CME LIVE OPEN ACCESS

Journal of
Biomedical and Engineering Research

Research Article ISSN: 3065-8780

Synthetic Image Generation for Rare Tumor Detection Using GANs and Stable Diffusion

Ayush Rane*

Monte Vista High School, California, United States.

Corresponding Author: Ayush Rane, Monte Vista High School, California, United States.

Received: 2026 Mar 27 **Accepted:** 2026 Apr 17 **Published:** 2026 Apr 30

Abstract

Rare tumor types present a challenge for machine learning models due to class imbalance and limited sample sizes. In this study, I explore the use of synthetic image generation using both Generative Adversarial Networks (GANs) [1,2] and pre-trained Stable Diffusion (SD) models [3,4] for augmenting a histopathologic dataset of tumor and normal tissue images. I evaluate the impact of synthetic data on a binary classification task using convolutional neural networks. My results show that while synthetic images exhibit low structural similarity (SSIM) to real tumor images, classifier performance improves when augmented with GAN or SD-generated images compared to a baseline CNN trained only on the original dataset, highlighting the potential benefit of synthetic data even when visual similarity is not ideal [5,6]. Metastatic lymph node images are used as an example of a rare tumor type, illustrating the clinical challenge of underrepresented tumor classes in histopathology datasets.

Keywords: Synthetic Images, Machine Learning, Histopathology

1. Introduction

1.1 Background

Histopathologic image datasets are often heavily imbalanced, with rare cancer types underrepresented [1]. This imbalance limits the effectiveness of machine learning classifiers, which tend to underperform on minority classes. Generative models such as GANs have shown promise in synthesizing realistic images to resolve class imbalance [2,3]. Recently, pre-trained models like Stable Diffusion have emerged as an alternative, capable of generating visually realistic images without adversarial training like GANs [4,5]. Metastatic lymph node sections serve as a concrete example of a clinically important but under-sampled category. Detecting these tumors early is critical for patient outcomes, yet data scarcity makes classifier training challenging. To simulate this underrepresentation, I artificially reduced the number of tumor images, creating a rare class scenario. While the focus of this study is on model performance, using metastatic lymph nodes highlights a real-world issue that synthetic image augmentation could help address. This study investigates the effectiveness of GANs and SD in generating synthetic tumor images to augment an artificially imbalanced dataset, and evaluates their impact on classification performance.

1.2 Related Work and Literature Gap

GANs have been applied in medical imaging to address class imbalance, including for histopathologic tumor detection. They can improve classifier performance for rare tumor types by increasing minority class representation. However, GANs require iterative adversarial training, which is computationally intensive and may be unstable [2,3]. Pre-trained models like Stable Diffusion have demonstrated high-quality image generation without adversarial training [4,5]. Despite their success in natural image generation, their application to medical imaging, particularly for augmenting rare tumor datasets, is limited. Most prior work has focused on generating synthetic images without comparing classifier performance across different synthetic data methods or against a baseline CNN trained only on real images. My work addresses this gap by directly comparing GAN and SD-generated synthetic images and evaluating their impact on classifier performance compared to a baseline CNN classifier trained without synthetic augmentation.

2. Methods

2.1 Dataset

I used the Histopathologic Cancer Detection dataset from Kaggle, consisting of 220,025 labeled images (89,117 tumor, 130,908 normal) [7]. To simulate rare tumor occurrence, I randomly sub-sampled the tumor class to 100 images and normal class to 1,000 images.

2.2 Generative Adversarial Network (GAN)

A convolutional GAN was trained on the sub-sampled dataset (100 tumor, 1,000 normal images) to generate synthetic tumor samples. The generator accepted 100-dimensional random noise vectors and upsampled them to produce 128×128 resolution synthetic images. Training was performed for 3,000 epochs with Adam optimization (learning rate = 0.0002, β_1 = 0.5), employing label smoothing and dropout for stability. A total of 300 synthetic tumor images were produced. The Structural Similarity Index Measure (SSIM)

was computed at each epoch between generated and real tumor images to monitor similarity [2,3,5].

2.3 Stable Diffusion (SD)

Synthetic tumor images were also generated using the pre-trained Stable Diffusion v1.5 model [4]. Images were sampled through 20 denoising steps, with a guidance factor used to strengthen alignment to histopathological features. Outputs were resized to 128×128, yielding 100 synthetic tumor images.

Synthetic images were limited to 100 due to high computational costs. SSIM was calculated on the final generated set against real tumor images to assess similarity. No additional fine-tuning on histopathology data was performed.

