Sample Generating for Generalisation

Amin Moradi[1 [s2588862]]

Leiden University, Netherlands

Abstract. The ability to generalise from only one single data point is effortless for human cognition and the more informative the a data point is, the more we can learn from it. In particular, humans have the ability to generalise well with very few data. In this paper we examine the ability of neural networks in generalising from only one sample data point by unconditional permutation and data augmentation techniques. We developed an image generator and image preprocessing pipeline using SinGans[10] and Data Augmentation techniques and compared our generated samples effectiveness in classification accuracy against Mini-Imagenet dataset using a Convolutional Neural Network (CNN).

Keywords: Sample Generating · Generative Adversarial Networks(GANs) · Generalisation.

1 Introduction

Generative Adversarial Networks (GANs) have shown significant improvements in probability distribution estimation on large scale and high dimensional data. We have seen amazing results on super-resolution[7], image translation[14] and photo realistic image generation[8] as they trained on domain-specific datasets. We have also seen generative networks as a solution to data starvation problems. On the Although, conditional GANs require relatively large datasets to train and this can not fully address the problem in cases that only a very few data points are available. Tamar Rott et.al. [10] introduced a new unconditional GAN that is able to generate, extend and harmonise large scale images with only one single sample data. Although GANs have been proved effective for measuring generalisation[3] by estimating the probability distribution, data augmentation techniques also conduct a significant improvement in the training process and overfitting prevention. We have seen data augmentation in most major Deep Networks like AlexNet[4], VGG[11] and ResNet[1]. By using multiple augmented views of the same data point as input, CNNs are forced to learn consistencies in their internal representations. This results in a visual representation that improves generalisation[13]. As mentioned by Yashua Benjo et.al. [5] pixel permutation can not induce generalisation. Although generalisation even with large datasets is an unsolved problem, we can measure the effectiveness of one data sample point for generalisation.

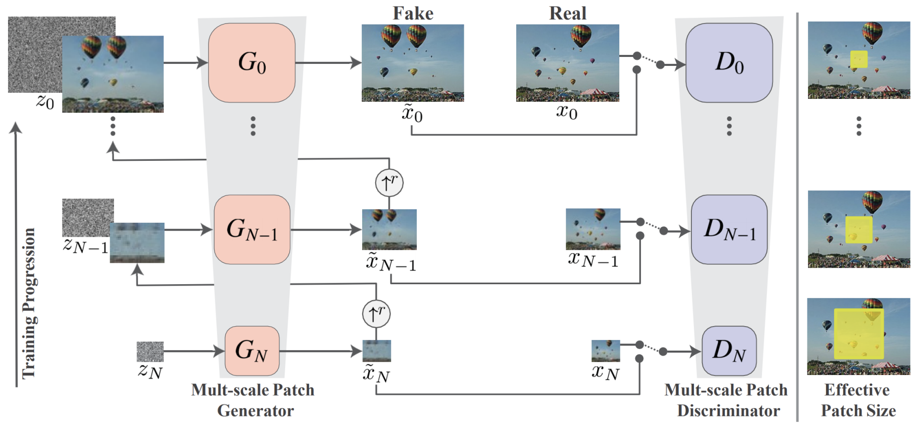

Fig. 1. Single GAN training architecture.[10]

2 Related Work

Generating sample points has been a common approach to inefficient datasets and models like Active GANs[3] have proven this method by calculating the degree of uncertainty for a sample with respect to the hyper-plane of the pre-trained classifier which states that the more informative a sample point the higher the probability that it will improve the classification. Along with texture and object permutation of GANs, regular data augmentation has shown significant improvement in the convergence of CNNs and pruning networks for overfitting. Michael Laskin et.al. showed that data augmentations have a direct impact on the generalisation of Reinforcement Learning agents in non-deterministic environments[6]. ResNets, VGG and AlexNet have used data augmentation as it can perform as a regularizer in overfitting in neural networks[4].
[1]

3 Preliminaries

GAN is a method of DL which estimating the probability distribution implicitly. We simultaneously train two different networks competitively; a generator network G and a discriminator network D. GANs behaviour can be considered as a game theoretic scenarios in a way that the generator network will play against an adversary. The generator g, produces sample $x = g\left(z; \theta^{(g)}\right)$. On the other hand, the adversary network distinguishes the generated value by countering the probability of the new sample being from the training set. The most convenient approach to create the optimisation problem is to evaluate GANs as zero-sum games. This idea infers that there exists a function $v\left(\theta^{(g)}, \theta^{(d)}\right)$ which determines

