Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020]
IEEE Xplore Part Number: CFP20M19-ART; ISBN: 978-1-7281-7089-3

Age and Gender Predictions using Artificial Intelligence Algorithm

Anirudh Ghildiyal
Department of Computer Science and Engineering
Graphic Era Deemed to be University
Dehradun, India
a.ghildiyal0506@gmail.com

Sachin Sharma
Department of Computer Science and Engineering
Graphic Era Deemed to be University
Dehradun, India
sachin.cse@geu.ac.in

Ishita Verma
Department of Computer Science and Engineering
Graphic Era Deemed to be University
Dehradun, India

Urvi Marhatta
Department of Computer Science and Engineering
Graphic Era Deemed to be University
Dehradun, India

2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) | 978-1-7281-7089-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICISS49785.2020.9316053

Abstract- Gender is still a central aspect of personality, and in social life it is still an important factor. Gender and age projections for artificial intelligence can be used in many areas, such as the development of smart human-machine interfaces, fitness, cosmetics, e-commerce, etc. The prediction of age and gender is an ongoing and active research question for individuals from their facial images. A number of approaches to solving this issue have been suggested by the researchers, but the criteria and actual performance are still insufficient. This paper proposes a mathematical approach to recognition patterns in order to solve this problem. The Convolution Neural Network (ConvNet / CNN) deep learning algorithm is used as a feature extractor in the proposed solution. CNN takes input images and assigns value to and can distinguish between various aspects / objects (learnable weights and biases) of the image. ConvNet needs much less pre-processing than other classification algorithms. While the filters are hand-made in primitive methods, ConvNet can learn these filters / features with adequate training. In this research, face images of individuals have been trained with convolution neural networks, and age and sex with a high rate of success have been predicted. More than 20,000 images are containing age, gender and ethnicity annotations. The images cover a wide range of poses, facial expression, lighting, occlusion, and resolution.

Keywords- Facial Images; Gender Prediction; Age Prediction; Convolutional Neural Network; Deep Learning.

I. INTRODUCTION

The aim is to predict the age of individuals using image data sets. A growing number of applications, especially after the increase in social networks and social media, are being concerned with automatic age classification. Age and gender are the two most fundamental facial qualities in social interaction. In smart applications, such as access control, human computer interaction, enforcement, marketing intelligence and visual supervision, etc, it is important to make age evaluations using one facial image. Machine learning: supervised learning, image recognition, and deep learning: a ground-breaking neural network and profound learning are the most common technologies used in this paper. Supervised learning can be described as a machine learning technique in which the input is mapped to the output using input-output pair training data. TensorFlow is an open-source library used for math, data flow and specific machine learning applications. Convolutional Neural Network (CNN) is one of the most prevalent algorithms that has gained a high reputation in image feature extraction.

CNN is a Deep Learning algorithm, which takes an image as input and take on different aspects / objects and can be distinguished from one image (learnable weights and biases). With adequate training ConvNet can learn filters/features, that are handmade in primitive methods since ConvNet requires less pre-processing. The working of the ConvNet architecture is similar to that of neurons in the human brain. About Dataset: a broad facial dataset of long age (range 0-116 years old) is a UTKFace data set. The data collection consists of more than 20,000 facial images with age, gender and ethnicity annotations. The images cover a wide range of poses, facial expression, lighting, occlusion, resolution. It can be used for a variety of tasks, for example face detection, age estimation, age progression, position of landmarks etc.

The survey is focused on the age detection of the neural network (CNN) image dataset architecture. The problem may then be treated as a classification concern with 3 convolution layers and 2 completely interconnected layers with a final output layer. Estimating the exact age by regression is a challenging process. Since the important modules that are used in many computer vision applications such as interaction between human and computer, safety system and visual monitoring, age prediction system has been growing rapidly. The value of an age prediction is shown by several examples. For example, there is an age to get alcohol, drive vehicles, travel alone outside the country, smoke cigarettes, etc. The problem is, however, that human capacities are poor and unreliable in age prediction. So, it would be necessary to reject underage individuals with computer vision systems. Hotels, airports, busses, casinos, government buildings, universities, hospitals, movie theatres, etc. are currently using automated age and gender prediction systems for improving protection and mitigating possible threats or poverty. Age prediction methods are also used in healthcare systems, knowledge recovery, academic studies, and Electronic Customer Relationship Management (ECRM) applications, which distributes customers to a range of aged groups including teenagers, teens, adults and senior citizens. Moreover, it may allow businesses to identify products and services according to their age groups that increase income and make more money to collect those customers' daily lives information including behaviours, preferences, practices, 4 priorities etc. For instance, clothing shops that sell men's or women's fashion according to their age groups; restaurants wish to know the most common meals per group of age; many businesses want those audiences to advertise according to their age groups.

