International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 490-495 www.ijaem.net ISSN: 2395-5252

Future Face Predictor using Generative Adversarial Networks

Dr. Shanta Sondur, Ronit Bhatia, Anirudh Iyer, Aniket Shinde
Dept. of Information Technology VESIT, University of Mumbai, Mumbai, India
Dept. of Information Technology VESIT, University of Mumbai, Mumbai, India
Dept. of Information Technology VESIT, University of Mumbai, Mumbai, India
Dept. of Information Technology VESIT, University of Mumbai, Mumbai, India
Corresponding author: Dr. Shanta Sondur

Date of Submission: 01-08-2020 Date of Acceptance: 15-08-2020

ABSTRACT - This report presents a generalized and effective methodology for predicting how a missing person would look after a given number of years into the future. The disappearance of people is a matter of grave concern which occurs at an alarming rate in metropolitan cities like Mumbai. A missing person is defined as a person who has disappeared and whose status as alive or dead can't be confirmed, as their location and fate are unknown. This paper aims to devise a method which predicts how a missing person will look after a certain number of years into the future given a photograph of the person.
The purpose of this project is to help the Police authorities in finding people who have been missing for a long period of time, for whom it's near impossible to find a recent image, while keeping a log of all the missing persons . In addition to this, their appearance may change as they age, which adds to the problem.
The image of the missing person is predicted by the use of generative adversarial networks- a class of machine learning frameworks.

I. INTRODUCTION

The purpose of this research is to assess how Generative Adversarial Networks - a class of machine learning frameworks, can be used to predict the future images of individuals given a current or past image of the person. The objective of this project is to produce an image that is highly accurate and resembles to a great extent with the current appearance of the subject. Findings from several research papers from reputed journals form the basis for this research.

The disappearance of people is unfortunately something that occurs on a daily basis all around the world. In most cases, people go missing for a large number of years which makes it difficult for the authorities to track them down and find them as their appearances change as well over time due to ageing. This impact is profound when a person goes missing in their childhood and is not found for quite some time. Our method aims to assist the authorities in finding a person who has been missing for a long time by predicting how the person might look like in the future given an image of the person taken at the time of missing

A. Generative Adversarial Networks

Generative Adversarial Networks (GANs) are a powerful class of neural networks that are used for unsupervised learning. It was developed and introduced by Ian J Goodfellow in 2014. GANs are basically made up of a system of two competing neural network models which compete with each other and are able to analyze, capture and copy the variations within a dataset.
A generative adversarial network has two parts :

- Generator : It generates fake samples of data (images in our case) and tries to fool the discriminator. A randomly sampled noise vector is given as the input to the generator.
- Discriminator : The output of the generator is given to the discriminator which has to determine whether the image came from the actual dataset or it came from the generator and is thus a fake image. The Discriminator outputs a single scalar value D(x) for an image X which determines how likely it is that the image X is a real image coming from the dataset.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 490-495 www.ijaem.net ISSN: 2395-5252

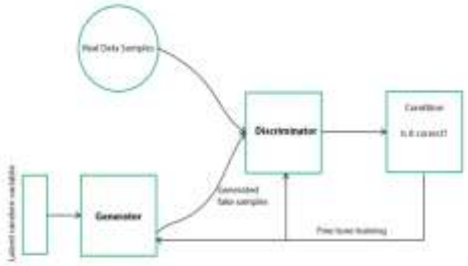

Fig. 1. GAN Structure

Both the Generator and Discriminator are neural networks which run in competition with each other.

The generative model captures the distribution of data and is trained in such a manner that it tries to maximize the probability of the Discriminator in making a mistake. The Discriminator on the other hand is based on a model that estimates the probability that the sample that it got is received from the training data and not from the Generator.

