International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 01 | Jan -2020 ISSN: 2582-3930

Research Paper on Stacked GAN
Technical Advancement Tests and Public Awareness Survey

Kshitij Bombarde

Keraleeya Samajam's Model College, Dombivli East, Mumbai, Maharashtra, India

ABSTRACT

Although the Generative Adversarial Networks have shown a remarkable progress in performing different type of tasks such as Text Synthesis and Audio Files Generation they still face obstacles in generating exceptional grade images. Stack GAN ++ which focuses on generating more crisp and clear high quality images with use of Multi Layered GAN stacked upon one another like Stack GAN v1 but it consist of more than one Generators and Discriminators and have shown more stable result than the Stack GAN v1 in generating more stable and high Quality images when trained on Single Class Dataset. This technology has been around us since 2017 and is currently group of researchers and data scientists are researching on it worldwide. The Proposed Research focuses on experimenting with Stacked GAN v2 by training the Existing model on newer datasets and more complex datasets and to evaluate how technologically advance Stack GANs are ? We also focus on finding how much people in IT and computers domain are Aware of this technology.

KEY TERMS

Generative Learning Models, Stack Generative Adversarial Networks (SGAN), Multi Level GAN, Distributed Approximation, Realistic Photo Image Generation, Sentence To Picture Processing, Generator Algorithm, Discriminator Algorithm

1. INTRODUCTION

Generative Adversarial Networks (GANs) a robust category of neural networks that square measure used for unsupervised learning. it had been developed and introduced by Ian J. Good fellow in 2014. GANs essentially created from a system of 2 competitive neural network models that contend with one another and ready to analyze, capture and duplicate the variations among a dataset.

It has been ascertained that that the majority of the neural nets are often simply fooled into incorrectly classifying things by adding noise into the parent datasets or the datasets on that ab initio model was trained on to. The model when adding impurities into it shows the next chance of wrong prediction than once it foretold properly. [GAN, GeeksforGeeks]

2. LITERATURE REVIEW

Generative adversarial networks (GANs) are studied thoroughly within the past decade. Arguably the revolutionary techniques within the space of pc vision like plausible image synthesis, image to image transition, facial attribute alteration and similar domains. Despite the many progress achieved within the computer vision field, applying GANs to real-world issues still poses important challenges, 3 of that we have a tendency to specialize in here:

1. prime quality image synthesis

2. various image synthesis

3. Stable Learning.

Through associate degree in-depth review of GAN-related analysis within the literature**,** we offer associate degree account of the architecture-variants and loss-variants, that are planned to handle these 3 challenges from 2 views. we have a tendency to propose. Loss-variants and architecture-variants for classifying the foremost common GANs, and discuss the potential enhancements with specializing in these 2 aspects. whereas many reviews for GANs are given so far, none have centered on the review of GAN-variants supported their handling the challenges mentioned on top of. during this paper, we have a tendency to review and critically discuss the knowledge and awareness among people about the Stacked GANS and how well the technology works with huge datasets. Generating photo-realistic pictures from text is a crucial downside and has tremendous applications, as well as photo-editing, package, etc. Recently, Generative Adversarial Networks (GAN) have shown promising ends up in synthesizing real-world pictures. Conditioned on given textual content descriptions, conditional Comparison of the projected StackGAN and a vanilla one-degree GAN for generating 256x256 pics.

1. Given text : descriptions, Stage-I of StackGAN sketches rough shapes and primary shades of gadgets, yielding low-decision snap shots.

2. Stage-II of StackGAN takes Stage-I results and textual content descriptions as inputs, and generates high-resolution images with photograph-practical info.

3. Results by a vanilla 256×256 GAN that merely adds a lot of upsampling layers to progressive GAN-INT-CLS. it's unable to come up with any plausible pictures of 256×256 resolution.

GANs area unit able to generate pictures that area unit extremely associated with the text meanings.

