

**A Deep Learning Model for Classification of  
Gender and Age from Facial Images**

**A PROJECT REPORT**

*Submitted by*

**FATHIMA NOVRIN (TKM20MCA2017)**

to

**The APJ Abdul Kalam Technological University**

*In partial fulfilment for the award of the degree of*

**MASTER OF COMPUTER APPLICATIONS**



**THANGAL KUNJU MUSALIAR COLLEGE OF ENGINEERING  
KERALA**

**MARCH 2022**

## DECLARATION

I undersigned hereby declare that the mini project report “**A Deep Learning Model for Classification of Gender and Age from Facial Images**”, submitted for partial fulfillment of the requirements for the award of degree of Master of Computer Applications of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by me under supervision of Dr. Nadera Beevi S. This submission represents my ideas in my own words and where ideas or words of others have been included, I have adequately and accurately cited and referenced the original sources. I also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in our submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

KOLLAM  
12/07/2022

FATHIMA NOVRIN

**Thangal Kunju Musaliar College of Engineering**  
**Dept. of Computer Applications**



**C E R T I F I C A T E**

This is to certify that, this report entitled **“A Deep Learning Model for Classification of Gender and Age from Facial Images”** is a bonafide record of the work submitted by **FATHIMA NOVRIN(TKM20MCA-2017)**, under our guidance and supervision, in partial fulfilment of the requirements for the award of the Degree of **Master of Computer Applications** in **APJ Abdul Kalam Technological University**. This report in any form has not been submitted to any other University or Institute for any purpose.

Internal Supervisor

Head of the Department

External Examiner

## **ACKNOWLEDGEMENT**

First and foremost I thank GOD almighty and our parents for the success of this project. I owe sincere gratitude and heart full thanks to everyone who shared their precious time and knowledge for the successful completion of our project.

I am extremely grateful to **Dr. Nadera Beevi S** project guide, for providing me with best facilities.

I profusely thank all other faculty members in the department and all other members of TKM college of Engineering, for there guidance and inspirations throughout my course of study.

I owe thanks to my friends and all others who have directly or indirectly helped me in the successful completion of this project.

**FATHIMA NOVRIN**

## **Abstract**

Gender classification and Age identification play an important role in our social lives. Gender is central characteristics of personality, and it is essential in our life. Age is important for our identity. Security, biometric system, and treatment are part of gender classification and age prediction. Age prediction can help to authorize people from buying adult products or other kind of restricted goods. In this project, for classification of gender and prediction of age from pictures deep learning model is used. The objective of this study is to create a model for gender classification and an age estimation using convolutional neural networks and ResNet50. The image's feature extraction and categorization are included by CNN. Feature extraction gives the features corresponding to gender and age from the face pictures whereas the classification classify the image into correct age and gender. ResNet50 is the convolutional network that have 50 layers. Age prediction is the regression problem and prediction of gender is a binary classification problem. The model is evaluated using the UTKFace dataset, a sizable face dataset with a broad age range. Deep learning algorithm is used to obtain higher accuracy and lower MAE, also MAE of the both algorithm is compared to obtain which algorithm more efficient.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Objective.....	2
<b>2</b>	<b>Literature Survey</b>	<b>3</b>
<b>3</b>	<b>Methodology</b>	<b>5</b>
3.1	Existing system .....	5
3.2	Proposed system .....	6
3.3	Materials and methods.....	6
3.3.1	Dataset .....	6
3.3.2	Pre-processing.....	8
3.3.3	CNN.....	9
3.3.4	ResNet50.....	11
3.4	Software requirements.....	11
3.4.1	Tensorflow.....	12
3.4.2	Python.....	12
3.4.3	Keras.....	13
3.4.4	OpenCV.....	14
3.5	working flow.....	14
<b>4</b>	<b>Results And Discussion</b>	<b>16</b>
4.1	Result for age and gender prediction (CNN). .....	16
4.2	Result for age prediction(ResNet50).....	19
<b>5</b>	<b>Conclusion</b>	<b>21</b>
5.1	Future Enhancement .....	22
	<b>References.....</b>	<b>23</b>

# List of Figures

- 3.1 UTK face dataset examples .....6
- 3.2 Distribution plot of age group .....7
- 3.3 Count plot of gender .....7
- 3.4 Examples of dataset after pre-processing.....9
- 3.5 CNN architecture.....10
- 3.6 ResNet50 architecture .....11
- 3.7 Block diagram .....15
  
- 4.1 Accuracy graph of gender prediction .....17
- 4.2 Loss graph of gender prediction.....17
- 4.3 Loss graph of age prediction. ....18
- 4.4 sample images of output.....18
- 4.5 Model Evaluation .....19
- 4.6 sample images of output.....20

# Chapter 1

## Introduction

Age and gender are regarded as crucial biometric characteristics for identifying humans. The process of identifying a human's face in a photograph, classifying them as male or female, and determining their age is known as gender and age prediction. The main important facial characteristics in social life are age and gender. These two characteristics are essential to our social lives. The characteristics of a human's face determine their age, gender, emotions, and ethnicity. Gender and age prediction has become essential for an increasing variety of applications, particularly in light of the growth of social media and social networks. In real life, several techniques, such as those used in product sales, forensics, entertainment, and biometrics etc., are closer to gender and age prediction.

Our social lives heavily depend on how old we are and what gender we are. The world has various welcomes for men and women for every language, and talking to elderly sometimes involves using a different vocabulary than talking to kids. These practises heavily rely on one's capacity to infer age and gender from a person's appearance, which are personal characteristics. Many app developers are passionate about automated age identification, particularly given the expansion of social media platforms. In day to day interactions, gender and age are the most important facial characteristics. The features of a person's face can be used to infer information about their ethnicity, identity, age, and gender. Gender and Age identity are two of these characteristics that might be crucial. It is required to determine the age and gender of humans from their photographs. However, various issues with age and gender identification remain open concerns for academics. Despite advancements, the computer vision community continues to develop the state-of-the-art by introducing new methodologies. However, unfiltered real-life face photographs cannot yet accurately forecast age and gender in a way that meets the demands of commercial and practical applications. A firm and accurate approach is therefore very necessary for age and gender identification.

