International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 06 | June 2021 www.irjet.net p-ISSN: 2395-0072

Own Emoji Creation Using GAN(Generative Adversarial Networks)

Prasanna Sambare[1]
Information Technology
Usha Mittal Institute of Technology
Mumbai, Maharashtra

Namrata Tamhankar[2]
Information Technology
Usha Mittal Institute of Technology
Mumbai, Maharashtra

Sneha Visave[3]
Information Technology
Usha Mittal Institute of Technology
Mumbai, Maharashtra

Prof. Prachi Dhannawat[4]
Information Technology
Usha Mittal Institute of Technology
Mumbai, Maharashtra

***

***Abstract*—**Emojis play very important role in today's digitized communication. Avatars are the ways to indicate nonverbal cues. There are various techniques and ways to make more immersive communication. One of the most efficient way of communication is by means of using emojis instead of typing long text. In the past experiments was demonstrate by Egaokun which is Automatic Avatar Building Tool. Tool finds the face within the digital image and positions a grid on specific facial markers which can then be used to creatively manipulate the facial expression. In this project by using Machine learning and Al we reviewed to build a system that will detect human face expression with the help of face reorganization algorithm in which we used python library for real-time computer vision for face detection and to create avatar/emoji and then will convert that expression into corresponding emojis or avatar's using Generative adversarial network (GANs) which is one of the most important research avenues in the field of artificial intelligence and The Tied Output Synthesis(TOS) method. It used to convert face expression into corresponding emoji or avatar's. The face processing technology allows ease and rapidity of use and allow automation of such functionalities as characterization or morphing of the facial image. Use of generator and discriminator for training network makes it more errorless and efficient to give augmented output with zero generalization error. the method which we are using for domain transfer is able to generate identifiable avatar that are coupled with a valid configuration vector.

***Index Terms*—GAN, unsupervised parameters, StyleGAN, OpenCV, TOS, Domain Transfer**

I. Introduction

The past several years machine learning and artificial intel- ligence is becoming a growing field with a great number of meaningful applications and valuable research topics which impacting on different aspects of our daily life. With ad- vancements in computer vision and Machine Learning In our day to day life we are interacting through different medium through chats, email in which emojis take part important role. We aim to build a system which will create emojis based on human facial expression and will map corresponding Emojis or Avatars.

The objectives of this thesis are to generating computer avatars based on the user's appearance. In which we are providing different features like providing hat, changing hair colour etc. This features will give the variations to the human Avatar. The aim of this work is to learn to map an input image to two tied outputs which is vector in some parameter space and the image generated by this vector.

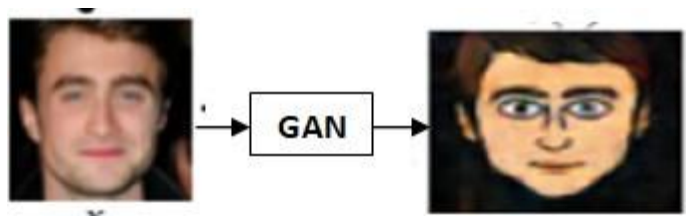

Fig. 1. Conversion of facial image into corresponding Avatar using GAN[2]

II. Related Work:

Egaokun: An Avatar Creation System: Egaokun: An Avatar Creation System: Egaokun Automatic Avatar[1] Building Tool, proposes to customize avatars by using face recognition tech- nology to process raw images of the face. Egaokun system detects the face within the provided image and positions a grid on specific facial area which can then be used to creatively manipulate the facial features. The latter feature may find use in applications (for example, in playful situations) or where the user wishes to increase similarity by cari- caturization, or accentuate emotions in the original picture. A important design feature of the Egaokun system[1] is rapidity of use. In a few seconds a user's picture is filtered and the area containing the face is extracted and assigned with an adaptable grid, facial attributes are classified and semantic labels attached to the face and finally the system provides an interesting looking avatar body to the user. To overcome the disadvantages of existing systems i.e time consuming, untrained parameters etc., our proposed system manages to overcome all this difficulties.

