International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 11, pp: 567-570 www.ijaem.net ISSN: 2395-5252

Realistic Image Outpainting Model

Gunjan Chugh[1.], Ishani Bhatia[2.], Lovlesh Bhatt[3], Riya[4.]

[1.]Assistant ProfessorDepartment of Information Technology Dr.Akhilesh Das Gupta Institute of Technology and Management, New Delhi(India).

[2.]Student(B.Tech), Department of Information Technology Dr.Akhilesh Das Gupta Institute of Technology and Management, New Delhi(India).

[3].Student(B.Tech), Department of Information Technology Dr.Akhilesh Das Gupta Institute of Technology and Management, New Delhi(India).

[4.]Student(B.Tech), Department of Information Technology Dr.Akhilesh Das Gupta Institute of Technology and Management, New Delhi(India).

Date of Submission: 15-12-2020 Date of Acceptance: 30-12-2020

ABSTRACT—Image out painting is a technique used in the recursive painting of images. It generates a new image that aims to restore the missing and damaged part of the image. It is a widely used computer vision algorithm in deep learning. This technique extrapolates the image without distorting the pixels of the image. This technique is widely used in various novel and existing applications for example- for panorama creation, vertically filmed video expansion. Our project focuses on building a three-phase Generative Adversarial Network (GANs) that enhances the quality of the image. The output image appears realistic to the human eye. The image generated after recursive painting has enhanced quality and high pixel resolution as compared to the input image. The model is trained on 36,500 images of baseline width is 128× 128 RGB encodings. The context encoders used in this model adversarially draw the missing region of the image from the surrounding image pixels. This model uses local discriminators combined with the global discriminator and produces very convincing results. Various masking and preprocessing techniques are used to decrease the loss during processing.

Keywords—Generative Adversarial Network (GANs), Computer Vision Algorithm

I. INTRODUCTION

This project is based on an advanced computer vision algorithm that recursively paints the damaged and lost part of the image. It extrapolates the image with various processing and masking techniques that resulted in high pixel resolution and improved quality image as compared to the original input image. This algorithm is used as a filter in various photoshop applications. Several visualizations of 35,000 (approx.) images and 100 validation images. Preprocessing and masking of images to improve the result of the model. Calculating the evaluation metrics and comparing the root mean square loss.[1] Boost the performance of the model, the generator by augmenting with perceptual, style, and total variation losses. Accordingly, we present a deep learning approach. for training the model to get image boundaries. We use a three-step training schedule to stably train our model. [7]We also use discriminators to improve the quality of our output. When our model is trained, it can outpaint color images recursively, thus allowing for recursive outpainting. Realistic Image outpainting has not been discussed more, but a similar task called image inpainting has been widely studied. In Image inpainting, it paints within the image. Image outpainting is quite a difficult task, as it requires extrapolation to unknown areas in the image with less neighboring information. The result of this model appears realistic to the human eye as there is not much difference in extrapolating image boundaries One common method for achieving this in image inpainting involves using GANs.

II. ALGORITHM

The GAN training involves training both the discriminator and the generator model in parallel. The main focus for Generative Adversarial Networks is to generate data from the beginning[2].

A generative adversarial network (GAN) is a popular approach to generating new data that follow similar distribution as those in the training sets. For many data generation tasks (e.g., face image generation and visual manipulation), GANs have obtained significantly improved results compared to previous approaches, such as variational autoencoders and deep belief networks. Specifically, a GAN contains two networks, a generator G and a discriminator D, which are trained together (e.g.,

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 11, pp: 567-570 www.ijaem.net ISSN: 2395-5252

alternatively). The two networks are adversaries to each other: D aims to tell real inputs from generated ones from G, whereas the goal of G is to maximize the error of D.

The first papers to show image outpainting used a data-driven approach merged with a graph representation of the source image. Other researchers were also able to achieve realistic results, we anticipate to apply adversarial training for possibly even better outcomes.

