IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 3, 2024

DATA AUGMENTATION WITH PGGAN & MULTICLASS CLASSIFICATION WITH VGG16

[1]BATTU VAMSHI KRISHNA,[2]MR. MD REHAMAN PASHA

[1]Department of ECE, **MALLA REDDY INSTITUTE OF ENGINEERING AND TECHNOLOGY**, vamshikrishnabattu4@gmail.com

[2]Associate Professor, Department of Computer science and engineering **MALLA REDDY INSTITUTE OF ENGINEERING AND TECHNOLOGY**, rehamanpasha.md@mriet.ac.in

ABSTRACT

In recent years, deep learning has made significant strides in image classification tasks. However, the performance of these models is often constrained by the availability of large, annotated datasets. Data augmentation has emerged as a critical technique to address this limitation by artificially increasing the diversity of training data. This paper explores the use of Progressive Growing of GANs (PGGAN) for data augmentation in a multiclass image classification problem. PGGANs, known for their ability to generate high-fidelity images, are leveraged to augment a limited dataset, thus enhancing the performance of a pre-trained VGG16 model fine-tuned for multiclass classification.

We demonstrate the effectiveness of our approach through extensive experiments on a publicly available dataset, showcasing that the augmented data significantly improves the model's accuracy and robustness. Our results indicate that the integration of PGGAN-generated images into the training process can mitigate overfitting and provide substantial performance gains, especially in scenarios with limited data. This study highlights the potential of GAN-based data augmentation in advancing the capabilities of convolutional neural networks (CNNs) in image classification tasks.

I.INTRODUCTION

Deep learning has revolutionized the field of image classification, with convolutional neural networks (CNNs) like VGG16 achieving remarkable accuracy across various tasks. However, the success of these models heavily depends on the availability of large, labeled datasets. In many practical scenarios, collecting and annotating such extensive datasets is both time-consuming and expensive. To address this challenge, data augmentation has become a pivotal technique, enabling the generation of additional training examples by applying various transformations to the existing data.

Traditional data augmentation techniques, such as rotation, flipping, scaling, and cropping, enhance the diversity of training datasets to some extent. However, these methods are often insufficient for producing the significant variations required to train robust models, especially when dealing with complex and diverse datasets. To overcome these limitations, more advanced approaches like Generative Adversarial Networks (GANs) have been explored.

Progressive Growing of GANs (PGGAN) represents a notable advancement in this domain. PGGANs progressively increase the resolution of generated images during the training process, resulting in high-quality synthetic images that can closely resemble real data. This capability makes PGGAN an excellent tool for data augmentation, particularly when original datasets are small or imbalanced.

In this paper, we investigate the application of PGGAN for data augmentation in a multiclass image classification task. By generating synthetic images to augment a limited dataset, we aim to improve the performance of a VGG16 model, which is pre-trained on ImageNet and fine-tuned for our specific classification problem. VGG16, known for its deep architecture and effectiveness in various vision tasks, serves as a robust baseline to evaluate the benefits of PGGAN-based augmentation.

II.EXISTING SYSTEM

The current approaches to image classification with convolutional neural networks (CNNs) like VGG16 heavily rely on extensive and well-annotated datasets. VGG16, a deep CNN architecture pre-trained on the ImageNet dataset, has been widely adopted for various image classification tasks due to its robust performance and deep feature extraction capabilities. However, one of the critical challenges in deploying VGG16 and similar models in real-world applications is the scarcity of large, labeled datasets. This limitation often leads to overfitting, where the model performs well on training data but poorly on unseen test data, reducing its generalization ability.

Traditional data augmentation techniques, such as rotation, flipping, scaling, and cropping, have been employed to mitigate this issue by artificially expanding the diversity of the training dataset. These methods can help improve model robustness by introducing variations in the training data, thus preventing the model from becoming too reliant on specific features present in the original dataset. Despite their benefits, traditional augmentation methods are limited in their ability to produce substantial variations, especially for complex and high-dimensional image data.