Collecting personal information on a large scale is not always easy, because not everyone has the time to complete many surveys at once and since it is a sensitive topic. Collecting facial images can be a simpler process to complete.

However, potentially be made easier. Installing several surveillance cameras throughout the facilities might be done, and it should be rather simple for businesses and organisations to collect facial images of their guests for security and research purposes. Due to this, businesses and organisations may urgently want reliable solutions that can accurately anticipate visitor identifying details from simply the faces. Researchers have been delving deeper into the problem as the outcome of the practical implications of deep learning and computer vision in recent years.

In this paper, Using CNN architecture, which combines classification and regression skills, the problems of gender classification and age prediction from face photos are resolved. Also Resnet50 architecture used to compare which model gives better result.

## **1.1 Objective**

The objective of this project is to detect age and gender through the given data set. Unfiltered face age and gender predictions categories free-range real-world face images into preset age and gender. A deep learning model can implemented to estimate the gender and age .Given an input, the images will be detected which is then compared to the images in the dataset Some of the preceding techniques have the limitation of having a greater MAE but a lower accuracy, namely for the task of age estimate. As a result, the Deep Learning technique is employed to achieve reduced MAE and greater accuracy. CNN is used for the classification and regression. Then predict the age and gender corresponding to that image. ResNet50 is used to predict the age and to compare which algorithm have lower MAE. This project will be evaluated on dataset which is a sizable face collection with a broad age distribution (range is 0 -110 years old).

## Chapter 2

# Literature Survey

Several important studies have been conducted for gender classification and age prediction using face pictures. Previous approaches for this area were primarily concerned with the extraction and computation of face characteristics.

In [5], deep neural network, which is computationally cheap and gives high accuracy on numerous competing datasets, was employed. In [6], a deep CNN network with locally linked hidden layers was proposed. For training, the datasets CAS-PEAL and FEI were utilised. An alternative to using 100 patches to cover the whole part was suggested by Zukang Liao et al. [7]: 9 overlapping patches per shot. The nine-patched technique produced gender prediction accuracy of 95.072 percent on the Labelled Face in the Wild (LFW) dataset and 78.63 percent on the Adience dataset. For age classification, accuracy was 40.25%. Rothe et al. [8] without employing facial landmarks. The IMDB-WIKI dataset, the regression method via deep classification, and obtaining accuracy of 64.0% are their primary achievements. In [9], A hybrid architecture that combines a CNN and ELM was introduced. The average accuracy for predicting age and gender using the datasets MORPH-II and Adience benchmark is about 52% and 88%, respectively.

A computationally effective CNN model was created for the mobile platform to estimate age and gender in [10]. In [2], Using the LBP and the LBP Histogram model, the authors created an android app. As simple CNN was implemented for apps in [11] Age top-1 accuracy with the Adience dataset is 44.26 %, while gender top-1 accuracy is 85.16 percent when utilising LMTCNN-2-1. Additionally, a smartphone-based solution that uses eye scans to estimate gender in [12]. The accuracy obtained from the iPhone and Oppo photos using the VISOB dataset was 78 percent and 86.2 percent, respectively.

Most recently, other analyst have implemented a CNN model for gender prediction that has average accuracy and incurs little calculation expense. Using popular datasets like IMDBWIKI, LFW, and Adience, they compared their approach against other well-known CNN architectures. dataset [13]. In [14] They used a simple model to estimate people's ages and genders. For an embedded system, they have applied multi-task learning with different types of datasets, including the FGNET, MORPH-II, and MegaAge-Asian datasets. In [15], They created an age prediction approach based on a simple CNN and Data Augmentation using the same datasets, MORPH-II and FG-NET. Regression and

classification algorithms were combined to create their suggested mixed attention strategy. In [16], Using CNN, developed to identify the sex, age and emotion. To train their model, they employed a haar cascade classifier and 10000 photos. In [17], an unfiltered picture was used to train a deep learning classifier to estimate age and gender. Two components of their CNN design extract features and categorise data. With the use of the IMDB-WIKI, MORPH-II, and OIC-Adience datasets, they trained their network to predict age and gender more accurately. When analysing and summarising the aforementioned studies, it becomes clear that their implementations have several shortcomings, such as CNN architecture, large CNN models, high CPU costs, dataset processing, and so forth. According to previous study, there are two main factors that influence how well CNN predicts age and gender:

- Design alterations.
- Dataset preparation

## Chapter 3

# Methodology

### 3.1 Existing System

Table I shows a comparative assessment of different age and gender prediction algorithms suggested in recent years. Also attempted to determine the merits and demerits of previous efforts done in this topic throughout the years.

References	Database	Advantages	Limitations
Chang et al. (2011)	FG-NET database, MORPH Album 2 database	The information of relative order between ages is more reliably employed than conventional ways of using by OHRank.	Achieves higher MAE (more errors) for age detection than other neural network based methods.
Karimi & Tashk (2012)	FGNET dataset	Achieved high accuracy for gender recognition. Achieved suitable performance even if the utilized images were subjected to intrusive noises.	The task for age recognition produced lower accuracy and the ages were divided into groups for classification
Levi & Hassner (2015)	Adience dataset	Reduced number of parameters thus reducing chances of overfitting. Also has reasonable accuracy for gender recognition task.	Lower accuracy for the age recognition task due to its simple design.
Samek et al. (2017)	Adience dataset	Achieved reasonable accuracy for gender recognition.	Produced a lower accuracy for the age recognition task
Zhang et al. (2017a)	IMDB-WIKI, ImageNet dataset	Achieved high accuracy for gender classification task. Works well for high-resolution facial images.	Produced a lower age detection accuracy.
Srinivas et al. (2017)	Wild East Asian Face Dataset (WEAFD)	Produced reasonable accuracy for gender classification task. One of a kind dataset consisting primarily of labeled face images of individuals from East Asian countries were designed for classification task.	Achieved very low accuracy for the age detection task.

