Image-to-Image Translation with GAN for Synthetic Data Augmentation in Plant Disease Datasets

Haseeb Nazki*, Jaehwan Lee*, Sook Yoon**, Dong Sun Park***

Abstract

In recent research, deep learning-based methods have achieved state-of-the-art performance in various computer vision tasks. However, these methods are commonly supervised, and require huge amounts of annotated data to train. Acquisition of data is an additional costly effort, particularly for the tasks where it becomes challenging to obtain large amounts of data considering the time constraints and the requirement of professional human diligence. In this paper, we present a data level synthetic sampling solution to learn from small and imbalanced data sets using Generative Adversarial Networks (GANs). The reason for using GANs are the challenges posed in various fields to manage with the small datasets and fluctuating amounts of samples per class. As a result, we present an approach that can improve learning with respect to data distributions, reducing the partiality introduced by class imbalance and hence shifting the classification decision boundary towards more accurate results. Our novel method is demonstrated on a small dataset of 2789 tomato plant disease images, highly corrupted with class imbalance in 9 disease categories. Moreover, we evaluate our results in terms of different metrics and compare the quality of these results for distinct classes.

Keywords : Deep Learning |Image Synthesis | Image-to-Image Translation | Data Augmentation | Data Balancing | Generative Adversarial Network

I. INTRODUCTION

Deep learning [1] based methods have consistently shown improvements in the state-of-the-art every year, and for many computer vision tasks they even surpassed human performance [2]. Learning from small and imbalanced data sets is a new challenge for several deep learning applications today. To circumvent these costs and train on smaller datasets, methods like transfer learning, domain adaptation [3] and data augmentation [4] have been followed. While transfer learning and domain adaptation are prevalent, they are not as easily appropriate for tasks where no large public datasets or pre-trained network parameters of a close domain are available, e.g. in disease detection and classification in plants and medical images. For this reason, we concentrate on data augmentation to deal with small amounts of data, specifically data augmentation using images synthesized from a generative model. Data augmentation is commonly used to synthetically generate additional training data. A major shortcoming with the standard data augmentation methods is that they require knowledge of the underlying task to perform well and introduce additional hyperparameters into the deep learning setup. With the goal to alleviate these issues, we assess a data augmentation strategy using

Generative Adversarial Networks (GANs). While GANs have shown potential for image synthesis in many fields when trained on large datasets, their potential given small datasets particularly from the agricultural domain has not been analyzed yet. We want to evaluate if GAN based data augmentation using state-of-the-art methods, such as the Image-to-image translation using GANs, is a practicable strategy for small datasets.

Imbalanced learning arises when some types of data distribution significantly dominates the instance space compared to other data distributions. Lately, theoretical study and applied applications for this problem have fascinated a growing concern and attention from both industry as well as academia [5,6,7]. An example of one such applications is in the role of deep learning manipulation of plant disease detection and classification algorithms. The data samples for various kinds of diseases are normally very limited compared to normal cases; therefore, the ratio of the minority class to majority class can be substantial. On the other hand, it is significantly important to predict the presence of diseases, or further classify diverse types of diseases as accurately as possible for proper treatment at an early stage

We test the algorithm on our tomato plant disease data set summarized in Figure 1. The images used for our analysis are collected in different farms, under different conditions. Collecting plant disease data is a complex and expensive procedure that requires the collaboration of people from different fields and at contrasting stages. Researchers often encounter the challenge of imbalance in classes which has been a universal problem in machine learning and computer vision. An efficient way to synthesize images and supplement training set may help boost accuracy in various fields. Using data augmentation techniques for enlarging training set and class balancing has been reported in various literatures [7,27,28]. However, the diversity and variation that can be achieved from small modifications of the images (such as translation, rotation, flip and scale) is relatively minor. This motivates the use of synthetic data examples where the generated samples enable the introduction of more

diversity and can possibly supplement the dataset further, to improve the training process and accuracy.

It has been shown that by rendering photorealistic synthetic images and performing a set of transformations on those rendered images, they can be used to train an object detector with good performance [8]. Similarly, in the medical domain, it has been shown that by training a deep neural network on high-quality rendered 3D images from other computer vision tasks and fine-tuning it towards medical data, the over-all network performance can be improved when data is scarce [9]. This shows that data augmentation by using a generative model can be used to improve the training of deep learning methods.