[3] A crucial implementation of image inpainting using deep learning by Pathak et al. familiarized the idea of a Context Encoder that is a CNN trained used to generate missing portions of an image using GAN also. The outcomes presented were comparatively realistic, but still had scope for visual enhancement. Iizuka et al. enhanced the Context Encoder method for image inpainting by adding a discriminator that only took as input the inpainted area and its instant.

Using these methods, the researchers were able to attain extremely accurate outcomes with a fraction of the training required. Besides, the researchers provided numerously quantifiable metrics that will be useful for assessing our models' performance.

This paper builds upon an initial effort in generative video models. Though, previous work has focused typically on minor patches, and assessed them for video clustering. Here, we develop a generative video model for natural scenes using beaches dataset using [4]state-of-the-art adversarial learning approaches. However, here we are interested in creating short videos with accurate progressive semantics by outpainting each video frame, rather than detecting or retrieving them.

III. DATASET

[2]In this model, we expect to overfit on a single 128 × 128color images of the beaches. We use a 128 × 128 image as opposed to the 512 × 512 image size to speed up training. For this research, we use the equivalent single image for training and testing. Our primary dataset for image and video outpainting is composed of 3500 256×256 images from the Beaches dataset. We down-sampled these images and frames of images to 128×128. This dataset is composed of a diverse set of beaches and sceneries.

Figure 1: Dataset Images

IV. PROJECT WORK

A. Preprocessing of the data

[5] To train our model on the beaches dataset, we use a pre-processing pipeline**[1].** We use a training image It, then we normalize the images to In ∈ [0, 1]128×128×3. **[1]**In this, we define a mask M, where M belongs to {0, 1} 128×128 such that Mij = 1 − 1[32 ≤ j < 96] to mask the center portion of the image. After that, we take mean pixel intensity μ, over the unmasked region In (1 − M). Then, we set the outer pixels of each channel to the average value μ. We define I'm = μ · M + In (1 − M). In the last step of pre-processing, we merge In with M to produce Ip that belongs to [0, 1]128×128×4. Thus, as the result of pre-processing it outputs (In, Ip).

B. Training Pipeline

In this paper we have used DCGAN architecture similar to that used by Iizuka et al. In this the generator is used in the form of an encoder-decoder CNN, while the discriminator uses stridden convolutions to repeatedly downsample an image for binary classification. For each iteration of training, we randomly sample a minibatch of training data. As shown in Figure 2, we preprocess it to get In and Ip, as previously described. We run the generator on Ip to get the outpainted image Io = G(Ip) that belongs to [0, 1]128×128×3. Afterward, we use a discriminator to classify the ground truth (In) and outpainted image (Io). We compute losses and update parameters according to our training schedule.

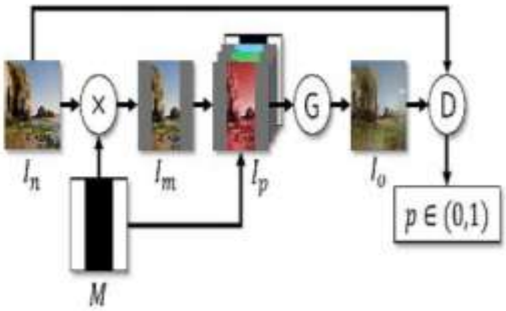

Figure 2: Image Training Pipeline

C. Training Agenda

In our training agenda, we focused on utilizing three-phase training to balance the training. There are three loss functions that we have used : LMSE (In, Ip) = ‖M (G(Ip) − In) ‖ (1) LD (In, Ip) = − [log D(In) + log (1 − D(G(Ip)))] (2) LG (In, Ip) = LMSE (In, Ip) − α · log D (G (Ip)) (3). In phase 1 of training, the generator updates the weights of the generator by equation 1 above for T1 iterations.[5] In phase 2 we use a discriminator where it updates discriminator weights per equation 2 above for T2 iterations. In the last phase, all the training is

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 11, pp: 567-570 www.ijaem.net ISSN: 2395-5252

performed for T3 iterations, where both discriminator and generator are trained.

