IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 6, June 2022

DOI: 10.17148/IARJSET.2022.9633

Mitigating unbalanced classes using CycleGAN

Mathias Russel Rudolf Richard[1], Dr. R. Savitha[2]

[1,]Students, Master of Computer Applications, RV College of Engineering, Bengaluru, India

[2]Associate Professor, RV College of Engineering, Bengaluru, India

Abstract : Deep learning methodologies have been used to create applications with utility in fields such as computer vision, natural language processing, speech recognition, etc. To train such neural networks that work on these specific tasks, the process of the collection of data becomes paramount. As massive datasets become more prevalent, the models need to be unbiased and detect the minority of classes in the dataset without population bias. To mitigate this problem, the Cycle Generative Adversarial Neural networks can be utilized. Depending on the datasets, one can determine whether minority classes would lie in space where they could be transformed into the different class. For image datasets CycleGANs can be used for image synthesis.

Keywords: Generative Adversarial Networks, CycleGANs, Unbalanced distribution of classes, Deep learning, Synthesis

I.INTRODUCTION

Advancements in Deep Learning techniques are being used to create softwares with utility in fields such as computer vision, natural language processing, speech recognition, etc. The advent of this was in computer vision, as a requirement for more accuracy models is of need, big data pipelines are being formed to get hold of massive datasets.
Massive availability of data has given rise to unbalanced classes in datasets to a greater degree, to solve the problem of model learning inefficiently from this unbalanced data, one could use CycleGANs to generate new data different from only a single labeled data instance. Deep learning models learnt on naive datasets perform poorly and biases might come forefront from such models, more data and better generalizations of models could be obtained using this methodology
This leads into the problem that the population distribution of the classes of the dataset will seep into the collected dataset. Even with naive data augmentation techniques, the distribution still remains intact. Leading to the model making biased choices for optimizing the loss function.

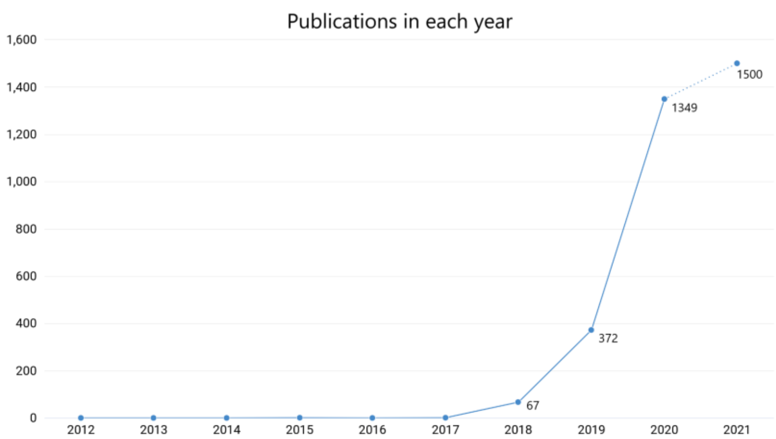

Fig 1: Publication of papers related to synthetic media from 2012 to 2021, sourced from https://app.dimensions.ai on June 3, 2022.

The above figure shows the generation of media spearheaded by modern deep learning algorithms, this shows that there's a scope for realistic synthetic images to be used for classification models since they are almost naked to the human eyes from its differences and would allow to increase the size of the image datasets. Generative models in deep learning are a relatively newer domain with generative adversarial neural networks and its variants being the utilized algorithms to create generative models
This problem can be partially solved by the utilization of CycleGANs. The advent of highly sophisticated generative algorithms such as GANs allows the synthesis of images that might make the distribution of the images in the collected dataset more uniform. Thus allowing for the model to learn the features of each class instead of the distribution of the classes and superseding the intention of the task or overfitting to the dataset and distribution.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified ⁙ Impact Factor 7.105 ⁙ Vol. 9, Issue 6, June 2022
DOI: 10.17148/IARJSET.2022.9633

Google has the resources to train large scale models using their distributed systems network of Tensor Processing Units (TPUs) and one of their major business model centers around. Amazon depends on large recommendation systems to some degree in some of their major products and which also needs them to have a scalable training architecture. Since their recent advancement with DALL-E 2 GAN which is a text to image GAN, OpenAI might have already implemented such a system but if so then its an internal system. These vendors and their current solution for a similar problem could be utilized for mitigating the unbalanced distribution of data in image datasets.