Recently introduced by Goodfellow et al., Generative Adversarial Networks (GANs) deliver a striking method of learning a generative model by training a deep neural network [10]. GANs have demonstrated potential in tasks such as state-of-the-art image generation [11], or synthetic data generation and translation [12], [13]. The idea of using GANs in the setting of data augmentation has also seen some progresses in research [12]. The main conception in [12] is to render synthetic images with matching labels and refine those synthetic images. This GAN uses the statistics from real, unlabeled images of the same domain while preserving the label information of the rendered images to produce realistic, refined images, which can further be used as training data for a supervised deep learning network. However, while GANs have shown remarkable results when trained on large datasets, it is still a topic of active research how GANs act when trained on a small amount of data, as most GAN-related research emphasizes the use of large datasets.

One of the advantages of Deep Learning is its ability to manipulate the raw data directly without using hand-crafted features. Recently, Deep Learning delivered exclusive results in automatic quantification and recognition of plant diseases based on image processing methods and techniques. The success of Deep learning is mainly related to the access to huge amounts of data for training a

deep model and high computing power provided by Graphics Processing Units (GPUs) which make it possible to train these deep models and impose the parallelism of data computing

In the current work we use deep learning methodology where we focus on the task of synthetic plant disease generation using GAN as a data augmentation technique. We synthesize plant disease samples of the minority class from a finite number of raw images using GANs which can be used further for augmentation and class balancing of the training set for any related computer vision task. The remainder of this paper is ordered as following. Section II presents the related works in the field of generative adversarial networks. In section III we discuss the image synthesis and augmentation technique including the architecture for the generative model. In section IV, we evaluate the performance of the technique using various metrics. Finally, a conclusion is presented in Section V.
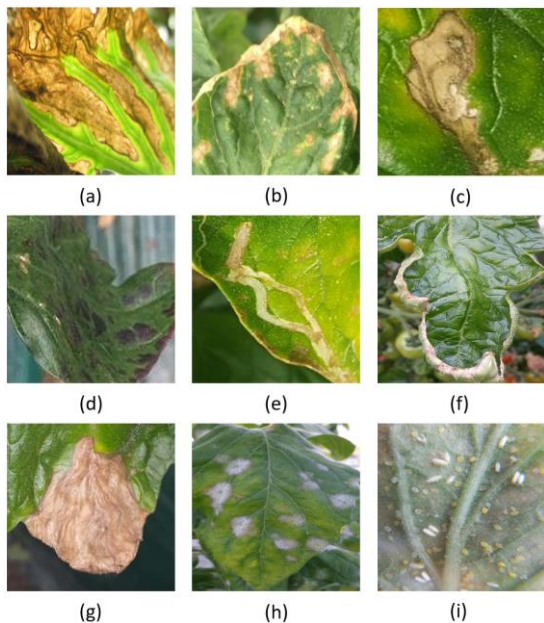

Fig. 1. A representation of diseases and pests that effect tomato plants. (a) Canker, (b) Gray mold, (c) Leaf mold, (d) Low temperature, (e) Miner Canker, (f) Nutritional excess or deficiency, (g) Plague, (h) Powdery mildew, (i) Whitefly. The images are collected under different variations and environmental circumstances.

II. RELATED WORK

A. Data Balancing and Data Augmentation

Data balancing in deep learning applications like plant diseases classification is a critical area that has been studied and surveyed through the years and is driven by the necessity of a healthy agricultural output. However, some essential elements to be considered are cost-effectiveness, user-friendliness, accuracy, and sensitivity. Model performance for tasks like classification, detection and recognition of these diseases can be improved using data augmentation which can overcome the problem of inadequate data and imbalanced distribution. Experiment from the studies in [15] found out that most of the learning algorithms are designed around the notion that training sets are well balanced in distribution, which most of the time is not correct. The authors in [15] went to prove that in the case of feed-forward neural networks, class imbalance does hinder its performance especially when the class complexity increases. However, it is an unsolved problem of how to generate (sample) data from the 'true distribution' of given limited training data such that generated samples are used to eliminate the existence of class imbalance in our dataset.

