A Project Report

on

FACIAL RECOGNITION AND AGING-EFFECT

Submitted in partial fulfilment of the

requirement for the award of the degree of

Bachelor of Technology

GALGOTIAS UNIVERSITY

(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

Under The Supervision of
Ms. Garima Pandey:
Assisstant Professor

Submitted By

Harshita Sengar
19SCSE1010725

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING /
/ DEPARTMENT OF COMPUTER APPLICATION
GALGOTIAS UNIVERSITY, GREATER NOIDA
INDIA
MAY, 2023

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

We hereby certify that the work which is being presented in the project, entitled **"FACIAL RECOGNITION AND AGING-EFFECT"** in partial fulfilment of the requirements for the award of the B.Tech(CSE) submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of Jan, 2023 to August, 2023, under the supervision of Ms. Garima Pandey, Assisstant Professor, Department of Computer Science and Engineering, of School of Computing Science and Engineering , GalgotiasUniversity, Greater Noida

The matter presented in the project has been submitted by me for the award of any other degree of this or any other place.

Harshita Sengar, 19SCSE1010725

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Ms. Garima Pandey

Assisstant Professor

CERTIFICATE

The Final Project Viva-Voce examination of Harshita Sengar: 19SCSE1010725 has been held on

____and his/her work is recommended for

the

award of B. Tech (CSE).

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Program Chair **Signature of Dean**

Date: May, 2023

Place: Greater Noida

Abstract

In computer vision, generating images of faces that appear to age is a challenging challenge. This is known as a conditional image generation problem because the generated image is dependent on a specified face image. Until recently, generative models were incapable of producing realistic, high-resolution images.

Nonetheless, a generative model known as a generative adversarial network (GAN) has demonstrated remarkable abilities in generating realistic images in both unconditional and conditional settings.

Despite this, generating images that display the ageing process while maintaining the identity of the person in the source image is a challenging task due to two contradictory constraints: the generated image must maintain the identity of the person in the source image while also possessing the characteristics of the target age.

In this study, the authors employed a GAN in a conditional setting, along with a custom loss function devised to strike a balance between preserving the individual's identity and generating characteristics of the target age.

The test has shown stepped forward overall performance each in maintaining the individual’s identification and category accuracy of generated pics with inside the goal magnificence as compared to preceding acknowledged approach to this problem.

Table of Contents

List of Table

List of Figures

Acronyms

GAN	Generative Adversarial Network
RF	Random Forest
SVM	Support Vector Machine

CHAPTER-1

Introduction

Aging is a natural process that changes a person's appearance over time. Although some facial features are preserved during the aging process, it can be difficult to recognize a later person just by looking at a young person. In many cases, it becomes very important to recognize someone in a later age while pictures of earlier ages are available. For example, the identification of long-hidden criminals. In some cases, reverse aging is also required. This work presents a new method for generating facial images of someone of different ages.

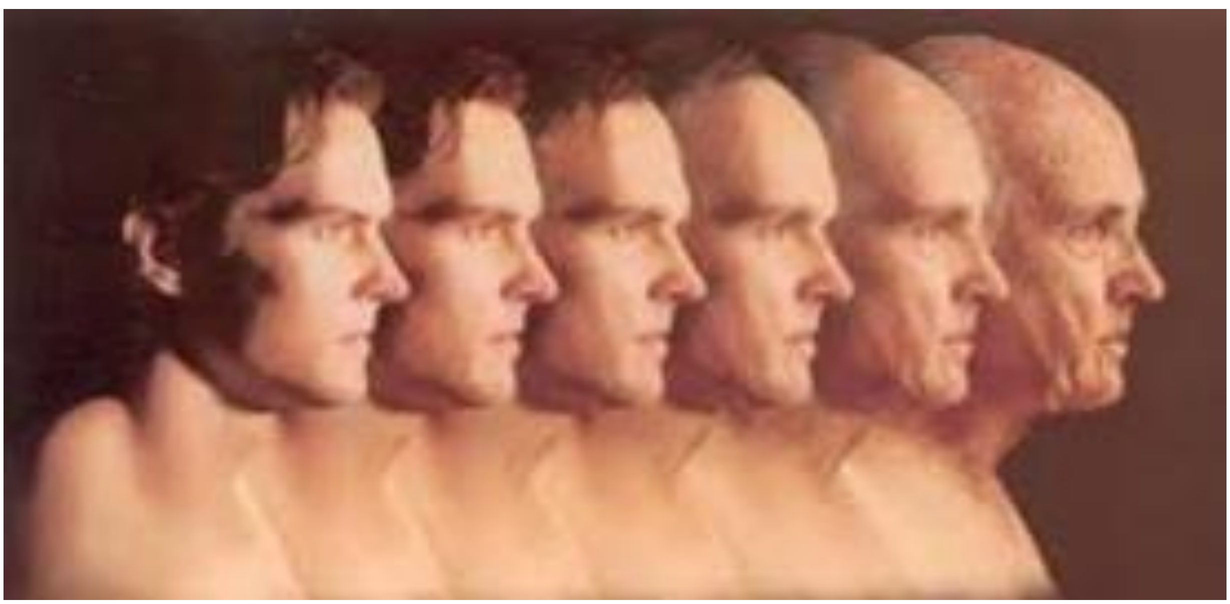

Figure 1: Images of a person at different ages generated by the proposed model. The first column displays the image provided by the user, while the remaining columns comprise images generated for each age group.

Figure 2: An example of our project result

The total age range is divided into his six age groups. Figure 1 & Figure 2 shows these examples of images generated at different ages of her of a particular person. This task has many real-world applications. This is useful for forensic applications. It can be used to enhance and modify photos of suspects, victims, and missing persons to identify them. This technique was used by police and investigators in his with manual modifications made with the help of human artists, computer graphics, physiologists, and many others. This long and complicated process can be simplified by algorithms that can handle aging. Another important application is entertainment. Aging is part of the virtual special effects of the film and Magazine. The actor's approach changes from old to young and vice versa in' a very realistic way. This is done using a frame-by-frame manual age synthesis process. Converting images is difficult, but this his process is repetitive and very time consuming. So, if computers can automatically convert characters in movies with aging effects, the process becomes much easier.