II. LITERATURE SURVEY

An architecture for face image classification named unsupervised CNN was introduced by S. U. Rehman et al. [2].

A CNN that handles multitask (i.e. Facial detection and emotional classification) is made by merging CNN with

other modules and algorithms. A hybrid deep CNN and RNN (Recurrent Neural Network) model was introduced by N. Jain et al. [4]. This model aims to improve the overall result of face detection. MI Facial Expression and JAFFE dataset were used to evaluate the model. A convolutional network architecture was proposed by G. Levi et al. [5] that classified the age with small amounts of data. The Audience Benchmark was used to train the model. A system in which a real time automatic facial expression system was designed was proposed by S. Turabzadeh et al. [6].

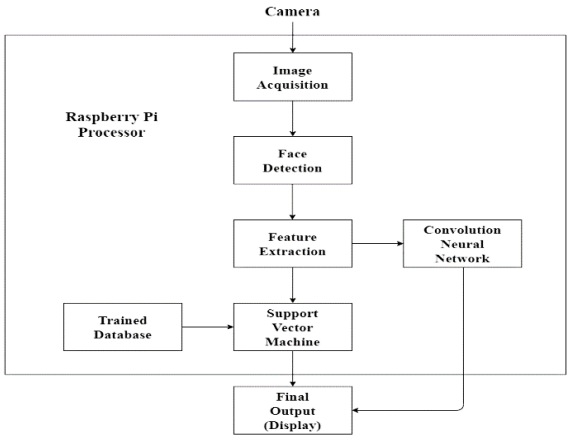

Fig. 1. Age detection using CNN.

The software was introduced and tested on an integrated device, which was the first move for a social robot to recognize a certain facial expression chip. MATLAB was first used for device construction and simulation, then based on an embedded system. The complexity of automatically predicting age, gender and ethnicity on the East Asian population using the CNN was studied by N. Srinivas et al. [3]. A fine-grained ethnicity has predictions based on a refined categorization of the human population (Chinese, Japanese, Korean, etc.). Previous results suggest that the most critical job is to predict the fine-grained ethnicity of a person, followed by age and lastly gender. An automated recognition system for age, gender and emotion was presented by A. Dehghan et al. [7] that was trained using deep neural network. At the ImageNet LSVRC-2010 contest, A. Krizhevskyetal. [8] presented a paper which suggested segregation of 12 lakhs images into 10 times of 100 different categories using a deep convolutional neural network. The results which were obtained suggested that supervised learning can deliver exceptional accuracies. Some datasets have annotations on the face images which are not considered to be of any use for face recognition. Some papers have also used RNN but it is not applicable for our research as the RNN takes text or speech as an input whereas this research work required an image to be as the input. Hence, CNN is chosen over RNN in this paper [9]. Some papers also suggest the use of unsupervised CNN, but, for this paper, proposed work used supervised learning algorithm and UTKFace dataset [10], [11].

III. SOFTWARE REQUIREMENTS AND TECHNOLOGIES

To predict the age, a convolutional neural architecture of the network (CNN) [Fig. 1] is used. This CNN uses 3 layers of convolution and 2 layers with one final output layer.

This problem can be interpreted rather than regression as a classification problem. It is a difficult job to estimate the exact age by means of regression. Simply by looking at the face even people cannot predict age. In an age range, like 20-30 or 30-40, research work seeks to predict the age. It is hard to predict how a person's age depends on many factors from a single image.

ANALYSIS

Age and gender detections are methods of automatically discerning the age and gender of a person solely from a photo of their face. Typically, the age and gender detection implemented in two-stages:

Stage #1: Detection of the faces in the input images and videos.

Stage #2: Apply age detector algorithm on the extracted face from the Region of Interest (ROI).

To detect faces in Stage 1, any face detector algorithm that can draw boundary box around faces in the input image. When the boundary co-ordinates are produced, the object can be moved to next stage, i.e. Stage 2, identifying the age of the person.

