Application of GAN for Reducing Data Imbalance under Limited Dataset

Gaurav Adke
Michelin India Private Limited, Pune, India

Keywords: Generative Adversarial Networks, Non-conformity Diagnosis, Unbalanced Dataset, Data Augmentation.

Abstract: The paper discusses architectural and training improvements of generative adversarial network (GAN) model for stable training. The advanced GAN architecture is proposed combining these improvements and it is applied for augmentation of a tire joint nonconformity dataset used for classification applications. The dataset used is highly unbalanced with higher number of conformity images. This unbalanced and limited dataset of nonconformity identification poses challenges in developing accurate nonconformity classification models. Therefore, a research is carried out in the presented work to augment the nonconformity dataset along with increasing the balance between different nonconformity classes. The quality of generated images is improved by incorporating recent developments in GANs. The present study shows that the proposed advanced GAN model is helpful in improving the performance classification model by augmentation under a limited unbalanced dataset. Generated results of advanced GAN are evaluated using Fréchet Inception Distance (FID) score, which shows large improvement over styleGAN architecture. Further experiments for dataset augmentation using generated images show 12% improvement in classification model accuracy over the original dataset. The potency of augmentation using GAN generated images is experimentally proved using principal component analysis plots.

1 INTRODUCTION

Deep learning algorithms in computer vision domain can get highly suffered with limited data. An accuracy of the deep learning model can get further degraded with imbalance dataset. Nonconformity detection in an automated inspection process is a task where the model needs to identify nonconforming samples in input images and classify them as per the class of the nonconformities. Collection of a dataset to train such model is a time-consuming process, as the samples are needed to be acquired from the relevant inspection line over the period of time. Another limitation of this collected dataset is that it can be highly imbalanced with a large number of samples of a normal or conforming class. This is obvious since any production line is designed to produce conforming samples. It is highly impractical and expensive to generate conforming samples from the production line to balance the dataset.

Standard image augmentation techniques have been developed to enhance the available dataset. These techniques apply label invariant and semantically preserving transformations to original images. Examples of such techniques are zooming in and out, random flips, random shifts, rotations, brightness variations etc. (Shorten and Khoshgoftaar, 2019). Since augmented images are in general mere modifications of real images, they are of limited help to capture complete probability distribution of input dataset (Antoniou et al., 2017). Moreover, application of these techniques is problem dependent. Considering these limitations of standard augmentations and the requirement to improve accuracy of classification models for nonconformity detection tasks, generative adversarial networks (GAN) (Goodfellow et al., 2014) are studied to tackle data augmentation challenges. GANs are primarily trained with the implicit objective of capturing a distribution of real data. This property of GAN is particularly beneficial for augmentation tasks as generated samples would cover maximum underlying distributions of real datasets. It can also lead to reduced overfitting in the classification model (Zhao et al., 2020b).

The research work presented in this paper describes exploration of recent state-of-the-art improvements in GAN algorithms to tackle low and

Adke, G.
Application of GAN for Reducing Data Imbalance under Limited Dataset.
DOI: 10.5220/0010782800003124
In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022) - Volume 4: VISAPP*, pages 60-68
ISBN: 978-989-758-555-5; ISSN: 2184-4321

unbalanced datasets at hand. These improvements cover changes in GAN architecture, loss function, data augmentation, regularization techniques. The work is focused on capturing fine details in generated images with larger variations. This objective is particularly challenging for a low number of training images.

The paper is organized as follows. Section 2 describes methodologies used to improve baseline StyleGAN architecture. Details of experiments, with proposed advanced GAN used to generate augmentation images, are presented in section 3. Section 4 concludes the article. To the best of the author's knowledge, this study is a first attempt to incorporate recent developments in generative adversarial networks to tackle data imbalance issues in low dataset scenarios.

2 RELATED WORK

Generative models such as Generative Adversarial Networks (GAN) are capable of generating sample images which follow similar distributions as the input real dataset (P_{data}) (Goodfellow et al., 2014). GAN is a deep neural network-based model, primarily used for creating synthetic images following a distribution of the training data. Basic architecture of GAN is shown in figure 1 below. It contains two models: Generator and Discriminator. The main objective of the generator model is to learn to match the distribution of real data and create samples similar to it. On the other hand, the discriminator tries to judge the samples provided to it as real or fake.

