International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 03 | March - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

GenAI for Synthetic Data Creation using StyleGAN2-ADA

VOGGU SAI MOHANA*[1], NADENDLA VENKATESWARARAO[2], MITHINTI VARSHINI[3], NANDIGALA RAVITEJA[4], TALLAPOGU ABHISHEK[5]

[1]*Student, Department of CSE(AIML),Bapatla Engineering College, Bapatla 522101,AP, India*
[2]*Assistant Professor, Department of CSE(AIML), Bapatla Engineering College, Bapatla 522101, AP,India*
[3]*Student, Department of CSE(AIML), Bapatla Engineering College, Bapatla 522101, AP, India*
[4]*Student, Department of CSE(AIML), Bapatla Engineering College, Bapatla 522101, AP, India*
[5]*Student, Department of CSE(AIML), Bapatla Engineering College, Bapatla 522101, AP, India*

---***---

Abstract - Deep learning models for medical image classification require large, balanced datasets, yet medical data is often scarce due to privacy and acquisition costs. While Conditional GANs (cGANs) and pixel-based mixup improve dataset size, they often produce anatomically unrealistic images that lead to unstable training. Building up framework , this study proposes an enhanced augmentation pipeline using StyleGAN2-ADA. We introduce a label-aware generative mixing approach where the generator is conditioned on soft labels derived from a Softmax distribution to ensure smoother class transitions. Our framework is evaluated on the using ResNet-50 and EfficientNet architectures. Beyond standard metrics, we prioritize the Macro-F2 score and False Negative Rate to better reflect clinical requirements for infection detection. Results demonstrate that our StyleGAN-based augmentation achieves a good accuracy and significantly reduces false negatives compared to traditional cGAN and pixel-space methods

Keywords: Deep Learning, Data Augmentation, Synthetic Images, CGAN's, Medical Images, StyleGAN2-ADA.

1. INTRODUCTION

Deep learning models have made tremendous success in the recent years in different fields, mainly due to the availability of large and good datasets. Yet, in numerous practical applications particularly in sensitive domains healthcare and medical imaging attaining sufficient data is still a major hurdle. The effectiveness of machine-learning model trained over real data fails due to data privacy issues, a high cost of acquisition and labeling, and insufficiency of rare or critical cases.

Synthetic data generation has come up as a feasible solution to these problems. Researchers can expand datasets with data that closely mimics reality without leaking privacy via the creation of artificial data.

Generative Adversarial Networks(GAN's) belongs to class of generative models which consists of strong ability to learn complex data distributions and generate realistic images. A style-based GAN architecture called StyleGAN2 is widely known for producing the realistic and high resolution images. This new or creative unique design allows a good control over image features. Data augmentation has emerged as a standard solution. Geometric and photometric transformations — flipping, rotation, intensity scaling — reduce overfitting but do not introduce new anatomical patterns. More sophisticated methods such as Mixup interpolate between image pairs with soft labels derived from a Beta distribution, improving classifier calibration and robustness. However, pixel-wise blending of anatomically distinct CT scans produces clinically implausible intermediate images, potentially introducing misleading training signals.

Carlesso et al. [1] proposed GeMix, which addresses this limitation by replacing pixel-level interpolation with GAN-based image synthesis conditioned on Dirichlet-sampled soft labels. GeMix uses a trained StyleGAN2-ADA [5] generator to produce anatomically coherent synthetic images, demonstrating consistent improvements in macro-F1 over pixel-level mixup across ResNet-50, ResNet-101, and EfficientNet-B0 architectures on the COVIDx-CT-3 dataset.

2. LITERATURE SUREVY

Geometric and Photometric Transformations: This includes a number of standard (and strong!) overfitting reduction approaches including rotations, crops, intensity augmentations. Nevertheless, these approaches do not produce new anatomical configurations or unusual pathologies.

Region-Removal Techniques: Methods such as Cutout black out random patches of an image to push our model to focus on the global context.

CutMix: This takes this further by mixing patches from one image into another, with the labels weighted according to area.

CGAN: Conditional Generative Adversarial Networks (cGANs) have emerged as a powerful framework for controlled data synthesis by incorporating auxiliary information such as class labels into both the generator and discriminator. By modeling the conditional distribution of images given disease categories, cGANs enable the generation of class-specific synthetic samples that closely resemble real medical images. This capability is particularly valuable in healthcare applications where labeled datasets are limited and class imbalance is prevalent. Numerous studies have demonstrated that cGAN-generated images can significantly improve the performance of deep learning models for diagnosis, segmentation, and classification tasks. However, challenges such as training instability, mode collapse, and limited ability to capture complex anatomical variations remain active research areas.

