AGE AND GENDER DETECTION USING CONVOLUTIONAL NEURAL NETWORK

A PROJECT REPORT

Submitted by

HEMANATHAN C 190701067

HARISH RAJAA M 190701063

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE,

ANNA UNIVERSITY : CHENNAI 600 025

JUNE 2022

BONAFIDE CERTIFICATE

Certified that this project report titled "**AGE AND GENDER DETECTION USING CONVOLUTIONAL NEURAL NETWORK**" is the bonafide work of "**HEMANATHAN C(190701067)**, **HARISH RAJAA M(190701063)**" who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

SUPERVISOR

Department of Computer and Engineering

Rajalakshmi Engineering College , Chennai - 602 105

Submitted to Project Viva-Voce Examination held on ______________________

Internal Examiner **External Examiner**

ABSTRACT

Nowadays research has explored to extracting auxiliary information from various biometric techniques such as fingerprints, face, iris, palm, voice etc. This information contains some features like gender, age, beard, moustache, scars, height, hair, skin color, glasses, weight, facial marks, tattoos etc. All this information contributes more and more during identification. The major changes that come across face recognition is to find age & gender of the person. This paper contributes a significant survey of various face recognition techniques for finding the age and gender. The existing techniques are discussed based on their performances. This paper also provides future directions for further research. Age and gender, two of the key facial attributes, play a very foundational role in social interactions, making age and gender estimation from a single face image an important task in intelligent applications, such as access control, human-computer interaction, law enforcement, marketing intelligence and visual surveillance. Age and gender recognition start primarily with face detection. The detection is a technique in which various factors are recognized based on the input and according to the requirement. Our system consists of several deep convolutional neural networks that are trained to recognize the face from an input image and predict its age and gender. The deep convolutional neural networks are not only inexpensive but also provide good results.

ACKNOWLEDGEMENT

First, we thank the almighty God for the successful completion of the project. Our sincere thanks to our chairman **Mr. S. Meganathan, B.E., F.I.E.,** for his sincere endeavor in educating us in his premier institution. We would like to express our deep gratitude to our beloved Chairperson **Dr. Thangam Meganathan, Ph.D.,** for her enthusiastic motivation which inspired us a lot in completing this project, and Vice-Chairman **Mr. Abhay Shankar Meganthan**, **B.E., M.S.,** for providing us with the requisite infrastructure. We also express our sincere gratitude to our college principal, **Dr.S.N.Murugesan M.E., PhD.**, and **Dr. P. Revathy M.E., Ph.D., Head of the Department of Computer Science and Engineering,** and our project guide **Ms. M. Diviya M.E.,** for her encouragement and guiding us throughout the project. We would like to thank our project coordinator, **Dr. Preethi Ph.D.,** for her encouragement and guidance towards the successful completion of this project and to our parents, friends, all faculty members, and supporting staff for their direct and indirect involvement in the successful completion of the project for their encouragement and support.

HEMANATHAN C

HARISH RAJAA M

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Over the past several years, the quantity of data mainly images that are getting uploaded to the Internet has grown at a large rate. This recent growth in the data has encouraged computer scientists to solve problems in computer vision. This has many applications. You may have seen recent applications of this. A sentiment analysis system that analyzes human sentiments based on the facial feature and there are several other applications like movie recommendation system based on the human age. Until now only the analysis of images like, how many faces are in the picture and where the faces are located are carried out. But now the research on the characteristics which the face possesses is ongoing and has a broader scope. The goal of this project does exactly that by attempting to detect the age and gender of the faces in an image. Researches on this technology are ongoing and have a very broad scope. Its applications have high potential which can make an impacton society. Detecting age and gender from an image is a challenging problem than many other tasks in the field of computer vision. For predicting anything we need to train our model with sample data. The difficulty lies in the nature of the sample data that is needed to train these types of systems. In this era of internet forgeneral object classification tasks we often have access to millions of images. But the data needed for supervised learning should be labeled data. That means the images should be labeled with age and gender. Finding this type of data is challenging since they are very small in number when compared to labeled data. For labeling data, the real problem is that we don’t have access to some personal information of the people like their date of birth and, they may not be accurate. So, for our work in detecting age and gender, we are using the IMDb-wiki dataset. The IMDb dataset contains 460,723 facial images with age and gender labels and the Wikipedia dataset includes

62,328 images. So, All-in-all we have got access to over 5 lakh images with labels.