A noise vector is used as an input to the generator for creating new samples. This noise is drawn from random normal distribution. The generator learns to map normal noise to features in output images. Both generator and discriminator models are modelled as convolution neural networks for image generation tasks (Radford et al., 2016). The generator has up-convolution layers which output images given the noise vector as input, whereas the discriminator has down-convolution layers which outputs a probability for the input being real. GAN training is an adversarial fight between generator and discriminator, where each one tries to defeat the other. Eventually the discriminator gets better in identifying real and fake samples; and the generator gets better in creating samples which are difficult to be distinguished from the real ones by the discriminator.

Since the introduction of GAN in 2014, many studies have attempted to use GAN for data generation tasks (AlQahtani et al., 2019). Aggarwal et al ((Aggarwal et al., 2021) have reviewed applications of GAN in augmentation of medical and pandemic applications. It is presented that fake image generation using GAN can help to increase datasets along with preserving privacy of patients and reducing extra cost of medical imaging processes. Gao et al (Gao et al., 2020) have used GAN for augmenting machine nonconformity diagnostic datasets. They have demonstrated improvements in classifier accuracy with GAN generated datasets. GAN is used for anomaly detection by Ackey et al (Akcay et al., 2018). For identifying abnormal/ nonconforming samples, their model has resulted in 92% of area under the curve of the receiver operating characteristics curve. Ma et al. (Ma et al., 2020) have explored 3D generation capabilities of GAN for labelled dataset augmentation for Augmented Reality applications. Many interesting applications of GAN have been explored by researchers in the areas of image preprocessing, inpainting, super resolutions, image background domain change etc (Li and Wand, 2016; Pathak et al., 2016; Ledig et al., 2017; Taigman et al., 2017).

Various studies have been carried out to understand GAN training behavior and improve its stability and output quality. (Karras et al., 2018; Karras et al., 2019; Karras et al., 2020b) have researched upon generating high resolution images with improved images quality. They have achieved an FID score as low as 2.84 for FFHQ dataset (Karras et al., 2019) and 2.32 for LSUN car dataset (Kramberger and Potocnik, 2020). The styleGAN architecture was extended to use label conditioning during generation by Oeldorf et al (Mirza and Osindero, 2014; Oeldorf and Spanakis, 2019). A labelled image dataset is used to train conditional GAN while the generator is fed with random labels along the noise vector during training. They could achieve an FID score of 101.9 when trained as a conditioned dataset. GAN training stability is an active area of research with numerous works carried out on regularizing techniques (Lee and Seok, 2020; Kurach et al., 2019). Zhang et al (Zhang et al., 2020) proposed consistency regularization for trained GAN, where the discriminator is regularized to produce consistent predictions for similar images with semantic preserving augmentations. This ensures that the discriminators focus on structural details in images and better gradient flows to the generator. Mescheder et al (Mescheder et al., 2018)

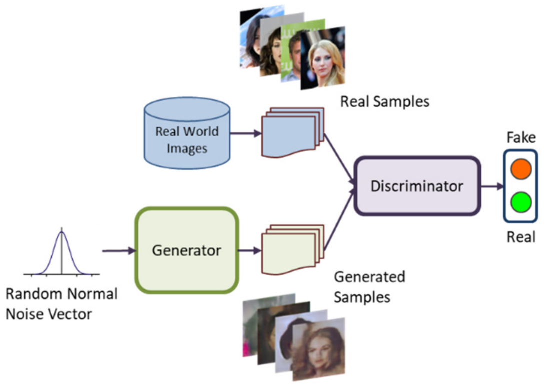


Figure 1: Basic GAN model is shown with example image taken from CelebA dataset (Liu et al., 2015).

have proposed a gradient-based penalty for the discriminator to ensure it follows Lipschitz continuity. This helps in producing a smoother prediction landscape for the discriminator with small steps of gradient for better convergence. Karras et al (Karras et al., 2020b) suggested to regularize the generator with perceptual path length. This ensures untangled and smoother mapping of latent vector to image features. Various research is focused on challenges of low training data by augmentation (Zhao et al., 2020a; Karras et al., 2020a; Sinha et al., 2021) and regularization (Tseng et al., 2021). These are discussed with further details in the methodology section.

3 METHODOLOGY

The main objective of presented work is to produce good quality images of nonconformities, which will be helpful for the downstream task of image classification. GAN architecture used for the current task is based on StyleGAN proposed by Karras et al (Karras et al., 2019). The following GAN model and training improvements are incorporated during the current study.