Data augmentation is the process of generating additional training data from the available existing data [3]. Typically, this is done by using annotation-preserving transformations on the input data, such as randomly rotating, deforming or translating the image. Through the random nature of data augmentation, it can be used to potentially generate an 'infinite' amount of training data by augmenting the existing data. For medical image analysis, data augmentation such as elastic deformation has been used with much success in combination with convolutional neural networks for medical image segmentation [8]. Although data augmentation is an effective way of dealing with the issue of small amounts of training data, it is not universally applicable, as prior information of target domain and task is essential to find a good data augmentation technique. Furthermore, the parametrization of data augmentation methods introduces additional set of vital hyperparameters, which can have a significant

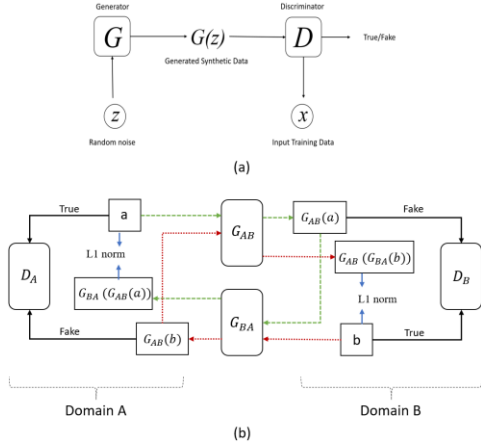impact on the error made by the deep learning method.

Fig. 2. (a) General structure of a Generative Adversarial Network. The generator G takes a noise vector z as input and output a synthetic sample G(z), and the discriminator takes both the true sample x and synthetic input G(z) as inputs and predicts whether they are real or fake. (b) CycleGAN Framework. A and B are two dissimilar domains. Two generators translate an image from one domain to another. A a discriminator for each domain judges if an image belongs to that domain. The two cycles of data flow are shown, the green one (dashed) performs a sequence of domain transfer A → B → A, while the blue one (dotted) is B → A → B. L1 loss is functional on input a (or b) and the reconstructed input $G_{BA}(G_{AB}(a))$ (or $G_{AB}(G_{BA}(b))$) to impose self-consistency.

B. Generative Adversarial Networks

Generative Adversarial Networks (GANs) [10] are a promising approach for training a model that can synthesize images. GAN models have generated better synthetic images than previous generative models, and since have become one of the most prevalent research areas. GANs have achieved great acceptance in the computer vision community and different variations of GANs recently proposed have produced high quality realistic natural images [11].

As shown in Figure 2(a) Generative Adversarial Net (GAN) [10] consists of two separate neural networks: a generator G that takes a random noise vector z and further

generates synthetic data $G(z)$; a discriminator D that takes an input x or $G(z)$ to output a probability $D(x)$ or $D(G(z))$ and indicate whether the input is derived from the synthetic distribution $G(z)$ or from the true data distribution, as depicted in Figure 2. The first GAN [10] proposed, uses fully connected layer as its building block. Later, DCGAN [22] effectively demonstrated the use of fully convolutional neural networks achieves better performance, and since convolution and transposed convolution layers [23] have become fundamental components in many GAN models. The novel process of training the generator and discriminator is by arranging a two-player min-max game between the generator and the discriminator, where the generator G tries to generate reasonably realistic data representations to fool the discriminator while the discriminator D tries to differentiate between real and synthetic data [10]. The value function thus formulated is optimized as follows:

$$\min_G \max_D V(D,G)$$
$$= \mathbb{E}_{x\sim p_{d(x)}}[log\,D(x)] \qquad (1)$$
$$+ \mathbb{E}_{z\sim p_{z(z)}}[log(1-D(G(z)))]$$

where $p_d(x)$ denotes true data distribution and $p_z(z)$ denote the noise distribution. Once the discriminator is trained much better than the generator, it can reject the samples from generator with a high confidence (i.e. close to 1), and thus the loss $log(1-D(G(z)))$ would saturate and hence G would not learn anything from zero gradient. To prevent this, instead of training G to minimize $log(1-D(G(z)))$, we train it to maximize $log\,D(G(x))$ [9]. Furthermore, this new loss function for G provides the same direction of gradient and does not saturate.