1.1 Image Processing

Image Processing (IP) is a computer technology applied to images that help us to process, analyze and extract useful information. It is among rapidly growing technologies and has evolved widely over the years. Today, several companies and organizations of different sectors use image processing for several applications such as visualization, image information extraction, pattern recognition, classification, segmentation, and many more.

Primarily, there are two methods for image processing: analogue and digital image processing. The analogue IP method is applied to hard copies like scanned photos and printouts, and the outputs here are usually images. In comparison, the Digital IP is used in manipulating digital images by using computers; the results here are traditionally information connected with that image, such as data on features, characteristics, bounding boxes, or masks.

Some familiar use cases that leverage ML image processing techniques:

- **Medical Imaging / Visualization**: Help medical professionals interpret medical imaging and diagnose anomalies faster.
- **Law Enforcement & Security**: Aid in surveillance & biometric authentication.
- **Self-Driving Technology**: Assist in detecting objects and mimicking human visual cues & interactions.
- **Gaming**: Improving augmented reality and virtual reality gaming experiences.
- **Image Restoration & Sharpening**: Improve the quality of images or add popular filters etc.
- **Pattern Recognition:** Classify and recognize objects/patterns in images and understand contextual information.
- **Image Retrieval**: Recognize images for faster retrieval from large

datasets.

1.2 Transfer Learning

Transfer learning is the process of reusing a pre-trained model on a problem being worked on. It has recently gained a lot of attention in the data science world and is becoming very popular in deep learning day by day because it can train deep neural networks with comparatively lesser amount of data. This is very useful in data science and several other analysis fields as well since most real-world problems nowadays typically do not have millions of labelled data points to train such intricate models. In transfer learning, the system/machine utilizes the knowledge and features gained from a previously trained model to improve performance and analysis of another. For example, training a classifier model to predict whether an image contains animals, you could use this knowledge to recognize the kind of an animal it is, like, herbivores, carnivores, or omnivores. The knowledge attained from a pre-trained model is applied to a different problem that is related to it in at least a few facets. In accordance, the transfer of weights takes place from one task to a new task based on what the machine had learned. Instead of starting the entire learning process from scratch, which is extremely time and power consuming, we start with patterns learned from a previous problem that solved a related task.Transfer learning is largely employed in computer vision and natural language processing problems due to the compulsory requirement of huge amount of computational power. It becomes quite a powerhouse when combined with neural networks, that require data and computational power in greatly huge amounts. In computer vision related tasks, neural networks try to detect edges in the earlier or input layers, unique shapes in the middle layer or layers and some problem-based features in the latter layers. In transfer learning only the latter layers are retrained, keeping the input and middle layers unchanged. The main advantages of employing transfer learning are as follows: saving a lot of training and validating time, works even for small amounts of data, and brings out better performance and results in neural networks (majority of the cases).