3.1 StyleGAN

StyleGAN is an extension of progressive GAN architecture proposed by same authors (Karras et al., 2018). Progressively growing the generator helps to produces high resolution images with improved quality. It segregates low level features training from high level training, thus capturing fine details in high resolution images. StyleGAN appends the mapping network to the progressive network. The mapping network is used to transform input latent noise into intermediate vectors. This helps in reducing entangled features in generated images. These intermediate vectors are injected in the generator network at different stages to have better control on generated images. The injection happens through Adaptive Instance Normalization (AdaIN)

layers to match the style of generator feature maps as per input vector. Stochastic variation in output images is achieved by adding random noise at each stage. The discriminator network is a mirror copy of the generator where image size is progressively reduced. Style mixing regularization is performed by injecting different noise vectors at various stages of the generator. An overview of StyleGAN is shown in figure 2.

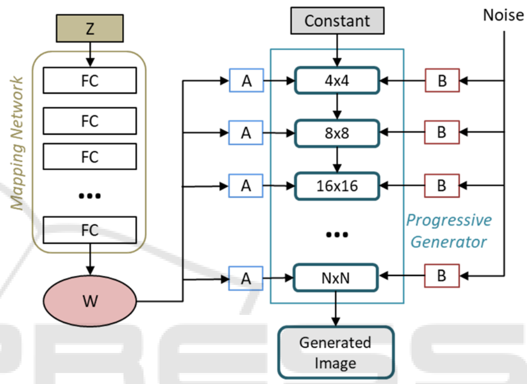


Figure 2: StyleGAN model with progressive generator and mapping network. Layers “A” are affine transformation and layers “B” are noise scaling operations.

3.2 U-NET Discriminator

The discriminator used in StyleGAN architecture classifies the global image as real or fake. Hence the loss gradients produced are of limited use to generate locally coherent structures in images. Schoenfeld et al. (Schonfeld¨ et al., 2020) have proposed a U-Net based discriminator. A schematic of this architecture is shown in Figure 3 below.

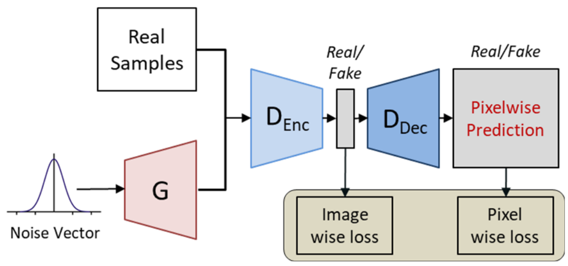


Figure 3: U-net GAN model.

The U-net GAN is capable of providing both global and pixel level feedback to train the generator.

AN encoder model of the discriminator provides global level information of input images, while a decoder model provides per-pixel information. Per-pixel information is useful for generating images with semantic relatedness as per real distribution as well as capturing fine intricate details in images as observed in our study. Skip connections between the encoder and decoder models transfer both high-level and low-level details of images.

The StyleGAN architecture model developed for the study is extended to incorporate the U-net structure. The discriminator of StyleGAN and the loss functions were modified accordingly as per U-net GAN. The generator of the architecture remains unchanged.

3.3 Data Augmentation in Training GAN

GAN-generated image quality can significantly deteriorate with a limited amount of training data. The discriminator may easily overfit by memorizing the salient features from the training dataset, whereby it stops providing meaningful gradients back to train the generator. This leads to poor quality of generated images and mode collapse (Bau et al., 2019). In literature, lots of studies are carried out to apply augmentation for training GAN (Karras et al., 2020a). When the conventional data augmentation is applied only to real images, the generator may produce samples similar to real, as well as transformed, images. This leads to undesirable distributions in generated samples. Instead, augmentation can be applied to both real and generated images. This would result in a discriminator which is better in classifying augmented images only. Consequently, it may not properly identify non-augmented generated images due to disconnected gradient flows after transformations.

A solution to this is the use of differential augmentation (Zhao et al., 2020a; Karras et al., 2020a). As the name suggests, all transformations performed on both real and fake images are differentiable, which helps in uninterrupted passing of gradients from the discriminator to the generator. This by and large trains the discriminator to identify unaltered images from the desired target distribution and maintains a precise training process for the generator. Differentiability of augmentations is achieved by using standard primary operations offered by deep learning frameworks.