Construction:

Both the discriminator and generator models' layer descriptions are detailed in Tables 1 and 2. It corresponds to the StarGAN implementation's generator and the discriminator levels [13]. Across all layers besides the output layer, instance normalization has been implemented. As the discriminator, leaky ReLU was used. In the second table, it represents the number of domains, NC represents the dimension of the domain label, Conv represents convolution with parameters (output channels, kernel size, stride, and padding), and InstNorm represents the instance of the normalization structure L. ReLU.

Layer	Input size → Output size	Layer type
Hidden Layers	(r,c,3) → ($\frac{r}{2}$,$\frac{c}{2}$,64)	Conv(64, (4×4), 2, 1), L. ReLU
	($\frac{r}{2}$,$\frac{c}{2}$,64) → ($\frac{r}{4}$,$\frac{c}{4}$,128)	Conv(128, (4×4), 2, 1), L. ReLU
	($\frac{r}{4}$,$\frac{c}{4}$,128) → ($\frac{r}{8}$,$\frac{c}{8}$,256)	Conv(256, (4×4), 2, 1), L. ReLU
	($\frac{r}{8}$,$\frac{c}{8}$,256) → ($\frac{r}{16}$,$\frac{c}{16}$,512)	Conv(512, (4×4), 2, 1), L. ReLU
	($\frac{r}{16}$,$\frac{c}{16}$,512) → ($\frac{r}{32}$,$\frac{c}{32}$,1024)	Conv(1024, (4×4), 2, 1), L. ReLU
	($\frac{r}{32}$,$\frac{c}{32}$,1024) → ($\frac{r}{64}$,$\frac{c}{64}$,2048)	Conv(2048, (4×4), 2, 1), L. ReLU
Outout Layer (D_{adv})	($\frac{r}{64}$,$\frac{c}{64}$,2048) → ($\frac{r}{64}$,$\frac{c}{64}$,1)	Conv(1, (3×3), 1, 1)
Outout Layer (D_{cls})	($\frac{r}{64}$,$\frac{c}{64}$,2048) → (1,1,n_d)	Conv(n_d, ($\frac{r}{64}$ × $\frac{c}{64}$), 1, 0)

Table 1. Construction of Generator

Layer	Input size → Output size	Layer type
Down-sampling	(r,c,3+n_c) → (r,c,64)	Conv(64, (7×7), 1, 3), InstNorm, ReLU
	(r,c,64) → ($\frac{r}{2}$,$\frac{c}{2}$,128)	Conv(128, (4×4), 1, 3), InstNorm, ReLU
	($\frac{r}{2}$,$\frac{c}{2}$,128) → ($\frac{r}{4}$,$\frac{c}{4}$,256)	Conv(256, (4×4), 1, 3), InstNorm, ReLU
Bottlenack	($\frac{r}{4}$,$\frac{c}{4}$,256) → ($\frac{r}{4}$,$\frac{c}{4}$,256)	Residual Connection: Conv(256, (3×3), 1, 1), InstNorm, ReLU
	($\frac{r}{4}$,$\frac{c}{4}$,256) → ($\frac{r}{4}$,$\frac{c}{4}$,256)	Residual Connection: Conv(256, (3×3), 1, 1), InstNorm, ReLU
	($\frac{r}{4}$,$\frac{c}{4}$,256) → ($\frac{r}{4}$,$\frac{c}{4}$,256)	Residual Connection: Conv(256, (3×3), 1, 1), InstNorm, ReLU
	($\frac{r}{4}$,$\frac{c}{4}$,256) → ($\frac{r}{4}$,$\frac{c}{4}$,256)	Residual Connection: Conv(256, (3×3), 1, 1), InstNorm, ReLU
	($\frac{r}{4}$,$\frac{c}{4}$,256) → ($\frac{r}{4}$,$\frac{c}{4}$,256)	Residual Connection: Conv(256, (3×3), 1, 1), InstNorm, ReLU
	($\frac{r}{4}$,$\frac{c}{4}$,256) → ($\frac{r}{4}$,$\frac{c}{4}$,256)	Residual Connection: Conv(256, (3×3), 1, 1), InstNorm, ReLU
Up-sampling	($\frac{r}{4}$,$\frac{c}{4}$,256) → ($\frac{r}{2}$,$\frac{c}{2}$,128)	DeConv(128, (4×4), 2, 1), InstNorm, ReLU
	($\frac{r}{2}$,$\frac{c}{2}$,128) → (r,c,64)	DeConv(64, (4×4), 2, 1), InstNorm, ReLU
	(r,c,64) → (r,c,3)	Conv(3, (7×7), 1, 3), Tanh