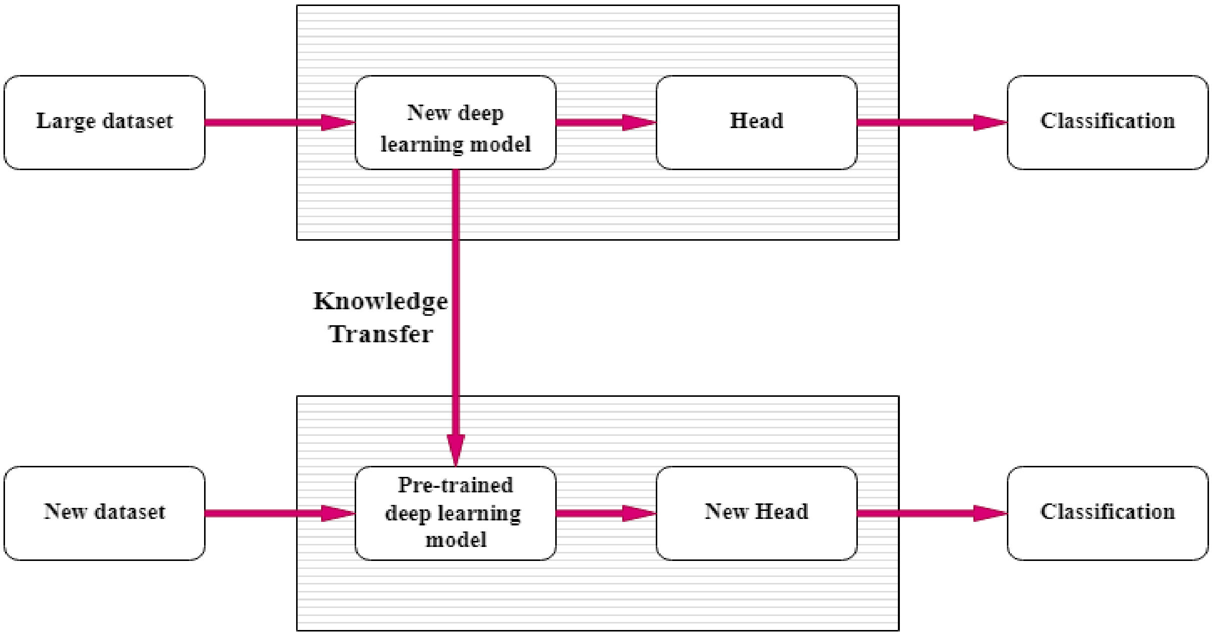

Figure 1.2: - Transfer Learning Methodology

1.3 Convolutional Neural Network (CNN)

Neural networks are a subset of machine learning that are inspired by the human brain, mimicking the way that biological neurons communicate to one another. Artificial neural networks are comprised of node layers, containing an input, one or more hidden and an output layers. Each artificial neuron syncs to another and also has an associated weight and threshold. If the output of any individual neuron is above the restricted threshold value, then that node is activated, thus sending data to the next layer of the network. If not, no data is passed along to the consequent or subsequent layers of the network.

Convolutional neural network is a subset of deep learning that has recently been drawing attention in works related to radiology and has become quite influential in many domains, like computer vision. Convolutional neural network is composed of layers, such as convolution layers, pooling layers, and fully connected layers, and is designed to spontaneously and adaptively learn spatial hierarchies of features through a backpropagation algorithm. CNNs are usually employed for classification and analysis tasks. Before CNNs were developed, several time-consuming feature extraction methods were used to identify features and objects in images. However, CNNs now provide a more feasible approach to image classification, making use of principles from linear algebra, to identify patterns within an image.

1.3.1 CNN Architecture

There are three types of layers that make up the CNN which are the convolutional layers, pooling layers, and fully connected (FC) layers. When these layers are stacked, a CNN architecture will be formed. In addition to these three layers, there is a dropout layer too.