Lately, numerous applications have applied the GAN into their framework e.g. in medical imaging [16,17] however, the application in plant disease imaging remains unexplored. Most studies have employed the image – to – image translation techniques to create label-to-segmentation translation [18,19,20], segmentation-to-image translation or cross modality translations [21]. Some studies have

been encouraged by the GAN method for image inpainting. In this research, built on Generative Adversarial Networks (GANs) for image-to-image translation, we suggest a method for data augmentation to generate new samples via adversarial training, to increase the data manifold for the estimate of the 'true distribution' that may lead to better margins between distinct categories of data.

In all the above approaches including the other approaches for various tasks in deep learning in computer vision so far, the availability of huge amount of data plays a crucial role for improved classification accuracy. Through the procedure given in this work, we can achieve an equivalent performance by starting with small number of original images and gradually generating more plant disease images synthetically using Generative Adversarial Networks (GANs) and feeding those samples to our training set.

III. METHOD

When using deep convolutional networks with several layers or dealing with insufficient number of training samples, we always observe a chance of overfitting. Also, the problem of imbalanced data is frequently associated with misclassification, where the minority class tends to be falsely classified as compared to the majority class. Problem arises when data of minority classes contain valuable information and hence become the focus of attention. Such errors in classification lead to errors in decision-making specifically in prediction accuracy of minority class. The standard solution to reduce overfitting and misclassification is data augmentation that artificially enlarges and balances the dataset and improves the classification results. We augment our data in two ways: 1) Classical processes or combination of processes, such as random rotation, shifts, shear and flips, etc. 2) Generating samples learned from the existing data using generative modelling. We describe our method of generating synthetic images of tomato plant diseases using Generative Adversarial Networks (GANs). We start with an overview of this techniques applied to our dataset.

Driven by the success of Generative Adversarial Networks (GANs) in image generation [10,22,36], existing unsupervised mapping methods such as CycleGAN [20] learn a generator which produces images in one domain given images from the other. Without the use of pairing information, there can be many possible mappings that could be inferred. To reduce the space of the possible mappings, these models are typically trained with a cycle-consistency constraint which imposes a strong connection across domains, by necessitating the mapping of image from the source domain to the target domain and back to the source will result in the same image that we started with. This framework has been shown to learn substantial mappings across image domains and proved effective in a variety of related applications [24,25,26].

Our GAN network architecture is based on CycleGAN [20] using both adversarial loss and cycle-consistency loss [20] to generate fake tomato plant disease images. Given two domains A and B, we assume there exists a mapping between their elements, possibly many-to-many. The objective is to unveil these mappings using unpaired samples from each domain distributions $p_d(a)$ and $p_d(b)$. This can be demonstrated as a conditional generative modeling problem where we try to approximate the true conditionals $p(a|b)$ and $p(b|a)$ while using samples from the true marginals. A significant postulation that can be drawn here is that elements in domains A and B are highly correlated; otherwise, it would be questionable that the model would uncover any meaningful relationship without any pairing information.

Such a model approximates these conditionals using two mappings $G_{AB}: A \rightarrow B$ and $G_{BA}: B \rightarrow A$, parameterized by deep convolutional networks, which satisfy the following constraints: a) The output of each mapping matches the empirical distribution of target domain, when focused over the source domain. This is represented as a marginal matching process. b) Mapping elements from one domain to the other, and then back, should yield a sample close to the element that we started with initially, thus exhibiting cyclic-consistency.

The first constraint is satisfied using the generative adversarial networks framework (GAN) (Goodfellow et al., 2014). Mappings G_{AB} and G_{BA} are given by deep networks trained to fool adversarial discriminators D_B and D_A, respectively. Applying this matching on target domain B, marginalized over source domain A, involves minimizing an adversarial objective with respect to G_{AB}:

$$\mathcal{L}_{GAN}^B \ (G_{AB}, D_B)$$
$$= \mathbb{E}_{b \sim p_{d(b)}}[log \, D_B(b) \qquad (2)$$
$$+ \mathbb{E}_{a \sim p_{d(a)}}[log(1 - D_A(G_{AB}(a)))]$$

as the discriminator D_B trains to maximize it. A parallel adversarial $\mathcal{L}_{GAN}^A \ (G_{BA}, D_A)$ is fixed for marginal matching in the opposite direction. Cycle-consistency imposes the objective where initializing with a sample a from A, the reconstructed $a' = G_{BA} \ (G_{AB}(a))$ remains close to the original a. Closeness between a and a' is typically measured with L1 or L2 norms. When using the L1 norm, cycle-consistency starting from A, it can be formulated as:

$$\mathcal{L}_{CYC}^B \ (G_{AB}, G_{BA})$$
$$= \mathbb{E}_{a \sim p_{d(a)}} \| \ G_{BA}(G_{AB}(a)) - a \ \| \qquad (3)$$