Table 2. Construction of Discriminator

1.1 Problem Formulation

Face aging was also referred to as age synthesis and age progression. It is the base of cross-age face detection. All previous approaches to this topic can be divided into two major categories- prototyping approaches and modelling approach. Prototyping approaches tries to estimate the average face of a certain age group. The difference between these faces gives the age pattern of that age group which is then used to generate the aged face of a given input. Such a method does not put any constraint to keep the identity of the original person. For that reason, the generated images look unrealistic. These rules-based methods are fast but not very effective. The modelling approach focuses on skin, muscle, and skull surface features of different age groups. These models are supervised and require images of her of the same person at different ages and are expensive to collect. Recent advances in neural networks have opened a very interesting line of research in computer vision. Using neural networks to generate natural images has been researched in recent years, but until recently he was unable to generate images with good resolution using neural networks. The key model that made possible is the Generative Adversarial Network (GAN) and Random Forest. Variational autoencoder) does not use L2 leakage. L2 loss tends to produce blurry images. Instead, GANs focus on generating realistic images, and is its distinct loss. And it helped to recreate realistic images in every detail. Previous GANs were able to generate random images from the data distribution of trained images. Subsequently, conditional generative adversarial networks (cGAN) provided control over the generated images. Take a class label as input and generate an image for that class. Isola et al. al. introduced an encoder/decoder architecture to tune his images for specific classes as well as classes. This approach has been successfully used in various face attribute editing tasks such as changing hair color or adding sunglasses to a given image. L2 loss tends to produce blurry images. Instead, GANs focus on generating realistic images, and is its distinct loss. And it helped to recreate realistic images in every detail. Previous GANs were able to generate random images from the data distribution of trained images. Subsequently, a conditional generative adversarial network (cGAN) provided control over the generated images. Take a class label as input and generate an image for that class. Isola et al. al. The encoder/decoder architecture introduced to tune his image specific to the class as well. This approach has been successfully used in various face attribute editing tasks such as changing hair color or adding sunglasses to a given image. The aim of this research is to design a model who can create a human image of any age and create it projected image at different specified ages. In this work, in the conditional setting we used the concept of GAN facial aging. Our proposed face generation method figure subject to two constraints. First, the suggested the method has a single generator for generating faces different age groups. Second, it preserves identity source subject when transferring to the desired age. We proposed a custom loss function along with GAN loss import these two constraints. Experiment shows improvement in creating better images with identity preserved resource persons. First, we describe the concept of generative adversarial network, which is the backbone of our model. We then present our face generation model a picture in another age of man. Then we discuss about the influence of another component of our goal function. Finally, we describe the training procedures model.

A. Generative Adversarial Network:

Generative adversarial network (GAN) and Random Forest are one of currently the most effective and popular generative model available. A GAN has two components – a generator and discriminator. The GAN generator is trained to match data distribution of the generated image to the target image distribution. This is done with a minimum of two players max game between generator and discriminator. The discriminator is trained to learn the distribution real data and distinguish it from generated data. At the same time, the generator is trained to generate more realistic images to fool the discriminator. Train improves the performance of both the generator and discriminator.

B. Face-Aging Model:

Our objective is to train a multi-domain image-to-image translation model. Along with the adversarial loss, we applied the custom loss to make the image looks like the target domain and keep the identity of the person of the source. The generator takes the class of the target image c, along with the input image x to generate the image; G (x, c) → y. The model generates images of all classes from a given image.

C. Network Architecture:

After the introduction of the generative adversarial network with low-resolution images, Radford et. al. [7] introduced a deep convolutional generative adversarial network. We followed the architecture from this model with other changes and variations introduced in later papers. Both generator and discriminator use the module of the form Convolution-Batch Normalization-ReLU

[8]. D. Training the Wasserstein GAN [2], [3] has shown that it is difficult to optimize the GAN objective proposed in the original GAN paper [7]. To stabilize the GAN training and to generate high-quality image it proposed a modified training objective.

1.2.1 Tool and Technology Used

We have implemented the model with OpenCV in python. All experiments were done in the NVIDIA 1060 GPU. To train the model we need facial images with the age as a label. For this, we have used our own face as a dataset. This dataset contains more than 50 images of people in the age range 0 to 116. We divided the images into 6 age groups. They are: 0-10, 10-14, 15-18, 18-20, Dividing the data into age groups leads to an unequal amount of data into the different age group. To balance this, we have manually augmented the data by cropping the exact facial area with a face detector.

There are a variety of available tools and technologies for facial detection and aging effects. Popular ones which we have included area as follows include:

OpenCV: OpenCV is a free and open-source library of machine learning and computer vision algorithms. Face detection, face recognition, and age estimation modules are pre-installed. Apart from that it also has modules such as object detection and recognition, feature extraction, camera calibration, and more. OpenCV is written in C++ and has interfaces for various programming languages, including Python, Java, and C#. It is widely used in research, industry, and academia for developing computer vision applications.

Some key features and capabilities of OpenCV include:

Image and video I/O: OpenCV support reading and writing images and videos in various formats.

Image processing: It offers a rich set of functions for image filtering, transformation, color manipulation, blending, and morphological operations.

Feature detection and extraction: OpenCV provides algorithms for detecting and extracting features, such as corners, blobs, and edges.

Object detection and tracking: It includes pre-trained models and methods for object detection, tracking, and recognition, including popular techniques like Haar cascades and deep learning-based approaches.

Camera calibration: OpenCV provides tools for calibrating cameras, estimating camera parameters, and performing geometric transformations.

Machine learning integration: OpenCV can integrate with machine learning libraries such as TensorFlow and PyTorch for tasks like image classification, object detection, and semantic segmentation.

GUI and visualization: OpenCV include graphical user interface (GUI) components for creating interactive computer vision applications. It also offers functions for visualizing images, drawing shapes, and annotating visual data.

Parallel and optimized processing: OpenCV leverages hardware acceleration and multi-threading to optimize performance on various platforms, including CPUs, GPUs

Fig 3. OpenCV

DLib: DLib is a machine learning and computer vision C++ library. It includes pre-built models for face recognition, face landmark recognition, and age estimation and other related modules,

Some key features and capabilities of DLib include:

Face recognition: DLib includes a state-of-the-art face recognition algorithm that can identify individuals in images or videos. It uses deep metric learning to learn a feature representation of faces and can achieve high accuracy even with limited training data.

Object detection: DLib provides tools for detecting objects in images, including HOG-based detectors and deep learning-based approaches like YOLO and SSD. It also includes a pre-trained object detector for faces.

Image segmentation: DLib offers algorithms for segmenting images into regions or objects, based on features like color, texture, and shape.