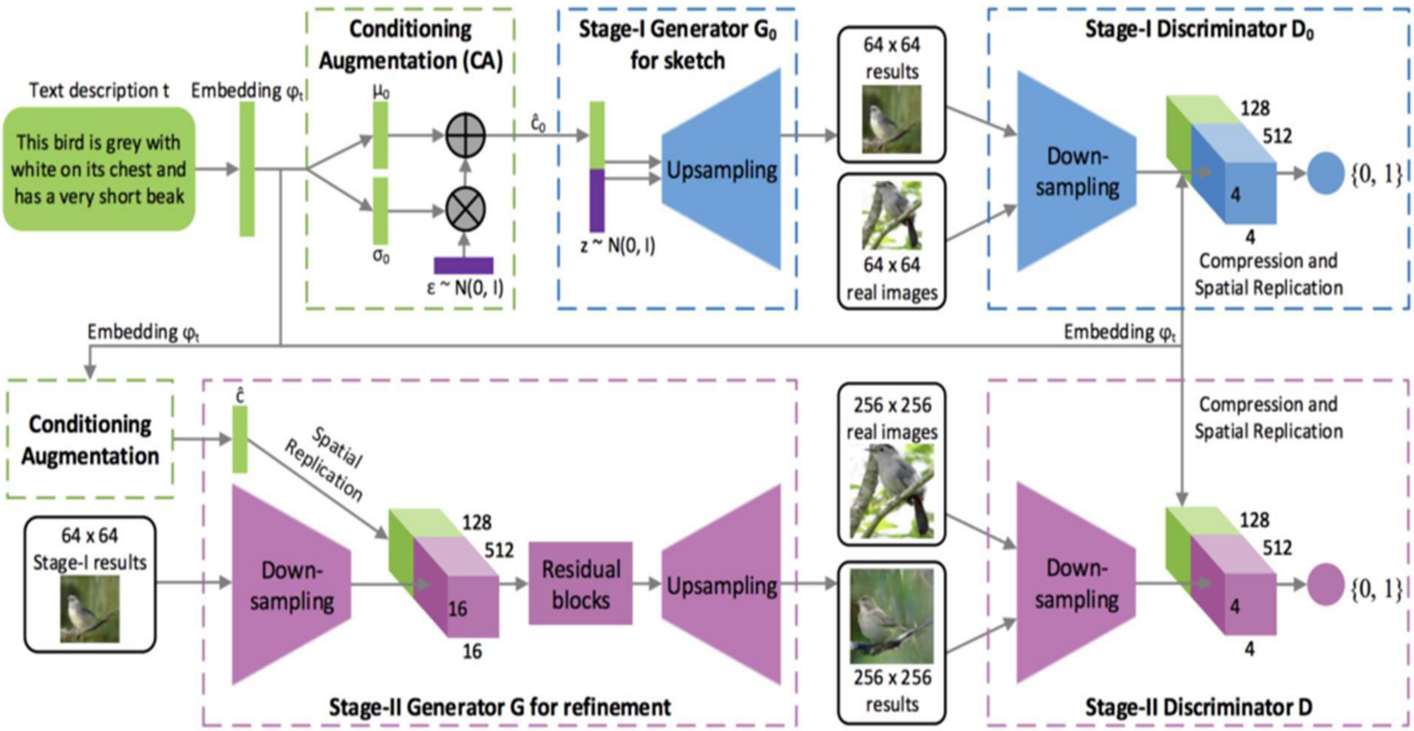
[Stacked GAN++, Arxiv.org, 28 Jul 2018]

3. GENERATIVE ADVERSARIAL NETWORKS

To generate high-resolution pictures with photo-realistic

details, we tend to propose a straightforward nevertheless effective Stacked Generative Adversarial Networks. It fragments the text to image synthesis method into 2 stages

- **Stage-I GAN**: it sketches the primitive form and basic colours of the item conditioned on the given text



[Stacked GAN++ Proposed Architecture , Arxiv.org, 28 Jul, 2018]

description, and attracts the background layout from a random noise vector, yielding a bad res

$$\mathcal{L}_D = \mathbb{E}_{(I,t)\sim p_{data}}[\log D(I, \varphi_t)] + \mathbb{E}_{s_0 \sim p_{G_0}, t\sim p_{data}}[\log(1 - D(G(s_0, \hat{c}), \varphi_t))],$$

$$\mathcal{L}_G = \mathbb{E}_{s_0 \sim p_{G_0}, t\sim p_{data}}[\log(1 - D(G(s_0, \hat{c}), \varphi_t))] + \lambda D_{KL}(\mathcal{N}(\mu(\varphi_t), \Sigma(\varphi_t)) \,||\, \mathcal{N}(0, I)),$$

image.