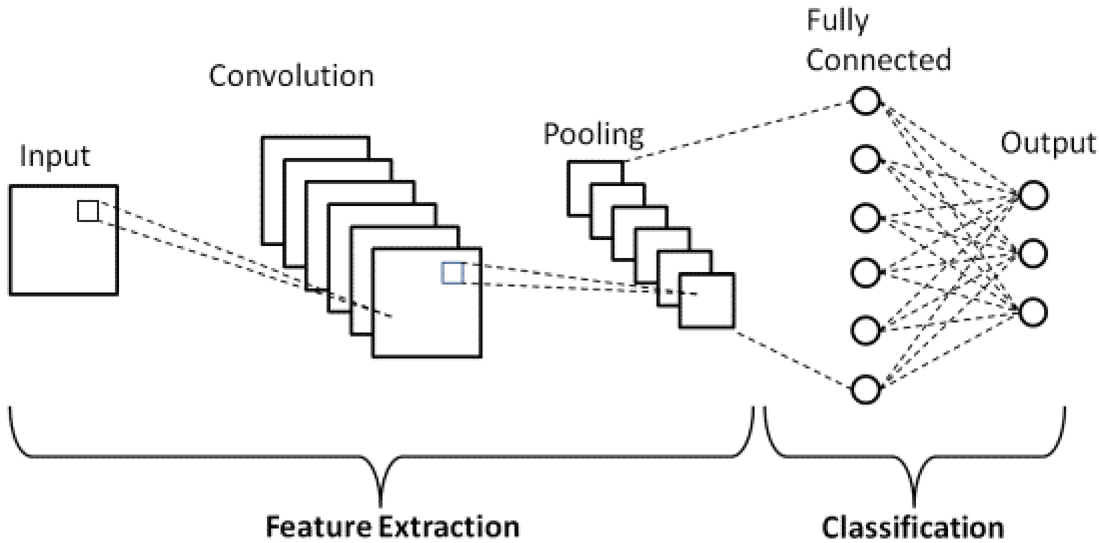

Figure 1.3: - CNN Architecture

1.3.1.1 Convolutional Layer

This layer is the first layer that is used to extract the various features from the input images. In this layer, the mathematical operation of convolution is performed between the input image and a filter of a particular size MxM. By sliding the filter over the input image, the dot product is taken between the filter and the parts of the input image with respect to the size of the filter (MxM).

The output is termed as the Feature map which gives us information about the image such as the corners and edges. Later, this feature map is fed to other layers to learn several other features of the input image.

1.3.1.2 Pooling Layer

In most cases, a Convolutional Layer is followed by a Pooling Layer. The primary aim of this layer is to decrease the size of the convolved feature map to reduce the computational costs. This is performed by decreasing the connections between layers and independently operates on each feature map. Depending upon method used, there are several types of Pooling operations. In Max Pooling, the largest element is taken from feature map. Average Pooling calculates the average of the elements in a predefined sized Image section. The total sum of the elements in the predefined section is computed in Sum Pooling. The Pooling Layer usually serves as a bridge between the Convolutional Layer and the FC Layer

1.3.1.3 Fully Connected Layer

The Fully Connected (FC) layer consists of the weights and biases along with the neurons and is used to connect the neurons between two different layers. These layers are usually placed before the output layer and form the last few layers of a CNN Architecture.

In this, the input image from the previous layers are flattened and fed to the FC layer. The flattened vector then undergoes few more FC layers where the mathematical functions operations usually take place. In this stage, the classification process begins to take place.

1.3.1.4 Dropout Layer

Usually, when all the features are connected to the FC layer, it can cause overfitting in the training dataset. Overfitting occurs when a particular model works so well on the training data causing a negative impact in the model's performance when used on a new data.

To overcome this problem, a dropout layer is utilised wherein a few neurons are dropped from the neural network during training process resulting in reduced size of the model. On passing a dropout of 0.2, 20% of the nodes are dropped out randomly from the neural network.

1.3.2 Activation Functions

Finally, one of the most important parameters of the CNN model is the activation function. They are used to learn and approximate any kind of continuous and complex relationship between variables of the network. In simple words, it decides which information of the model should fire in the forward direction and which ones should not at the end of the network.

It adds non-linearity to the network. There are several commonly used activation functions such as the ReLU, Softmax, tanH, and the Sigmoid functions. Each of these functions has a specific usage. For a binary classification CNN model, sigmoid and softmax functions are preferred and for multi-class classification, generally softmax is used.

1.3.2.1 ReLU

The rectified linear activation function or ReLU is a non-linear function or piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. It is the most commonly used activation function in neural networks, especially in Convolutional Neural Networks (CNNs) & Multilayer perceptrons (A perceptron is a neural network unit that does certain computations to detect features or business intelligence in the input data). Mathematically, it is expressed as: $f(x)=max(0,x)$

1.3.2.2 Sigmoid

It takes a real-valued number and "squashes" it into a range between 0 and 1. However, a very undesirable property of a sigmoid is that when the activation is at either tail, the gradient becomes almost zero. If the local gradient becomes very small, then in backpropagation it will effectively "kill" the gradient. Also, if the data coming into the neuron is always positive, then the output of the sigmoid will be either all positives or all negatives, resulting in a zig-zag dynamic of gradient updates for weight. Mathematically, it is expressed as: σ(κ) = 1/(1+e¯κ).