Machine learning: DLib provides a range of machine learning algorithms, including support vector machines (SVMs), deep neural networks, and k-nearest neighbours (k-NN). It also includes tools for training and evaluating machine learning models.

Optimization: DLib includes tools for optimizing functions and solving optimization problems, such as linear programming and quadratic programming.

GUI and visualization: DLib provide a graphical user interface (GUI) for visualizing and annotating images, as well as tools for displaying and manipulating data.

Cross-platform support: DLib is cross-platform and runs on Windows, Linux, and macOS.

DLib is widely used in industry and academia for developing computer vision and machine learning applications, particularly in the areas of face recognition and object detection.

Fig 4. Dlib

TensorFlow: Google developed the open-source machine learning framework TensorFlow. Face detection, face recognition, and age estimation are pre-built models. It provides a flexible and comprehensive ecosystem of tools, libraries, and resources for building and deploying machine learning models.

Key features and capabilities of TensorFlow include:

Computational graph: TensorFlow uses a dataflow graph to represent the computations in a machine learning model. This graph allows for efficient parallel execution and automatic differentiation for backpropagation during training.

Deep learning: TensorFlow offers extensive support for deep learning models, including neural networks with various architectures such as feedforward networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformers. It provides a wide range of layers, activation functions, and optimization algorithms for building and training deep learning models.

Eager execution: TensorFlow supports both static and dynamic computation graphs. With eager execution, developers can execute TensorFlow operations immediately, enabling a more intuitive and interactive programming experience.

Model deployment: TensorFlow allows for deploying models to a variety of platforms and devices, including desktops, mobile devices, edge devices, and cloud infrastructure. TensorFlow Serving and TensorFlow Lite are specialized frameworks for serving and deploying TensorFlow models.

Model optimization: TensorFlow provides tools for optimizing and improving the performance of machine learning models. This includes techniques like quantization, pruning, and model compression to reduce the model size and improve inference speed.

Fig 5. TensorFlow

Installation: -

Install 3rd-package dependencies of python (listed in requirements.txt)

```
tensorflow-gpu==1.4.1
  scipy==1.0.0
  opencv-python==3.3.0.10
  numpy==1.11.0
  Pillow==5.1.0
pip install -r requirements.txt.
```

Other libraries

CUDA 8.0

Cudnn 6.0

PyTorch: Facebook developed the open-source machine learning framework PyTorch. Face detection, face recognition, and age estimation are pre-built models. FaceApp is a mobile application that applies various facial effects, including aging effects, to photographs using neural networks. Aging Booth is a mobile application that simulates the aging process on a person's visage using face detection and aging algorithms. It is designed to provide a flexible and efficient platform for building and training deep learning models.

Key features and capabilities of PyTorch include:

1. Dynamic computational graph: PyTorch uses a dynamic computational graph, which allows for more flexibility and ease of use when building and debugging deep learning models. With dynamic graphs, users can define and modify the computation graph on the fly during model training.

2. Neural network building blocks: PyTorch provides a wide range of neural network building blocks, including various layers, activation functions, and optimization algorithms for building and training deep learning models. It also includes pre-trained models for computer vision and natural language processing tasks.

3. Automatic differentiation: PyTorch includes automatic differentiation capabilities that allow for computing gradients of tensors with respect to other tensors. This is a key feature for training deep learning models using backpropagation.

4. Distributed training: PyTorch provides support for distributed training of deep learning models across multiple GPUs and even multiple machines, using frameworks like Distributed Data Parallel (DDP) and Horovod.

5. Model deployment: PyTorch provides tools for deploying models to a variety of platforms, including desktops, mobile devices, edge devices, and cloud infrastructure. PyTorch Mobile is a specialized framework for deploying PyTorch models on mobile devices.

6. Integration with other libraries: PyTorch integrates well with other popular Python libraries, such as NumPy, SciPy, and scikit-learn, making it easy to use in a wide range of data science workflows.

PyTorch is widely used in industry and academia for developing deep learning models, particularly in the areas of computer vision, natural language processing, and speech recognition. Its dynamic computational graph, flexible architecture, and strong community support make it a popular choice among researchers and developers alike.

Fig 6. PyTorch

CHAPTER-2

Literature Survey

Ageing affects face recognition, particularly in false-alarm errors, where elderly subjects are more likely than young adults to recognise new images as "old." To determine whether this difference in face recognition ability between young and elderly subjects is influenced by the age of the face, an experiment was conducted in which both young and elderly participants studied and recognised faces of varying ages. The results demonstrated that age-related deficits in recognition accuracy were reduced for older faces, and that this effect extended to face-picture recognition. However, face age had no effect on the age-related increase in false recognitions of faces. This suggests that the ability to accurately recognise features and the tendency to generate false alarms are distinct processes that are affected differentially by ageing.

This article from [2] discusses the phenomenon of cognitive ageing and emphasises several significant findings that must be considered in order to comprehend the causes of age-related differences in cognitive performance. Contrary to popular belief, the article notes that age-related cognitive declines are relatively significant and commence in early adulthood. In addition, these declines are observed in a variety of cognitive abilities and are not always accompanied by an increase in inter-individual variability. These results indicate that cognitive ageing is a complex and multifaceted process that cannot be solely attributed to age-related brain alterations. Alterations in motivation, experience, and environment may instead play a role in shaping age-related differences in cognitive functioning. Overall, the article emphasises the need for additional research to better comprehend the mechanisms underlying cognitive ageing and to develop interventions that can mitigate its negative effects.