Karras et al. (Karras et al., 2020a) have studied types of transformations which do not cause leaking in generated images. Their results show that using invertible transformations like pixel blitting, geometric, and color transforms have an improved effect on generated images in terms of measurement metrics. These transformations are applied with nonzero probability (preferably lower than 0.8) to use non-augmented images as well during the training.

3.4 Loss Functions

The selection of loss function in the current study is mainly governed by the presence of mode collapse in generated images. Mode collapse is a situation where the discriminator is overfitted to few features in real image distributions. Hence, the generator tends to produce images which are only suitable in fooling the discriminator on those features. Consequently, the generator loses the capability to produce variations in the images. In the presence of limited data, the possibility of mode collapse increases. This issue is mainly tackled by use of Wasserstein loss with gradient penalty (Gulrajani et al., 2017) (WGAN-GP). It trains the discriminator to reduce Wasserstein distance between generated distribution (P_g) of produced samples and real distribution (P_r) of real samples. WGAN-GP loss term is also appended with a consistency term (Wei et al., 2018) to enforce Lipschitz continuity near real data manifold. Wasserstein loss is implemented in non-saturating form (Goodfellow et al., 2014) as mentioned below.

Critic (discriminator) loss:

$$E_{x\sim \mathrm{Pr}}[D(x)] - E_{z\sim \mathrm{Pg}}[D(G(z))] \tag{1}$$

Generator loss:

$$E_{z\sim \mathrm{Pg}}[D(G(z))] \tag{2}$$

In WGAN-GP, the discriminator is referred to as "critic", since it does not classify images as being fake or real. Critic gives a score for images as being real or fake. Here, critic is required to follow 1-Lipschitz continuity to make sure a loss evaluated on critic output follows Wasserstein distance metric (Gulrajani et al., 2017). Use of the gradient penalty as given by equation below, enforces Lipschitz continuity by making norm of gradients of critic output with respect to an input less than one.

Gradient Penalty term:

$$GP = E_{x\sim \mathrm{Pr,Pg}}[(\|\nabla_x D(x)\|_2 - 1)^2] \tag{3}$$

Consistency term:

$$CT = E_{x\sim \mathrm{Pr}}[(\|\nabla_x D(x)\|_2 - 1)^2] \tag{4}$$

Total critic loss is formulated as below:

$$E_{x\sim \mathrm{Pr}}[D(x)] - E_{z\sim \mathrm{Pg}}\big[D\big(G(z)\big)\big] + \lambda * GP + \lambda_1 * CT \tag{5}$$

Here, $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda_1}$ are scaling factors for gradient penalty term and consistency term respectively. It is recommended by authors to scale GP term by a value of 10 and CT term by 2 in critic loss calculation.

3.5 Regularizations

Regularizing techniques are used in GAN training for improving stability and convergence. These methods can be subdivided based upon their implementation on weights of network, their gradients and layer outputs. A majority of regularizing techniques is applied on the discriminator (Lee and Seok, 2020). Very few techniques like perceptual path length regularization are applied on generator weights (Karras et al., 2020b). Current work focuses on regularizing the discriminator mainly for training stability and alleviating the mode collapse issue. Consistency regularization (Zhao et al., 2020b) is applied to the discriminator to impose equivariant behaviour for applied differential augmentation. It is applied through CutMix augmented images (Schonfeld¨ et al., 2020). These images are created by merging crops of real and fake images. The consistency loss term, as given in equation 6, ensures that the difference between a discriminator prediction for CutMix image and a mix of predictions of its independent crops is minimal. This loss term is added in WGAN-GP loss mentioned above.

$$L_{consi} = \left\| D(CutMix(x, G(z))) - CutMix(D(x), D(G(z))) \right\|^2 \quad (6)$$

Gradient penalty terms, as described in the previous section and as incorporated in loss evaluations, also provide a regularizing effect by keeping gradient under unity and applying Lipschitz continuity. During training, the exponential weight averaging track of the generator weights is saved. While generating images for augmentation, these averaged weights are used. It produces better quality images, as averaged weights are insensitive towards outlier and noisy iterations during training.