The GANs are formulated as a minimax game, where the Discriminator is trying to minimize its reward V(D, G) and the Generator is trying to minimize the Discriminator's reward or in other words, maximize its loss. It can be mathematically described by the formula below:

$$\min_G \max_D V(D,G)$$

$$V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))]$$

Fig. 2. GAN objective function

where,
G = Generator
D = Discriminator
Pdata(x) = distribution of real data
P(z) = distribution of generator
x = sample from Pdata(x)
z = sample from P(z)
D(x) = Discriminator network
G(z) = Generator network

The above given objective function is minimized for the generator and maximized for the discriminator.

We want the Discriminator to recognise real images X as 'real' and output a high value close to 1 for the given objective function. Similarly, the Discriminator should be able to recognise fake images G(z) as 'fake' and should output a low value close to 0 for the given objective function. Training a GAN has the following two parts :

- Part 1 : The Discriminator is trained while the Generator is idle. In this phase, the network is only forward propagated and no back-propagation is done. The Discriminator is trained on real data for n epochs, and see if it can correctly predict them as real. Also, in this phase, the Discriminator is also trained on the fake generated data from the Generator and see if it can correctly predict them as fake.
- Part 2 : The Generator is trained while the Discriminator is idle. After the Discriminator is trained by the generated fake data of the Generator, we can get its predictions and use the results for training the Generator and get better from the previous state to try and fool the Discriminator.

B. StyleGAN - Style Generative Adversarial Networks

Style GAN uses the baseline progressive GAN architecture and proposed some changes in the generator part of it.The discriminator architecture is quite similar to baseline progressive GAN.

- Baseline Progressive Growing GANs: Style GAN uses baseline progressive GAN architecture which means the size of generated image increases gradually from a very low resolution (4×4) to high resolution (1024 x1024). This is done by adding a new block to both the models to support the larger resolution after fitting the model on smaller resolution to make it more stable.
- Bi-linear Sampling: The authors of paper use bi-linear sampling instead of nearest neighbor up/down sampling (which was used in previous Baseline Progressive GAN architectures) in both generator and discriminator. They implement this bi-linear sampling by low pass filtering the activation with a separable 2nd order binomial filter after each of the upsampling layer and before each of the downsampling layer.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 490-495 www.ijaem.net ISSN: 2395-5252

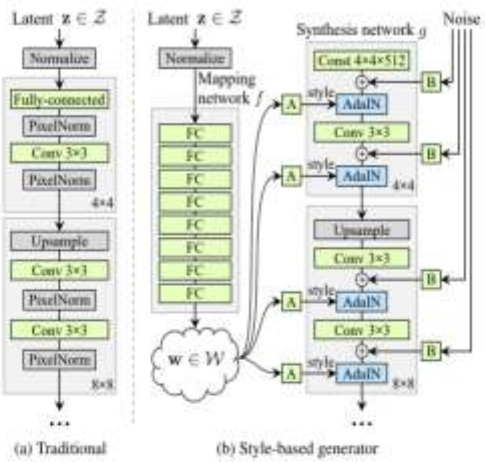

Fig 2. StyleGAN

- The input to the AdaIN is y = (ys, yb) which is generated by applying (A) to (w). The AdaIN operation is defined by the following equation:

$$AdaIN(x_i, y) = y_{s,i}\left(\left(x_i - \mu_i\right)/\sigma_i\right) + y_{b,i}$$

where each feature map x is normalized separately, and then scaled and biased using the corresponding scalar components from style y. Thus the dimensional of y is twice the number of feature maps (x) on that layer. The synthesis network contains 18 convolutional layers 2 for each of the resolutions (4×4 – 1024×1024).

