International Journal of Computer Applications (0975 – 8887)
Volume 185 – No. 29, August 2023

Convolutional Neural Networks for Prediction of Age and Gender

Sonika Malik
Assistant Professor
MSIT, GGSIPU
New Delhi

Suyash Awasthy
Student
MSIT, GGSIPU
New Delhi

Shivansh Sondhi
Student
MSIT, GGSIPU
New Delhi

Rachit Singh
Student
MSIT, GGSIPU
New Delhi

ABSTRACT
Automatic age and gender prediction from face images has recently attracted important attention due to its wide range of operations in multitudinous facial analyses. We show in this study that exercising the Caffe Model Architecture of Deep Learning Framework; we were suitable to greatly enhance age and gender recognition by learning representations using deep convolutional neural networks (CNN). The designed methodology preprocesses the input image before performing point birth using the convolutional neural network (CNN) strategy.

This network excerpts dimensional characteristics from the source face image, followed by the point selection strategy. The proposed system is estimated using bracket rate, perfection, and recall using Adience dataset, and real-world images parade excellent performance by achieving good prediction results and calculation time.

Keywords
CNN, Caffe architecture, Gender, Age, Deep learning

1. INTRODUCTION
In recent times, numerous real-life operations like social media, security control, advertising, and entertainment have made use of information contained in a mortal face. Automatic age as well as gender prediction from facial image plays a vital part in interpersonal communication and is always a significant area for experimenters of computer vision. Face age and gender recognition are a veritably important aspect of face analysis that has piqued the interest of experimenters in areas similar as demographic information collection, surveillance, mortal-computer commerce, marketing intelligence, and security [1].

Different methodologies have been available to identify gender grounded on mortal biometric traits, behavior, characteristics, and actions. A face provides distinguished information about a person that includes age, gender, expression, mood, race, etc. Gender identification from a person's face image is a delicate operation in the computer vision community, image analysis, and artificial intelligence that recognizes gender grounded on virility and femininity. It's a double bracket problem which assigns a gender class to an existent.

Gender identification is one part of facial analysis which focuses on classifying the images under a controlled terrain. There's a need for gender brackets under an unbridled terrain which is proposed in [2]. The gender of a person provides supplementary information that helps to recoup presto and accurate information using mortal examination whereas it is a grueling problem for computers. Exploration sweats are taken to automatically prognosticate the age from the face of a person.

The proposed system focuses on carrying age-specific characteristics from face image, followed by age bracket. The age of a human can be estimated using aging cues present in the face image. Skin changes also help in perceiving the age of the grown-ups.

Age identification is a complex process that depends on gender, race, race, life, make- up, and other external factors. Accurate facial age prediction remains grueling as the exact age differs from prognostic age. Some public age recognition datasets include groups similar as child, teenager, adolescent, intermediate, and elderly citizens. The proposed system obtains the input through the real- time camera. Preprocessing is carried out to make it ready for further processing.

CNN is performed on the preprocessed image to recoup the important features. The gender has 2 classes (manly and womanish), and age of a person is classified into 8 age classes as '(0 - 4)', '(6 - 10)', '(12 - 17)', '(18 - 24)', '(25 - 32)', '(35 - 42)', '(45 - 60)', '(60 - 100)'. CNN is employed for point birth, which learns applicable characteristics by reacquiring distinctive features [2].

2. BACKGROUND
Various studies and work have been done previously on this topic which are included in this section. Also, topics like face identification, gender and age identification and others have been included here for better understanding of the project.

2.1 Face Discovery and Identification
It is an important module of any face recognition system which should be more accurate and faster. Face discovery algorithms are inspired substantially from object discovery approaches. Region- grounded object discovery classifies the generated object proffers. Each suggestion is classified as a face or no-face using a classifier. HyperFace is a hierarchical multitask training armature to conduct face identification, landmark mapping, posture prediction, and gender recognition. Region-grounded processing is briskly. R-CNN employs the region offer network (RPN), a bit like CNN. It predicts whether there's a sliding on the last point chart object or not and also predicts the boundary of those objects. RPN aids in the reduction of gratuitous face recommendations and the improvement of their position. Face discovery is generated at every place in a point space at a particular scale using sliding window approaches. It is grounded on the feed-forward convolutional network. It has a shallow sludge that can read object groups and perform discovery at multiple scales. Several facial tasks, similar as facial trait conclusion, face verification, and face recognition, need the recognition and labeling of facial milestones [6].