The current study on image augmentation using GAN generation utilizes the above-mentioned improvements to produce better quality images. A discriminator of a styleGAN model is modified to U-NET architecture to capture pixelwise details. Differential augmentation is implemented to address low training dataset availability. An improved WGAN-GP loss term is used to reduce the mode collapse issue and generate images with increased variations. A regularization effect is achieved by adding consistent loss term and gradient penalty term in loss evaluations. Finally, the generator with exponential moving averaged weights is used to generate images for augmentation. Hereafter, this improvised GAN architecture is referred as Advanced GAN in the remaining article.

4 EXPERIMENTS AND RESULTS DISCUSSION

The applicability of the proposed advanced GAN is evaluated using a tire joint conformity dataset. Images are generated using multiple experiments with combinations of nonconforming and conforming images from the dataset. Data augmentation is carried out in three approaches. The summary of all approaches followed for image generation is given in Table 1. In the first approach, an individual GAN model is trained for each nonconformity category. Then these trained models are used to generate augmented images of each nonconformity independently. In the second approach, a GAN model is trained on images from all categories. Augmented images are produced using style merging on the trained generator (Karras et al., 2019). Latent vectors of two different nonconforming images are injected at different resolutions of the styleGAN generator. This way of style injection produces images changing from nonconformity to another. Consequently, we can have a dataset where we can convert an image from one nonconformity category to another. The third approach trains a separate GAN model on a set of normal images and nonconforming images of a single category. This model can be used to insert the nonconformity, with which it is trained, into a normal image by using style merging. Latent vector interpolation is also used with the second and third approach of data augmentation for transition image generation from one category to another.

The proposed advanced GAN algorithm is developed in Python 3.6 with TensorFlow 2.1.0 framework. The training of all models is carried out in Microsoft Azure Machine Learning Services. Single NVIDIA Tesla K80 GPU is used for computation. The final size of images generated is 256x256 pixels. Quality of generated images is evaluated using Fréchet inception distance (FID) (Heusel et al., 2017a). The effectiveness of augmentation is checked using a classification model trained to classify images either from each nonconformity or conformity (OK) category. The classification model is a convolution neural network-based model.

Table 1: Description of different approaches followed for data augmentation.

		Generation Methods		
Approach	Description	Single noise vectors	Style merging	Latent interpolation
1	Individual GAN model for each defect	✓	--	✓
2	Single GAN model for all defective images only	--	✓	✓
3	Separate GAN model for each defect and normal images	--	✓	✓

The proposed advanced GAN model is compared with basic styleGAN model architecture. Their performance is evaluated using FID. Note that a lower FID score is related to better generated image quality and improved variation. Both architectures are trained on the same tire joint nonformity datasets and results are compared. Table 2 shows their comparison.

Table 2: Performance comparison of StyleGAN (Karras et al., 2018) and proposed Advanced GAN (*NC – Nonconformity).

	FID Scores		
GAN Architecture	**NC 1**	**NC 2**	**NC 3**
StyleGAN	165.6	162	161.1
Advanced GAN	96.3	93.8	95.7

These results show a large improvement in the FID score for advanced GAN as compared to the styleGAN model. Results also show the usefulness of advanced GAN in improving generation quality under a limited number of images available for training. An improvement in the results is contributed by architectural and training changes carried out in Advanced GAN. Implementation of differential augmentation and consistency regularization has helped in tackling limited dataset regimes. It also stabilizes training for better convergence. The UNET discriminator provides pixelwise feedback which helps in improving generated image quality and hence helps in reducing the FID score. Exponential weight averaging of the generator weights further reduces the FID score by smoothening training oscillations and diminishing outlier noisy iterations.

To study the consequence of augmentation, initially the classifier model is trained on all real images without any GAN generated images. Standard augmentations like horizontal flip, crop and translate are used in classifier model training for all experiments. The classifier model is tested on real images only, extracted randomly from the original dataset. Real images are split by 10% for testing and 90% for training and validation. Comparison of different experiments on augmentation is done using accuracy of the trained classifier model. Accuracy is evaluated on a test dataset and reported as an average of all test samples over all classes. Table 2 provides details of all experiments carried out using generated images along with real images. Results presented here are averaged over multiple classification models trained on the same dataset to reduce variance.