[1] Source code available at
https://github.com/maminio/sample-generating

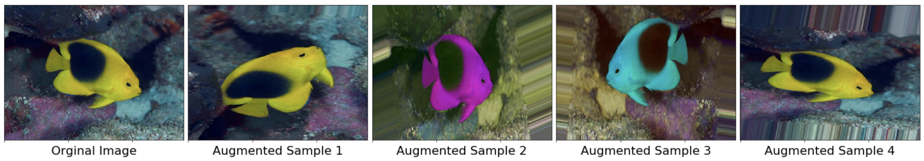

Fig. 2. Random augmentation of the original image.(First stage of the pipeline.)

the payoff of the discriminator and $-v\left(\theta^{(g)}, \theta^{(d)}\right)$ as the payoff of the generator network. In other words, each network tries to maximise its payoff. Therefore the objective function looks as follows:

$$\min_{G} \max_{D} V(D, G)=E_{x \sim p_{data}(x)}[\log D(x)]+E_{z \sim p_{z}(z)}[\log (1-D(G(z)))] \quad (1)$$

SingleGAN is methods of image generation from only one image by unconditionally sampling over its probability distribution and constructing a pyramid of GANs at different scales. The network G learns to generate samples at each scale in which all the overlapping patches cannot be distinguished from the patch in the down-sampled image. In the training process each trained GAN in the pyramid is stored in a fixed state where the loss of the nth GAN is as follows:

$$\min_{G_{n}} \max_{D_{n}} \mathcal{L}_{\text{adv}}\left(G_{n}, D_{n}\right)+\alpha \mathcal{L}_{\text{rec}}\left(G_{n}\right) \quad (2)$$

Where $\mathcal{L}_{\text{adv}}$ is the adversarial loss between the distribution of patches in both the downsampled image and the generated sample and $\mathcal{L}_{\text{rec}}$ the reconstruction loss.

4 Implementation

In this section, we will explain in detail our approach in building data augmentation and sample generation pipeline.

Augmentation
We used data augmentation best practices for image classification proposed by[12]. Our method aims to expand the augmentation horizon to unconditional generative sampling while preserving the structural feature representation of each generated sample point with regards to the real-image. To do so, we pass each image to augmentation pipeline to shift colour channels, crop images, rotation and adding noise.

CoSinGAN/SinGAN Training
After the augmentation stage each image is then passed to the generative network for training. As SinGAN[10] training uses a hierarchical network architecture the

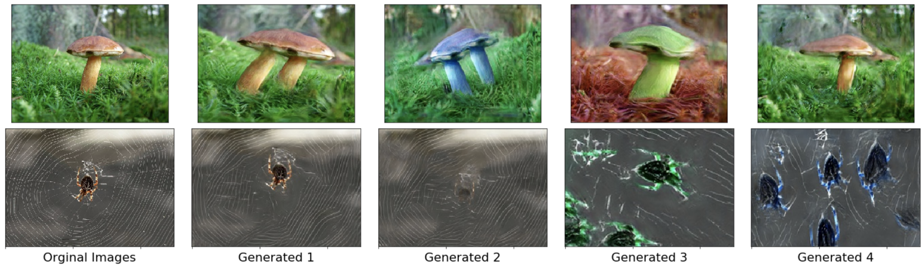

Fig. 3. Generated samples from CoSinGAN.

training process is computationally exhaustive and the pipeline will be impractical. We, therefore, used Concurrent Single GAN(CoSinGAN)[2] which improved the training process of SinGANs significantly by progressively training multiple stages concurrently and taking advantage of learning-rate decay technique at lower stages.

Generating
The last preprocessing stage where the trained network generates multiple new samples of the input image.

5 Experiments and Results

In this section, we will explore experiments and the effect of GANs in increasing accuracy of classification. We designed our experiments to calculate the difference between data augmentation and generative augmentation on learning from only one single image in a five-category classifier. We used Mini Imagenet [9] as our reference dataset with 600 images in each of the 100 categories and selected one sample from each of Fish, Corn, Mushroom, Spider, Bird categories. We trained each augmented sample with learning rate scaling of 0.2 and 7 training stages on CoSinGAN and generated 480 images and 480 augmentations on each of the single sample points. Figure 3 show 4 generated samples after training from the original image. We then trained our 5 Category Mini-Image net and the other two augmented and generated data using a VGG16 network. All three training were processed by the exact same hyper-parameters and random seeds.