II.LITERATURE SURVEY

The paper is authored by Jiang, Y., Chang, S., & Wang, Z. It is titled "TransGAN: Two Pure Transformers Can Make One Strong GAN, and That Can Scale Up"[1]. It consists of a Reformer based discriminator and generator to achieve super-resolution in images.
The paper is authored by Karras, T., Laine, S., & Aila, T. It is titled "A style-based generator architecture for generative adversarial networks"[2]. The paper uses a different approach than the standard approach for the creation of GANs using a style based approach for generators.
The paper is authored by Hirose, S., Wada, N., Katto, J., & Sun, H. It is titled "ViT-GAN: Using Vision Transformer as Discriminator with Adaptive Data Augmentation."[3] The paper proposes the utilization of Vision transformer as its discriminator in a GAN for reducing the total model size while maintaining the same model efficiency and carbon footprint.
The paper is authored by Frégier, Y. and Gouray, J.B. It is titled "Mind2Mind: transfer learning for GANs."[4] The authors propose the use of pre-trained layers of both the generator and discriminator to reduce the training time of GANs.
The paper is authored by Kwon, Y.H. and Park, M.G. The paper is titled "Predicting future frames using retrospective cycle GAN".[5] The authors propose the use of a CycleGAN to predict frames in video data by the use of two separate discriminators.
The paper is authored by Kishore, P. S. R., Bhunia, A. K., Ghose, S., & Roy, P. P. It is titled "User constrained thumbnail generation using adaptive convolutions.".[6] The authors propose the use of specific metrics to generate characteristic images of videos as their thumbnails.
The paper is authored by Mansourifar, H., Chen, L., & Shi, W. It is titled "Virtual big data for GAN based data augmentation.".[7] The authors introduce a methodology to steamline big data pipelines to improve data passing to GAN for their training.
The paper is authored by Yuan, Y., Liu, S., Zhang, J., Zhang, Y., Dong, C., & Lin, L. It is titled "Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks."[8] The authors utilize unsupervised learning to discard pair based learning of image transformation for super-resolution based data generation.
The paper is authored by Brock, A., Donahue, J., & Simonyan, K. It is titled "Large scale GAN training for high fidelity natural image synthesis."[9] The authors train a massive GAN model and account for how as models get bigger in scope their instability trends increase.
The paper is authored by Adler, J., & Lunz, S. It is titled "Banach wasserstein GAN."[10] The authors propose a general framework for WGANs in higher dimensional space for automated feature selection in N-dimensional manifolds.
The paper is authored by Karras, T., Aila, T., Laine, S., & Lehtinen, J. It is titled "Progressive growing of GANs for improved quality, stability, and variation."[11] The authors propose an incremental approach of simultaneously making generators and discriminators to change parameters for faster convergence.

III.IMPLEMENTATION

The proposed solution involves creation of pairs of images of different distributions, this is then trained on a standard CycleGAN. The six term loss function is optimized for image- to-image translation. In the final production model. This will be then used for generation images which are in minority, such as for a rare disease.
The fundamental transformation to translate from one image to another. One has to denoise the data. Standard denoising autoencoders can be utilized for such tasks. One has to also observe whether the classes are relatively similar to being within the first place. For example, Equine to Equine like animals can be transformed relatively easily whereas an Equine to feline would be an extremely challenging task for the model to learn and might not even be possible if the models are too small thus, having less learnable parameters.
Since leaves are similar to each other. The author has decided to try this implementation on the well known Cassava Leaf Disease. In this dataset the images are paired with each other and transformed from healthy to sich leaves and vice versa. Even though the difference is subtle the generated leaves are sick.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 6, June 2022

DOI: 10.17148/IARJSET.2022.9633

Fig 2: The input image is a healthy leaf, the generated image is an ill foliage and the cycled image is supposed to be the closest to the identity loss of the healthy leaf

In Fig. 2, The input image is a healthy leaf, the generated image is an ill foliage and the cycled image is supposed to be the closest to the identity loss of the healthy leaf. These parameters and features will be learnt by the vision based classification models.[12] Since GANs in and of themselves don't have any metric for a convergence of accuracy. One has to test whether the relatively complex data generation process yields to a better accuracy by training image classification models over the non-transformed dataset, naively transformed dataset and CycleGAN generated images amalgamated in the dataset respectively.