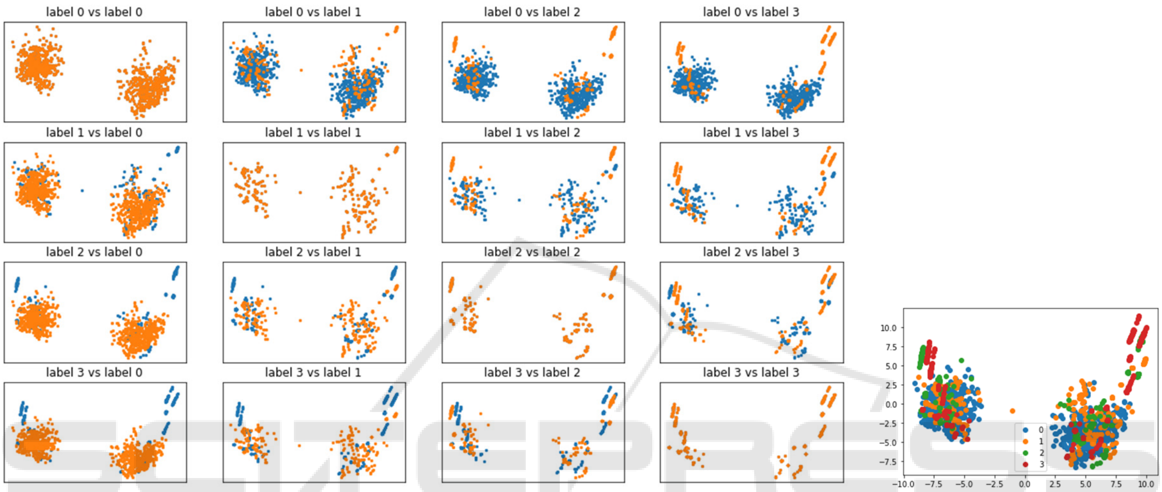
The dataset used for this study is collected in two stages from a production line. In the first stage, a total of 1183 samples were collected. In the second stage, 1108 additional samples were collected, making the total count 2291 samples. GAN models are initially trained on the first stage real dataset and generated images are used for augmentation. Later, all real images from both stages are used for the training of GAN. The effectiveness of augmentation is evaluated separately for each set of generated images from the two stages.

All approaches presented in Table 1 are used to generate images for each stage. Experiments in Table 3 indicate that augmentation by GAN produced images has always enhanced the performance of the classification model. In the first stage of dataset collection, classification accuracy was too low due to insufficient data. Even in this low dataset scenario, the advanced GAN architecture presented here was able to get trained with sufficient convergence and helped in improving classification accuracy by augmentation. Classification accuracy of the increased dataset of the second stage was further enhanced by images generated using real images from both stages.

The effectiveness of GAN augmentation is visualized using Principal component analysis (PCA)

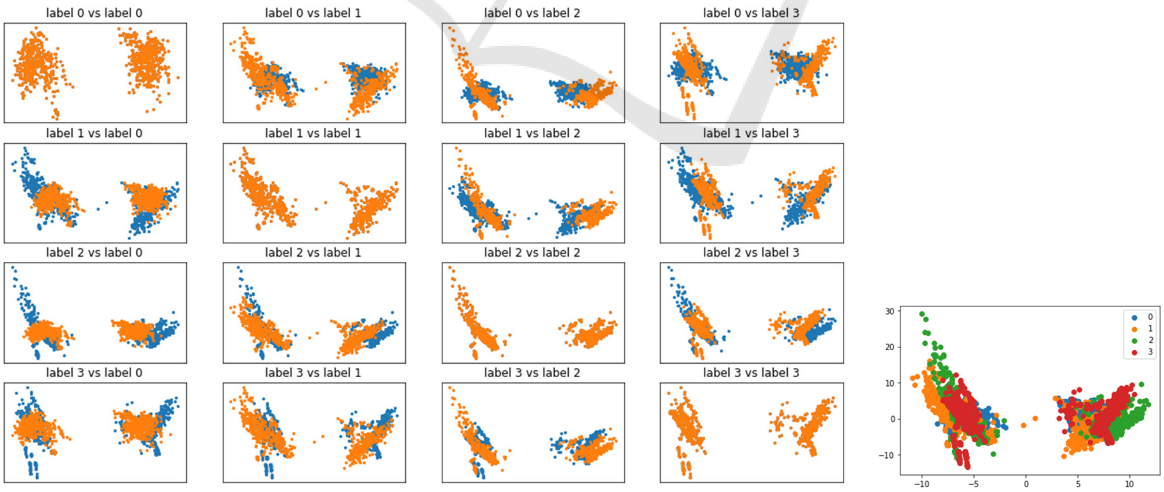
Table 3: Evaluation details of classification model with original and augmentation datasets.