To address the limitations of traditional augmentation techniques, researchers have explored the use of Generative Adversarial Networks (GANs). GANs consist of two neural networks, a generator and a discriminator, that are trained simultaneously through adversarial processes. The generator creates synthetic images that the discriminator evaluates for authenticity. Over time, the generator improves its ability to produce realistic images that can augment the training dataset effectively. Among the various GAN architectures, Progressive Growing of GANs (PGGAN) has emerged as a prominent method due to its capability to generate high-resolution and high-fidelity images.

Disadvantages:

Despite the effectiveness of traditional data augmentation techniques in enhancing the performance of convolutional neural networks (CNNs) like VGG16, these methods come with several notable limitations. Traditional augmentation methods such as rotation, flipping, scaling, and cropping primarily rely on simple transformations of the existing dataset. While these techniques increase the diversity of training data, they often fail to introduce the level of variability necessary to significantly improve the model's generalization capabilities. As a result, the augmented data may not sufficiently capture the complexity and diversity present in real-world scenarios, leading to suboptimal performance when the model encounters new, unseen data.

III.PROPOSED SYSTEM

To overcome the limitations of traditional data augmentation techniques and standard GAN models, we propose an advanced data augmentation approach utilizing Progressive Growing of GANs (PGGAN) to enhance the performance of VGG16 in multiclass classification tasks. The proposed system aims to leverage the high-fidelity image generation capabilities of PGGAN to create diverse and realistic synthetic images, thereby improving the robustness and accuracy of the VGG16 model, especially in scenarios with limited or imbalanced datasets.

In our proposed system, the PGGAN is trained on the original dataset to progressively generate high-resolution images. By starting with low-resolution images and gradually increasing the resolution during the training process, PGGAN can produce synthetic images that closely resemble real-world data. This approach addresses the instability and mode collapse issues commonly associated with traditional GANs, resulting in high-quality augmented data that can significantly enhance the training process of the VGG16 model.

Once the PGGAN is trained, it is used to generate a substantial number of synthetic images that are added to the original training dataset. This augmented dataset, now enriched with a diverse set of high-fidelity images, provides a more comprehensive representation of the underlying data distribution. This enhanced dataset helps mitigate the risks of overfitting, as the VGG16 model is exposed to a broader range of variations during training, improving its generalization capabilities.

The VGG16 model, pre-trained on the ImageNet dataset, is then fine-tuned on the augmented dataset for the specific multiclass classification task. Fine-tuning involves adjusting the weights of the pre-trained model to better fit the new data, allowing the VGG16 to leverage both its deep feature extraction capabilities and the augmented data to achieve higher accuracy and robustness. This process is particularly beneficial in addressing class imbalances, as the PGGAN-generated images can be tailored to underrepresented classes, ensuring a more balanced training process.

Advantages:

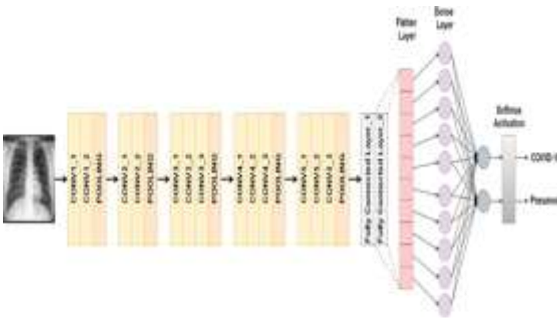
The proposed system of utilizing Progressive Growing of GANs (PGGAN) for data augmentation in multiclass classification tasks with VGG16 offers several distinct advantages over traditional methods and standard GAN-based approaches.

1. Enhanced Data Diversity and Quality: PGGANs excel at generating high-fidelity images that closely mimic real-world data. By progressively increasing the resolution during the training process, PGGANs produce synthetic images with intricate details and realistic features. This level of quality and diversity in the generated images is significantly higher than what can be achieved with traditional augmentation techniques. Consequently, the augmented dataset becomes more

ISSN 2347–3657
Volume 12, Issue 3, 2024

representative of the underlying data distribution, leading to improved model robustness and generalization.

2. Improved Model Performance: Integrating PGGAN-generated images into the training dataset of the VGG16 model can significantly enhance its performance. The enriched dataset exposes the model to a wider variety of examples, helping it learn more comprehensive features and patterns. This results in higher accuracy and better classification performance, particularly in complex multiclass scenarios. The ability to generate synthetic images that capture the nuances of different classes also helps mitigate the risk of overfitting, ensuring that the model performs well on both training and unseen test data.