And similarly, for cycle-consistency starting

from B. The full objective is given by:

$$\mathcal{L}_{GAN}^A \ (G_{BA}, G_{AB}) + \mathcal{L}_{GAN}^B \ (G_{AB}, D_B)$$
$$+ \alpha \mathcal{L}_{CYC}^A \ (G_{AB}, G_{BA}) \qquad (4)$$
$$+ \alpha \, \mathcal{L}_{CYC}^B \ (G_{AB}, G_{BA})$$

where α is a hyper-parameter to stabilize marginal matching and cycle-consistency. The success of CycleGAN [20] can be attributed to the complementary roles of cycle-consistency and marginal matching constraints in its objective. Marginal matching incites the generation of realistic samples in either domain. Cycle-consistency encourages a robust relationship between these domains. It also helps to prevent multiple items from one domain map to a single item from another domain, similar to the troublesome problem of mode collapse in adversarial generators [27].

Architecture: In this section, we explore the architecture for generative model that generates better tomato plant disease images on an unaligned dataset in the following aspects: (a) using WGAN loss [29] in the adversarial part and (b) using a generator with skip layers to increase multi-scale invariance.

For the generator we experiment with different

Fig. 3: Synthetic images from the evaluated model. The real images shown at the leftmost column are inputs that the synthetic images are based on.

structures. As demonstrated in the next section a widely adopted variant known as the U-net architecture [30] performs well under our data and settings. Deeper networks proposed in CycleGAN [20] are known to better capture high-level features and concepts, however, the vanishing gradient problem affects the convergence rate as well as the quality of convergence. Several works emerged to overcome this issue among which U-Net [23] is of particular interest. U-Net incorporates longer skip connections between its encoder-decoder architecture to preserve low-level features and hence the quality of convergence.

Motivated by [20], we use a patch-based discriminator D and train it iteratively along with G. Patch based discriminator ensures preservation of high-frequency details that are usually lost when we only use L1 loss. All the convolutional layers in D use a filter size of 4 × 4. This novel combination of the U-net generator and Patch-wise discriminator architecture enables efficient learning and improved convergence quality.

IV. EXPERIMENTS AND RESULTS

In this section we elaborate on our final experimental setup for our task of synthetic tomato plant disease generation:

Data: We collected the images for our dataset (Fig. 1) effected with several common diseases and pests in tomato plants under various circumstances like illumination, temperature, season, humidity, and places where they were taken, using simple camera devices. For that purpose, we have braced our dataset with images having distinct features and environmental conditions. These conditions help to estimate the process of infection and deduce how a plant is affected by the disease or pest(its origin or possible developing cause).

After collecting these images for the dataset, we manually annotated each image containing the disease or pest by categorizing each image into its respective class. The judgement from the professionals in the area is a must, depending on the infection status in different diseases that look similar, where the information for identifying the disease was required. This serves as our ground truth. When gathering the images, we find that the best way to get more exact information is to capture the samples containing the ROIs as the main part of the image. This marks the solution to the problem formulation of synthesizing images using GANs for the infected part of the plant.
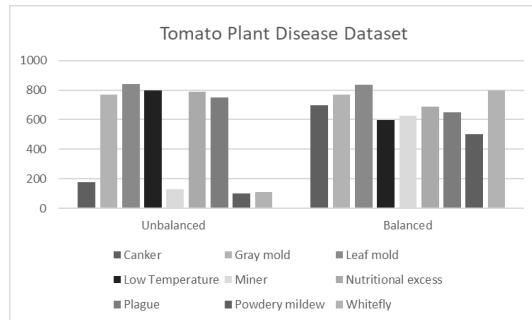

Fig. 4. Our dataset before (left) and after class balancing using synthetic data augmentation (right) using GANs in image-to-image translation setting.