Using deep autoencoder networks, the paper [3] proposed a method to convert high-dimensional data to low-dimensional codes. To reconstruct high-dimensional input vectors, autoencoder networks are trained with a multilayer neural network containing a small central layer. However, the network's efficacy is contingent on the initial weights. In order to resolve this issue, the paper describes an efficient method for initialising the weights that enables deep autoencoder networks to learn low-dimensional codes that are more effective than principal components analysis at reducing the dimensionality of data. This method is particularly advantageous when dealing with high-dimensional data, such as images, that must be represented in a condensed and meaningful manner. Experiments on multiple benchmark datasets demonstrate that the proposed method outperforms conventional methods in terms of both accuracy and efficiency. Overall, this work presents a practical solution for dimensionality reduction in high-dimensional data and makes a significant contribution to the field of deep learning.

Combining self-organizing maps (SOM) and convolutional neural networks (CNN), the authors of [4] proposed a hybrid neural-network approach for human face recognition. The SOM reduces the dimensionality of image samples and provides invariance against minor changes, whereas the CNN provides partial invariance against translation, rotation, scaling, and deformation. In a series of hierarchical layers, CNN is designed to derive increasingly large features. Using the Karhunen-Loeve transform in lieu of the SOM and a multilayer perceptron (MLP) in place of the CNN, the system is compared. The results are demonstrated using a database containing 400 images of 40 individuals with varying expressions, poses, and facial characteristics. The authors analyse computational complexity and discuss how the trained recognizer could be expanded to include new classes. The system described in this paper compares favourably to other face recognition methods.

The purpose of [5] was to investigate the reliability of eyewitness identification in a simulated criminal scenario. Three groups of eyewitnesses observed two targets and were subsequently assessed for identification accuracy 2, 21, or 56 days later. A live "show-up" or photographs were used to assess the subjects, and only one of the two targets was actually present in the 5-man array during testing. According to [5], there was a correlation between delay and the number of false alarms, with lengthier delays resulting in more false alarms. The testing method affected the number of matches, with live appearances yielding more accurate identifications than photographs. More than sixty percent of the subjects selected one of the four distractor individuals, indicating a high rate of misidentification. Only 28% of the subjects made no identification mistakes. These findings in [5] have significant implications for the use of eyewitness identification in criminal investigations, as they suggest that eyewitnesses may make erroneous identifications even with relatively brief delays. Additionally, the study highlights the potential limitations of using photographs to identify suspects.

Chapter-3

Functionality

Python was used to implement the model using PyTorch [4]. NVIDIA 1060 GPU was used for all investigations.

A. Dataset

Age-labeled facial images are necessary for training the model. UtkFace [5] dataset was utilized for this purpose. This dataset includes over 23,000 images of individuals ranging in age from 0 to 116 years old. The images were separated into six age categories. 0-18, 19-30, 31-40, 41-50, 51-60, and 60+. There are unequal amounts of data in each age category after dividing the data into age categories. In order to counteract this, we manually augmented the data by limiting the exact facial area using a face detector.

B. Experimental Setup

Model input image dimensions are 128 by 128 pixels. We trained the model using the Adam optimizer [26] as opposed to the gradient descent optimizer. It has been demonstrated to be a superior option for working with GANs. The Adam optimizer has two parameters, beta1 and beta2, that determine the gradient descent's momentum. We trained the model for fifty iterations with a group size of eight. The method lists few additional hyper-parameters for the model.

C. Evaluation Matrix

The generated image is difficult to evaluate in any context. There is no widely accepted method for evaluating the quality of generated images. GAN is a generative model that has been used in recent years to generate various types of images. Various publications have reported performance in various methods. The majority are human evaluations [10]. A volunteer is shown both a computer-generated and an actual image for a brief period of time. The objective is to identify which is generated and which is genuine. This score is converted to accuracy to reflect the image's performance. This evaluation is neither trustworthy nor comparable.

With the intent of preserving the original person's identity, we have generated images of various age groups in this work. Face Aging [7] had previously introduced a face recognition score evaluation matrix to evaluate the performance of preserving the identity of the original person.

A pre-trained visage verification algorithm is used for this purpose. OpenFace [8] has been used for visage verification. OpenFace is an open-source library for face recognition and verification. The face verification algorithm receives two images as input and returns a score indicating the degree of similarity between the two faces. The lower the score, the higher the algorithm's confidence that both images contain the same person. Typically, a distance of less than 0.7 is deemed adequate for positive verification. We have calculated the distance between all generated images of different ages and the original images. For all images where this distance was less than 0.70, we treated it as a positive verification and calculated the model's overall accuracy.

As we have generated images of various age groups, the image must maintain the characteristics of the respective age group. We have trained a classifier of these age categories for this evaluation. We calculated the predicted class for each image based on the generated image and the overall accuracy of each class.

D. Face Recognition Score

As mentioned in the previous section, the face recognition score was calculated using the pre-trained OpenFace library. We have compared the outcome to known prior outcomes. Antipox et al. [7] proposed a method for face ageing that uses explicit Identity preserving optimization to maintain the individual's identity. They did not concentrate on the classification rate or the number of images with content**.**

Chapter – 4

Steps to Run the Project

1. Installation

Install 3rd-package dependencies of python (listed in requirements.txt)

tensorflow-gpu==1.4.1

scipy==1.0.0

opencv-python==3.3.0.10

numpy==1.11.0

Pillow==5.1.0

pip install -r requirements.txt.

Other libraries

CUDA 8.0

Cudnn 6.0

2. Download datasets.

We use the Cross-Age Celebrity Dataset for training and Evaluation. More details about this dataset, please refer to (http://bcsiriuschen.github.io/CARC/). After face detection, aligning and centre cropping, we split images into 5 age groups: 11-20, 21-30, 31-40, 41-50 and 50+.

3. Test on saved models

Download the trained face aging model(https://1drv.ms/u/s!AlUWwwOcwDWobCqmuFyKGIt4qaA) and place models files in checkpoints/0_conv5_lsgan_transfer_g75_0.5f-4_a30.

Test images are in images/test, and some training images that belong to 11-20 age group are in images/train.