2.2 Gender Identification
Gender authentication may be done using a variety of data, including face photos, hand skin prints, and physiological movements, which contains a bean on gender discovery

systems exercising face prints. Gender identification may be divided into two orders, according to (i) geometric acquainted recognition and (ii) texture acquainted recognition. Golomb et al. proposed work on mortal gender discovery that relies on neural networks. In gender discovery, neural networks were generally employed for point reclamation and categorization. Back propagation neural networks are used for gender recognition. Likewise, CNN has lately been set up to be effective in carrying exclusionary features and distinguishing genders. SVM, LDA, and AdaBoost are many of the bracket algorithms employed in visual gender discovery [2].

2.3 Age Identification

The person's face carries a great deal of information, including individuality, emotion, station, maturity, position, race, race, and gender, which provides a detailed study of age modeling approaches using face prints. Kwon and Lobo suggested a strategy for classifying prints into distinct age orders grounded on face characteristics by calculating rates of different criteria. This strategy, still, may not be applicable for photos with a lot of oscillations in position, lighting, emotion, or blockage. The birth of features is an important step in prognosticating mortal age. Active appearance model (AAM), original double patterns (LBP), anthropometric features, and biologically inspired features (BIF) are some of the point birth approaches that have been developed [5].

2.4 Deep literacy styles

The original deep literacy technology employed in a ML algorithm was the deep neural network (DNN). Still, DNN has an over-fitting problem and takes much too long to train. During literacy, DNN was enhanced by exercising limited Boltzmann machines (RBM) and a deep belief network (DBN). DBN literacy is quicker than DNN due to the addition of RBM. The RBM are piled DBM with unguided connections across the situations [9].

2.5 Feature-Based Methods

A direct appearance grounded system called top element analysis (PCA) was proposed. PCA is infelicitous for classifying because it maintains uninvited intra-person differences when used for biometrics. Babu et al. proposed another direct appearance grounded system that classifies objects into sets of measurable object features called direct discriminant analysis (LDA). LDA has been more sensitive towards the training set's specific selection, performing in lower issues than PCA. To depict a different face expression, Donato employed independent element analysis features using support vectors. Several experimenters use it to dissect faces and facial expressions. Kernel PCA (KPCA) was proposed by Tanaka et al., a non-parametric fashion on the data to determine direction and minimize high confines. Several natures - inspired ways, similar as PSO, GA, and ACO, have lately been employed for point selection. In comparison to the former ways, GWO is a new methodology grounded on wolf chasing strategy. Wolf communities are created arbitrarily, which might lead to a lack of variation among wolves throughout the hunt process.

This has a significant influence on the eventual result's global confluence rate and effectiveness. Therefore, a new approach is proposed to overcome this debit [7].

2.6 Literature Survey

In this section, we compactly review the age and gender bracket literature and describe both the early styles and those that are most affiliated to our proposed system, fastening on age and gender bracket of face images from unconstrained real-world surroundings. Nearly all of the early styles in age and gender groups were handcrafted, fastening on manually negotiating the facial features from the face and substantially providing a study on constrained images that were taken from controlled imaging conditions. We have reviewed the significant exploration papers in the field published during 2010 – 2020, substantially from the times of 2020 and 2019 with some papers from 2021. We explain CNN in depth, which is the most popular deep literacy algorithm by describing the generalities, propositions, and state of-the-art infrastructures.

In a 2016 study conducted by Gil Levi and Tal Hassner, Department of Mathematics and Computer Science and Open University of Israel, they used CNN to predict age and gender from face images, and achieved an accuracy of 77% for age prediction and 81% for gender prediction [1].