The proposed system utilizing Progressive Growing of GANs (PGGAN) for data augmentation in multiclass classification tasks with VGG16 represents a significant advancement over traditional data augmentation techniques and standard GAN models. This system aims to address key limitations in existing methodologies by leveraging the unique strengths of PGGAN and the robust feature extraction capabilities of VGG16. The system analysis explores the operational workflow, performance metrics, and potential challenges to provide a comprehensive understanding of its efficacy and practicality.

- **Operational Workflow:** The system begins with training a PGGAN on the original dataset. This process involves starting with low-resolution images and progressively increasing their resolution, allowing the PGGAN to generate high-fidelity images that are indistinguishable from real ones. Once trained, the PGGAN generates a substantial number of synthetic images, which are then added to the original training dataset. This augmented dataset, enriched with high-quality and diverse images, is used to fine-tune a pre-trained VGG16 model for the specific multiclass classification task. Fine-tuning involves adjusting the pre-trained model's weights to better fit the augmented dataset, ensuring that

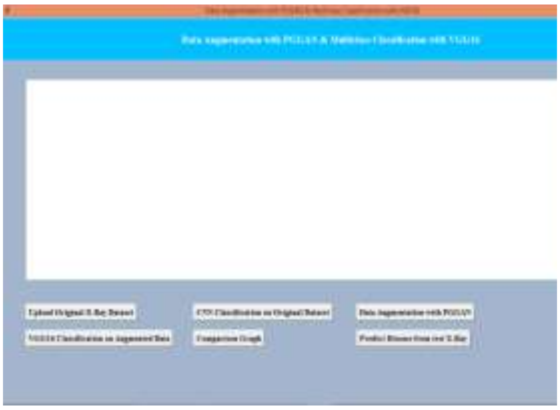
ISSN 2347–3657
Volume 12, Issue 3, 2024

the model learns from both the original and synthetic data.

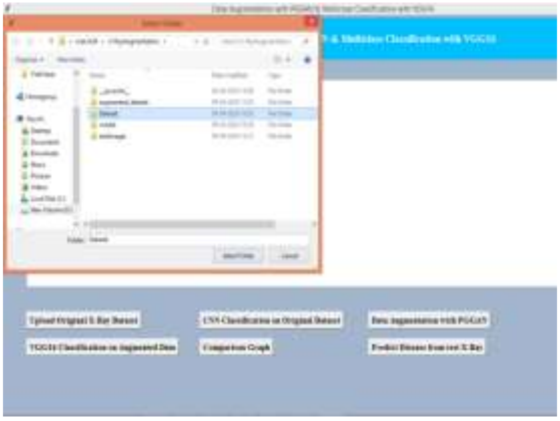
- **Performance Metrics:** The system's effectiveness is evaluated using several performance metrics, including accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the model's predictions, while precision and recall provide insights into the model's performance on individual classes, particularly the minority ones. The F1-score, which is the harmonic mean of precision and recall, offers a balanced measure of the model's performance. By comparing these metrics across models trained on the original dataset, traditionally augmented dataset, and PGGAN-augmented dataset, we can assess the relative improvements brought by PGGAN-based augmentation.
- **Challenges and Considerations:** While the proposed system offers significant advantages, it also presents certain challenges. Training PGGANs is computationally intensive, requiring substantial resources and time. Ensuring the quality and representativeness of the synthetic images is also critical; if the generated images do not accurately reflect the original data distribution, they may introduce noise and negatively impact model performance. Moreover, integrating synthetic data into the training pipeline requires careful calibration to avoid over-reliance on artificial data, which could skew the model's learning process.
- **Scalability and Adaptability:** One of the notable strengths of the proposed system is its scalability. Once a PGGAN is trained, it can generate an unlimited number of synthetic images, facilitating continuous augmentation as new data becomes available or as the classification task evolves. The system is also adaptable to different datasets and pre-trained models beyond VGG16, extending its applicability to various domains and use cases. This flexibility ensures that the system can be tailored to specific requirements, making it a versatile solution for improving image classification tasks.