* Running the script to get aged faces

python test.py

4. Train from scratch

Firstly, download the pre-trained Alex net model(https://1drv.ms/u/s!AlUWwwOcwDWobkptownyu5fjlfU) and age classfication model(https://1drv.ms/f/s!AlUWwwOcwDWocX-Z0IJft_VbcoQ). Then unzip these files and place model files in checkpoints/pre_trained.

```
python age_lsgan_transfer.py \
  --gan_loss_weight=75 \
  --fea_loss_weight=0.5e-4 \
  --age_loss_weight=30 \
  --fea_layer_name=conv5 \
  --checkpoint_dir=./checkpoints/age/0_conv5_lsgan_transfer_g75_0.5f-4_a30 \
  --sample_dir=age/0_conv5_lsgan_transfer_g75_0.5f-4_a30
sh age_lsgan_transfer.py
```

You can change the hyperparameters to adapt to your own dataset.

5. Experiment results

Fig 7. Result Outcome

Chapter – 5

Methodology

The basic idea of the generating artificial network, a key component of our model, gets clarified first.. Finally, we present our future generation's approach.the visage images of individuals of various ages. The impact of the numerous objective function components is then discussed at last to explain the model's training processes.

A. Network of Creative Writing Challengers

In attempting to trick a discriminator, the programmer learns to produce pictures that approach the target domain, which has been trained to differentiate between genuine and synthetic images. Training is provided to the generating device and detector iteratively using adversarial training, where the generator attempts to improve its performance by generating better images and the discriminator attempts to improve its ability to distinguish between actual and generated images. This procedure is repeated until the generator can produce images that the discriminator cannot distinguish from actual images.

GAN wants to create an actual, unique image that matches the distribution of the domain under all circumstances.

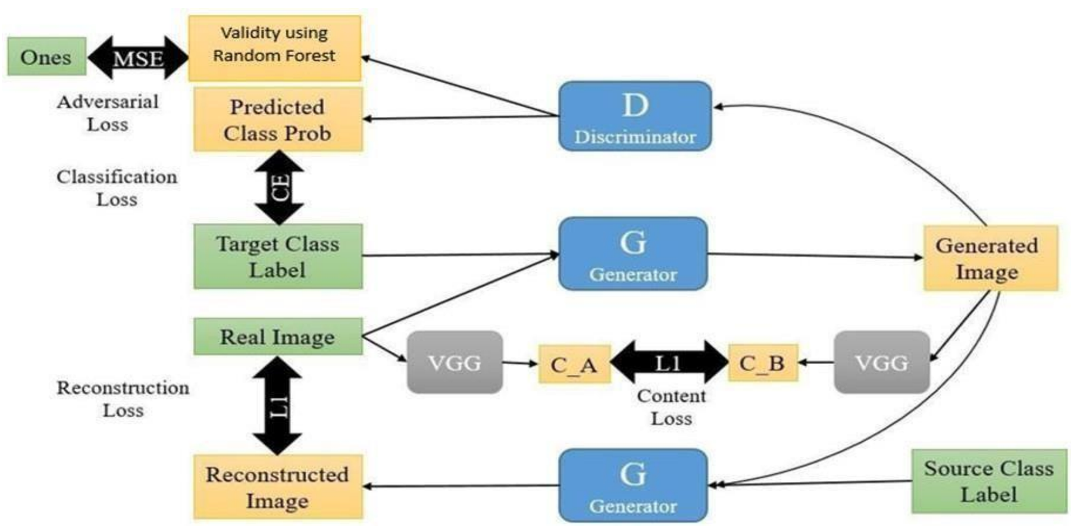

Fig. 8 Random Forest Training

The illustration in Figure 8 describes RF development. GAN can also be utilized in conditional contexts by providing the generator with additional input. The constrained generated artificial network (cGAN) was introduced by Mirza et al. in [9]. By providing additional information, such as the target class or attribute, GAN can be utilized in the conditional configuration.
It popularized the image-to-image translation dimension of the generative model. These models generate a target image using an input image as a condition. Isola et al. [10] first proposed the network for image-to-image conversion which generates an excellent quality photograph. Because an image is used as a conditional variable, this type of model is generated using encoder-decoder architecture.

The compression algorithm receives the image as contributions and transforms it into a matrix with fewer parameters, which is then used as input by the receiver. It is sampled by the receiver in order to generate the desired image. Occasionally, the algorithm for decoding accepts an unpredictable vector as input in addition to the encoding device matrix. The purpose of the unpredictable vector is to imbue the generated image with inconsistencies.

B. Aging Face prototype

Our goal is to educate a model for picture-to-picture transformation across multiple domains. In addition, the contentious and customary losses were implemented to make the image resemble the target domain while preserving the identity of the source. The discriminator functions as a supplementary classifier [2], enabling a single discriminator to evaluate images from multiple fields. A discriminator has two outputs to this objective: class probability and software validation. Similar to traditional GAN, validity is the random score of an image living authentic or fraudulent, category probability is the expected value of a derived image being a part to a particular area, as suggested using a discriminator.

Domains Classifier Changes: The objective of this framework is to offer a picture of the area of focus that is recognised as a representation associated with the domain... We already have implemented For this, the discrimination algorithm produces an additional result and an assessment loss. The classifier is given instructions on how to correctly identify pictures during distinction training.

Reconstruction Loss: By minimizing this hostile loss teaches the generator how to create a realistic image. that corresponds to the target image's statistics division. The classification reduction creates visuals which resemble those of the subject class. However, none of these losses impose any restrictions on keeping the source image's details or preserving the source image's subject's identity. To add this constraint, we optimize two additional generator losses.

The reconstruction loss imposes half of Cycle Gan’s objective [11].

Since the curriculum level is a parameter to a generator, the two separate ones are unnecessary to impose this loss. This loss provides an indirect requirement on the image that is produced so it can be transformed back to the initial picture with the identical class level.