TABLE I. A comparison of several approaches for predicting gender and age

### 3.2 Proposed System

According to the preceding discussion and Table 1, the prior techniques all share a similar flaw of poorer accuracy and increased MAE, primarily for the age estimation task. The gender classification accuracy is low of the previous approach

To address these issues, the deep learning method is utilized to construct the model that identify the gender and age. CNN is used in facial image prediction to estimate both age and gender. Resnet50 is used to forecast age and compare which algorithms have the lowest MAE.

### 3.3 Materials and Methods

#### 3.3.1 Dataset

The age range of the large-scale face dataset UTKFace is quite broad, range is 0 -110 years old. 22,000 face images with attributes like gender, age and ethnicity are included in the collection. Each image shows distinct face emotions, unique stances, and close, distinct facial features. The resolutions of the images also vary. For more accurate feature recognition, this dataset offers the original photos in a cropped format. The images contain a broad variety of positions, facial expressions, lighting, occlusion, resolution, and so on. Using this dataset, a range of tasks, including face identification, age estimate, landmark localization and age progression, may be carried out. Figure 3.1 displays samples from the UTK face dataset.



Fig 3.1 UTK face dataset samples

Figures 3.2 and 3.3 shows the distribution of the dataset

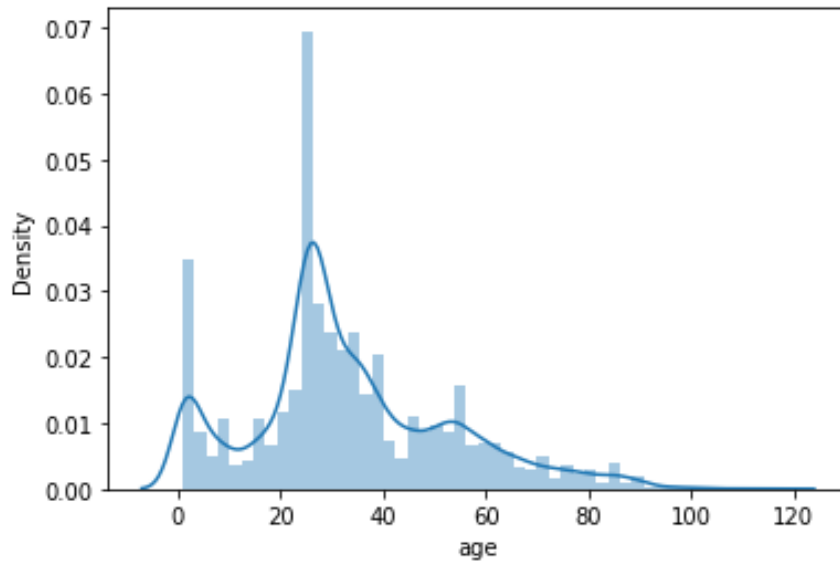


Fig 3.2 Distribution plot of age group (1-120 years)

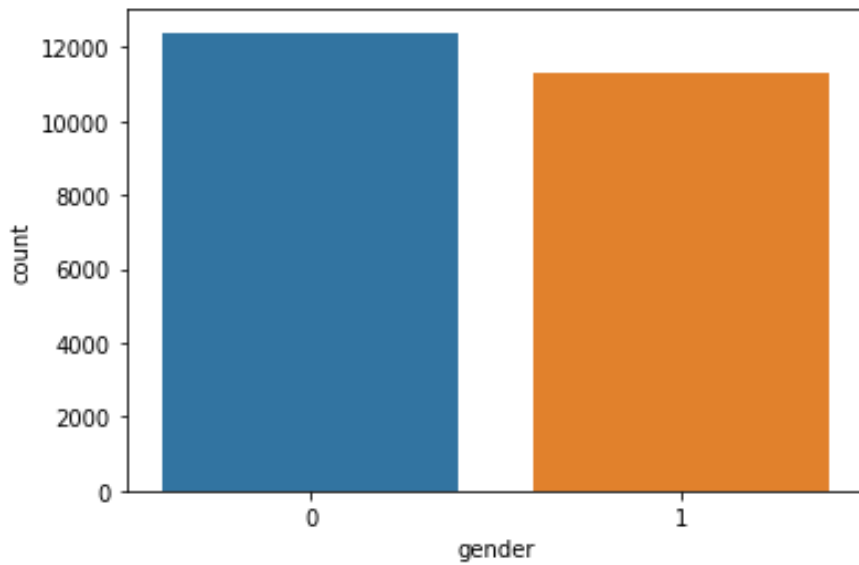


Fig 3.3 Count plot of gender(0-male,1-female)

### **3.3.2 Pre-processing**

Pre-processing occurs before data is trained and evaluated. Pre-processing includes numerous stages such as resize the image, converting the image to array, transforming the image to a greyscale image, and normalising.

- An essential stage in computer vision's pre-processing is image scaling. Because of the effectiveness of training models. With a reduced image size the model will perform better. The image is being resized into  $128 \times 128$  pixels.
- The following step is to convert all of the images in the dataset to an array. The image is saved as an array.
- Converting image into grey scale image. Instead of working directly on colour images, grayscale representations are typically used for extracting descriptors since doing so is more straightforward and requires less computational power.
- Data normalisation, commonly referred to as data pre-processing, is an essential component of data mining. It refers to the process of converting data, or turning the original data into a format that allows for effective data processing. Reduce or even eliminate duplicate data is the main goal of data normalisation.

After the pre-processing step the images in the dataset are shown in Fig 3.4



Fig 3.4 Samples of dataset after pre-processing

### 3.3.3 CNN

A common model in machine learning is the neural network (NN). It was created using neurons to model the architecture of the human brain. The neuron in a neural network is capable of receiving input and producing an output by applying that input to a function. These operations are known as activation operations. There are three levels in a neural network: input, hidden, and output layers. There are many neurons in each layer of the neural network, and the neurons in the input layer stand in for the number of characteristics selected. The total number of hidden layers varies according on the model, data volume, and issue complexity. The quantity of features does not affect the neurons number in the hidden layer, and as a consequence, the logistic functions are able to see the probability score for each class at the output layer. The feature produces a nonlinear network by applying biases or weights to the hidden layer and activation. The back propagation technique can accurately forecast the weights that will be applied at the hidden layer.