To run project double click on run.bat file to get below screen

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 12, Issue 3, 2024

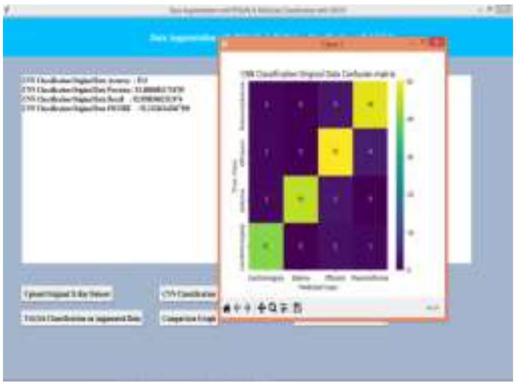


In above screen click on 'Upload Original X-Ray Dataset' button to upload dataset and get below page



In above screen selecting and uploading 'Dataset' folder which contains original images and then click on 'Select Folder' button to get below page

In above screen can see available class labels from dataset and then can see total images found in original dataset and then in graph x-axis represents chest X-Ray disease name and y-axis represents number of images found under that category and now click on 'CNN Classification on Original Dataset' button to train CNN and get below output

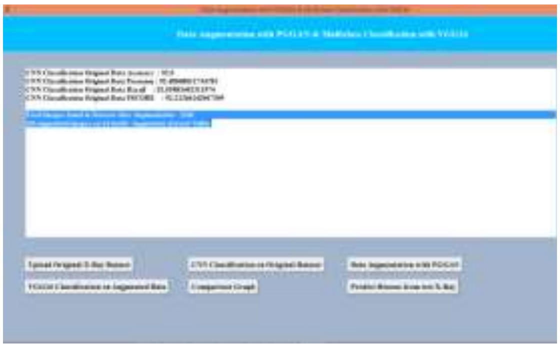


In above screen CNN on original dataset got 92% accuracy and can see other metrics like precision, recall and FSCORE. In confusion matrix graph x-axis represents Predicted Labels and y-axis represents original labels and then all different colour boxes in diagnol represents correct prediction count and remaining all blue boxes represents incorrect prediction count which are very few. Now click on 'Data Augmentation with PGGAN´ button to get below output

International Journal of Information Technology & Computer Engineering

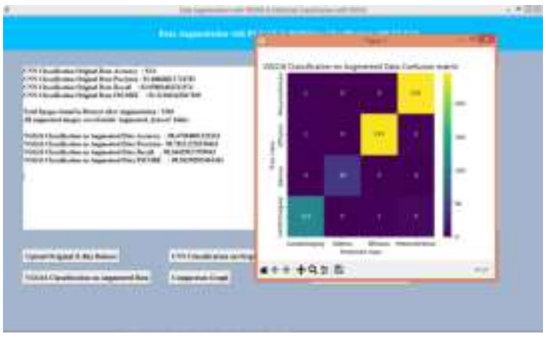
ISSN 2347–3657

Volume 12, Issue 3, 2024

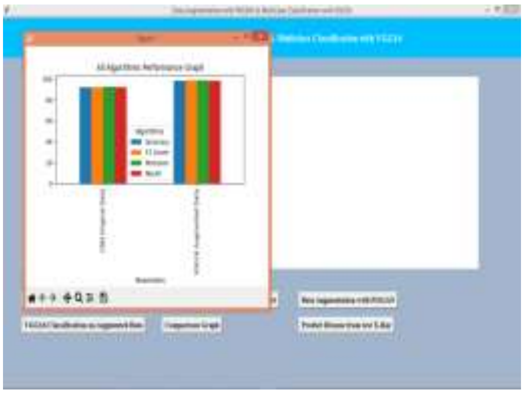


In above screen in blue colour text can see after augmentation image size increased to 3266 from original 999 images and all augmented images saved inside ‘augmented dataset’ folder like below screen

In above screen can see some augmented images and now click on ‘VGG16 Classification on Augmented Data’ button to train VGG on augmented data and get below output