- Removing traditional (Latent) input: Most previous style transfer models use the random input to create the initial latent code of the generator i.e. the input of the 4×4 level. However the style-GAN authors concluded that the image generation features are controlled by w and AdaIN. Therefore they replace the initial input with the constant matrix of 4x4x512. This also contributed to an increase in the performance of the network.
- Addition of Noisy:Input A Gaussian noise (represented by B) is added to each of these activation maps before the AdaIN operations. A different sample of noise is generated for each block and is interpreted on the basis of scaling factors of that layer.
- There are many aspects in people's faces that are small and can be seen as stochastic, such as freckles, exact placement of hairs, wrinkles, features which make the image more realistic and increase the variety of outputs. The common method to insert these small features into GAN images is adding random noise to the input vector.
- Mixing Regularization: The Style generation used an intermediate vector at each level of the synthesis network which may cause the network to learn correlation between different levels. In order to reduce the correlation, the model randomly selects two input vectors (z1 and z2) and generates the intermediate vector (w1 and w2) for them. It then trains some of the levels with the first and switches (in a random split point) to the other to train the rest of the levels. This switch in random split points ensures that the network doesn't learn correlation very much.

II. PROBLEM STATEMENT

The current face prediction applications are quite specific in their usage and there are quite a few security flaws in the existing applications . Our project is aimed at providing the user (Police Authorities in this case) with a centralised portal where tabs can be kept on all the missing persons for which the future images have been predicted.Further,the users can be flagged as missing or found as per the investigations that are being carried out.

The problem statement states that given a photograph of a missing person, predict the image of how the person may look after a given number of years.

III. EXPERIMENTS

A. Dataset Used

Flickr-Faces-HQ (FFHQ) is a high-quality image dataset of human faces, originally created as a benchmark for generative adversarial networks (GAN).The dataset consists of 70,000 high-quality PNG images at 1024×1024 resolution and contains considerable variation in terms of age, ethnicity and image background. It also has good coverage of accessories such as eyeglasses, sunglasses, hats, etc. The images were crawled from Flickr and automatically aligned and cropped using dlib.

B. Implementation Details

The whole system consists of two parts:

a)Website

b)Backend Server.

The security authorities first register themselves and authenticate themselves as security users. After authenticating themselves, they can create new entries for each missing person which includes his personal details and his image and feed into our backend database.Here details of all the users,entries submitted by the users which includes all the details of lost persons are stored.Now, from this our Python API would take the image and process it using our GAN Architecture and return the aged image which is then sent back to our website.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 490-495 www.ijaem.net ISSN: 2395-5252

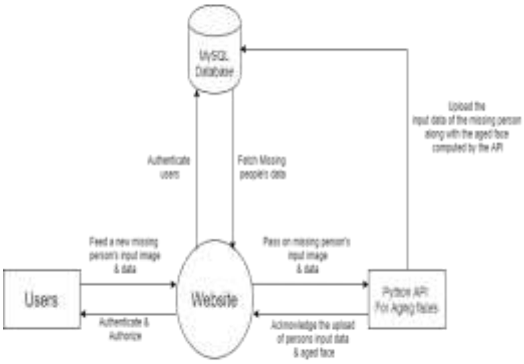

Fig 3. Overall Architecture of System

After drawing the whole architecture , all the implementation details of each component need to be studied.First we will look at the GUI i.e. frontend of our system to understand how the website works,and then we will look at our Backend design which includes our MySQL Database and our Python API which utilizes the GAN architecture for producing good high quality age-progressed images.

A. PYTHON AGING API

Our aging api comprises two parts:- a)Encoding of input images into numpy vectors b)Predicting future image using pretrained vectors for age and SVM.

In the first part we first take the input image and pass it to the resnet encoder to get latent code estimation of our image.Then this estimate is passed to a StyleGAN based GAN model whose Generator produces the image based on the latent code estimate.Then the input image and this generated image is passed to a pretrained VGG Classifier which extracts the feature vectors from intermediate layers of both images.So the difference between these two vectors is calculated as loss which is then back propagated to the initial code and Stochastic Gradient Descent is run over it to optimize the latent code and minimize the loss to get good estimation of feature vector.