The most widely used deep learning algorithm in computer vision is CNN. It is computationally effective and automatically recognises pronounced traits. The layers in a CNN are identical to those in a typical neural network, however the hidden layer in a CNN is made up of several internal layers, including convolutional, pooling, fully connected, and normalising layers. CNN, which takes the shape of a matrix to interpret an image that

is an array of pixels, plays role in image processing. The filter that directs away for the feature map combines with the image matrix that enters the convolution layer. A picture can be in colour or in grayscale (RGB), which is black and white. Convolution includes the sliding of the filter over the input. The total of the element-wise multiplication of the pixel values between the input and filter determines the first-pixel value in the feature map matrix of size equal to the filter. By swiping the core over each pixel in the input picture matrix, pixel values in the feature map are correspondingly filled. By altering the filter, operations such as identification, edge detection, and sharpening may be applied to the input. By padding the input picture with zeroes at the edges to match the kernel, the future map is appropriately produced, and The stride regulates the filter sliding in this layer.

Data's dimensionality is decreased after convolution by the pooling layer. Every feature map is condensed by the pooling layer using several pooling techniques, including maximum pooling, average pooling, and sum pooling.

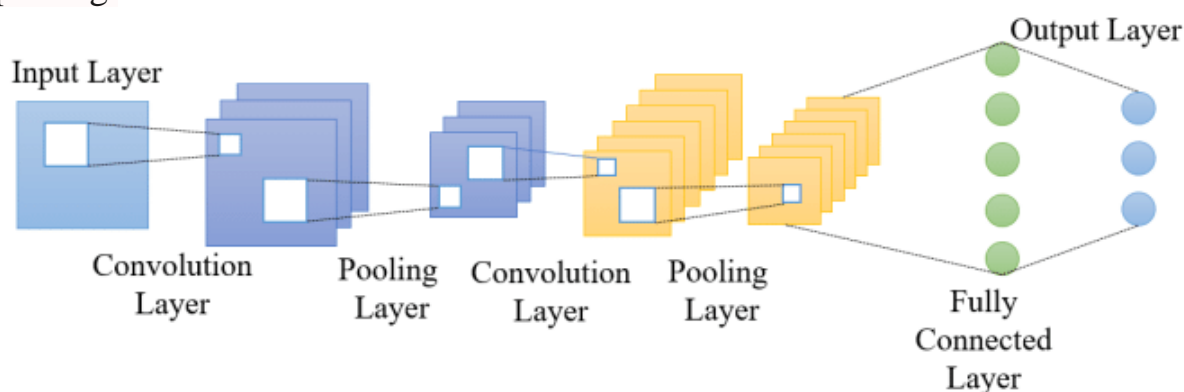


Fig 3.5 CNN Architecture

### CNN Architecture (5 Layers)

As seen in Figure 3.5, the CNN architecture is made up of many types of layers.

- **Convolutional layer:** convolutional layer is the backbone of CNN. photos are scanned by pixel by pixel to produce a feature map that will be used to define classes.
- **Pooling layer:** Pooling is otherwise known as data down sampling since it reduces the overall size of the pictures. Each feature's input from each convolutional layer is confined to only the most important data. Numerous convolutional layer and pooling may need for the iterations.
- **Fully connected input layer:** The flattening of the pictures is another name for this. Output of the previous layer are flatten and given as the input to the next layer

- **Fully connected layer:** When the analysis of feature is done and then compute, random weights are added by this layer to the inputs and label are predicted.
- **Fully connected Output layer:** last layer, and it returns the output of the labels computed for classification and gives a output class to the pictures.

### 3.3.4 ResNet50

Residual neural network is a deep learning model that is part of CNN. It works well with huge datasets that contain above 1000 images. It is a deep learning neural network with 50 layers and approximately 23 million number of training parameters. AlexNet, GoogleNet, and VGG19 are some training models that are similar. ResNet-50 distinguishes itself by having similar connections. This model is divided into five phases, each with three layers comprising convolutions. This network, like the classic NN, The output of one layer is provided as the input for the following layer and also offers by ideally being two to three levels away. Hence it is known as identity linkages.

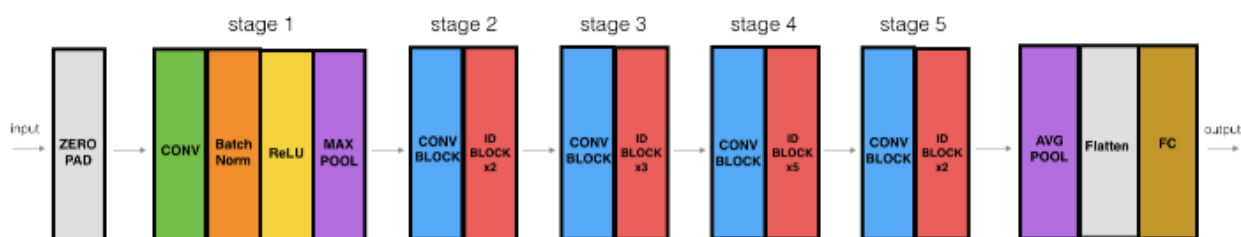


Fig 3.6 ResNet50 Architecture

As illustrated in picture 3.6, the model have 5 stage each stage have a identity block and convolutional block. All the convolution block contains 3 convolution layers, as do the identity blocks. Over 23 million trainable parameters are available in this model

## 3.4 Software Requirements

Pre-processing, splitting the data, constructing the model, testing the model, and finally implementing the model all require the following requirements.