2.4 Classifier

A CNN classifier was trained on three datasets: the original sub-sampled dataset containing 100 tumor and 1,000 normal images, a GAN-augmented dataset consisting of 100 real tumor images combined with 300 synthetic images and 1,000 normal images, and an SD-augmented dataset containing 100 real tumor images combined with 100 synthetic images and 1,000 normal images. The CNN architecture consisted of two convolutional layers with max pooling, followed by a dense layer and a sigmoid output. To address class imbalance, class weights were computed during training. Each dataset was split into training (80%), validation (10%), and test (10%) sets, with the following numbers:

- Original sub-sampled dataset: Training 880 images (80 tumor, 800 normal), Validation 110 images (10 tumor, 100 normal), Test 110 images (10 tumor, 100 normal)

- GAN-augmented dataset: Training 1,120 images (320 tumor, 800 normal), Validation 140 images (40 tumor, 100 normal), Test 140 images (40 tumor, 100 normal)

- SD-augmented dataset: Training 960 images (160 tumor, 800 normal), Validation 120 images (20 tumor, 100 normal), Test 120 images (20 tumor, 100 normal)

Classifier models were trained for 10 epochs with a batch size of 32. Figures 1 and 2 visualize the progressions of train and validation losses over epochs for all 3 approaches.

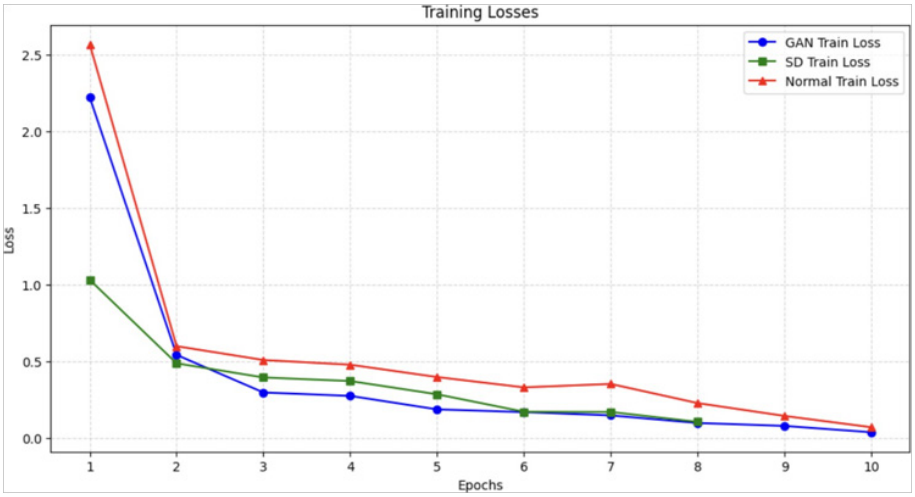

Figure 1: Training Losses Across Epochs

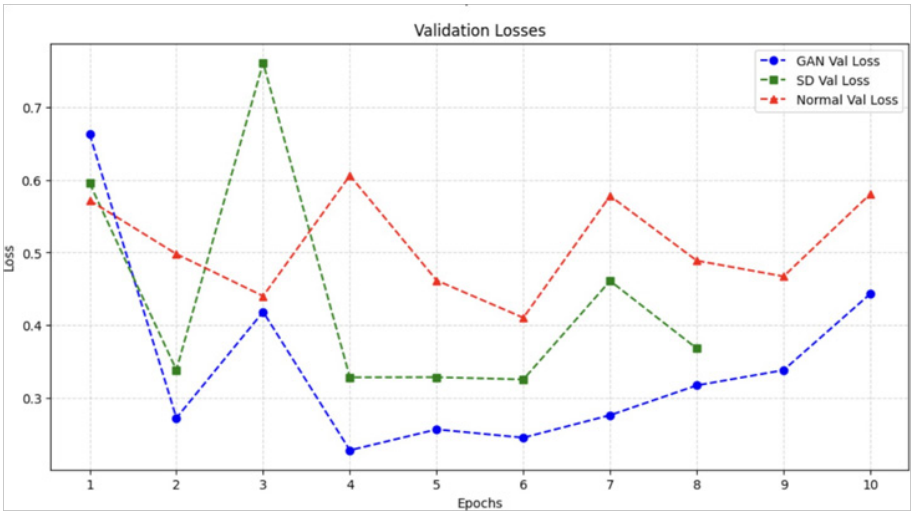

Figure 1: Training Losses Across Epochs

2.5 Data Preprocessing

All images were resized to 128×128 pixels and normalized to a range for standardization [1,1]. Color channels were converted to RGB to match original histopathology images. Corrupted or missing files were removed during preprocessing to reduce noise. No additional image augmentations, such as flips or rotations, were applied. Diversity in the training set was instead enhanced through synthetic images generated by GANs and Stable Diffusion models. This preprocessing ensures the CNN can learn meaningful patterns from both real and synthetic tumor images.

3. Results

3.1 Evaluation

Classifier performance was measured using accuracy, F1-score, and ROC AUC (See Figure 3). SSIM quantified structural similarity between synthetic and real tumor images [5]. While SSIM provides an estimate of visual similarity, it does not fully capture the effectiveness of synthetic images for classifier learning.

Model	Synthetic Images	Avg SSIM	Test Accuracy	Tumor F1-score	ROC AUC
GAN	300	0.052–0.094	0.9	0.81	0.916
SD	100	0.0825	0.88	0.57	0.856
Original	0	N/A	0.85	0.3	0.692

Figure 3: Performance Metrics

GAN-generated images improved classifier performance substantially, despite low SSIM [5,6].