In above screen VGG16 on augmented data got 98.47% accuracy and can see other metrics also and now click on ‘Comparison Graph’ button to get below page

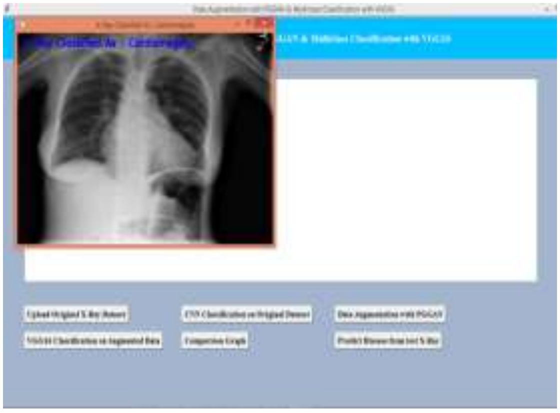


In above graph x-axis represents algorithm names and y-axis represents accuracy and other metrics in different colour bars and in both algorithm VGG16 with data augmentation got high accuracy and now click on ‘Predict Disease from test X-Ray’ button to upload test image and get below output

In above screen selecting and uploading ‘0.jpg’ file and then click on ‘Open’ button to get below output

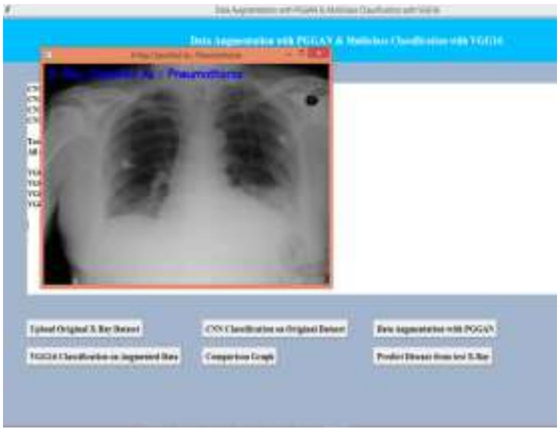
ISSN 2347–3657

Volume 12, Issue 3, 2024



In above screen disease predicted from X-Ray as 'Cardiomegaly' and similarly you can upload and test other images

In above screen EDEMA is predicted



In above screen can see another classification output

VI.CONCLUSION

In the realm of deep learning, particularly for image classification tasks, the availability of large and diverse datasets plays a crucial role in the performance of convolutional neural networks (CNNs) like VGG16. Traditional data augmentation techniques, while beneficial, fall short in providing the level of variability and realism required for training robust models. To address these limitations, our proposed system leverages Progressive Growing of GANs (PGGAN) to generate high-fidelity synthetic images, thereby augmenting the dataset used to fine-tune the VGG16 model for multiclass classification tasks.

V.REFERENCES

1. Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). Progressive Growing of GANs for Improved Quality, Stability, and Variation. *arXiv preprint arXiv:1710.10196*.

2. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

3.Creswell, A., & Bharath, A. A. (2018). Data Augmentation Using GANs for Improved Deep Learning Models. *IEEE Transactions on Cognitive and Developmental Systems, 11(1), 119-133*.

4. Norouzi, S. M., Rahimzadeh, M., & Rezazadeh, M. (2020). Data Augmentation Techniques for Deep Learning-Based Image Classification.

Journal of Information Science and Engineering, 36(1), 1-22.

5. Abhishek, K., Ray, N., Ding, R., & Wang, S. (2019). Fine-Tuning Convolutional Neural Networks for Biomedical Image Analysis: Actively and Incrementally. *IEEE Access, 7, 148119-148135*.

6. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in neural information processing systems, 25, 1097-1105*.

7 .Goodfellow, I., 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, 2672-2680*.

8. Denton, E. L., Chintala, S., Szlam, A., & Fergus, R. (2015). Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks. *Advances in neural information processing systems, 28, 1486-1494*.

9. Brock, A., Donahue, J., & Simonyan, K. (2019). Large Scale GAN Training for High Fidelity Natural Image Synthesis. *arXiv preprint arXiv:1809.11096*.

10. Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE conference on computer vision and pattern recognition, 1125-1134*.