- **Stage-II GAN**: it corrects defects within the low-resolution

image from Stage-I and completes details of the item

by reading the text description once more, manufacturing a high-resolution realistic photo image.

Instead of immediately producing a high-decision picture conditioned on the text description, we simplify the project to first generate a low-resolution photograph with our regularization parameter that balances , which specializes in drawing handiest tough shape and accurate colours for the object. Let T be the textual content embedding of the given description, that's generated by way of a pre-skilled encoder in this paper. The Gaussian conditioning variables for text embedding are sampled to seize the that means with versions

Contioned on random variable z, Stage 1 Trains Discriminator D1 and Generator G1 simultaneously. z is the unwanted data.

regularization parameter that balances.

Stage-II GAN, Low-resolution pics generated through Stage-I GAN commonly lack brilliant item components and might incorporate shape distortions. Some information within the text may be not noted inside the first stage, that is critical for generating photograph-sensible images. Our Stage-II GAN is constructed upon Stage-I GAN results to generate high-decision photographs. It is conditioned on lowresolution pictures and additionally the text embedding once more to correct defects in Stage-I outcomes

The Stage-II GAN completes formerly unnoticed textual content statistics to generate extra picture-practical info. Conditioning on the low-decision result and Gaussian latent variables, the discriminator D and generator G in Stage-II GAN are skilled via as an alternative maximizing LD and minimizing LG . Different from the authentic GAN formula, the random noise z isn't always used in this level with the assumption that the randomness has already been preserved by means of s0.

Gaussian conditioning variables used on this level utilized in Stage-I GAN share the identical pre-educated textual content encoder, generating the equal text embedding ?T. However, StageI and Stage-II Conditioning Augmentation have distinctive fully related layers for producing different manner and fashionable deviations. In this way, Stage-II GAN learns to seize beneficial statistics within the text embedding this is disregarded via Stage-I GAN.

[Stacked GAN Architecture, Arxiv.org, Jul 8, 2018]

4. Methodologies & Approaches

4.1 Hybrid Research Model

A particular model could embrace descriptive and analytical aspects as represented higher than, however models could favor one side or the opposite. The logical relationships of a descriptive model may be analyzed, and inferences may be created to reason regarding the system. even so, logical analysis provides totally different insights than a quantitative chemical analysis of system parameters.

We First Carried out a survey of people using online form creator and data collection service collect.chat and collected data from people about the awareness in people and then referring to previous papers we have

organized the code and conducted experimentation on the existing code.

4.2 Research Approach

Firstly we created a survey form using collect.chat an online form designing and data collection service which allows users to collect data from the research subjects and then the collected data can be exported into the .csv format. Secondly we Tweaked the initial Model a bit and also trained it on new datasets other than CIFAR-10 which it was previously trained on and tried generating some images which our research subjects wanted to generated with the help of Stacked GAN's and Descriptive text.