III. GENERATIVE ADVERSARIAL NETWORKS:

Machine learning deals with the various types of the networks, one of them is Generative Adversial Network (GAN) which is basically a deep learning modeling technique. As our paper describes unsupervised creation of parameterized avatars using GAN, we are using conditional GAN as described in paper [3]. Generative network deals with the two types of models in it i.e. Generator and Descriminator. Our aim is to produce the exact avatar replica of human face, we have used face detection technique of using Open Source Computer vision library provided by Intel[4]. Since the generator will play the most important role in training the network and creating the avatar with the help of activation functions and external datasets that we have taken from kaggle.com (freely available on internet,having size of 2.2 GB). Output from the generator will act as input to the descriminator. Descriminator will basically tries to discriminate between the original input and the output from the generator[5]. For working of these two models simultaneously we are using Tied Output Synthe- sis(TOS) [2] method which is combination of cross domain transfer and unsupervised domain adaptaion.

IV. FACE DETECTION USING OPENCV

Human interaction with the web camera can be enabled by simply using OpenCV library of python[4]. OpenCV uses a Haar Cascade classifier which is a type of face detector. It is one of the pretrained model based on the positive and negative images, which detects face from the provided image.Given an image, which can come from a file or from live video, the face detector identifies each image location and classifies it as face or not.

V. STYLE GAN:

For the generation of quality pictures we are accompanying our Generative Adversarial Network with the StyleGAN gen- erator[6]. StyleGAN generator starts from a learned constant input and adjusts image at each convolution layer in the network based on the provided code, therefore directly con- trolling the image features at different scales.With the help of external dataset(from kaggle.com, 2.2 GB) combined with the network, StyleGAN architecture produces unsupervised high level attributes in generated pictures.

Fig.2 Comparison of traditional and style-based genera- tor[6]. Both the generators use normalization method for input data preparation from the provided dataset. The goal of mapping network is to generate input vector into intermediate vector whose element control has different visual features. For this mapping purpose it uses 8 fully connected layers. Output from those 8 layers denoted by 'w' is passed through Affine transformation(A) means any transformation that preserves collinearity (i.e., all points lying on a line initially still lie on a line after transformation) and ratios of distances (e.g., the midpoint of a line segment remains the midpoint after transformation) and then to the synthesis network model which uses adaptive instance normalization(AdaIN)[6]. The AdaIN operation is defined as follow-

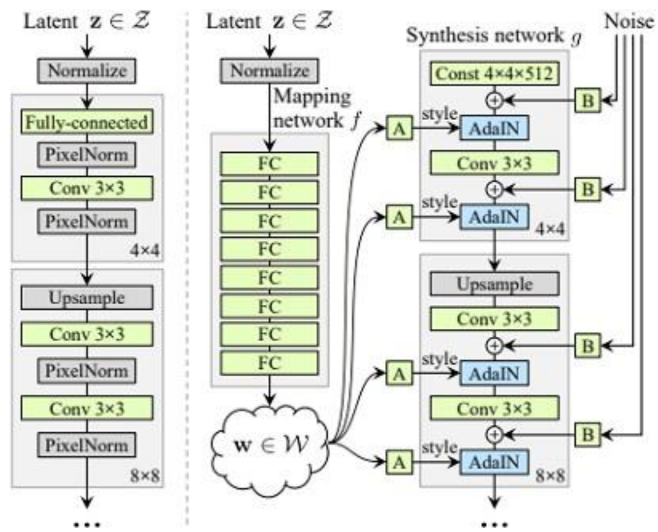

Fig. 2. (A)Traditional(B)Style-based generator

$$\mathrm{AdaIN}(\mathbf{x}_i, \mathbf{y}) = \mathbf{y}_{s,i}\frac{\mathbf{x}_i - \mu(\mathbf{x}_i)}{\sigma(\mathbf{x}_i)} + \mathbf{y}_{b,i},$$

Where xi is input map and y is generated style.This ultimately increases performance of overall network. Gaussian noise (B) is added to each activation map and interpreted which ultimately helps it to produce realistic images.