In 2017, Benyamin Ghojogh, Saeed Bagheri Shouraki, Hoda Mohammadzade, Ensieh Iran Mehr used CNN to predict age and gender from a combination of facial images and textbook data (similar as social media biographies), and achieved an accuracy of 83% for age prediction and 82% for gender prediction [2].

In 2018 Koichi Ito, Hiroya Kawai, Takehisa Okano, Takafumi Aoki, used CNN to predict age and gender from images of faces and achieved 81% accuracy for age prediction and 86% for gender prediction [3].

In general, the accuracy of age and gender prediction using CNN has improved over time, as the performance of CNN has increased and further data has come available for training. Unborn exploration in this area may concentrate on perfecting the accuracy of age and gender prediction using CNN, as well as exploring the use of other machine literacy ways for this task.

In 2022, some of the challenges in using CNN for age and gender prediction include the need for large quantities of labeled data and the difficulty of directly predicting age, which can vary significantly within a single demographic group [7].

3. PROPOSED WORK

The proposed work includes details and theory about the project material like Convolutional Neural Network, Network Architecture (Caffe Net model), training and testing of the model using Caffe Net Architecture and the details about the datasets and the Adience dataset which is used in the making of this project. This helps in understanding the overview of the project and the paper.

3.1 Convolutional Neural Network (CNN)

CNNs were first developed and used around the 1980s. The most that a CNN could do at that time was recognize handwritten integers. It was substantially used in the postal sectors to read zip codes, pin codes, etc. The important thing to flash back about any deep literacy model is that it requires a large quantum of data to train and requires a lot of computing resources. This was a major debit for CNNs at that period and hence CNNs were only limited to the postal sectors and it failed to enter the world of machine literacy. In deep learning, a convolutional neural network (CNN) is a class of deep neural networks, utmost generally applied to dissect visual imagery. Now when we suppose a neural network, we suppose about matrix multiplications but that is not the case with ConvNet. It uses a special fashion called Convolution. Now in mathematics, complication is a fine operation on two functions that produces a third function that expresses how the shape of one is modified by the other.

3.2 Network Architecture

Caffe is a deep knowledge frame made with expression, speed, and modularity in mind. It's developed by Berkeley AI disquisition (BAIR) and by community contributors. Yangqing Jia created the design during his PhD at UC Berkeley. The BSD 2- Clause license governs Caffe's release [10]. Speed makes Caffe perfect for disquisition trials and sedulity deployment. Caffe can exercise over 60M images per day with a single NVIDIA K40 GPU. That is 1 ms/image for conclusion and 4 ms/ image for knowledge and more recent library performances and attack are hastily still. We believe that Caffe is among the fastest ConvNet prosecutions available.

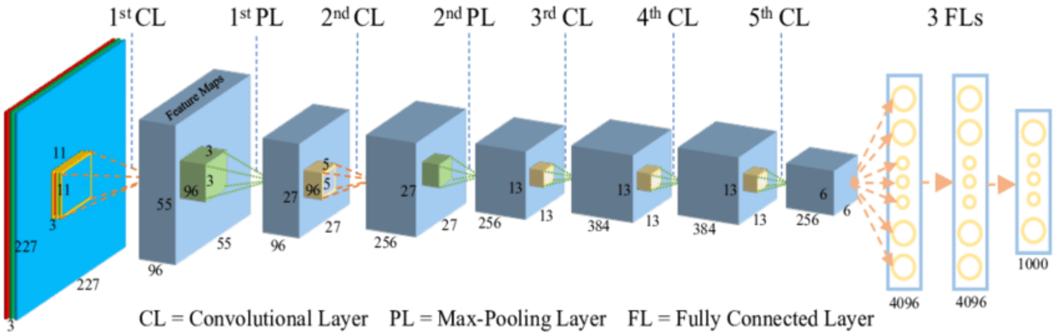

Figure 1. CaffeNet architecture

3.3 Training and Testing

The training is done using object recognition. Object recognition is a general term to describe a collection of affiliated computer vision tasks that involve relating gender or other parameters in digital photos. Predicting the class of a single object in a picture is known as an image bracket. Object localization refers to relating the position of one or further objects in an image and drawing a limiting box around their extent. These two tasks are combined in object discovery, which locates and categorizes one or further effects in an image [13].