	Validation accuracy	Validation Loss
SinGAN Generated(ours)	**%35.77**	**11.4**
Random Augmentation	%26.83	25.7
Mini-Imagenet dataset	%68.88	3.5

Table 1. Accuracy comparison on validation set between SinGAN generated samples and the whole Mini-Imagenet over 5-categories and 40 epochs of training.

As you can see in table 1, SinGAN generated images outperform random augmentation and compared to mini-imagenet dataset, we were able to achieve %35.77 accuracy with only one image.

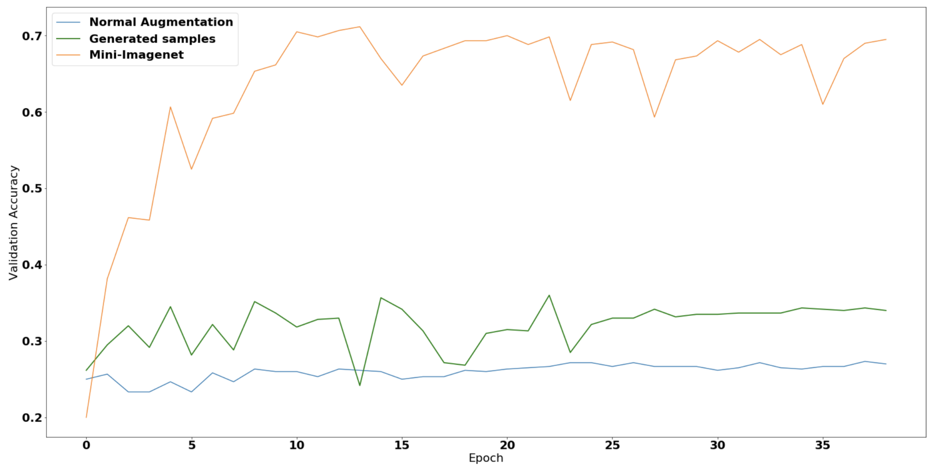

Fig. 4. Test result of trained networks on unseen labelled data.

6 Conclusion and Discussion

In this work, we designed and implemented a practical data preprocessing pipeline which was able to get an accuracy of %35.77 on the validation set of mini-imagenet over 5-category with only one single image. Our experiments show the potentials of unconditional image generation in image classification and training generalisation. During our training, we observed that some images of the reference dataset are more effective in representing their respective category. We observed that images that are more feature representative of their own class are more suited for image generation. This means we can generalise better if we remove images that are less representative and generate more sample data from the generatively suited images. Although this experiment showcased the potentials of GANs, they are computationally very expensive in comparison to other generative sampling methods. Models like CoSinGANs has enabled us to take image generation one step further by creating more optimised and computationally efficient architecture. Few-shot learning has always been tackled by Meta-Learning or Transfer learning, which simply adapts neural network weights to a new domain by performing very few gradient steps. Our approach showed that without any prior knowledge and with just data augmentation and generation we increased accuracy significantly.

References

1. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
2. Hinz, T., Fisher, M., Wang, O., Wermter, S.: Improved techniques for training single-image gans (2020)
3. Kong, Q., Tong, B., Klinkigt, M., Watanabe, Y., Akira, N., Murakami, T.: Active generative adversarial network for image classification. Proceedings of the AAAI Conference on Artificial Intelligence **33**, 4090–4097 (Jul 2019). https://doi.org/10.1609/aaai.v33i01.33014090, http://dx.doi.org/10.1609/aaai.v33i01.33014090
4. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. pp. 1097–1105 (2012)
5. Lake, B.M., Salakhutdinov, R., Tenenbaum, J.B.: Human-level concept learning through probabilistic program induction. Science **350**(6266), 1332–1338 (2015)
6. Laskin, M., Lee, K., Stooke, A., Pinto, L., Abbeel, P., Srinivas, A.: Reinforcement learning with augmented data (2020)
7. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4681–4690 (2017)
8. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015)
9. Ravi, S., Larochelle, H.: Optimization as a model for few-shot learning (2016)
10. Shaham, T.R., Dekel, T., Michaeli, T.: Singan: Learning a generative model from a single natural image. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 4570–4580 (2019)
11. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
12. Wong, S.C., Gatt, A., Stamatescu, V., McDonnell, M.D.: Understanding data augmentation for classification: when to warp? In: 2016 international conference on digital image computing: techniques and applications (DICTA). pp. 1–6. IEEE
13. Xie, Q., Hovy, E., Luong, M.T., Le, Q.V.: Self-training with noisy student improves imagenet classification. arXiv preprint arXiv:1911.04252 (2019)
14. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. pp. 2223–2232 (2017)