Training details: Using a single Nvidia 12 GB Titan X GPU, we train the model for an average of 180 epochs for each. The generation of synthetic images took 0.0176 ms per image. We use the Adam solver with a batch size of 1. We keep the learning rate constant for the initial 100 epochs and then linearly decay this learning rate to zero over the next epochs. Weights are initialized from a Gaussian distribution with mean 0 and standard deviation 0.02. We employed this model to train each tomato plant disease class distinctly in our dataset. Figure 3 presents some examples of the synthesized tomato plant disease samples from each class. To keep the classes balanced, we sampled the required number of synthetic images for classes with insufficient data only (Figure 4).

Metrics for Quantitative evaluation: It is well known that GAN results are hard to evaluate as the applications are generally on the edge of art and technology.

We evaluate our results with the metric FID score [31] that correlates well with human judgment. FID score is known to capture the similarities of generated image to real one better than the Inception Score [14]. Even

Fig. 5: We compare the effect of using a CycleGAN with U−Net generator that incorporates longer skip connections between its encoder−decoder architecture to preserve low−level features and hence an improved visual and perceptual quality in the generated. From left to right: input image, translated image using CycleGAN, translated image using CycleGAN with U−net generator.

though this metric assists us to avoid depending on human evaluations as it associate well with our subjective judgment of image quality [14], it is recommended to use a large sample

size to calculate this metric, otherwise the true realization of the generator can be underestimated. Considering that there are not many samples in our test data to evaluate and easily tell the difference, our results are mainly based on manual inspection of the visual fidelity of the generated images as reported in Fig. 3 and Fig. 4.

Apart from the metric mentioned above, we also evaluate using the state−of−art evaluation model namely neural image assessment (NIMA) [32]. The NIMA [32] estimates the aesthetic qualities in aspects of photography skills and perceptual relevance. A Larger value of NIMA, and smaller FID value denote better quality, respectively. The corresponding results verify that the generated synthetic plant disease samples can be utilized over any CNN−based architecture for diverse computer vision tasks like detection and semantic segmentation to achieve state−of−the−art results in that field without the requirement of exclusively huge datasets.

Model / Metric	CycleGAN	CycleGAN + U-net
NIMA score	4.584 ± 1.860	**4.612 ± 1.859**
FID score	57.76	**56.34**

Table 1.　Quantitative results for healthy tomato plant leaves−to−leaves infected with powdery mildew obtained using our tomato plant disease dataset. The ground truth powdery mildew images produce a NIMA score of 4.72 ± 1.83. NIMA: higher is better, FID.

Results: In Table 1 we evaluate on the generated images for the class powdery mildew disease from our tomato plant disease dataset. Using a set of 3,172 images of healthy tomato plant leaves we translate them using both CycleGAN and CycleGAN with U−net generator trained on 1,289 powdery mildew images as illustrated in Fig.5. From both Table 1 and Fig. 5, we conclude that CycleGAN with U−net generator successfully captures the low−level details to a much larger extent when compared to the baseline CycleGAN implementation. Hence this method can synthesize translations with an improved perceptual quality and shape preservation. It also produces more realistic textures as compared to the baseline.

We also report the metric scores for some classes in our dataset set suffering from acute shortage of samples in Table 2. Compared with

Class / Metric	Whitefly	Canker	Leaf Miner	Powdery mildew
NIMA score	4.76 ± 1.73	4.56 ± 1.86	4.42 ± 1.92	4.61 ± 1.86
FID score	49.85	41.74	27.40	56.34

Table 2.　Table of FID, NIMA scores for samples generated using 4 classes from our tomato plant disease dataset. The ground truth images produce a NIMA score of Whitefly (4.79 ± 1.72), Canker (4.55 ± 1.89), Leaf Miner (4.41 ± 1.93), Powdery Mildew (4.72 ± 1.83). NIMA: higher is better, FID: lower is better.

ground truth images, generated images for these classes have, on average, 0.6% lower average NIMA scores than their ground truth images. It shows that the generated images have much similar quality to the ground truth images. Fig. 4 displays the class distribution of our tomato plant disease dataset after using synthetic data augmentation for a successful class balancing and augmentation.

V. CONCLUSION

In conclusion, we presented a method that can be used for generating synthetic images for data augmentation and balancing to improve the performance for various deep learning challenges suffering with limited available plant disease data. We demonstrated this technique over our tomato plant disease dataset and synthesized fake diseased image samples. We also introduce a U-net generator in CycleGAN for improved perceptual quality in the generated samples. We believe that other problems can benefit from using these synthetic samples, and that the presented approach can lead to a stronger and more robust support system.