As such, we can distinguish between these three computer vision tasks:

3.3.1 Image Classification

Predict the type or class of an object in an image.
Input: An image with a single object, such as a photograph.

Output: A class label (e.g., one or more integers that are mapped to class labels).

3.3.2 Object Localization

Locate the presence of objects in an image and indicate their location with a bounding box.

Input: An image with one or more objects, such as a photograph.

Output: One or more bounding boxes (e.g., defined by a point, width, and height).

3.3.3 Object Detection

Locate the presence of objects with a bounding box and types or classes of the located objects in an image.

Input: An image with one or more objects, such as a photograph.

Output: One or more bounding boxes (e.g., defined by a point, width, and height), and a class label for each bounding box.

3.4 Datasets

Adience dataset serves as a benchmark for face photos and is inclusive of various real-world imaging conditions like noise, lighting, pose, and appearance. The photos were gathered from Flickr albums and made available under a Creative Commons (CC) license. It is roughly 1GB in size and contains a total of 26,580 images of 2,284 subjects over eight age ranges (as previously indicated) [12].

Figure 2. Visualization of Adience dataset on the Deep Lake UI

3.4.1 Adience Dataset in Python

Instead of downloading the Adience in Python, you can effortlessly load it in Python via Deep Lake open-source with just one line of code.

Load Adience Fold Faces Dataset Subset in Python

```
import deeplake ds =
deeplake.load('hub://activeloop/adience')
```

3.4.2 Data Fields

- Images: tensor containing the image
- Ages: tensor containing ages (label) of a corresponding image
- Gender: tensor containing gender of each image
- x: part of bounding box of the face in the original

Flickr image

- y: part of bounding box of the face in the original Flickr image
- dx: part of bounding box of the face in the original Flickr image
- dy: part of bounding box of the face in the original Flickr image
- tilt_ang: pose of the face in the original Flickr image
- iducial_yaw_angle: pose of the face in the original Flickr image
- fiducial_score: score of the landmark detector

Table 1. Number of People Tested

Gender /Age	0-4	6-10	12-17	18-24	25-32	35-42	45-60	60-100
Male	5	8	13	16	10	7	6	5
Female	3	7	14	20	7	8	5	3
Both	8	15	27	36	17	15	11	8

3.4.3 Adience Data Splits

- The Adience data set training set is composed of 19370 images.

4. EXPERIMENTAL RESULTS

The system is executed using the Caffe open- source frame. Training was performed on an Amazon GPU machine with 536 CUDA cores and 4 GB of video memory. Training each network took about four hours, predicting age or gender on a single image using the network requires about 200ms. vacationing running times can conceivably be substantially better by running the network on image batches.

Figure 3 Misclassifications

Misclassification is the act of wrongly saying that someone or something is in a particular group or is a particular type. Generally, the proposed system outperforms the reported state-of- the- art on both tasks with considerable gaps. Also apparent is the donation of the over-sampling approach, which provides a fresh performance boost over the original network. This implies that better alignment (e.g., frontalization) may give a fresh boost in performance.

We give many examples of both gender and age misclassifications. These show that numerous of the miscalculations made by our system are due to extremely grueling viewing conditions of some of the Adience standard images. The utmost notable is miscalculations caused by blur or low resolution and occlusions (particularly from heavy makeup). Gender estimation miscalculations also constantly do for images of babies or veritably youthful children where appalling gender attributes are not yet visible.

Figure 4 Accurate Prediction of the subjects

Accuracy-

Accuracy refers to the closeness of the measured value to a standard or true value. Below is the accuracy matrix of our age and gender detection project.