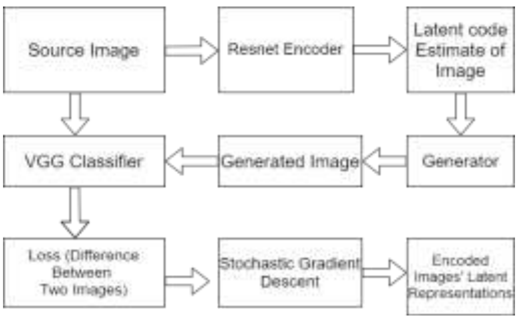

Fig 4. Encoding Image Architecture

In the second part we utilize the learnt vector representations of image to get a future prediction of our input image.In this we have age based vector representations of dataset which are learnt from Azure-Cognitive API.Support Vector Machines(SVM) algorithm is used to learn the structure of our dataset from these vector representations of our dataset which creates a feature model.Then we pass the latent representations learnt from the first part along with age condition which translates along the learnt hyperplane of SVM to get a prediction of how the source image would look like in future.

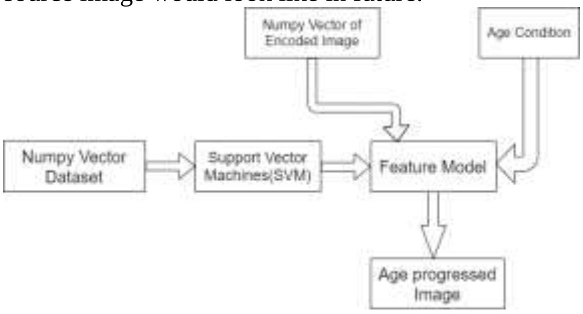

Fig 5. **Age Progression from vectors**

A. MySQL Database

Our database consists of two tables namely Faces and Users.Their schema can be described as-
Faces(id,uploaded_by,firstName,lastName,phone_no_of_ward,email_of_ward,age,gender,date_lost,upload_time,input_image,output_image,found(y/n))
Users(email,firstname,lastname,hashed_password,date_of_birth,gender,mobile_no,profilepic,id-proof,type)
The relationships between these two entities can be shown as-

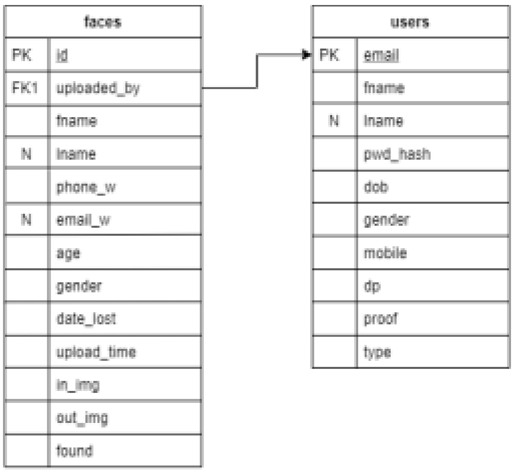

Fig 6. ER diagram of database

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 4, pp: 490-495 www.ijaem.net ISSN: 2395-5252

B. Website (Front-End)

The Website front-end provides approved users with the following 5 modules :

C. Registration & Login Module
D. New Entry Module
E. User Uploads' Database Module
F. Universal Database Module
G. Approve New Users Module

All these modules require the user to be approved by an already approved user (Through Module 4).

- **Registration & Login Module** includes the Login and Registration pages. During registration, users need to upload a certificate and a display picture, which will enable the verifiers (existing users) to verify the new user easily.
- **New Entry Module** enables users to upload a new missing person's data along with his/her photo through an HTML form. This form includes the first name, last name, age when lost, gender, date lost, phone number of his / her ward, email of the ward and finally, the last known image of the person who is missing.
- **User Uploads' Database Module** shows all the faces a user has uploaded in a tabular format, sorted by timestamp of upload. It includes a link to the **full detail page**, in which the user can edit the details and update the found-status of the person.
- **Universal Database Module** is similar to the user uploads' database module, except that it shows all the uploads by all the users sorted by timestamp of upload.
- **User Approval Module** enables the users to approve new users.