We will work to explore the ability of this model and the corresponding results. And we will further research to verify that the generated synthetic plant disease samples can be utilized over any CNN-based architecture for diverse computer vision tasks like recognition, detection and semantic segmentation to achieve comparable results in that field without the requirement of exclusively large datasets.

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436-444, May 2015

[2] K. He, X. Zhang, S. Ren, and J. Sun, " Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," 2015 IEEE International Conference on Computer Vision (ICCV). *IEEE Computer Society*, 2015, pp. 1026-1034

[3] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning" MIT Press, 2016, http://www.deeplearningbook.org. (accessed May 2019)

[4] O. Ronneberger, P. Fischer, and T. Brox, " U-Net: Convolutional networks for biomedical image segmentation," in Medical Image Computing and Computer-Assisted Intervention - *MICCAI 2015*. Springer, Cham, pp. 234-241, 2015

[5] Hughes, D.P.; Salathe, M. "An open access repository of images on plant health to enable the development of mobile disease diagnostics" arXiv:1511.08060v2, 2016

[6] Cire ̧san, D.C., Meier, U., Masci, J., Gambardella, L.M., Schmidhuber, J. "High performance neural networks for visual object classification" arXiv preprint arXiv:1102.0183, (2011)

[7] Krizhevsky, A., Sutskever, I., Hinton, G.E. "Imagenet classification with deep convolutional neural networks" *NIPS*. pp. 1097-1105, 2012

[8] A. Rozantsev, V. Lepetit, and P. Fua, "On rendering synthetic images for training an object detector," *Computer Vision and Image Understanding (CVIU)*, vol. 137, pp. 24 - 37, 2015.

[9] G. Riegler, M. Urschler, M. Ruther, H. Bischof, and D. ̈ Stern, ̌ "Anatomical landmark detection in medical applications driven by synthetic data," *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pp. 85-89. Dec 2015

[10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," *Advances in Neural Information Processing Systems 27 (NIPS)*. Curran Associates, Inc., pp. 2672-2680, 2014

[11] T.Karras, T. Aila, S. Laine, and J. Lehtinen, " Progressive growing of GANs for improved quality, stability, and variation," *Proceedings of the Sixth International Conference on Learning Representations (ICLR)*, 2018.

[12] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from simulated and unsupervised images through adversarial training," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2242–2251, 2017

[13] T. Neff, C. Payer, D. Stern, and M. Urschler, "Generative Adversarial ˘ Network based synthesis for supervised medical image segmentation," *Proceedings of the OAGM&ARW Joint Workshop*, pp. 140–145, 2017

[14] Tim Salimans, Ian J. Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen, "Improved Techniques for Training GANs" CoRR abs/1606.03498, 2016

[15] Japkowicz, N. "The class imbalance problem: Significance and strategies" *Proceedings of the 2000 International Conference on Artificial Intelligence (ICAI′ 2000)*. Citeseer, 2000

[16] P. Costa et al., " Towards adversarial retinal image synthesis," arXiv preprint arXiv:1701.08974, 2017.

[17] D. Nie et al., Medical "Image Synthesis with ContextAware Generative Adversarial Networks", Springer International Publishing, Cham, pp.417–425, 2017.

[18] Z. Yi, H. Zhang, P. Tan, and M. Gong, "DualGAN: Unsupervised Dual Learning for Image−to−Image Translation," ArXiv:1704.02510, Apr. 2017.

[19] M.−Y. Liu, T. Breuel, and J. Kautz, " Unsupervised Image−to−Image Translation Networks, " ArXiv:1703.00848, 2017.

[20] J.−Y. Zhu, T. Park, P. Isola, and A. A. Efros, " Unpaired Image−to−Image Translation using CycleConsistent Adversarial Networks," ArXiv:1703.10593, 2017.

[21] Y. Choi, M. Choi, M. Kim, J.−W. Ha, S. Kim, and J. Choo, " StarGAN: Unified Generative Adversarial Networks for Multi−Domain Image−to−Image Translation, " ArXiv:1711.09020, 2017.

[22] A. Radford, L. Metz, and S. Chintala, " Unsupervised representation learning with deep convolutional generative adversarial networks, " arXiv preprint arXiv:1511.06434, 2015.