Table 2. Accuracy Table

Gender	Accuracy
Male	85.70%
Female	88.05%
Total	**86.86%**

5. CONCLUSION AND FUTURE WORK

Though multitudinous former styles have addressed the problems of age and gender type, until recently, much of this work has concentrated on constrained images taken in lab settings. Analogous settings do not adequately reflect appearance variations common to the real- world images in social websites and online magazines. Internet images, still, are not simply more challenging, they are also abundant. The easy vacuity of huge image collections provides modern machine knowledge predicated systems with effectively endless training data, though this data is not always suitably labeled for supervised knowledge. Taking illustration from the related problem of face recognition we explore how well CNN performs on these tasks using Internet data.

We give results with a spare deep- knowledge architecture designed to avoid over-fitting due to the limitation of limited labeled data. With the backing of machine literacy technology, it's come easy to seek out relation and patterns among colorful data. The work done in this design substantially concentrates around detecting the age and gender of a mortal and counting the number of mortal faces present in the frame. This model provides high delicacy. Data visualization helps in analysis of the dataset.

Using this analysis, we revealed intriguing characteristics that helped in understanding the dataset in a better and easy way. In future, using face age, mortal expression bracket to prop face recognition, facial complaint discovery, ameliorate gests with images, film land of social media, and much further than this. Then we can also consider a deeper CNN armature and a more robust image processing algorithm for exact age estimation. Also, the apparent age estimation of human’s face will be intriguing exploration to probe in the future.

6. REFERENCES

[1] Age and Gender Classification Using CNN: Gil Levi and Tal Hassner Department of Mathematics and Computer Science the Open University of Israel 2016

[2] A Fusion-based Gender Recognition Method Using Facial Images: Benyamin Ghojogh, Saeed Bagheri Shouraki,

Hoda Mohammadzade*, Ensieh Iran Mehr 2017

[3] Age and Gender Prediction from Face Images Using CNN: Koichi Ito, Hiroya Kawai, Takehisa Okano, Takafumi Aoki 2018

[4] Age and Gender Prediction using Deep CNN: Insha Rafique, Awais Hamid, Sheraz Naseer, Muhammad Asad, Muhammad Awais, Talha Yasir 2019

[5] Facial Data Based Deep Learning: Emotion,Age and Gender Prediction: Abhijit Roy 2020

[6] CNN Based Features for Age and Gender Estimation and Gender Classification: Mohammed Kamel Benkaddour University Kasdi Marbah, Department of Computer Science and Information Technology FNTIC Faculty, Ouargla, Algeria 2020

[7] Prediction of Age and Gender Based on Human Face Images Based on Deep Learning Algorithm: S. Haseena, S. Saroja, R. Madavan, Alagar Karthick, Bhaskar Pant, Melkamu Kifetew 2022

[8] Y. Kumar, A. Koul, P. S. Sisodia et al., "Heart failure detection using quantum-enhanced machine learning and traditional machine learning techniques for Internet of artificially intelligent medical things," *Wireless Communications and Mobile Computing* 2021.

[9] J. Shafi, M. S. Obaidat, P. V. Krishna, B. Sadoun, M. Pounambal, and J. Gitanjali, "Prediction of heart abnormalities using deep learning model and wearable devices in smart health homes," *Multimedia Tools and Applications* 2022.

[10] G. Guo and G. Mu, "A framework for joint estimation of age, gender and ethnicity on a large database," *Image and Vision Computing*, vol. 32, no. 10, pp. 761–770 2014.

[11] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA 2015.

[12] S. G. Bhele and V. H. Mankar, "A review paper on face recognition techniques," *International Journal of Advanced Research in Computer Engineering & Technology* 2012.

[13] S. Chidambaram, S. S. Ganesh, A. Karthick, P. Jayagopal, B. Balachander, and S. Manoharan, "Diagnosing breast cancer based on the adaptive neuro-fuzzy inference system," *Computational and Mathematical Methods in Medicine* 2022.

[14] D. A. Asanov, *Algorithms and Methods in Recommender Systems*, Berlin Institute of Technology, Berlin, Germany 2011.

[15] H. Sikkandar and R. Thiyagarajan, "Soft biometrics-based face image retrieval using improved grey wolf optimization," *IET Image Processing* 2020.