MODULES

- NUMPY: NumPy is a mathematical package, used in scientific computing and data manipulation. It is an open source library available in python.
- PANDAS: Pandas library is used for data manipulation and analysis. It supports reading and writing excel spreadsheets and whole lot of manipulation.
- CV2: OpenCV is a high-performance library used in computer vision and Digital Image Processing,
- MATPLOTLIB: Matplotlib is used for plotting graphs and chart in python for various types of data.
- OS: The OS module helps to read, write and alter data in python. It is a way of using the operating system dependent functionality.
- PIL: PIL library is used for image manipulation in python. Using this library, an image from dataset can be opened and resizing it is performed.
- SCIPY: SciPy library contains different modules for statistics, linear algebra, optimization and integration.
- KERAS: Keras is a high-level open source, python-based, neural network API. It makes prototyping simple and fast.
- TENSORFLOW: Tensorflow is a free open-source data flow software library that supports a number of tasks. It is used in neural network applications.

TECHNOLOGIES

Image Processing: To get an enhanced and a clear image or to extract certain features from an image processing operation are done. An input image is taken and it is enhanced or certain features are extracted associated with the image, and the output the according to the requirements. Image processing is a rapidly growing industry since most images requires enhancing or only a certain area of interest is required for further process, making it the core of any research area.

There are three steps in image processing: -

- Importing the image.
- Analyse and manipulate the image according to the requirement.

- The final image is in which result can be altered image or report that is based on image analysis.

For image processing, two techniques are used, digital image processing and analog image treatment. For hard copies, such as pictures and photographs, analog image treatment is usually used. When using these visual methods, image analysts use a set of basics of perception. Digital image processing techniques help to manipulate computer-based digital images. When using digital technologies, the three general phases to which all types of data have to undergo are pre-processing, improvement, display and extraction.

Computer Vision: computer vision is an artificial intelligence field which trains computers to understand and interpret the visual world. With digital images of cameras and videos, machines can recognize and classify objects with accuracy and then respond to what they see.'

IV. SOFTWARE DESIGN

- UML DIAGRAMS

A diagram that represent a system visually, along with main actors, roles, actions, classes etc. to make a better understanding, maintain, alter or document information about the system is called UML diagram.

- CLASS DIAGRAM

Fig. 2 illustrates the class diagram of age prediction. Fig. 3 presents the use case diagram.

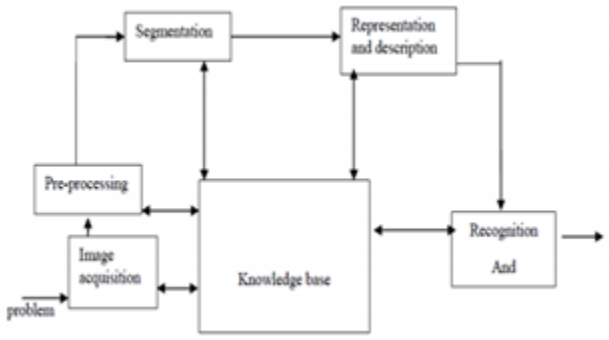

Fig. 2. Class diagram.

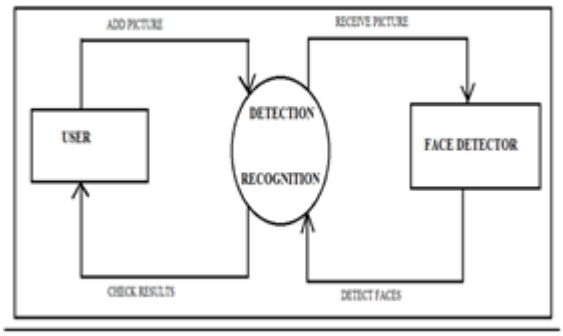

Fig. 3. Use case diagram.

The ideal approach is Age Prediction, since the proposed model expect an actual number as the output. However, it is difficult to estimate age with precise regression. And people cannot determine age on the basis of a person's views precisely. Nonetheless, in their twenties or thirties there may be a chance to predict the age. It is therefore prudent to regard this topic as a classifying topic in which proposed work is used to determine the age group in which the individual is present. For instance, in the 0-2 range is a single grade, 4-6 is a different grade etc. This must be borne in mind that estimating an age with a single image is not easy to solve as it depends on various variables, and in different parts of the world, the people of the same age will look very different.

In addition, before making predictions, use of face alignment are evaluated and found that the predictions had improved for some cases but that they were worse for some people. If non-frontal ears are considered, then it may be a smart idea to use alignment.

V. IMPLEMENTATION AND RESULT ANALYSIS

HARDWARE REQUIREMENTS

- System : Intel i5 2.1 GHZ
- Memory : 4 GB
- Hard Disk : 1TB

SOFTWARE REQUIREMENTS

- Operating System : Windows 10 and above
- Domain : Machine Learning
- Scripts : Python
- Tool : Anaconda Navigator, Jupiter Notebook IDE
- Libraries : Numpy, pandas, math, cv2, matplotlib, seaborn, os, Image, scipy, sklearn, keras and tensorflow.