1.3.2.3 Softmax

The softmax activation function is the generalized form of the sigmoid function for multiple dimensions. It is the mathematical function that converts the vector of numbers into the vector of the probabilities. The softmax activation function is commonly used as an activation function in the case of multi-class classification problems in machine learning. The output of the softmax is interpreted as the probability of getting each class.

The mathematical expression of softmax activation function is,

$$softmax(Z_i) = \frac{\exp\,(Z_i)}{\sum \exp\,(Z_i)}$$

1.4 Datasets

Datasets that we are using are downloaded from www. data.gov.in and www.Kaggle.com , but these data samples are not used directly as input to the machine learning algorithms. This dataset contains 23705 tuples with 5 attributes (age , ethiniciy , gender , image , pixels). Dataset comprises age, gender, images, and pixels in .csv format. Age and gender detection according to the images have been researched for a long time. Different methodologies have been assumed control over the years to handle this issue. Presently we start with the assignment of recognizing age and gender utilizing the Python programming language.

Keras is the interface for the TensorFlow library. Use Keras on the off chance that you need a profound learning library that allows simple and quick prototyping (through ease of use, seclusion, and extensibility). Support both convolutional networks and repetitive organizations, just as blends of the two. Run flawlessly on CPU and GPU.

Table 1 shown the snapshot of one of the files present in our datasets.