### **3.4.1 TensorFlow**

TensorFlow is a set of processes with accessible, high-level APIs that enable the creation of machine learning models in a range of languages by both experts and beginners. As a result, moving from model creation and training to deployment is made significantly simpler for developers.

Use of Tensorflow:

#### **1. Image recognition:**

**a. Finding features of pixel**– Each image is a container for pixels, which are a collection of integers. The colour depth is represented by these numbers.

**b. Create a training picture.** – To train a model, categorise the photos into separate sections. For example, to improve comprehension, label a picture as 'vehicle,' 'bike,' and so on. model is trained with a huge range of photos for improved performance.

**c. Train the model to classify images**– Train a network with numerous photos to create a label as an output from the provided image as an input.

**d. Give an unknown input**– Test the model by feeding it a fresh picture that can be classified into any of the predefined categories.

### **3.4.2 Python**

Python is a high-level, interpreted, interactive, and object-oriented scripting language. The goal of Python's design is to be very readable. In addition to having fewer syntactical features than other languages, it frequently uses English terminology rather than punctuation.

High abstraction is a feature of the computer language Python. Through the use of excessive indentation, its design approach prioritises code readability. Its object-oriented methodology and language features are intended to aid programmers in creating clean, comprehensible code for both little and big projects.

some of the key advantages of Python:

- **Python can be interpreted** – This is comparable to PHP and PERL. The interpreter handles Python at runtime. Your programme does not have to be built before being used.
- **Python is Interactive** – To write your programmes, you may sit at a Python prompt and interact directly with the interpreter.
- **Python is Object Oriented program**– Python supports the OOPS concept or approach, which wraps into objects.
- **Python is a language for Beginner's** – Python is very useful for the beginners, as It enables the development of a wide range of programmes, including simple text editors, web browsers, and games.

#### **Characteristics:**

- It supports functional, structured, and OOPS paradigms.
- It may be used as a scripting language or converted to byte-code for the development of huge applications.
- It allows dynamic type verification and provides very high-level dynamic data types. It supports automatic garbage collection.
- It is simple to integrate with ActiveX, C, C++, COM, CORBA and Java.

#### **3.4.3 Keras**

The TensorFlow machine learning framework provides the foundation for the Python-based deep learning API known as Keras. Its creation was intended to facilitate speedy experimentation. When conducting research, it is essential to be able to transition swiftly from a concept to an output.

Keras characteristics:

- **Simple** - Keras reduces developer cognitive load so that you may concentrate on the key aspects of the issue while yet keeping things from being unnecessarily simplistic.
- **Flexible**- Basic actions should be quick and easy, while arbitrarily complicated workflows should be feasible via a clear route that builds on what you've learnt so far. Increased complexity disclosure is a principle that Keras upholds.

- Powerful - Keras offers performance and scalability that dominate the industry: it is utilised by YouTube, and NASA, among others.

### **3.4.4 OpenCV**

It is a free and open-source software library for computer vision and machine learning. OpenCV was created to offer a standard foundation for computer vision applications and to speed up the incorporation of machine perception into commercial goods. OpenCV is an excellent tool for image processing and computer vision tasks. It is an open-source library that may be used for tasks like as face identification, objection tracking, landmark detection, and many more. It supports a variety of programming languages, including Python, Java, and C++.

The OpenCV library is used in the project to resize images and generate feature vectors from them, which is accomplished by transforming the image data to numpy arrays.

### **3.5 Working flow**

The workflow of the age and gender prediction model using CNN is shown in Fig 3.7

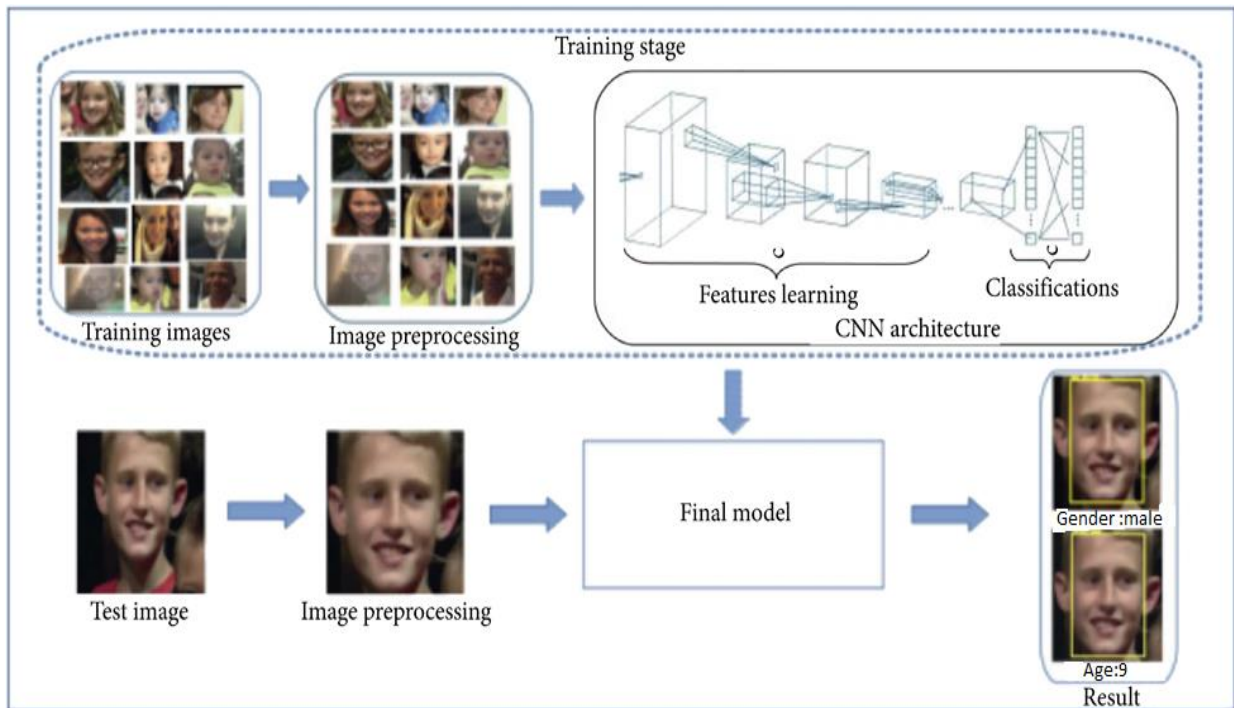


Fig 3.7 Block diagram

## Chapter 4

### Result and Discussion

The developed CNN model uses input of 22000 images of male and female with attribute age, gender, 21,000 images used for training and 1000 images is used for the validation purpose.