Even though this loss demonstrated a remarkable result according to reports within an assortment of uses, of papers, it was a significant setback.

Nonetheless, when the engine discovers an easy way that produces the picture returned from its starting nation, known as picture, might get converted with no the feature you want being preserved at the target level. We have implemented content loss to guarantee that the material of the original image is maintained in the image that was generated.

Content Loss. The intent of content loss is to require the generator to preserve the fundamental content of the source image in the generated image. It is the difference between the content of the source file and the generated image. comparable losses in comparable programmes, such as image super-resolution, have yielded positive results [4]. Extracting the data information via a deep convolutional neural network's learnt high-level representation. Deep CNNs trained on huge data sets are well-known to be efficient at identifying high-level image features. To detect content loss, we used the Image Network [6]-trained visual Geometric Group-19 [5] model for obtaining characteristics.

The generator's output is provided with target identifiers for subclass and real image. The picture that is generated is then sent along with the source class label to the detector and maker. Validity and predicted class probability are output by the discriminator. We compute two objective losses based on this information: both the conflict and classification losses. The algorithm that generates it results in the unprocessed images rebuild. It is to determine the difference involving the initial pic and the picture that was reconstructed with L1 loss. The real picture and its associated generated pictures are fed into the trained fully stored visual geometry group-19 algorithm to determine information degradation. What kind of damage occurred is then determined by combining the results of both models.

C. Construction of Networks

Author in [7] presented cGAN after the introduction of generative adversarial network for low-resolution images. This model's architecture was adopted, with modifications and variations incorporated into subsequent publications. [8] Both the generator and discriminator utilize a convolution group to renormalize the ReLU module.

D. Coaching: In order to sustain the GAN training and produce high-quality images, a modified training objective was proposed. Wasserstein [2, 3] demonstrated that optimizing the GAN objective proposed in the original GAN paper [7] is challenging. The proposed second WGAN is known as WGAN-GP. As an alternative to clipping, gradient plenty was utilized to enforce Lipchitz constant. A differentiable function f is 1-Lipschitz if the maximal value of its gradient everywhere is one.

Consequently, instead of clipping the model as WGAN does, WGAN- GP penalizes it if the gradient norm is not identical to its objective norm of one. During our model's instruction.

Generator with bypass link: The issue of excellent quality photographs is used in the process of translating a photograph into another photograph. In the majority of Both the source & target pictures have a comparable visual format when used in solutions despite their distinct what it looks like. In the app that we are developing, the identity of a person who differs at various points of ages must be preserved, which is why we developed the loss function. We believe that this property is the generator.

Numerous applications in this field employ encoder-decoder framework [9]. In this kind of network, the input is transmitted to the generator so that it can be compressed to create a low-dimensional approximation. This concept is referred to as block visualization. The data is further transmitted to the decoder, which resamples it to the image size. In numerous applications, the majority Both networks share some fundamental characteristics. Therefore, it would be beneficial to instantaneously transmit the same attribute between networks.

We adhered to the logic presented in U-Net [2] to provide the source who has the tools to spread this knowledge. In between encoder and decoder layers, this network utilizes skip-connection.

ResNet [2] popularized the idea of skipping the connection. It utilized a connection among its tiers that skips. Combining inputs of a layer with its output or the outputs of multiple layers is the concept. It transmits direct information between network layers to aid the model in understanding a transition's transition and change. A bypass level i to level (n-i), wherein n represents the total amount of levels within the a generator, are connected, exists in U-Net, is inserted between the encoder and decoder.

Chapter-5

Result and Discussion

This research contributes to the field of facial expression recognition by examining aging's effect on how one looks at detection. Our experiments demonstrate that age-related effects must be considered when attempting to recognise facial expressions. By considering this, the solution for expression recognition will be more general and efficient.

Fig. 6. Illustrations of pictures produced by the suggested technique.

The programme is written in Python and its graphical interface is provided by the Tkinter module. It uses the Random Forest Algorithm for image analysis, as well as Math and PyPDF2. The programme prompts the user to open the camera to capture and analyze facial pictures, retrieves the images on match found as in fig 6.. The preparatory phase involves tokenization, removal of stop words, and stemming. The algorithm then employs a statistical approach to frequency-inverse document frequency to characterize the document's characteristics and selects key words and phrases based on these characteristics. The overview is displayed in the interface segment. Using Rouge 2.0, which employs a synonym dictionary to capture semantic overlap, the efficacy of the programme is evaluated. Using the CNN/daily mail dataset, the results of the programme are compared to other traditional methods and baselines, and it is discovered that the programme significantly outperforms them with a higher score.

Chapter-6

Conclusion and Future Work

6.1 Conclusion:

Numerous applications exist for the ability to generate visage images of various ages based on a picture and the present age-factor. In this paper, the method for completing this endeavour is proposed. Our GANs-dependent procedure can produce facial pictures of different peer-groups. Experiments have proven the fact that our approach is superior and preserves the individual's identity in the generated image than other methods. In addition, our method produces images with age-appropriate characteristics, which is supported by the precision of the categorization. Suggested enhancements for this effort involve incorporating photos via greater pixels to enhance the picture quality. and enhancing the image recognition function.

6.2 Future Work:

In the field of facial detection and aging effects, there are several potential areas of future work and advancements. Here are a few ideas:

Improved Accuracy: Researchers can focus on improving the accuracy of facial detection algorithms, especially in challenging conditions such as low lighting, occlusions, or pose variations. This can involve developing more robust feature extraction techniques, leveraging deep learning models, or combining multiple modalities like infrared or 3D imaging.

Real-Time Performance: Enhancing the speed and efficiency of facial detection algorithms is crucial for real-time applications. Future work can involve optimizing algorithms and architectures to achieve faster processing times while maintaining high accuracy. This can enable applications like real-time facial recognition for security or video analytics.