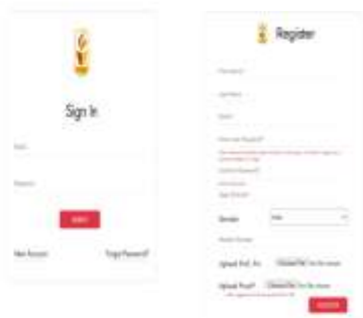

Fig 7: Registration & Login Module

Fig 8: New Entry Module

Fig 9: Personal / Universal Database Module

Fig 10: Lost Person Details View & Edit Module

Fig 11: User Approval Module

III. RELATED WORK

Various softwares and applications have been developed for future face prediction but surprisingly none has been made for a social security purpose to track lost children.Existing systems include a database of lost child but our system provides an upgrade on existing system with current image as well as an estimation of future image. Various softwares include one by Washington University which is an illuminationaware age progression.Other websites include changemyface.com which have been used by FBI and other US based security branches for predicting future faces but again it is a third-party tool. So our website provides a major upgrade on existing work by providing full integration of the prediction system with the website.

IV. RESULT ANALYSIS

To test how good the encoding is, we use Microsoft Azure Cognitive Services for face comparison.

Results show that an image with good high pixel resolution and less background clutter gives a confidence score of 90% and above for the similarity measure between two images. Therefore, good accuracy numbers are observed. Also, if the image of that person is a side face, accuracy decreases and tends to be between 75-90%. If the background clutter is large, encoded image accuracy again decreases to 75-85%. Image blurring can occur if resolutions are low, therefore recommended image size is 1024x1024.

V. CONCLUSION

The disappearance of people is a problem which occurs at an alarming rate on a day to day basis.The number of kidnappings in Mumbai has seen a steady uptick. Between 2018 and April 2019 an alarming 3,041 cases of kidnapping of young boys and girls came up before the Mumbai Police. Out of these,a significant number of cases remained untraceable.

So,We believe our project helps in tackling the problem to quite an extent. Due to a simple and effective User Interface,security personnel can easily create and find details of missing persons which aids them to greater extent.Also with proper security protocol applications,our system is safe from various security attacks.

When used in accordance with the investigations, the image generated by our project may immensely aid police authorities in searching for a missing person.

REFERENCES

[1]. Age Progression/Regression by Conditional Adversarial Autoencoder by Zhifei Zhang, Yang Song, Hairong Qi, The University of Tennessee, Knoxville, TN, USA https://arxiv.org/pdf/1702.08423.pdf

[2]. Facial Aging and Rejuvenation by Conditional Multi-Adversarial Autoencoder with Ordinal Regression by Haiping Zhu, Qi Zhou , Jumping Zhang, James Z. Wang , Shanghai Key Lab of Intelligent Information Processing, School of Computer Science, Fudan University, Shanghai, China.https://arxiv.org/pdf/1804.02740.pdf

[3]. Dual Conditional GANs for Face Aging and Rejuvenation by Jingkuan Song, Jingqiu Zhang, Lianli Gao, Xianglong Liu, Heng Tao Shen, Center for Future Media and School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China.https://www.ijcai.org/Proceedings/2018/0125.pdf

[4]. Face Aging with Identity- Preserved Conditional Generative Adversarial Networks by Zongwei Wang, Xu Tang, Weixin Luo, Shenghua Gaohttp://openaccess.thecvf.com/content_cvpr_2018/papers/Wang_Face_Aging_With_CVPR_2018_paper.pdf

[5]. DualGAN: Unsupervised Dual Learning for Image-to-Image Translation by Zili Yi, Hao Zhang, Ping Tan, Minglun Gong, Memorial University of Newfoundland, Canada

[6]. A Style-Based Generator Architecture for Generative Adversarial Networks by Tero Karras, Samuli Laine, Time Aila, NVIDIA https://arxiv.org/pdf/1812.04948.pdf

International Journal of Advances in Engineering and Management
ISSN: 2395-5252

IJAEM

Volume: 02 Issue: 01 DOI: 10.35629/5252

www.ijaem.net
Email id: ijaem.paper@gmail.com