The following libraries are used in the implementation.

NUMPY: NumPy is a python library, that gives users the ability to deal with single and multidimensional arrays and matrices, along with a large number of high-level operations that can be performed on them. Since an image is also a form of grid, it can be representing in a 2D metric.

PANDAS: Pandas library in python gives users the ability to handles data structures, handling large amount of data in form of excel or tables. It has a collection of operation to alter, manipulate this data.

MATPLOT LIBRARIES: Matplot library is used to visualize data in form of tables, charts, or graph. It is built on NumPy designed to work with the broader SciPy stack.

KERAS: Keras gives users the ability to work on neural networks that has the capability to run on top of TensorFlow, R, Theano or PlaidML. It helps user to implement these neural networks in a fast, reliable, modular and user-friendly way.

TENSORFLOW: TensorFlow is a data flow and differentiable programming Python software library for various tasks. It's a math library for software like neural networks for learning machines.

The following tools are used in the implementation.

Anaconda Navigator: Anaconda is the distribution of languages for programming in Python and R for scientific data processing, for example machine learning, predictive analysis, data science and data processing.

Jupyter Notebook IDE: Jupyther Notebook is web-based application that allows us to create and share live code, narrative text, visualisation, equations. The IDE also includes data cleaning and transformation, statistical modelling, numerical simulation, data visualisation etc.

Steps to follow:

1. Face detection with Haar cascade
2. Age Recognition with CNN

Step1: Face detection with Haar cascades: OpenCv has a built-in method known as Haar-cascade that one can import and can be used to detect faces.

Step 2: Gender and age recognition with CNN: CNN algorithm is used for age recognition. The CNN's output layer (probability layer) in this CNN consists of 5 values for 2 gender and 5 age classes ("1–14", "14–25", "25–40", "40–60", "60-").

PROCEDURE

- First we are changing the current directory to the path where our image dataset is been stored by using os.chdir( ) and then we get the list of all the files and directories in the specified path using os.listdir( ).
- Shuffle() is used to randomize all the image files.
- Using split() we get the ids of image files and then store it in age variable.
- A list named classes is created to store ages less than 14 as 0, between 14 and 25 as 1, between 25 and 40 as 2, between 40 and 60 as 3 and above 60 as 4.
- Using misc.imread( ) and cv2.resize( ), we read an image from each file as an array and resize its dimensions to 32x32 and then store it in a list names X_data.
- Using squeeze(), we remove single-dimensional entries from shape of array and store in the variable X.
- Next, we normalize the data by converting the datatype of variable X to float32 and dividing it by 255.
- Slice the list classes up to 10 items and convert class vector(integers) to binary class matrix using to_categorical() before training our model.
- The total length of dataset is 23708, out of which 15008 is used for training, 1700 for testing and 7000 for validating the data.
- There are two ways to build keras models: sequential and functional. In our research, we used sequential by which we create a model layer-by-layer.
- We need to describe the input type when using Conv2D as the first layer. This layer produces a convolution kernel which is transformed to generate a tensor of outputs with the layer data.
- MaxPooling2D layer is used for spatial data.
- Dropout layer is used to prevent a model from overfitting.
- Dense layer is a linear operation in which every input is connected to every output by a w eight.
- We summarize the model by using model.summary( ). Total parameters: 554,910, trainable parameters: 554,910 and Non-trainable parameters: 0.
- Next, we compile and fit the model.
- We then evaluate the model on test set to get the accuracy up to 0.6170588.
- Now the model is trained and its ready to predict the age of any random image from dataset.
- We plot a random sample of 10 test images, their predicted labels and ground truth.
- We get the output by displaying each image along with its title.

In this paper, Keras is used to work on Tensorflow. Keras provides various features like activation functions, layers, optimizers etc, and supports CNN. By using the appropriate keras class, you can construct different deep learning models. Keras helps the model to perform random transformations of image data and normalization operations by working with various attributes including height shift, width shift, rotary range, rescale, scope of zoon shear, etc. The device will automatically add the attributes in the image by means of these attributes. ConvNet is therefore being used.

Training Dataset: The sample of data on which the model will be trained, i.e to fit the different parameters.

Validation Dataset: The sample of data on which the model will be validated while training. It is used to fit the hyper-parameters of the classifier; hence overfitting is avoided. It is independent of the training dataset.

Test Dataset: The sample of data on which the model will be tested after model has been trained. We get to know the performance (accuracy, sensitivity, loss etc.) of out classifier. It is independent of training and validation dataset.