3. PROPOSED SYSTEM

There are two steps in the proposed system. Initially, an unsupervised training is done for a generator to model the data surface. Next step, train the conditional GAN. The learning of the base distribution of medical images is the primary objective at this point. The generators typically use GAN architecture which caters to their best performance with less data. This model maps a noise vector and class label to synthetic image. In GAN framework, The Discriminator will act as a binary classifier. It contests against the generator. It the generator. Thus, it learns to discriminate between real CT scans and synthetic fakes.

Input Data: Real images are resized (128 x 128) along with label encodes conditionally supervised. In the second stage of the process, they train the classifier, and then apply augmentation on their generator The model samples a "soft" label vector from a softmax distribution, instead of a standard label encoding. A class is chosen at random. The generator synthesizes an image that represents a mixture of these soft labels. This ensures that the resulting image maintains structural integrity.This generator produces an image which is a mixture of these soft labels. So that the resulting image does not become a ghostly overlay as happens in the traditional way. The last training set is Real+Generated, which combines real-world data and synthetic data obtained from GANs. Networks such as ResNet-50,etc are trained on this data.

4. METHODOLOGY

The proposed system leverages a two-stage pipeline that involves first establishing a high-fidelity generative base and then utilizing a novel soft-label interpolation scheme intended to bolster the classification for the medical images.

The system is trained on a dataset consisting of CT scan images from three different classes which are Pneumonia and Normal. All these images are cropped and resized as 128×128 datasets for training the generative model.

4.1 Conditional generative modelling

We use the architecture of StyleGAN2, designed to work well in limited data regimes, for modeling conditional distribution of CT scans.

4.1.1 Arrangement of system

We will denote to the starting point of their generative framework as.

$$S = \{G, R, E\} \qquad (1)$$

Where:

- G learns the synthesis of 128 x 128 resolution.
- R is the discriminator network that is trained to recognize a synthetic sample.
- E is the preprocessed **o**f dataset.

4.1.2 Objective Function

The networks are optimized via a supervised mini-max game:

min_G max_D Ex~p_data [logD(x|y)] Ex~p_z [log(1-D(G(Z|Y)))] (2)

- Ex~p_data[logD(x|y)]: The Discriminator seeks to maximize its ability of distinguishing between real vs generated images.
- Ez~p_z [log(1-D(G(Z|Y)))] : The Generator (G) tries to minimize the chance that the Discriminator (D) detects its fake images.
- y Represents Condition Encoded Label (for example [1,0,0] = a particular)

4.2 Latent Space

We will first introduce an equation that allows to generate the combination of the image. This eqation indicates how much the resulting image will like the one selected and how much it will like all the others.

$$c \sim z+ab(\{1,\ldots,K\}) \qquad (3)$$

4.2.1 Choosing Choices

The system picks one of the K classes (any one disease regarding pick) at random to be the "primary" focus of the new image

4.2.2 Class mixing - Vector(ϑ)

To tune the "intensity" of the mix, we define a weight vector using temperature-scaled softmax function.

The string of values in this vector controls how the created image will be towards the target class vs other classes.

$$z_j = \begin{cases} z + b & if\ j = c \\ z & otherwise \end{cases} \qquad (4)$$

- The final mixing weights are obtained using softmax

$$\theta j = \sum_{k=1}^{k} \frac{\exp\left(\frac{z}{T}\right)}{\exp\left(\frac{z}{t}\right)} \qquad (5)$$

- b is the bias applied to the dominant class,
- T is the temperature parameter controlling smoothness,

Lower T produces sharper (more dominant) mixtures, Higher T produces smoother distributions

4.2.3 Soft-Label Sampling (*l*)

Instead of hard-labeling we draw a soft-label vector from a softmax distribution based on the previously established vector:

$$\text{label} = softmax\,\frac{(z+b)}{T} \qquad (6)$$

- This samples a certain soft-label vector where (for example $l = [0.7, 0.2, 0.1]$).
- The concept of the diseases is blended together instead of the original pixel values as done in standard mixup.