	Description	Classification Accuracy
v01	Stage 1 real dataset	0.73
v02	Stage 1 GAN generated images augmentation	0.79
v03	Stage 2 real dataset	0.85
v04	Stage 2 real + stage 1 GAN generated images	0.89
v05	Stage 2 real + stage 2 GAN generated images augmentation	0.92
v06	Stage 2 real + all generated images augmentation	0.97



(A) (B)

Figure 4: PCA scatter plots of top two principal components for real images.



(A) (B)

Figure 5: PCA scatter plots of top two principal components for augmented image dataset.

in Figures 4 and 5. They show distribution of nonconforming images and conforming images in two dimensions. The top two principal components from PCA are plotted against each other for image samples. Figure 4 (A) shows comparisons of each class with the other for real images, while Figure 4

(B) shows a plot of distribution of all classes together. Similarly Figure 5 (A) shows comparison of class-wise PCA plots and Figure 5 (B) shows distribution of all classes for real images augmented with GAN generated images.

PCA plots of real images, as seen in Figures 4 (A) and (B), show that different nonconformity categories are difficult to distinguish from conforming images and other nonconformities. When the dataset is balanced by augmentation using GAN, as seen in Figures 5 (A) and (B), the PCA plot shows improved distinction between different image categories. From this visualization it can be asserted that lack of data leads to reduced generalization capabilities of the classification model in capturing overall distribution of the input data domain. This also results in lower performance of the image classification task. GANs are trained to capture implicit distribution of the input data on which it is trained. Accordingly, GAN generated augmentation images can be used to facilitate the classification model in capturing the input data distribution in an improved manner, thus improving its prediction accuracy and generalization towards unseen samples extracted from a sample space having same distribution.

5 CONCLUSION

The paper discusses incorporation of recent developments in GAN models for better generated image quality. Proposed advanced GAN architecture produces much lower FID scores than styleGAN, which indicates improved image quality and variation in generation. Various architectural and training improvements discussed in this article are useful for smoother convergence of GAN training. Hence proposed advanced GAN can generate varied images with fine details captured. Advanced GAN is particularly useful in situations of augmentation of limited and unbalanced datasets. An augmented balanced dataset has shown good improvement in accuracy of downstream tasks of image classification. Principal component analysis of the augmented dataset experimentally proves that generated images from proposed advanced GAN can be helpful to improve the distinction among different classification classes.

Experiments presented in this study were limited to images of size 256x256 pixels due to constraints of computing power and processing time. Effectiveness of augmentation by GAN generated images is high in case of smaller datasets. Its usefulness for large datasets needs to be studied as further work. Future scope of the present work involves incorporating GAN model improvements with styleGAN2 (Karras et al., 2020b) architecture and use style merged images for augmentation. Classwise augmentation can be tried for classes with worse classification recall.

ACKNOWLEDGEMENTS

I am thankful to Mr. Suhas Bindu for providing a relevant dataset which was used in developing and improving GAN algorithms. I am also thankful to Mr. Himanshu Pradhan and Mr. Saurabh Gupta for their valuable inputs during the current study. I would like to appreciate efforts taken by Ms. Kelly Merkel for organizing smooth content flow along with suggesting grammatical corrections in this article.

REFERENCES

Aggarwal, A., Mittal, M., and Battineni, G. (2021). Generative adversarial network: An overview of theory and applications. *International Journal of Information Management Data Insights*, 1:1.

Akcay, S., Abarghouei, A. A., and Breckon, T. (2018). Ganomaly: Semi-supervised anomaly detection via adversarial training. In *ACCV*.

AlQahtani, H., Thorne, M. K., and Kumar, G. (2019). Applications of generative adversarial networks (gans): An updated review. *Archives of Computational Methods in Engineering*, 28:525–552.

Antoniou, A., Storkey, A., and Edwards, H. (2017). Data augmentation generative adversarial networks. *ArXiv*, abs/1711.04340.

Bau, D., Zhu, J.-Y., Wulff, J., Peebles, W. S., Strobelt, H., Zhou, B., and Torralba, A. (2019). Seeing what a gan cannot generate. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4501–4510.

