and comparable results with GAN-based approaches determine the usability of the proposed approach in real-life use cases.

7 Conflict of Interest Statement

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report.

8 Data availability

The author confirms that all data generated or analyzed during this study are included in this published article. Furthermore, primary and secondary sources and data supporting the findings of this study were all publicly available at the time of submission.

References

1. Cap, Q.H., Uga, H., Kagiwada, S., Iyatomi, H.: Leafgan: An effective data augmentation method for practical plant disease diagnosis. IEEE Transactions on Automation Science and Engineering **19**(2), 1258–1267 (2022). https://doi.org/10.1109/TASE.2020.3041499
2. Castillo, C., De, S., Han, X., Singh, B., Yadav, A.K., Goldstein, T.: Son of zorn's lemma: Targeted style transfer using instance-aware semantic segmentation. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1348–1352. IEEE (2017)
3. Choi, Y., Choi, M., Kim, M., Ha, J.W., Kim, S., Choo, J.: Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 8789–8797 (2018)
4. Cicalese, P.A., Mobiny, A., Yuan, P., Becker, J., Mohan, C., Nguyen, H.V.: Stypath: Style-transfer data augmentation for robust histology image classification. In: Martel, A.L., Abolmaesumi, P., Stoyanov, D., Mateus, D., Zuluaga, M.A., Zhou, S.K., Racoceanu, D., Joskowicz, L. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. pp. 351–361. Springer International Publishing, Cham (2020)
5. Cubuk, E.D., Zoph, B., Mane, D., Vasudevan, V., Le, Q.V.: Autoaugment: Learning augmentation policies from data. arXiv preprint arXiv:1805.09501 (2018)
6. Cubuk, E.D., Zoph, B., Shlens, J., Le, Q.V.: Randaugment: Practical automated data augmentation with a reduced search space. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. pp. 702–703 (2020)
7. Darma, I.W.A.S., Suciati, N., Siahaan, D.O.: Neural style transfer and geometric transformations for data augmentation on balinese carving recognition using mobilenet. International Journal of Intelligent Engineering and Systems **13**, 349–363 (2020), https://api.semanticscholar.org/CorpusID:228916696
8. Food, of the United Nations, A.O.: About fao's work on plant production and protection (2024)
9. Gatys, L.A., Ecker, A.S., Bethge, M.: Image style transfer using convolutional neural networks. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2414–2423 (2016). https://doi.org/10.1109/CVPR.2016.265

10. Gong, C., Wang, D., Li, M., Chandra, V., Liu, Q.: Keepaugment: A simple information-preserving data augmentation approach. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1055–1064 (2021)
11. Ho, D., Liang, E., Chen, X., Stoica, I., Abbeel, P.: Population based augmentation: Efficient learning of augmentation policy schedules. In: International conference on machine learning. pp. 2731–2741. PMLR (2019)
12. Hong, M., Choi, J., Kim, G.: Stylemix: Separating content and style for enhanced data augmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 14862–14870 (June 2021)
13. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1125–1134 (2017)
14. Jackson, P.T., Atapour-Abarghouei, A., Bonner, S., Breckon, T.P., Obara, B.: Style augmentation: Data augmentation via style randomization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops (June 2019)
15. Khosla, C., Saini, B.S.: Enhancing performance of deep learning models with different data augmentation techniques: A survey. In: 2020 International Conference on Intelligent Engineering and Management (ICIEM). pp. 79–85. IEEE (2020)
16. Kulkarni, H., Khare, O., Barve, N., Mane, S.: Improved object-based style transfer with single deep network (2024)
17. Kumar, S.: Plant disease management in india: Advances and challenges. African Journal of Agricultural Research **9**(15), 1207–1217 (2014)
18. Kurzman, L., Vazquez, D., Laradji, I.: Class-based styling: Real-time localized style transfer with semantic segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. pp. 0–0 (2019)
19. Lim, S., Kim, I., Kim, T., Kim, C., Kim, S.: Fast autoaugment. Advances in Neural Information Processing Systems **32** (2019)
20. Mikołajczyk, A., Grochowski, M.: Data augmentation for improving deep learning in image classification problem. In: 2018 international interdisciplinary PhD workshop (IIPhDW). pp. 117–122. IEEE (2018)
21. Psychogyios, K., Leligou, H.C., Melissari, F., Bourou, S., Anastasakis, Z., Zahariadis, T.: Samstyler: Enhancing visual creativity with neural style transfer and segment anything model (sam). IEEE Access (2023)
22. Shijie, J., Ping, W., Peiyi, J., Siping, H.: Research on data augmentation for image classification based on convolution neural networks. In: 2017 Chinese Automation Congress (CAC). pp. 4165–4170 (2017). https://doi.org/10.1109/CAC.2017.8243510
23. Simonyan, K., Vedaldi, A., Zisserman, A.: Deep inside convolutional networks: Visualising image classification models and saliency maps. arXiv preprint arXiv:1312.6034 (2013)
24. Thapa, R., Snavely, N., Belongie, S., Khan, A.: The plant pathology 2020 challenge dataset to classify foliar disease of apples. arXiv preprint arXiv:2004.11958 (2020)
25. Virtusio, J.J., Talavera, A., Tan, D.S., Hua, K.L., Azcarraga, A.: Interactive style transfer: Towards styling user-specified object. In: 2018 IEEE Visual Communications and Image Processing (VCIP). pp. 1–4. IEEE (2018)
26. Xi, J., Ye, X., Li, C.: Sonar image target detection based on style transfer learning and random shape of noise under zero shot target. Remote Sensing **14**(24) (2022). https://doi.org/10.3390/rs14246260, https://www.mdpi.com/2072-4292/14/24/6260

27. Xu, Y., Goel, A.: Cross-domain image classification through neural-style transfer data augmentation. ArXiv **abs/1910.05611** (2019), https://api.semanticscholar.org/CorpusID:204509398
28. Yang, S., Xiao, W., Zhang, M., Guo, S., Zhao, J., Shen, F.: Image data augmentation for deep learning: A survey. arXiv preprint arXiv:2204.08610 (2022)
29. Zalouk, S., Ramjee, S.: Image data augmentation for plant leaf disease classification using neural style transfer
30. Zheng, X., Chalasani, T., Ghosal, K., Lutz, S., Smolic, A.: Stada: Style transfer as data augmentation. pp. 107–114 (01 2019). https://doi.org/10.5220/0007353401070114
31. Zhong, Z., Zheng, L., Zheng, Z., Li, S., Yang, Y.: Camstyle: A novel data augmentation method for person re-identification. IEEE Transactions on Image Processing **28**(3), 1176–1190 (2019). https://doi.org/10.1109/TIP.2018.2874313
32. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. pp. 2223–2232 (2017)