1	age	ethnicity	gender	img_name	pixels
2	1	2	0	201612192	129 128 128 126 127 130 133 135 139 142 145 149 147 145 146 147 148 149 149 150 153 153 153 152 153 153 153 153 151 149 147 146 146 144 143 140 134 129 129 126 122 121 119 118 118 99 53 130
3	1	2	0	201612192	164 74 111 168 169 171 175 182 184 188 193 199 200 199 200 196 198 192 193 188 187 186 187 188 183 182 178 177 175 174 176 174 172 165 158 153 147 142 138 131 125 120 114 110 110 111 111 104 17
4	1	2	0	201612192	67 70 71 70 69 67 70 79 90 103 116 132 145 155 161 166 169 175 177 178 179 180 183 186 187 188 192 194 198 203 206 213 214 216 220 219 215 213 211 211 210 207 206 203 200 197 194 192 65 66 69 70
5	1	2	0	201612201	193 197 198 200 199 200 202 203 204 205 208 211 211 211 209 205 204 204 204 205 205 207 207 207 207 208 208 208 208 207 207 205 204 203 202 202 202 205 209 214 221 226 227 224 223 234 243 250 1
6	1	2	0	201612201	202 205 209 210 209 209 210 211 212 214 218 219 220 221 221 220 221 221 221 220 220 222 222 222 220 220 218 218 218 217 217 218 218 219 219 218 220 222 219 223 225 221 210 169 117 154 214 227 2
7	1	2	0	201612201	195 198 200 200 198 198 199 199 198 197 197 198 203 204 203 199 197 194 193 193 194 193 193 193 194 194 194 193 193 195 196 200 202 200 199 197 197 197 197 199 206 214 218 219 212 205 204 215 1
8	1	2	0	201612201	208 216 217 219 222 223 222 221 220 220 221 221 220 220 220 220 220 220 220 218 217 217 216 216 216 214 214 212 213 212 209 208 207 207 207 207 206 207 208 208 208 205 200 198 194 194 186 2
9	1	2	0	201701091	99 142 169 177 179 181 183 186 187 186 191 190 193 196 197 200 198 198 199 201 202 202 199 202 202 201 202 204 203 201 198 198 196 194 193 192 192 191 191 190 193 189 185 182 179 179 159 116 10
10	1	2	0	201612192	127 127 133 140 143 148 152 157 160 165 172 178 182 187 189 191 199 206 206 202 195 194 202 196 191 190 188 184 181 181 180 181 182 180 180 182 181 180 177 173 168 162 152 145 143 145 139 66 12
11	1	2	0	201701091	199 211 211 214 216 216 219 221 222 224 219 214 210 202 194 192 182 172 171 171 165 167 161 155 155 152 151 148 149 146 142 136 136 132 128 126 126 128 129 126 127 127 132 130 127 124 123 123 2
12	1	2	0	201701091	136 138 145 143 152 160 162 169 178 182 192 198 203 208 211 214 218 218 220 224 225 225 226 230 232 234 235 236 237 237 237 237 234 234 234 233 231 228 227 224 222 219 218 215 213 209 203 195 1
13	1	2	0	201701091	253 253 253 253 252 251 250 209 228 250 221 194 180 213 184 151 171 176 179 182 190 189 186 183 182 182 178 178 179 178 177 174 170 160 147 156 137 136 139 133 132 185 237 237 236 176 136 134 2
14	1	2	0	201701091	223 222 226 227 229 228 228 226 222 214 199 157 134 147 157 158 159 156 156 156 156 156 157 157 153 147 143 136 132 132 163 200 209 206 213 212 211 211 214 210 205 199 191 192 212 199 191 215 2
15	1	2	0	201701102	181 185 186 185 182 180 180 177 183 177 175 173 172 170 168 174 175 175 173 173 176 181 182 182 183 187 189 189 190 189 190 190 191 190 189 186 184 181 185 191 200 200 205 208 207 208 204 196 1
16	1	2	0	201701102	167 172 175 174 172 170 172 173 171 172 172 168 171 174 177 179 179 180 180 180 181 178 178 181 180 178 174 174 174 174 174 174 174 175 175 175 176 176 176 178 177 178 175 156 140 119 156 251 1
17	1	2	0	201701102	99 122 138 147 148 149 149 150 153 157 162 166 171 178 184 189 193 196 198 199 201 207 213 214 217 221 225 226 228 228 231 231 232 230 229 229 228 227 226 227 225 223 224 219 212 212 206 188 10
18	1	2	0	201701102	49 63 84 95 105 108 112 115 122 124 132 135 134 130 120 124 116 113 111 113 106 111 116 111 112 111 109 108 112 111 109 105 103 96 95 91 83 77 77 79 76 71 56 46 38 21 5 3 49 70 83 98 106 109 118 12
19	1	2	0	201701091	226 225 223 166 136 167 170 173 176 177 180 182 182 184 184 183 182 180 178 178 178 177 175 174 170 169 169 170 167 168 166 163 161 159 156 152 150 144 139 115 166 217 218 216 212 208 209 206 2
20	1	2	0	201701161	238 187 173 156 140 132 124 239 157 54 106 172 193 196 203 199 204 193 189 188 180 174 168 156 155 148 141 133 131 126 123 121 120 119 118 114 107 100 90 80 66 56 34 15 77 95 90 114 212 186 165 1
21	1	2	0	201612192	151 155 159 171 179 193 198 204 208 208 210 211 213 216 217 220 220 225 227 230 226 223 222 220 217 216 214 212 210 213 213 212 209 209 208 206 198 196 191 180 164 156 133 78 66 90 92 89 154 152
22	1	2	0	201612192	43 84 97 105 101 98 96 105 108 109 115 116 120 121 125 127 133 134 135 137 136 135 139 140 147 147 144 144 146 150 148 136 131 125 121 121 118 113 111 109 106 104 102 102 98 99 96 88 46 83 91 102
23	1	2	0	201612192	217 219 46 27 47 67 106 135 140 144 149 151 152 153 153 154 157 163 171 177 183 190 194 191 188 190 195 198 201 202 205 209 213 219 222 223 219 221 214 204 185 151 129 103 95 208 223 223 218 21
24	1	2	0	201612192	56 20 45 67 90 99 111 129 147 154 162 171 177 180 183 186 187 189 189 189 189 188 188 188 188 187 186 184 183 183 181 181 177 174 169 166 164 162 160 159 157 154 150 147 144 147 136 119 44 23 47
25	1	2	0	201612192	39 60 62 53 65 81 88 99 109 114 119 122 124 127 132 136 139 143 150 152 157 157 160 157 160 167 169 168 168 167 161 156 156 155 148 145 146 142 139 140 134 130 126 122 118 113 107 102 40 53 58 51
26	1	2	0	201612192	184 183 189 186 185 190 197 202 193 185 173 162 159 156 154 154 156 159 161 165 168 169 173 176 179 180 180 181 184 186 185 184 183 172 157 123 112 161 123 201 202 205 194 145 66 53 62 73 184 18
27	1	2	0	201612192	161 162 161 163 167 173 177 181 184 186 191 195 198 199 200 202 203 205 207 207 210 212 215 217 217 218 218 220 221 222 224 225 225 224 224 219 215 213 204 199 190 161 107 112 160 202 219 227 1
28	1	2	0	201612192	193 201 197 190 188 193 198 197 197 196 197 196 196 197 196 198 199 202 203 206 207 208 210 210 210 208 210 211 213 212 209 214 214 215 217 221 222 223 225 226 227 225 225 225 228 235 237 233 2
29	1	2	0	201612192	175 210 221 227 231 233 234 237 238 238 234 233 230 228 228 227 226 225 225 226 226 226 226 227 227 227 226 225 225 225 225 225 224 222 220 218 216 213 210 209 207 201 203 199 191 167 116 73 18
30	1	2	0	201612192	163 160 150 139 83 119 143 152 163 170 172 175 181 183 188 192 191 192 192 193 193 192 180 176 179 180 177 174 173 174 175 174 168 157 145 134 135 121 136 138 125 103 107 120 127 130 120 102 16