5. Public Survey & Experiment

5.1 Public Survey

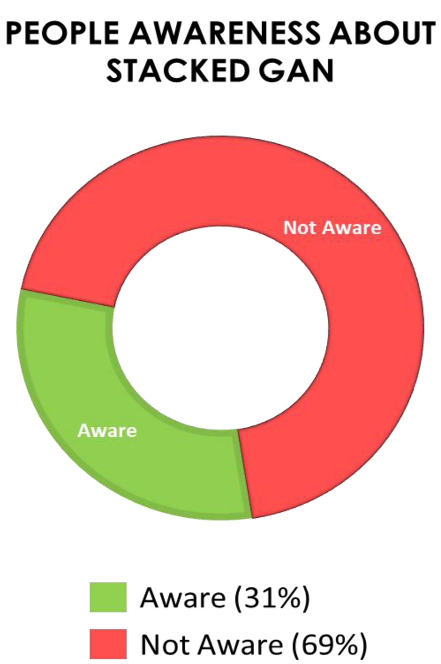
After creating our data collection utility or the survey bot we sent it to various people and collected data on various aspects of the Generative Adversarial Networks and Image Generation using Stacked GAN and Descriptive Text.

5.1.1 Questionnaire

- Are you aware of Stacked Generative Adversarial Networks?
- Have you heard of Image Generation using Artificial Intelligence & Machine Learning Before?
- How useful did you find the Idea of Generating Images using just Text?
- Would you use Stack GAN to generate novel Images for your own?
- What Image will you like generate using this Technology?
- Describe the Image that you would like to generate in at least 6 words.
-

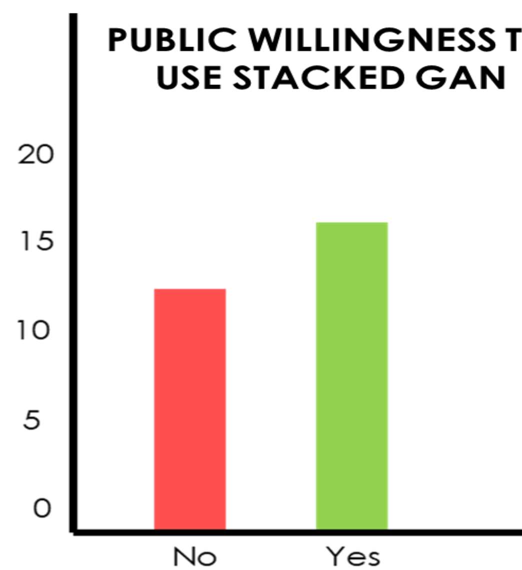
5.1.2 Results

When People were asked that if they knew about Stacked Generative Adversarial Networks or



SGAN, most of the people were unaware that such technology existed. The main reason is that this is a very new and experimental technology which is currently under research and many research professionals are working on it to get better results day by day. The Below Graph will show you how many people are aware about this technology.

When People were asked that how much they were aware of Artificial Intelligence (AI) and

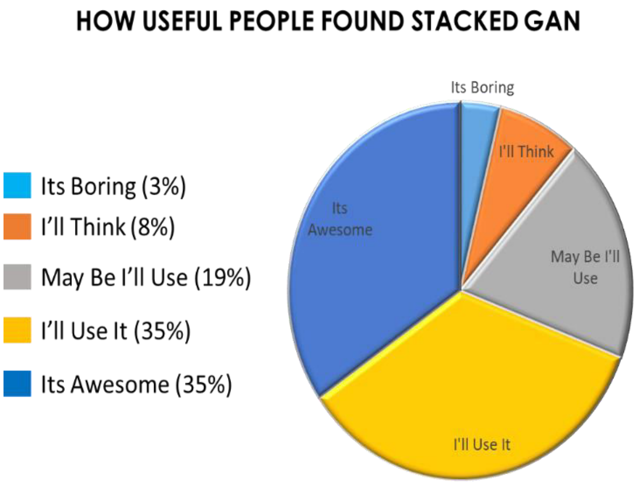


Machine Learning (ML) the results were quite impressive as compared to that of Stacked GAN.

We also asked would they like to use the Stacked GAN for themselves we found out that 28 people 55% people would like to use it for themselves for generating images. Below Graph shows how many people are willing to use Stacked GAN for themselves.

When People Were asked if they find this technology useful or not we found out that 70% people thought that this is a very helpful technology and 27% were not sure if they will use it or not and 3% people said they wont be using it it's a boring technology.

The Above graph shows how much useful this technology will be as reviewed by the people.