#### 4.1 Result for Gender and Age prediction (CNN)

First, the photos are scaled and shrunk to 128 x128. These photos are subsequently fed into the model for training. The output layer is build using sigmoid layer with two classes that represent the "Male" and "Female" classes. Binary cross entropy is used for calculating the loss function, because this is a binary classification problem. The Adam optimizer is used for optimization, with 0.001 as the learning rate.

The photos are scaled and shrunk to 128 by 128 to begin the process of gender identification. The loss function used is categorical cross entropy, because this is a multiclass classification, and the optimizer is Adam that is used in gender identification.

The accuracy obtained using CNN is shown in table II

	<b>Training (accuracy)</b>	<b>Training (loss)</b>	<b>Validation (accuracy)</b>	<b>Validation (loss)</b>
Gender	0.9660	0.0807	0.8891	0.4312
Age	0.447	3.867	0.420	6.7116

Table II. Accuracy score

Several iterations are required to get the model to predict properly. The first step is to create predictions based on the test findings. Iterations are performed to verify the loss and accuracy. As seen in figs. 4.1 and 4.2, accuracy began to improve after the first iteration, while loss decreased.

When the accuracy line is steady, it signifies that there is no need for more iterations to improve the model's accuracy. The output obtained by the model is displayed in Fig 4.4 along with the predicted and actual values