age_gender

1.5 Objectives

CHAPTER 2

LITERATURE REVIEW

Yunjo Lee, et.al proposed that the fMRI method is used to study upon age detection methods. The study involves a proper recording of the variations of people on the basis of their changes according to age, gender, identity and other features. The brain activation tasks related to face matching are performed and tested outside the scanner. There was a same result in face processing in older as well as young adults. The performance results high in both the cases having same facial viewpoints. The aging of the elders is not based on any one factor. It is combination of various factors that result in accountancy of such results. The results need to be kept a track on which are based on all credentials kept in certain environments.

R. Begg et.al explained the automatic recognition of walking changes because of aging through the artificial neural networks is the aim of the article. The balance control of the locomotors system is disturbed due to the gait factors which are caused through walking patterns which change according to the age. There are many advantages of such techniques. The standard back propagation, scaled conjugate gradient and the back propagation with Bayesian regularization were the three methods involved. The three networks came out with better results but the Bayesian regularization method was the one with greatest results in some fields. The neural networks thus are a great help for the age identification purposes.

Hang Qi et.al, proposed that various techniques have been arising for the detection of faces which can also identify the age of the person. Here, an automated system has been proposed which can classify the age and help distinguishing kids face from that of an adults face. There are three parts that the system encompasses. They are face detection, face alignment and normalization, and age classification. Face samples are created by the normal face detection and alignment methods. ICA is used for the extraction of the local facial components that are present in the images. This system has been proved to be much faster and the results are efficient. So this system can be used in future as a prototype.

Kensuke Mitsukura, et.al that on the basis of the color information the threshold value in multi-value images is considered. There is a lack of versatility when there is no change in the threshold of an image. Whenever there is an influence of any light conditions, the information of the color varies.It becomes prominent to decide the face. It is difficult to determine the face division standard. This is done for providing information to the Genetic Algorithm used in the method. Also a face decision method is proposed further which determines whether it is a decision method face or not. The identification of an individual is also very important. There is a use of the color maps for the differentiation of the detected faces. The features that are missed result in false identifications as well as the poor results.

Chao Yin et.al, the Conditional Probability Neural Network (CPNN) is a distribution learning algorithm used for the age estimation using facial expressions. It follows the three-layer neural network system in which the target values and the conditional feature vectors are used as an input. This can help it in learning the real ages. The relationship between the face image and the related label distribution through the neural network is used as the learning method for this system. The earlier method used proposed that the relationship is to be used according to the maximum entropy model. CPNN has proved to be providing better results than all the previously made methods. Through this method the results provided were very easy, there was less computational involved and the outcomes very efficient. Due to all such advantages it was preferred more than the others.