We also asked people that what images would they like to generate with Stacked GAN and Descriptive text and also requested them to describe their imagination in 6 words, below listed are some crazy responses form the research subjects we received.

- Girl with Dog ears and tongue.
- Cat with red wings and dragon head.
- A hamster with wings and long tail.
- A Really Fat Dog with Red Horns and Blue Tail.
- A Rabbit with pink ears, baby pink body n a white tail.
- Scenery with a sun, a lake, a small house, trees and mountains.

5.2 Experimentation

We started with downloading the existing Vanilla GAN code from the original creators and installing it on our machine and then executing

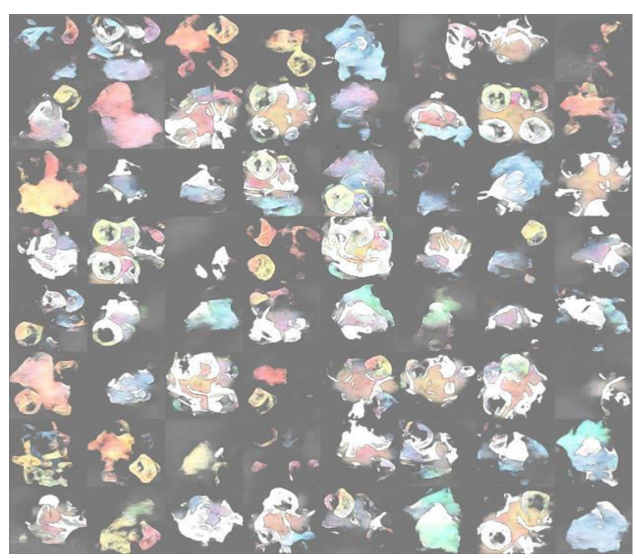
The Bird is small has read head with feathers with gradience red to gray from head to tail

A tiny flower with pinkish petals and huge central orange and dark stamen

and trying to generated

[Image Results from StackedGAN ++ Results, Arxiv.org, JUL 2018]

some images which the original creators made just to check if the system is installed correctly and the images were good enough considering it is Vanilla GAN, Then we used StackGAN++ trained on CIFAR 10 dataset which was successful in generating 256x256 images of

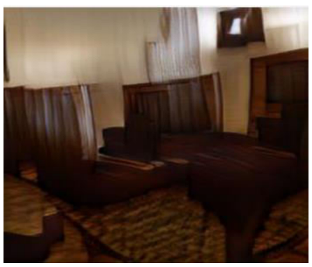


International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 01 | Jan -2020 ISSN: 2582-3930

flowers and birds.

Then we Tried Training our Data set on a single Class 150 element Image dataset which we found on the Internet, The Pokémon dataset. As it was a single class which was quite easy to train on and generate images and was giving quite simple images of Pokémon's the result were not upto the mark but for a new dataset the results were quite remarkable.

Below are the images of Pokémon's which we generated using Stacked GAN.

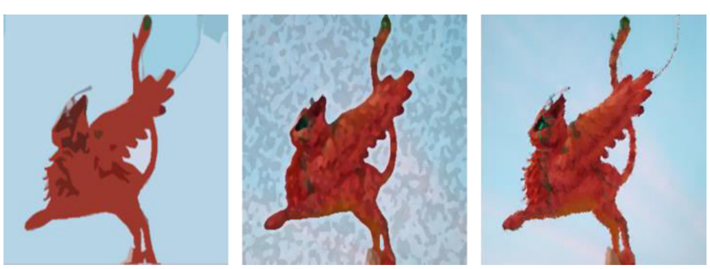


A living room with hard wood floors filled with furniture

A Pizza with lots of White Cheese on it

We also Trained Model on MS COCO Datasets which is a more complex data set with thousands of data images. When we tried to generate images using this model the generated some images were the real nightmare and some common objects gave remarkable results.