Age Progression/Regression: Aging effects can be further refined to create more realistic and accurate transformations. Researchers can explore advanced deep learning techniques, such as generative adversarial networks (GANs), to generate more natural-looking and personalized age progression/regression effects. The focus could be on capturing individual-specific aging patterns or considering additional factors like lifestyle, genetics, and health conditions.

Long-Term Age Progression: While current age progression techniques primarily focus on short-term aging effects, future work can involve simulating long-term aging changes. This could help in forensic investigations, missing persons cases, or simulating age-related effects for entertainment purposes. Capturing the complex and dynamic nature of long-term aging would require sophisticated modelling and incorporating factors like environmental influences, skin elasticity changes, and lifestyle choices.

Cross-Ethnicity and Cross-Cultural Aging: Current aging effects primarily focus on certain ethnicities or cultural backgrounds. Future work can involve developing more diverse and inclusive models that accurately capture the aging process across different ethnicities, races, and cultural groups. This would require diverse training datasets and models that can generalize well to various populations.

Facial Aging Prediction: Predicting future facial appearance based on present features can have several practical applications. By leveraging machine learning techniques and longitudinal datasets, researchers can develop predictive models to estimate how individuals might age over time. This could find applications in personalized healthcare, anti-aging treatments, or virtual makeovers.

Robustness to Adversarial Attacks: Adversarial attacks, where subtle modifications are made to input images to deceive facial detection systems, pose a significant challenge. Future work can focus on developing robust facial detection algorithms that are resilient to adversarial attacks. This involves exploring techniques such as adversarial training, robust feature representations, or incorporating security measures to ensure reliable performance in the presence of malicious attempts.

These are just a few potential areas for future work in facial detection and aging effects. As technology advances, researchers and practitioners can continue to explore these directions to enhance accuracy, realism, and applicability in various domains.

Reference

1. Fulton, A. and Bartlett, J.C., 1991. Young and old faces in young and old heads: the factor of age in face recognition. *Psychology and aging*, *6*(4), p.623.
2. Salthouse, T. A. (2004). What and when of cognitive aging. Current Directions in Psychological Science, 13, 140-144.
3. Hinton, G.E. and Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. *science*, *313*(5786), pp.504-507.
4. Lawrence, S., Giles, C.L., Tsoi, A.C. and Back, A.D., 1997. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*, *8*(1), pp.98-113.
5. Goldstein, A.G., 1977. The fallibility of the eyewitness: Psychological evidence. *Psychology in the legal process. New York: Spectrum*.
6. Mason, S.E., 1986. Age and gender as factors in facial recognition and identification. *Experimental Aging Research*, *12*(3), pp.151-154.
7. Johnson, M.H., Senju, A. and Tomalski, P., 2015. The two-process theory of face processing: modifications based on two decades of data from infants and adults. *Neuroscience & Biobehavioral Reviews*, *50*, pp.169-179.
8. Suo, J., Zhu, S.C., Shan, S. and Chen, X., 2009. A compositional and dynamic model for face aging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), pp.385-401.
9. Tazoe, Y., Gohara, H., Maejima, A. and Morishima, S., 2012. Facial aging simulator considering geometry and patch-tiled texture. In *ACM SIGGRAPH 2012 Posters* (pp. 1-1).
10. Bartlett, J.C. and Fulton, A., 1991. Familiarity and recognition of faces in old age. *Memory & Cognition*, *19*(3), pp.229-238.
11. George, P.A. and Hole, G.J., 1995. Factors influencing the accuracy of age estimates of unfamiliar faces. *Perception*, *24*(9), pp.1059-1073.
12. Agrawal, A.K. and Singh, Y.N., 2015. Evaluation of face recognition methods in unconstrained environments. *Procedia Computer Science*, *48*, pp.644-651.
13. Fontaine, X., Achanta, R. and Süsstrunk, S., 2017, March. Face recognition in real-world images. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1482-1486). Ieee.

14. Puthea, K., Hartanto, R. and Hidayat, R., 2017, November. A review paper on attendance marking system based on face recognition. In *2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)* (pp. 304-309). IEEE.
15. Durkin, E., 2019. New York school district’s facial recognition system sparks privacy fears. *The Guardian*.

Communicated

5th IEEE International Conference on Advances in Computing, Communication Control and Networking (ICAC3N-23) : Submission (318) has been edited. Inbox x

Microsoft CMT <email@msr-cmt.org>
to me ▾

11:56 AM (3 hours ago)

Hello,

The following submission has been edited.

Track Name: ICAC3N2023

Paper ID: 318

Paper Title: Facial Detection and the Effect of Ageing

Abstract:
In computer vision, generating images of faces that appear to age is a challenging challenge. This is known as a conditional image generation problem because the generated image is dependent on a specified face image. Until recently, generative models were incapable of producing realistic, high-resolution photo simulations. Nevertheless, a generative model known as a generative adversarial network (GAN) has demonstrated remarkable abilities both in absolute and constrained contexts for generating realistic images. Despite this, generating images that display the aging process while preserving The true identity of the individual in the image source is a difficult task due to two contradictory constraints: The image that is produced must maintain the person's identification from the primary image. while also possessing all the characteristics of age targeted. Regarding this study, the authors used a GAN in a conditional setting in conjunction with a custom loss function designed to strike a balance between preserving the individual's identity and generating characteristics of the target age

Created on: Fri, 12 May 2023 06:24:28 GMT

Last Modified: Fri, 12 May 2023 06:26:42 GMT

Authors:
- imrnnzr1@gmail.com (Primary)
- harshitasengar7597@gmail.com
- S.SRINIVASAN@GALGOTIASUNIVERASITY.EDU.IN

Secondary Subject Areas: Not Entered
Submission Files: BT4067_RP (2).pdf (1 Mb, Fri, 12 May 2023 06:24:21 GMT)

Submission Questions Response: Not Entered

Thanks,
CMT team.