VI. DOMAIN TRANSFER TECHNIQUES:

For converting the actual input image from user into the avatar, we require domain transfer techniques. We are formu- lating this problem with the help of two methods i) cross domain transfer and ii) unsupervised domain adaptaion as described in paper[2]. In the unsupervised domain transfer method, algorithm tries to trains the source domain i.e. input image and tests it on different target domains that we are providing as external datasets from Kaggle(Kaggle.com). The algorithm has a labeled dataset of the source domain and an unlabeled dataset of the target domain. The conventional approach to deal with this problem is to learn a feature map function that (i) enables accurate classification of images in the source domain and (ii) captures the meaningful invariant relationships between the source and target domains from the network. Whereas cross domain transfer works on changing the mapping functions until it trains the whole model which gives desired

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 06 | June 2021 www.irjet.net p-ISSN: 2395-0072

output.It learns a function that maps samples from the input domain X to the output domain Y. It was recently presented in [7], where a GAN based solution was able to convincingly transform face images into caricatures from a specific domain.

VII. Tied Output Synthesis Method:

This technique for converting source domain to output domain is the combination of above described methods that are
i) cross domain transfer and ii) unsupervised domain adaptaion [2,7] which gives best accuracy.

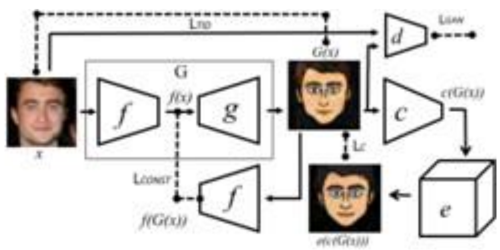

Fig. 3. The training constraints of the Tied Output synthesis method.The learned functions are c,d and G for given f.Mapping function e is known priori.

Above figure 3[2] describes the actual working of tied output synthesis method. It learns to adjust the mapping function f similar to the cross domain transfer. Where e is the prelearned function. And g and c are mapped together to trained the whole model using the ReLu function. This makes sense, since while e is a feedforward transformation from a set of parameters to an output, c requires the conversion of an input of the form g(f(x)) and f becomes invariant under
G. Other than the functions f and e, the training data is unsupervised and consists of a set of samples from the source domain X and a second set from the target domain of e, which we call Y1. The Tied Output Synthesis (TOS) method is also evaluated on a toy problem of inverting a polygonsynthesizing engine and on avatar generation from a photograph for two different CG engines[2].

VIII. Avatar creation:

Using the above stated methods, we convert facial characters from user input to the caricatures with the help of millions of random images of size 2.2 GB as a external datasets.Based on the coordinates of input image, the emoji were centered and scaled into 152 *152 RGB images, with the help of StyleGan which generates the quality pictures contains five convolutional layers, each followed by batch normalization and a leaky ReLU[2]. For the evaluation purpose we used CelebA dataset(200K images) which is available freely on internet.

GAN training could be accelerated greatly by devising better methods for coordinating Generator and Discriminator or determining better distributions to sample from during training

IX. Conclusion and future scope:

With the help of above stated techniques, this proposed system overcomes all the disadvantages of existing systems giving the accuracy of 84.4%. StyleGAN technique not only produces high-quality and realistic images but also allows for superior control and understanding of generated images, making it even easier than before to generate believable fake images. The TOS method that we present is able to generate identifiable emoji that are coupled with a valid configuration vector.

X. References:

[1] Michael Lyons, Andre Plante,Sebastien Jehan, Seiki Inoue and Shigeru Akamatsu, “Avatar Creation using Automatic Face Recognition”, 1998

[2] Lior Wolf, Yaniv Taigman, and Adam Polyak, “Unsupervised Creation of Parameterized Avatars”,2017

[3] ZHAOQING PAN, WEIJIE YU1, XIAOKAI YI1, ASIFULLAH KHAN2, FENG YUAN1, AND YUHUI ZHENG 1, “Recent Progress on Generative Adversarial Networks (GANs): A Survey”,2019

[4] Shervin EMAMI1, Valentin Petruț SUCIU, “Facial Recognition using OpenCV”, 2012

[5] Ian J. Goodfellow, Jean Pouget-Abadiey, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozairz, Aaron Courville, Yoshua Bengio, “Generative Adversarial Nets”,2014

[6] Tero Karras ,Samuli Laine,Timo Aila, “A Style- Based Generator Architecture for Generative Adversarial Networks”,2019

[7] Y.Taigman, A. Polyak, and L. Wolf, “Unsupervised cross domain image generation. In International Conference on Learning Representations (ICLR)”, 2017.