Sarah N. Kohail et.al proposed that the age estimation is now the current challenge being faced. Here, the article puts forward the approach of neural networks to estimate the age of humans. The main change that has been made in this method is the fine tuning of the age ranges. To learn the multi-layer perception neural networks (MLP) the facial features of the new images were extracted and recorded. The inputs were provided to the layer . The results have shown the MLP method as a good method with minimum errors in the results. These results can be used in many of the applications like age-based access control applications and also in the age adaptive human machine interaction.

Thakshila R.kalansuriya and Anuuja T. Dharmaratne et al. [2] proposed a age gender detection system Using artificial neural network which achieved 70.5% accuracy rate. where FERET and FGNET datasets are used.

M.R Dileepa and Ajit Dantib et al.[3] proposed a age gender prediction system using Neural network and sigma control limit which achieved 95% accuracy rate.

Sepidehsadat et al. [4] suggested that the utilization of Gabor filter will make it simpler for the network to focus on the face , because the output direction is perfectly matched with facial wrinkles, and wrinkles will become the input of CNN. The network focuses on providing useful features with 7% age accuracy and 2% gender accuracy.

Ramin Azarmehr , Robert Laganiere , Won-Sook Lee et al.[9] proposed a system using EDA achieved 99% and for better accuracy and performance use support vector machine (SVM) and demographic classification strategies.

Jang-Hee Yoo, So-Hee Park and Yongjin Lee et.al[10] proposed a age gender detection system achieved 72.53% for age detection and 98.90% for gender detection.

Octavio Arriagal and Matian Valdenegro –Toro and paul G Pl oger et al. [11] propose a Real –time emotion gender classification system using CNN which achieved 66% accuracy on emotion classification on FER-20133 dataset and with IMDB dataset achieved 95% for gender classification.

2.1 Existing System

Over the past several years, the studies are undergoing on the areas of facial feature extraction and predicting age and gender based on that. To address this concern scientists had come up with various approaches, and each of these approaches solved some critical problems that were raised in this field. For predicting age and gender accurately, even some of the minor differences in the images should be extracted carefully. Extracting facial features to that extent is challenging and only a few approaches focused on solving this problem. Those minor differences include size of the eyes, ears, mouth and the distances between them. most of early methods have focused on images which were maintained in lab conditions (like maintaining ideal lighting, angle of the image, etc.). very few methods have addressed the difficulties that arise when applying these methods to real world images.

2.1.1 Advantages of the existing system

1) Very high accuracy.

2) Prior to this, researchers were not clear about how transfer learning on a set images can be applicable to detect age and gender , because of differences in image size, domain differences, and the use of

local variations in texture to detect.

2.2 Proposed System

In this proposed system, we will develop a deep learning model using Convolutional neural networks. The model receives the image, passes it through different layers by reducing the size in each layer. We then train our model using the IMDb-wiki dataset which contains over 5 lakh images of human beings which include different ethnicity, color, and many more factors. The dataset also provides a label for each image. Once the model is trained, we can use the model for testing. We first detect the presence of a human in each image. Then, we process the image using the CNN to obtain all the human faces present in the image. Each face is then processed by a developed deep learning model to get the output label which essentially gives us the age and gender of each person in the image.

2.2.1 Advantages of the proposed system

1) Multi-label classification is performed.
2) Model has an accuracy of more than 89%, which is pretty much high in the case of a very large complex dataset with multiple labels to classify.

CHAPTER 3

SYSTEM DESIGN

3.1 Development Environment

3.1.1 Hardware Requirements

The hardware requirements may serve as the basis for a contract for the training of the models used and should therefore be a complete and consistent specification of the whole system. They are used by machine learning engineers as the starting point for the system design. It shows what the system does and not how it should be implemented.

Table 3.1: - Hardware Requirements

COMPONENT	SPECIFICATION
PROCESSOR	Intel Core i5.
RAM	8 GB DDR4 RAM
GPU	NVIDIA GEFORCE 960
MONITOR	15” COLOR
HARD DISK	10 GB
PROCESSOR SPEED	MINIMUM 500MHZ

3.1.2 Software Requirements

- Python 3.8
- Anaconda
- Jupyter Notebook

CHAPTER 4

PROJECT DESCRIPTION