[23] V. Dumoulin and F. Visin, " A guide to convolution arithmetic for deep learning, " arXiv preprint arXiv:1603.07285, 2016

[24] Tung, H.−Y. F., Harley, A. W., Seto, W., and Fragkiadaki, K. "Adversarial inverse graphics networks: Learning 2d−to3d lifting and image−to−image translation from unpaired supervision" *The IEEE International Conference on Computer Vision (ICCV)*, vol. 2, pp.4364−4372, 2017.

[25] Wolf, L., Taigman, Y., and Polyak, A. "Unsupervised creation of parameterized avatars" arXiv preprint arXiv:1704.05693, 2017.

[26] Hoffman, J., Tzeng, E., Park, T., Zhu, J.−Y., Isola, P., Saenko, K., Efros, A. A., and Darrell, T. Cycada:"Cycleconsistent adversarial domain adaptation" arXiv preprint arXiv:1711.03213, 2017.

[27] Li, C., Liu, H., Chen, C., Pu, Y., Chen, L., Henao, R., and Carin, L. Alice "Towards understanding adversarial learning for joint distribution matching" *Advances in Neural Information Processing Systems*, pp.5501−5509, 2017.

[28] Cire‚san, D.C., Meier, U., Masci, J., Gambardella, L.M., Schmidhuber, J. "High performance neural networks for

visual object classification" arXiv preprint arXiv:1102.0183, 2011

[29] Simard, P.Y., Steinkraus, D., Platt, J.C., et al. "Best practices for convolutional neural networks applied to visual document analysis" *ICDAR*. vol.3, pp.958–962, 2003

[30] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville. "Improved training of wasserstein gans" arXiv preprint arXiv:1704.00028, 2017.

[31] O. Ronneberger, P. Fischer, and T. Brox. "U–net: Convolutional networks for biomedical image segmentation" *International Conference on Medical Image Computing and Computer–Assisted Intervention*, pp. 234–241, 2015.

[32] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter. "Gans trained by a two time–scale update rule converge to a local nash equilibrium" *Advances in Neural Information Processing Systems*, pp.6629–6640, 2017.

[33] Talebi, H., Milanfar, P.: "NIMA: neural image assessment," *IEEE Trans. Image Process*, Vol.27, pp.3998–4011, 2018

[34] ChangSeong Kim, DongSuk Lee, Dong Sun Park, "Automatic Pedestrian Removal Algorithm Using Multiple Frames", *Smart media journal*, vol.4, no.2, pp.26–33, June. 2015

[35] Seo Jeong Kim, Jae Su Lee, Hyong Suk Kim, "Deep learning–based Automatic Weed Detection on Onion Field", *Smart media journal*, vol.7, no.3, pp.16–21, Sept. 2018

[36] Nazki, H., Lee, J., Yoon, S. and Park, D.S., "Synthetic Data Augmentation for Plant Disease Image Generation using GAN," *Proceedings of the Korea Contents Association Conference*, pp.459–460.

―――――――――― Authors ――――――――――

Haseeb Nazki

He completed his Bachelor of Technology in Electronics and Communication Engineering from Islamic University of Science and Technology, Kashmir, India in the year 2016. Recently he completed his master degree in Electronics and Information engineering from Chonbuk National University, Republic of Korea on February, 2019.

Jaehwan Lee

He received his M.S. degree in Electronic Engineering from Chonbuk National University, South Korea in 2014.

Sook Yoon

She is a Professor with the Department of Computer Engineering at Mokpo National University in South Korea. She received her Ph.D. degree in Electronics Engineering from Chonbuk National University, South Korea, in 2003. She was a researcher in Electrical Engineering and Computer Sciences at the University of California, Berkeley, USA, until June 2006. And she joined Mokpo National University in September 2006. She was a visiting scholar at Utah Center of Advanced Imaging Research, University of Utah, USA, from 2013 to 2015. Her main research interests include image processing and computer vision, object recognition, machine learning, and biometrics.

Dong Sun Park

He is a professor at the Chonbuk National University, Republic of Korea. He received his BS from Korea University, Republic of Korea in 1979, and MS and PhD degrees from the University of Missouri, United States in 1984 and 1990. He has

published many papers in international conferences and journals. He is a member of IEEE Computer Society. His research interests include computer vision and artificial neural network, especially deep learning.