MODEL FOR TRAINING THE DATA SET

SYSTEM TESTING

To check our trained model and validate fully integrated software product, is called System Testing. We evaluate end-to-end system specifications. Since software is one small part of a big computer-based system. We interface this software with other hardware and software systems. In System Testing execute a series of test on to exercise the full computer-based system.

- The total length of dataset is 30708, out of which 25008 is used for training, 2800 for testing and 8000 for validating the data.
- Variations in shape: One image had a shape (66, 46) whereas other had (102, 87).
- Multiple viewpoints: We have faces with whichever view possible.
- The system is tested based on all the possible angles of images and problems which may occur.

The proposed system works quite well with all the testing conditions mentioned in Fig. 4 with pretty well accuracy of age group and gender.

Fig. 4. Result analysis of proposed model.

VI. CONCLUSION AND FUTURE WORK

The proposed model was developed very carefully and while being effective, error-free. The proposed model estimates the age and gender of individuals during this study

by feeding the CNN image dataset, a deep learning algorithm, and by training in face recognition in a large database. In addition, the performance of the proposed model is good and better than many current models, but can be improved further by using more data, increasing data and improving network design. The model proposed also predicts the image's age and gender with little slip and angle problem. The completely automated face recognition program was not sufficiently accurate to achieve high recognition accuracy. It was mainly due to the fact that even a slight invariance in the face recognizing subsystem to the segmented facial image height, rotation or shift errors did not occur. The useful knowledge of a variety of subjects is explored in this paper, such as deep learning, the use of different libraries, such as Keras, Pil, Seaborn, and Tensorflow. The whole model is protected and explains the stages of development of the project and working together. Study work helps to learn how to test the project's various characteristics. In this article, there is sufficient room for further development. Moreover, it is possible to add a variety of attributes to this program, such as gender and age. If adequate data is available, a regression model may also be used outside of the age prediction classification. By further enhancing this research, camera footage can be used for safety purposes to predict real-time age and gender. For the crowd control application, the automated real-time system will be fine. Invariant appliances for facial detection and recognition. All other methods are based on the deformable prototype and main component analysis techniques and have shown good performance.

REFERENCES

[1]. Dataset downloaded from Kaggle website:- https://www.kaggle.com/age-groupclassification- with-cnn accessed on 20 September 2020.

[2]. Rehman, Sadaqat Ur, Shanshan Tu, Yongfeng Huang, and Zhongliang Yang. "Face recognition: A novel un-supervised convolutional neural network method." In *2016 IEEE International Conference of Online Analysis and Computing Science (ICOACS)*, pp. 139-144. IEEE, 2016.

[3]. Srinivas, Nisha, Harleen Atwal, Derek C. Rose, Gayathri Mahalingam, Karl Ricanek, and David S. Bolme. "Age, gender, and fine-grained ethnicity prediction using convolutional neural networks for the East Asian face dataset." In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pp. 953-960. IEEE, 2017.

[4]. Jain, Neha, Shishir Kumar, Amit Kumar, Pourya Shamsolmoali, and Masoumeh Zareapoor. "Hybrid deep neural networks for face emotion recognition." *Pattern Recognition Letters* 115 (2018): 101-106.

[5]. Levi, Gil, and Tal Hassner. "Age and gender classification using convolutional neural networks." In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 34-42. 2015.

[6]. Turabzadeh, Saeed, Hongying Meng, Rafiq M. Swash, Matus Pleva, and Jozef Juhar. "Real-time emotional state detection from facial expression on embedded devices." In *2017 Seventh International Conference on Innovative Computing Technology (INTECH)*, pp. 46-51. IEEE, 2017.

[7]. Dehghan, Afshin, Enrique G. Ortiz, Guang Shu, and Syed Zain Masood. "Dager: Deep age, gender and emotion recognition using convolutional neural network." *arXiv preprint arXiv:1702.04280* (2017).

[8]. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In *Advances in neural information processing systems*, pp. 1097-1105. 2012.

[9]. Piyush Juyal, Chitransh Kulshrestha, Sachin Sharma, and Tejasvi Ghanshala. "Common Bamboo Species Identification using Machine Learning and Deep Learning Algorithms." In *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-9 Issue-4, February 2020.

[10]. Piyush Juyal, and Sachin Sharma. "Estimation of Tree Volume Using Mask R-CNN based Deep Learning." In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1-6. IEEE, 2020.

[11]. Shuchi Bhadula, Sachin Sharma, Piyush Juyal, and Chitransh Kulshrestha. "Machine Learning Algorithms based Skin Disease Detection." In *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-9 Issue-2, December 2019.