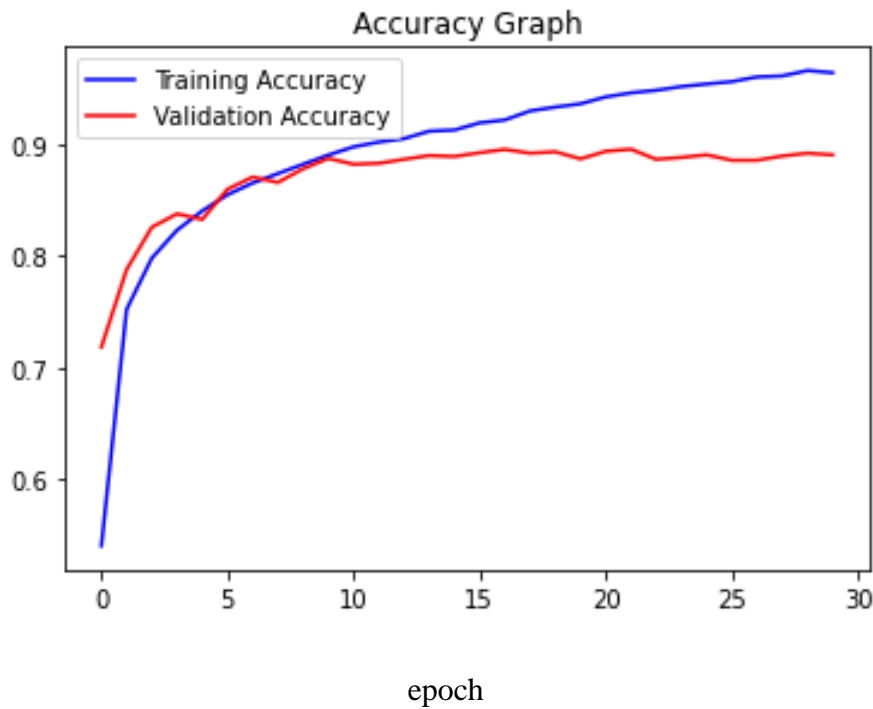


Fig 4.1 Accuracy graph of gender prediction

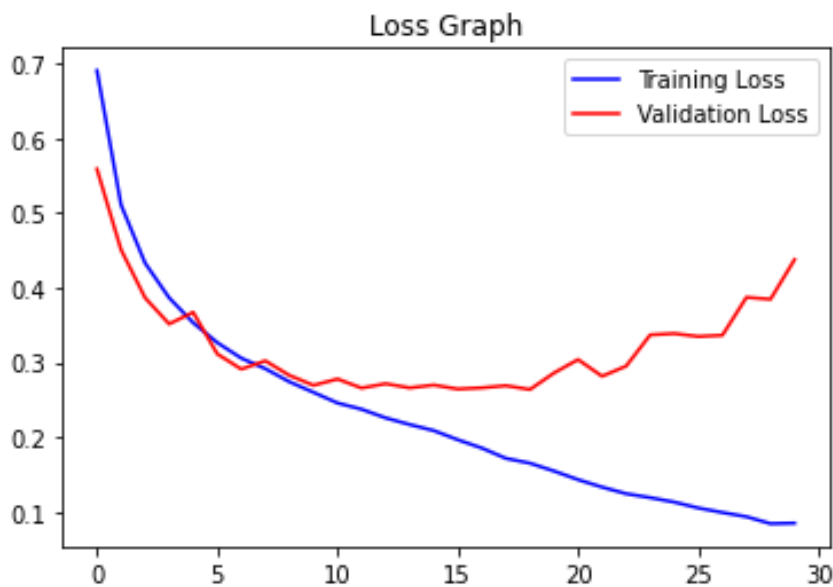


Fig 4.2 Loss graph of gender prediction

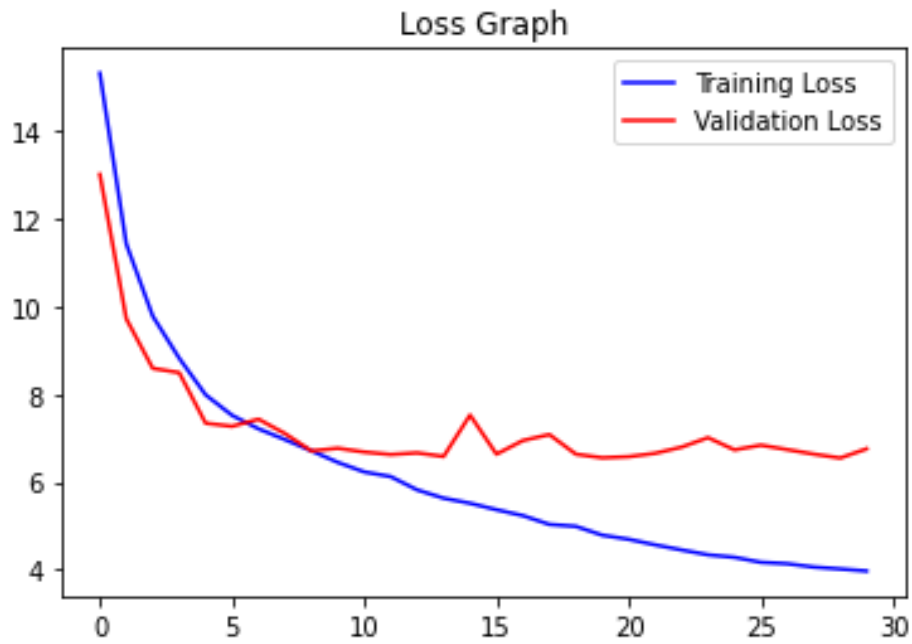


Fig 4.3 Loss graph of age prediction.

Original Gender: Female Original Age: 3  
Predicted Gender: Female Predicted Age: 2



Original Gender: Male Original Age: 40  
Predicted Gender: Male Predicted Age: 44



Original Gender: Female Original Age: 35  
Predicted Gender: Female Predicted Age: 31



Fig 4.4 sample images of output (along with their predicted genders and age given above)

The MAE obtained by the CNN model is 6.71 which is high as shown in table II so another deep learning algorithm called ResNet50 is used to make a comparison between both algorithms and to check which have lower MAE.

### 4.2 Result for Age prediction (ResNet50)

Experiments were carried out on grayscale photos for 10 epochs with a batch size of 64. After each epoch, MAE value was computed, the size of the discrepancy between an observation's real value and its forecast is referred to as the mean absolute error.

Training loss	Training MAE	Validation loss	Validation MAE
4.49	1.5	56.54	5.14

Table III. MAE score

```

begin training using onecycle policy with max lr of 0.0001...
Epoch 1/10
339/339 [=====] - 382s 1s/step - loss: 6.5890 - mae: 1.9414 - val_loss: 60.7641 - val_mae: 5.4747
Epoch 2/10
339/339 [=====] - 377s 1s/step - loss: 7.0695 - mae: 1.9969 - val_loss: 65.1609 - val_mae: 5.7024
Epoch 3/10
339/339 [=====] - 375s 1s/step - loss: 9.8141 - mae: 2.3303 - val_loss: 65.0700 - val_mae: 5.6989
Epoch 4/10
339/339 [=====] - 376s 1s/step - loss: 15.3725 - mae: 2.9043 - val_loss: 82.7730 - val_mae: 6.2997
Epoch 5/10
339/339 [=====] - 376s 1s/step - loss: 25.2150 - mae: 3.7122 - val_loss: 71.1842 - val_mae: 6.2638
Epoch 6/10
339/339 [=====] - 375s 1s/step - loss: 23.8586 - mae: 3.6119 - val_loss: 63.2216 - val_mae: 5.5238
Epoch 7/10
339/339 [=====] - 378s 1s/step - loss: 15.1470 - mae: 2.8718 - val_loss: 61.6565 - val_mae: 5.3570
Epoch 8/10
339/339 [=====] - 378s 1s/step - loss: 8.9929 - mae: 2.2054 - val_loss: 57.5433 - val_mae: 5.1520
Epoch 9/10
339/339 [=====] - 378s 1s/step - loss: 5.9445 - mae: 1.8149 - val_loss: 57.3098 - val_mae: 5.1390
Epoch 10/10
339/339 [=====] - 377s 1s/step - loss: 4.4921 - mae: 1.5685 - val_loss: 56.5438 - val_mae: 5.1437

```

Fig 4.5 Model Evaluation

For each of the training and validation operations, the MAE value reduces as the number of epochs of the model increases, as shown in Fig 4.5. This indicates that the model was adequately trained and verified.

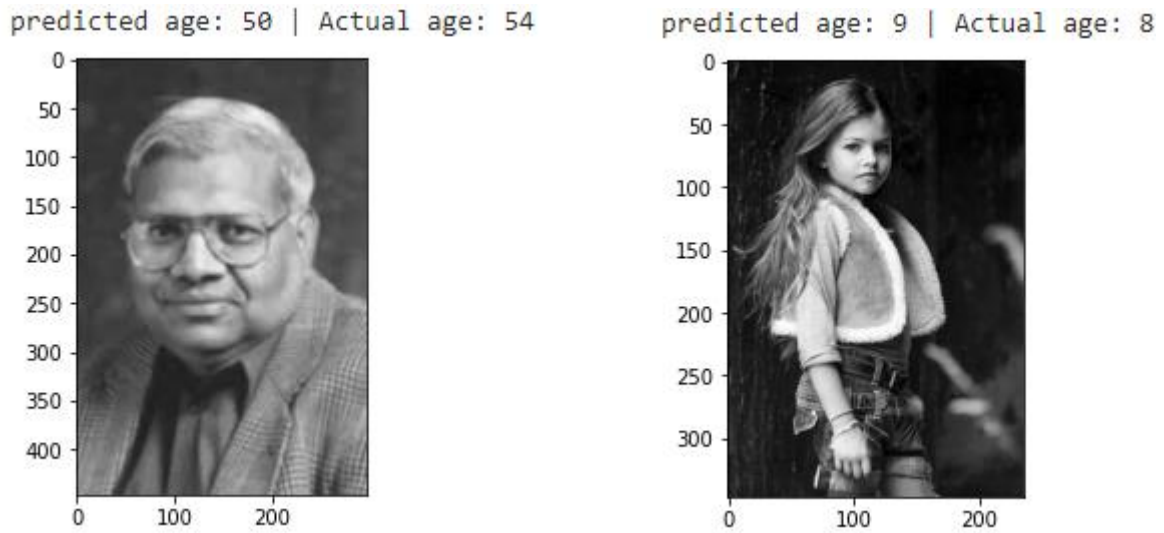


Fig 4.6 sample images of output (with their Actual age and predicted )

Hence, the MAE score for Resnet50 model is 5.14 which lesser as compared to the CNN model.