Gao, X., Deng, F., and Yue, X. (2020). Data augmentation in fault diagnosis based on the wasserstein generative adversarial network with gradient penalty. *Neurocomputing*, 396:487–494.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. C., and Bengio, Y. (2014). Generative adversarial nets. In *NIPS*.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. C. (2017). Improved training of wasserstein gans. In *NIPS*.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. (2017a). Gans trained by a two timescale update rule converge to a local nash equilibrium. In *NIPS*.

Karras, T., Aila, T., Laine, S., and Lehtinen, J. (2018). Progressive growing of gans for improved quality, stability, and variation. *ArXiv*, abs/1710.10196.

Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., and Aila, T. (2020a). Training generative adversarial networks with limited data. *ArXiv*, abs/2006.06676.

Karras, T., Laine, S., and Aila, T. (2019). A style-based generator architecture for generative adversarial networks. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4396–4405.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., and Aila, T. (2020b). Analyzing and improving the image quality of stylegan. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8107–8116.

Kramberger, T. and Potocnik, B. (2020). Lsun-stanford car dataset: Enhancing large-scale car image datasets using deep learning for usage in gan training. *Applied Sciences*, 10:4913.

Kurach, K., Lucic, M., Zhai, X., Michalski, M., and Gelly, S. (2019). A large-scale study on regularization and normalization in gans. In *ICML*.

Ledig, C., Theis, L., Huszar, F., Caballero, J., Aitken, A. P.,´ Tejani, A., Totz, J., Wang, Z., and Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 105–114.

Lee, M. and Seok, J. (2020). Regularization methods for generative adversarial networks: An overview of recent studies. *ArXiv*, abs/2005.09165.

Li, C. and Wand, M. (2016). Precomputed real-time texture synthesis with markovian generative adversarial networks. *ArXiv*, abs/1604.04382.

Liu, Z., Luo, P., Wang, X., and Tang, X. (2015). Deep learning face attributes in the wild. *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 3730–3738.

Ma, Q., Yang, J., Ranjan, A., Pujades, S., Pons-Moll, G., Tang, S., and Black, M. J. (2020). Learning to dress 3d people in generative clothing. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6468–6477.

Mescheder, L. M., Geiger, A., and Nowozin, S. (2018). Which training methods for gans do actually converge? In *ICML*.

Mirza, M. and Osindero, S. (2014). Conditional generative adversarial nets. *ArXiv*, abs/1411.1784.

Oeldorf, C. and Spanakis, G. (2019). Loganv2: Conditional style-based logo generation with generative adversarial networks. *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pages 462–468.

Pathak, D., Krahenb¨ uhl, P., Donahue, J., Darrell, T., and¨ Efros, A. A. (2016). Context encoders: Feature learning by inpainting. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2536–2544.

Radford, A., Metz, L., and Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks. *CoRR*, abs/1511.06434.

Schonfeld, E., Schiele, B., and Khoreva, A. (2020). A u-¨ net based discriminator for generative adversarial networks. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8204–8213.

Shorten, C. and Khoshgoftaar, T. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6:1–48.

Sinha, A., Ayush, K., Song, J., Uzkent, B., Jin, H., and Ermon, S. (2021). Negative data augmentation. *ArXiv*, abs/2102.05113.

Taigman, Y., Polyak, A., and Wolf, L. (2017). Unsupervised cross-domain image generation. *ArXiv*, abs/1611.02200.

Tseng, H.-Y., Jiang, L., Liu, C., Yang, M.-H., and Yang, W. (2021). Regularizing generative adversarial networks under limited data. In *CVPR*.

Wei, X., Gong, B., Liu, Z., Lu, W., and Wang, L. (2018). Improving the improved training of wasserstein gans: A consistency term and its dual effect. *ArXiv*, abs/1803.01541.

Zhang, H., Zhang, Z., Odena, A., and Lee, H. (2020). Consistency regularization for generative adversarial networks. *ArXiv*, abs/1910.12027.

Zhao, S., Liu, Z., Lin, J., Zhu, J.-Y., and Han, S. (2020a). Differentiable augmentation for data-efficient gan training. *ArXiv*, abs/2006.10738.

Zhao, Z., Zhang, Z., Chen, T., Singh, S., and Zhang, H. (2020b). Image augmentations for gan training. *ArXiv*, abs/2006.02595.