4.2.4 Synthetic Image Synthesis

$$\text{Image} = \text{generator}(\text{noise} + \text{label}) \qquad (7)$$

- Generator(G) takes the noise and the soft-label as an input to generate image
- The result is an image that portrays features of the dominating class while containing subtle features of the other classes as defined in the soft-label.

4.3 Classifier Training and Evaluation

Then, real images and generated images are aggregated to train the diagnostic models.

4.3.1 Dataset Combination

$$D_2 = D_1 + \text{generated images} \qquad (8)$$

- This equation simply indicates that we are augmenting our original real dataset with N synthetic pairs.

4.3.2 Evaluation

$$\text{Macro -F1} = \frac{1}{K}\sum_{i=0}^{k} F1_I \qquad (9)$$

- This calculates mean F1-score over all labels.
- The goal is to enlarge this number maximizing the True Positive Rate and minimizing False Negative Rate for disease detecetion

Where e precision and recall of $F1_i$ score are the harmonic mean for a certain class.

$$\text{F1} = 2 \times \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (10)$$

- **Precision:** Measures the accuracy of positive predictions.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (11)$$

- **Recall (Sensitivity):** Measures the model ability of recollecting all the positive actual cases.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (12)$$

5.ARCHITECTURE

This depicts the overall framework of the suggested architecture for medical image classification. It shows the process from image generation of GAN training the classifier on this dataset

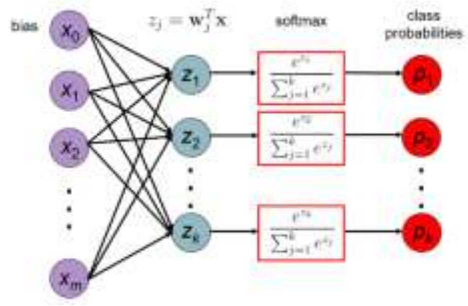

Figure 1: shows the generation of class probabilities

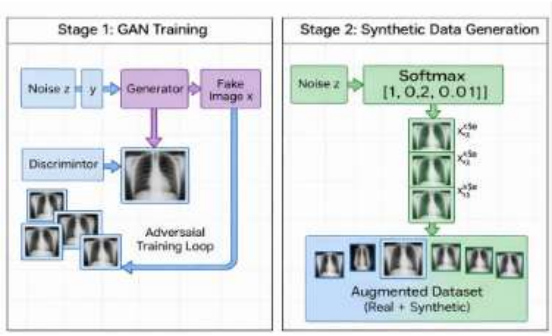

Figure 2: shows a two-stage process for GAN-based medical image augmentation.

6.RESULTS AND DISCUSSIONS

The classification model was enhanced after implementing the proposed technique of generating and augmenting synthetic data as per the experiments.

6.1 Synthetic Data Generation

This depicts that proposed generative framework can yield synthetic images to enhance the training dataset and the data scarcity

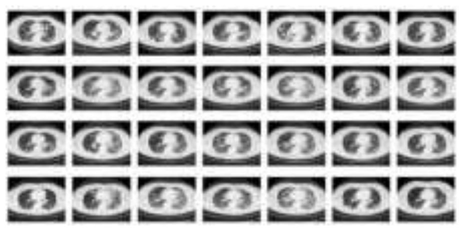

Figure 3 : Shows the generated synthetic images

6.2 Performance Comparison with Baseline Model

The new model does better than the baseline model(CGAN) in terms of its predictive ability.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 03 | March - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Metric	CGAN	Our Model
Accuracy	82.4%	90%
F1-Score	0.79	0.89
Precision	0.81	0.88
Recall (Sensitivity)	0.76	0.93

Table 1: shows a comparison of performance metrics between the baseline model and our proposed model.

6.3 Statistical Performance

The proposed model is shown to be effective from the result Statistical performance of the proposed model suggests strong and reliable classification capability. As also shown in the Table 2, the model reaches an accuracy along with very high macro-averaged evaluation metrics like F1-score, precision and recall indicating a good balanced behaviour across all classes.