We Tried Generating Images Suggested by the research subjects and below are the top 3 results



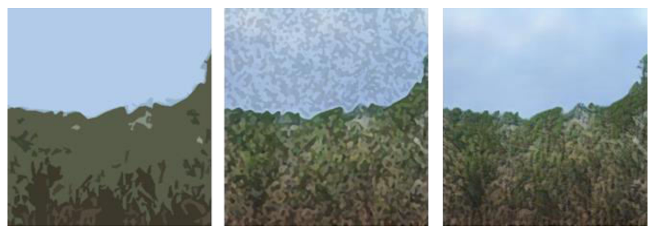
for the same.

Cat with red wings and dragon head flying in sky.

A Rabbit with pink ears, baby pink body n a white tail.

Scenery with a sun, a lake, a small house, trees and mountains.

Basically we can say that when trained on less data variation it gives more accurate results than



that of more complex data.

6. Discussion

The paper introduces an end to end generative architecture named SGAN, that efficiently uses the symbolic data from a trained discriminative model.

Our approach decomposes the laborious drawback of estimating image distribution into multiple comparatively easier tasks – every generating plausible representations conditioned on higher-level representations. The key plan is to use illustration discriminators at totally different coaching hierarchies to supply intermediate direction. We also propose a completely unique entropy loss to tackle the matter that conditional GANs tend to ignore the noise.

Our entropy loss can be used in different applications of conditional GANs, e.g., synthesizing totally different future frames given the same past frames, or generating a various set of pictures conditioned on an equivalent label map. we have a tendency to believe this can be an interesting analysis direction within the future.

7. Findings

- The Technology can be really very useful to solve some real world problems.
- Very Less People are aware about this technology
- People are Willing to Use This Technology.
- People are getting familiar with the modern concepts of A.I.
- The Technology is still an experimental technique and many people are working on it worldwide to make it better.
- As this is currently an experimental technology it is having lots of drawbacks.

8. Applications

- Generate New Datasets
- Generate Photorealistic Images
- Generating Cartoon Characters
- Text to Image Translation
- Content Creation
- 3D Object Detection
- Video Prediction

9. Conclusions

- GANs are cutting edge technology for generation of Novel Content.
- It can be useful in multiple verticals.
- This is currently an experimental technology.
- People are working on this technology worldwide to make it more precise and useful.
- The Results are good enough when trained on Specific Datasets but not good when trained on generalized datasets like MS COCO.
- This is a very new technology and still people are getting to know about it.
- Various researchers are working on Video Production using GANs
- Downsides can be if the users intention are not good then he can use this technology for creating.

10. References

[1] M. Arjovsky and L. Bottou. Towards principled methods for training generative adversarial networks. In ICLR, 2017.

[2] A. Brock, T. Lim, J. M. Ritchie, and N. Weston. Neural photo editing with introspective adversarial networks. In ICLR, 2017.

[3] T. Che, Y. Li, A. P. Jacob, Y. Bengio, and W. Li. Mode regularized generative adversarial networks. In ICLR, 2017.

[4] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In NIPS, 2016.

[5] E. L. Denton, S. Chintala, A. Szlam, and R. Fergus. Deep generative image models using a laplacian pyramid of adversarial networks. In NIPS, 2015. 1,

[6] C. Doersch. Tutorial on variational autoencoders. arXiv:1606.05908, 2016.

[7] J. Gauthier. Conditional generative adversarial networks for convolutional face generation. Technical report, 2015.

[8] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D.

Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio. Generative adversarial nets. In NIPS, 2014. 1, 2, 3

[9] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016. 4

[10] X. Huang, Y. Li, O. Poursaeed, J. Hopcroft, and S. Belongie. Stacked generative adversarial networks. In CVPR, 2017. 2, 3

[11] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In ICML, 2015. 5

[12] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In CVPR, 2017. 2

[13] D. P. Kingma and M. Welling. Auto-encoding variational bayes. In ICLR, 2014. 2, 3

[14] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther. Autoencoding beyond pixels using a learned similarity metric. In ICML, 2016. 3

[15] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. Photo-realistic single image super-resolution using a generative adversarial network. In CVPR, 2017. 2