V. RESULTS

A. Outpainting on our Dataset

In this, we fed images from the validation set through our outpainting pipeline. The final results are shown in Figure 4. As seen in the third example of Figure 3, the network does not merely copy visual features and lines during outpainting. Rather, it learns to hallucinate new features, as shown by the appearance of a house on the left-hand side.

Figure 3: Outpainting results for a sample of held-out images in the validation set.

B. Recursive Outpainting

Recursive outpainting is the process in which an already outpainted image is taken as input. Here we have taken an outpainted image which is sent again as input to the network after padding with the mean pixel value.[6] Here we have shown Figure 5 in which this process takes place recursively five times by expanding the image width. As noise tends to compound with each successive iterations, such that the model effectively learns the overall textures of the image and extrapolates the sky and landscape comparatively realistically.

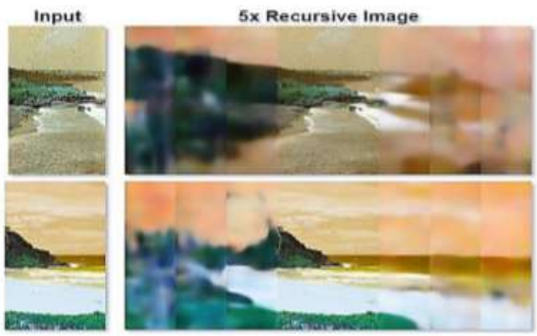

Figure 5: Each right image is the result of recursively outpainting the corresponding left image five times.

VI. ACCURACY

Its accuracy is about 70% off the Realistic Image Outpainting model using the Generative Adversarial Network(GAN) algorithm.

VII. CONCLUSION

We can successfully recognize the image by applying a deep learning approach. Three-phase training ends up being vigorous during Generative Adversarial Networks (GAN) training. The outcomes from training with only a global discriminator were equally realistic in the image. In the end, we initiate recursive outpainting for an image as a means of arbitrarily outspreading an image. Even though image noise compounded with consecutive iterations, the recursively-outpainted image persisted relatively realistic.

VIII. FUTURE WORK

There are numerous potential improvements for our image outpainting model. To stabilize training, the Wasserstein GAN algorithm could be incorporated into three-phase training

•Explore sequence models for video outpainting.

Video outpainting use the same algorithm as image outpainting that is generative adversarial networks but in very less portion as firstly the video is been separated into numerous frames and then the algorithm is applied and at the end, the outpainted frames are merged or concatenated with each other in the same sequence they were extracted.

•Incorporate perceptual and style loss.

To improve the performance of the model, the generator loss could be augmented with perceptual, style, and total variation losses.

•Experiment with partial convolutions.

REFERENCES

[1]. S. Iizuka, E. Simo-Serra, and H. Ishikawa. Globally and locally consistent image completion. ACM Transactions on Graphics (TOG), 36(4):107, 2017.

[2]. G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro. An Image in painting for irregular holes using partial convolutions. arXiv preprintAXiv.1804.07723, 2018.

[3]. Dr. Pathak P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros. Context encoders: Feature learning by inpainting.

[4]. Novel Framework for Image Inpainting". International Journal of Computer Trends and Technology (IJCTT) V14(3):141-147, Aug 2014. ISSN:2231-2803. www.ijcttjournal.org.

[5]. Kshitij Tripathi, Rajendra G. Vyas, Anil K. Gupta "Deep Learning through Convolutional

Neural Networks for Classification of Image: A Novel Approach Using Hyper Filter". IJCSE Vol.-7, Issue-6, June 2019. https://www.ijcseonline.org/

[6]. Charu Khare, Kapil Kumar Nagwanshi "Implementation and Analysis of Image Restoration Techniques". IJCTT - May to June Issue 2011.

[7]. Finn, Chelsea, Ian Goodfellow, and Sergey Levine." Unsupervised learning for physical interaction through video prediction." Advances in neural information processing systems. 2016.