## Chapter 5

### Conclusion

Prediction of gender and age of humans we meet in our daily lives is also crucial in our social lives. We humans can determine gender and age based just on their facial appearance. Applications like marketing intelligence visual surveillance, and medical diagnosis now require persons to be evaluated using face photos. This project used a deep learning algorithm to propose a simple CNNs model in gender classification and age prediction, trained using widely used datasets called UTKFACE, then applied an effective deep learning algorithm to obtain successful and reliable outcomes in accuracy of prediction. The most important discovery is that research using gender and age from facial recognition are incredibly popular and may be used in social media and advertising to build a sophisticated system that can generate accurate recognition results.

A reliable mechanism for determining gender and age is implemented. The CNN model's accuracy is 88%, and its MAE score is 6.71. ResNet50 model's MAE score is 5.14. The resulting findings are exact and transparent with accurate gender predictions; it is legitimate to categorise a person as either male or female. Due to the usage of bias, the algorithm displays accurate results when age is taken into account.

Github Link: <https://github.com/fathimanovrin/GENDER-AND-AGE-PREDICTION>

## **5.1 Future Enhancement**

The suggested model can yet be improved in the near future in a few specific areas. For instance, focus more on determining the gender of children because at this young age, both male and female people share a lot of visual traits. Additionally, this model has to be smarter when determining the age and gender of subjects in photographs that are obscured, partially visible, wearing hats, glasses, wigs, or other strange make-up, etc. Even the face representations of many regions of the world vary. We must thus make our model more capable of handling this.

## References

- [1] A. Garain, B. Ray, P. K. Singh, A. Ahmadian, N. Senu and R. Sarkar, "GRA\_Net: A Deep Learning Model for Classification of Age and Gender From Facial Images," in *IEEE Access*, vol. 9, pp. 85672-85689, 2021, doi: 10.1109/ACCESS.2021.3085971.
- [2] A. Salihbasić and T. Orehováckí, "Development of android application for gender, age and face recognition using opencv," in 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE, 2019, pp. 1635–1640.
- [3] M. K. Benkaddour, S. Lahlali and M. Trabelsi, "Human Age And Gender Classification using Convolutional Neural Network," *2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH)*, 2021, pp. 215-220, doi: 10.1109/IHSH51661.2021.9378708.
- [4] Mitigating Bias in Gender, Age and Ethnicity Classification: a Multi-Task Convolution Neural Network Approach, Springer professional, springer, 2018
- [5] A. Garain, B. Ray, P. K. Singh, A. Ahmadian, N. Senu and R. Sarkar, "GRA\_Net: A Deep Learning Model for Classification of Age and Gender From Facial Images," in *IEEE Access*, vol. 9, pp. 85672-85689, 2021, doi: 10.1109/ACCESS.2021.3085971.
- [6] K. Zeeshan, H. Kaleem, R. Malik, and S. Khalid, "Deepgender : realtime gender classification using deep learning for smartphones," *Journal of Real-Time Image Processing*, 2017
- [7] Z. Liao, S. Petridis, and M. Pantic, "Local deep neural networks for age and gender classification," arXiv preprint arXiv:1703.08497, 2017.
- [8] R. Rothe, R. Timofte, and L. V. Gool, "Deep expectation of real and apparent age from a single image without facial landmarks," *International Journal of Computer Vision*, vol. 126, no. 2-4, pp. 144–157, 2018.
- [9] M. Duan, K. Li, C. Yang, and K. Li, "A hybrid deep learning cnn–elm for age and gender classification," *Neurocomputing*, vol. 275, pp. 448–461, 2018.
- [10] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 6848–6856.

- [11] J.-H. Lee, Y.-M. Chan, T.-Y. Chen, and C.-S. Chen, "Joint estimation of age and gender from unconstrained face images using lightweight multi-task cnn for mobile applications," in 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, 2018, pp. 162–165
- [12] A. Rattani, N. Reddy, and R. Derakhshani, "Convolutional neural networks for gender prediction from smartphone-based ocular images," IET Biometrics, vol. 7, no. 5, pp. 423–430, 2018.
- [13] A. Greco, A. Saggese, M. Vento, and V. Vigilante, "A convolutional neural network for gender recognition optimizing the accuracy/speed tradeoff," IEEE Access, vol. 8, pp. 130 771–130 781, 2020.
- [14] H.-T. Q. Bao and S.-T. Chung, "A light-weight gender/age estimation model based on multi-taking deep learning for an embedded system," in Proceedings of the Korea Information Processing Society Conference. Korea Information Processing Society, 2020, pp. 483–486.
- [15] X. Liu, Y. Zou, H. Kuang, and X. Ma, "Face image age estimation based on data augmentation and lightweight convolutional neural network," Symmetry, vol. 12, no. 1, p. 146, 2020.
- [16] S. Manasa, J. S. Abraham, A. Sharma, and K. Himapoornashree, "Age, gender and emotion detection using cnn," International Journal of Advanced Research in Computer Science, vol. 11, no. Special Issue 1, p. 68, 2020
- [17] O. Agbo-Ajala and S. Viriri, "Deeply learned classifiers for age and gender predictions of unfiltered faces," The Scientific World Journal, vol. 2020, 2020.