SD-generated images provided moderate improvement. SSIM values indicate that the synthetic images do not closely resemble real tumor images visually, yet still benefit classification. These results suggest that classifier performance is not only correlated with visual similarity (SSIM). Even with low SSIM, GAN-generated images enabled the model to extract useful features for tumor detection [5,6]. This suggests that despite dissimilar images, having more images to train on overall, and the features extracted from it can lead to strong classifier accuracy.

4. Discussion

4.1 Model Performance Comparison

GAN-generated images substantially improved classifier performance for rare tumor detection, even though SSIM values indicate low visual similarity to real tumor images. CNNs are able to extract meaningful spatial and structural features from synthetic images, enhancing learning for underrepresented classes [2,5]. SD-generated images provided moderate improvement, likely due to the smaller synthetic dataset and single inference pass generation process. The original, sub-sampled dataset performed the worst, highlighting the critical role of augmentation. Comparing performance: GAN > SD > original across accuracy, F1-score, and ROC AUC. The Tumor F1-score increased from 0.30 (original) to 0.81 (GAN), showing strong benefits from GAN augmentation. SD-generated images raised the F1-score to 0.57, demonstrating that even smaller synthetic datasets provide some benefit, though less than GANs.

4.2 Clinical Relevance

The metastatic lymph node example underscores a broader clinical issue: rare tumors are underrepresented in datasets, leading to missed diagnoses or delayed detection. Augmenting datasets with synthetic images can improve classifier training, potentially supporting earlier intervention and better patient outcomes. Synthetic data generation also reduces reliance on costly collection of rare histopathology images, offering a cost-effective solution.

4.3 Visual Similarity vs. Utility

Visual similarity is not the sole determinant of synthetic data utility. Low SSIM does not prevent meaningful learning; synthetic images can encode statistical patterns and feature patterns that models utilize for improved classification. Evaluating synthetic image utility requires both quantitative metrics (accuracy, F1-score, ROC AUC) and qualitative analysis of model behavior.

4.4 Summary of Key Findings

Augmenting the dataset with GAN-generated synthetic tumor images resulted in the largest improvement in CNN classifier performance, while Stable Diffusion-generated images provided more modest gains. These results emphasize that both the choice of generation method and the number of synthetic images influence classifier performance. Importantly, even synthetic images with low structural similarity (SSIM) to real tumors can improve detection of underrepresented tumor types, addressing class imbalance that often limits early and accurate diagnosis in clinical histopathology.

4.5 Limitations

This study has several limitations that suggest directions for future work. The Stable Diffusion-generated dataset was relatively small, which may have constrained its impact on classifier performance. The CNN architecture was kept simple to evaluate the effect of synthetic images on classifier performance, and only SSIM was used to quantify image similarity. Additionally, the analysis was limited to a binary classification task, and interpretability of CNN decisions or clinical validation of synthetic images was not assessed. Despite these constraints, the study demonstrates

that synthetic image augmentation can enhance classifier performance, highlighting its potential utility for addressing underrepresented tumor types in clinical datasets. This can guide future research toward more complex architectures, larger synthetic datasets, and multi-class applications.

4.6 Future Directions

Future work could address these limitations by generating larger synthetic datasets and leveraging more advanced CNN architectures, such as Res Net. Expanding the study to multi-class tumor detection would provide a more comprehensive evaluation of synthetic image augmentation. Exploring alternative similarity metrics, including LPIPS or feature-space measures, may yield a better understanding of the utility of synthetic images. Finally, integrating synthetic data into clinical workflows for diagnosis and extending experiments to multiple histopathology datasets would enhance generalizability and clinical relevance.

5. Conclusion

Synthetic image generation, particularly using GANs, can enhance tumor detection in imbalanced histopathologic datasets, even when visual similarity to real images is low. Pre-trained SD provides a simpler alternative, though more images may be needed to match GAN performance. Using metastatic lymph nodes as an example illustrates the importance of addressing rare tumor underrepresentation. This work underscores the potential of synthetic data augmentation to improve rare tumor detection in clinical settings and provides directions for future research [1-6].

References

1. Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018, April). Synthetic data augmentation using GAN for improved liver lesion classification. In 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) (pp. 289-293). IEEE.
2. Changhee, H. A. N., Kohei, M. U. R. A. O., Shin'ichi, S. A. T. O. H., & Hideki, N. A. K. A. Y. A. M. A. (2019). Learning more with less: GAN-based medical image augmentation. Medical Imaging Technology, 37(3), 137-142.
3. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.
4. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10684-10695).
5. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 13(4), 600-612.
6. Xue, Y., Ye, J., Zhou, Q., Long, L. R., Antani, S., Xue, Z., ... & Huang, X. (2021). Selective synthetic augmentation with HistoGAN for improved histopathology image classification. Medical image analysis, 67, 101816.
7. Zhai, J., Shen, W., Singh, I., Wanyama, T., & Gao, Z. (2020). A review of the evolution of deep learning architectures and comparison of their performances for histopathologic cancer detection. Procedia Manufacturing, 46, 683-689.