Metric	Value
Accuracy	90%
F1-Score(Macro)	0.9075
Precision (Macro)	0.9067
Recall (Macro)	0.9084
Macro-F2 Score	0.9080
False Negative Rate	0.0916

Table 2 : summarizes the overall statistical performance

6.4 Confusion Matrix Analysis

This shows how the model is able to discriminate correctly among the three diagnosis categories indicating an efficacy of the proposed approach

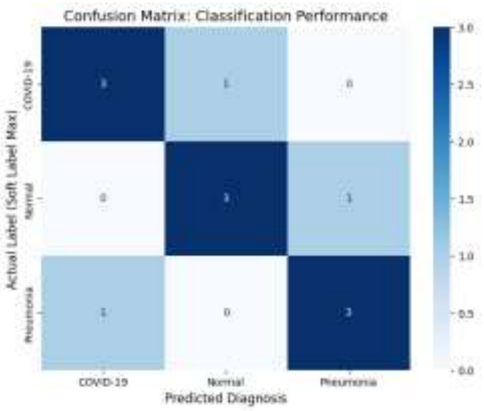

Figure 3 : shows the model's ability to distinguish the three diagnostic categories

6.5 Overall Performance

For the image classification results, we found that our proposed GAN based augmentation method could significantly outperform the baseline model (which was only trained on real life images). Using softmax soft labels we were able to generate synthetic images which enabled the single classifier to learn smoother decision boundaries across the three classes Cancer, Pneumonia, and Normal.

7. CONCLUSION

This study presents a comparative evaluation of generative data augmentation strategies for medical image classification, focusing on chest CT scans for infection detection. By transitioning from a traditional CGAN baseline to an advanced StyleGAN2-ADA framework, we demonstrated a significant improvement in diagnostic performance, achieving a peak accuracy of 90%. A key technical contribution of this work is the implementation of a temperature-scaled Softmax distribution for soft-label sampling. This method facilitated smoother class transitions and more anatomically coherent synthetic images compared to standard pixel-level blending techniques.Our findings indicate that generative augmentation is particularly effective at reducing the False Negative Rate (0.0916), thereby addressing a critical safety requirement in clinical settings where missed diagnoses must be minimized. Furthermore, the statistical performance across Macro-F1, Precision, and Recall confirms that the StyleGAN-augmented dataset allowed the model to learn more robust decision boundaries. Ultimately, this research confirms that combining high-fidelity style-based GANs with probabilistic label mixing offers a viable solution to the data scarcity and privacy challenges inherent in medical deep learning .

8.REFERENCES

1. H. Carlesso, M. E. Patulea, M. Garouani, R. T. Ionescu, and J. Mothe, "GeMix: Conditional GAN-Based Mixup for Improved Medical Image Augmentation," arXiv:2507.15577, Sep. 2025.
2. C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," Journal of Big Data, vol. 6, no. 1, 2019.
3. G. Litjens et al., "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60-88, 2017.
4. H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond Empirical Risk Minimization," in ICLR, 2018.
5. T. Karras, M. Aittala, S. Laine, et al., "Training Generative Adversarial Networks with Limited Data," in NeurIPS, 2020.
6. S. Yun, D. Han, S. J. Oh, S. Chun, J. Choe, and Y. Yoo, "CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features," in ICCV, 2019.
7. V. Verma et al., "Manifold Mixup: Better Representations by Interpolating Hidden States," in ICML, 2019.
8. T. DeVries and G. W. Taylor, "Improved Regularization of Convolutional Neural Networks with Cutout," arXiv:1708.04552, 2017.
9. D. Hendrycks et al., "AugMix: A Simple Data Processing Method to Improve Robustness and Uncertainty," in ICLR, 2020.

10. I. Goodfellow et al., "Generative Adversarial Networks," Communications of the ACM, vol. 63, no. 11, pp. 139-144, 2020.
11. H. Gunraj, L. Wang, and A. Wong, "COVIDNet-CT: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest CT Images," Frontiers in Medicine, vol. 7, 2020.
12. L. van der Maaten and G. Hinton, "Visualizing Data using t-SNE," Journal of Machine Learning Research, vol. 9, pp. 2579-2605, 2008.
13. R. R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," in ICCV, 2017.
14. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in CVPR, 2016..
15. Training Generative Adversarial Networks with Limited Data – Tero Karras, Miika Aittala, Samuli Laine, NeurIPS, 2020.
16. Image-to-Image Translation with Conditional Adversarial Networks – Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros, CVPR, 2017.
17. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks – Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros, ICCV, 2017.
18. Medical Image Synthesis for Data Augmentation – Saeed Shakeri Hossein-Abad et al., Journal of Medical Imaging, 2021.
19. PyTorch – Developed by Meta AI for deep learning model development.
20. StyleGAN2 – Developed by NVIDIA.