Data Augmentation of the sample in machine learning and deep learning, image and video data based models could boost in performance. One has to compare the scores of the chosen metric of each model with each other and recognized the best approach and its shortcomings

IV.ANALYSIS

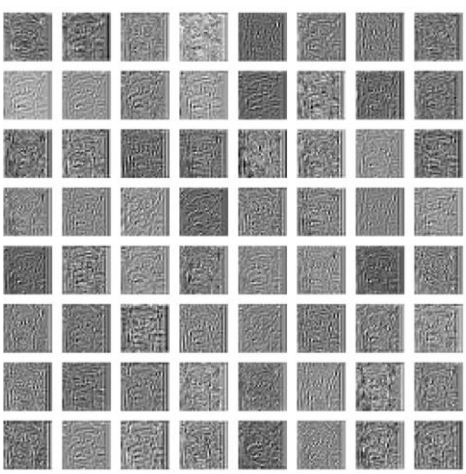

Figure 3: middle layers of the CycleGAN generator visualized
Figure 4: last few layers of the CycleGAN generator visualized

Figure 3 shows a more abstract representation of the features that are passed from the previous layer namely from the layer 23, which are then forward passed to layer 24 as shown. Figure 4 shows the features that the model’s last convolutional layer has learnt and which will then be passed to the next layer which are 2 sets of fully connected dense layers to generate images.

This methodology will allow for a more scalable data augmentation methodology with a slight high computation threshold for the creation of the GAN model. It is less time consuming and resource intensive than other alternatives for achieving similar quality of outcome. CycleGAN, being an unsupervised learning algorithm allows for lesser time to gather data for a supervised learning task

DOI: 10.17148/IARJSET.2022.9633

V.RESULTS AND CONCLUSION

In this paper, the problem tackled which revolves around the biased learning of modern vision classifiers is mitigated by data augmentation from CycleGANs, This allows for a wider flexibility of the model learning from the data, since it leaves no room for the model to learn anything from the very distribution of the data.
The CycleGAN is trained using the dataset which is to be classified and an unsupervised learning pairing is used for its training. In reference to the augmentation load on the GAN, the idea scales better than manual naming of entities by individuals while also generating a relatively significant amount of data since GAN. GANs can generate a potentially infinite number of data as long as the generation seed space (all possible number of seed vectors) is infinite. After the initial threshold cost of training the GAN, there'd be a very small scalable cost thus reducing the training impact on a large scale. This methodology can be used to generate more extensive and better quality datasets if data available is in relatively little quantity.

REFERENCES

[1] Jiang, Y., Chang, S., & Wang, Z. (2021). Transgan: Two pure transformers can make one strong gan, and that can scale up. Advances in Neural Information Processing Systems, 34.
[2] Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4401-4410).
[3] Hirose, S., Wada, N., Katto, J., & Sun, H. (2021, June). ViT-GAN: Using Vision Transformer as Discriminator with Adaptive Data Augmentation. In 2021 3rd International Conference on Computer Communication and the Internet (ICCCI) (pp. 185-189). IEEE.
[4] Frégier, Y. and Gouray, J.B., 2021, July. Mind2Mind: transfer learning for GANs. In the International Conference on Geometric Science of Information (pp. 851-859). Springer, Cham.
[5] Kwon, Y.H. and Park, M.G., 2019. Predicting future frames using retrospective cycle gan. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1811-1820).
[6] Kishore, P. S. R., Bhunia, A. K., Ghose, S., & Roy, P. P. (2019, May). User constrained thumbnail generation using adaptive convolutions. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1677-1681). IEEE.
[7] Mansourifar, H., Chen, L., & Shi, W. (2019, December). Virtual big data for GAN based data augmentation. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 1478-1487). IEEE.
[8] Yuan, Y., Liu, S., Zhang, J., Zhang, Y., Dong, C., & Lin, L. (2018). Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 701-710).
[9] Brock, A., Donahue, J., & Simonyan, K. (2018). Large scale GAN training for high fidelity natural image synthesis. arXiv preprint arXiv:1809.11096.
[10] Adler, J., & Lunz, S. (2018). Banach wasserstein gan. Advances in Neural Information Processing Systems, 31.
[11] Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). Progressive growing of GANs for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196.
[12] Dimitri Oliveira, Cassava Leaf Disease - CycleGAN data augmentation using TPUs, 2020, https://www.kaggle.com/code/dimitreoliveira/cassava-leaf-disease-cyclegan-data-augmentation