**Age and Gender Detection using Deep Learning Techniques**

Aadam Sunday Olatunji (Professor)1, Nawaf Mohmed Al Qurashi2, Saifallah Tarek Gebreel3, Turki Mutib Al Mutib4, Abdelhadi Al Amoudi5

**Abstract**

When looking at someone’s face, we can identify certain qualities as humans that tell us their gender, age, mood, etc. Having a way to recognize the age of somebody using an image can improve any security system tenfold. This project focuses on detecting the age and gender of a person using images. We want to tackle this issue to help improve security standards and to help with identifying an age range and gender for a suspect in an investigation perhaps. Also, it can be used as an amusing piece of technology that people use for entertainment. This is a project typically developed using Deep Learning with Supervised Learning on thousands of image data, which we will be applying as well as using the standard algorithm for dealing with images, Convolutional Neural Networks (CNNs), in the development of this project as they have become the most used in terms of dealing with image-based data. After reviewing the earlier work of some of our predecessors we realized that the scarcity of any previous work was a result of the unavailability of the needed technology to make this idea possible, we expect our work to improve upon it as technology improved and the software developed. We achieved an accuracy score of 0.9659 out of 1.0, while also contributing towards a safer environment for all using Artificial Intelligence, or at least providing another means of entertainment for others.

**Keywords:** Supervised Learning, Deep Learning, Convolutional Neural Network CNN, Artificial Intelligence.

# **1. Introduction**

This paper introduces an advanced approach for age and gender detection using convolutional neural networks (CNNs) exclusively based on image data. Accurate estimation of age and gender from images holds significant practical value across various domains. Our aim is to address these challenges by developing a robust CNN-based solution capable of handling variations in appearance and lighting conditions.[1]

By leveraging the power of CNNs, our approach leverages the ability to automatically learn meaningful features directly from raw image data. CNNs have proven highly effective in computer vision tasks, making them an ideal choice for age and gender detection. Our proposed method employs a deep learning architecture designed to capture hierarchical patterns and dependencies within facial images.[2]

Through extensive experimentation and evaluation, we aim to showcase the superiority of our CNN-based method in terms of accuracy and efficiency. By addressing the limitations of existing techniques, we contribute to advancing the field of age and gender detection, opening new possibilities for practical applications in diverse real-world scenarios.[3]

Real-world security needs an overall upgrade as every aspect of life is upgrading around it. So, having a widely available face recognition system that can detect the age and gender of the suspect would increase the sense of security for every business owner. And by using the latest advances in Deep Learning, we will arrive at a usable, reliable, and cost-free solution to be used everywhere.

The rest of this research is divided into the following sections: Section 2 provides an extensive review of related literature, as well as an outline of current research on the topic. The proposed techniques will be discussed in Section 3. Section 4 contains empirical studies that describe the dataset, experimental setup, and performance metrics used. Section 5 contains the results of this study as well as the related suggestions. Finally, Section 6 brings the study to a close by summarizing the important findings from the research.

# **2. Review of Related Literatures**

This study answers the increased need for such tools in the era of pervasive social media usage. This project focuses on building software for automatically determining the age and gender of message authors. For categorizing and identifying messages across various platforms, machine learning techniques were used. The process used by the proposed system includes data input, tokenization-based feature extraction, data cleaning using string to word vector conversion, and feature selection. Decision Trees, Naive Bayes, and Random Forest are employed as classification techniques. The study ends by examining the output class and evaluating the outcomes produced when these algorithms are used on datasets.[4]

The difficult problem of determining age and gender from face traits using computer vision and psychophysics is the focus of this work. The research has a high success rate of 95% when classifying age and gender categories using a feed-forward propagation neural network. Three-sigma control limits and hierarchical decision-making are used in the methodology. The effectiveness and precision of the suggested strategy are demonstrated by experimental findings utilizing photos from benchmark databases. Overall, the system successfully categorizes human faces, achieving precise detection and age and gender classification. The method shows off its efficiency, speed, and accuracy in face picture categorization.[5]

This study employs a method based on SIFT and morphological algorithms to determine age and gender. In the area, several methods for determining age and gender from picture characteristics have been developed. The research looks on age detection methods and records individual differences using fMRI. Brain activation experiments using face matching produce consistent findings in both older and younger people. Age detection includes several variables, and conclusions drawn from credentials stored in the system need to be carefully watched. The input picture is processed using the SIFT approach, then the SVM classifier is used to accurately classify features and estimate age and gender with 94% accuracy.[6]

In this publication, Yunjo Lee et al. suggested using the fMRI method to research age detection techniques. The study entails a thorough documentation of individual differences based on age, gender, identity, and other characteristics. The face matching brain activation tests are carried out and tested outside of the scanner. In terms of facial processing, both older and younger persons showed the same results. With identical facial perspectives in both scenarios, the performance is excellent. There is no single cause for the aging of the elderly. The accounting of such findings is the consequence of a mix of many elements. The outcomes, which are based on all credentials stored in certain contexts, need to be monitored.[7]

The Conditional Probability Neural Network (CPNN) for age prediction from facial expressions was proposed by Chao Yin et al. in this paper. The CPNN makes use of a three-layer neural network architecture and inputs conditional feature vectors and target values to efficiently learn real ages. Using the maximum entropy model, the system learns the relationship between face pictures and related label distribution. When compared to earlier approaches, CPNN has demonstrated greater performance, offering straightforward, computationally efficient, and highly effective solutions. Due to its positive traits, the study chose CPNN above alternative strategies.[8]

# **2.1 Post Literature Review Discussion**

The idea of face identification and age detection has been one that haunted the brilliant minds of the brightest scientists in the scene of Artificial Intelligence, and after reviewing the work of our predecessors we gained valuable information about what goes into building such a fascinating model that can have human like functionalities and can guess the age and gender of other humans.

We identified an opening for us to have our work leave a mark, that opening was represented in the scarcity of research done well on the topic and that the technology needed to make this model exist is new, and as we all know technology is always everchanging and evolving every day. We aim to improve on the existing work that we reviewed to the extent of our abilities and use the standard methodologies and tools used by professionals in the AI sphere such as Convolutional Neural Networks (CNNs) to make our idea a reality and of course building a capable, accurate, and up-to-date model.

# **3.0 Description of the Proposed model**

Using TensorFlow and the Keras Sequential API, we created the Convolutional Neural Network (CNN) classification, a deep learning model with a few layers, including convolutional, normalization, activation, pooling, normalization, and dense layers.[9]

The suggested model is made to categorize photos according to their attributes. Multiple layers in the model extract information from the photos and use those features to establish predictions. To find features, the model employs many convolutional layers that apply filters to the input picture dataset. Convolutional, Max Pooling, and Dropout layers are used to further process the detected features. These layers reduce the features' complexity and guard against overfitting.

The model also includes thick, fully linked layers that anticipate outcomes based on features extracted. The activation function connecting these layers to the layers above gives the model non-linearity. To prevent overfitting, dropout is utilized on the dense layers.

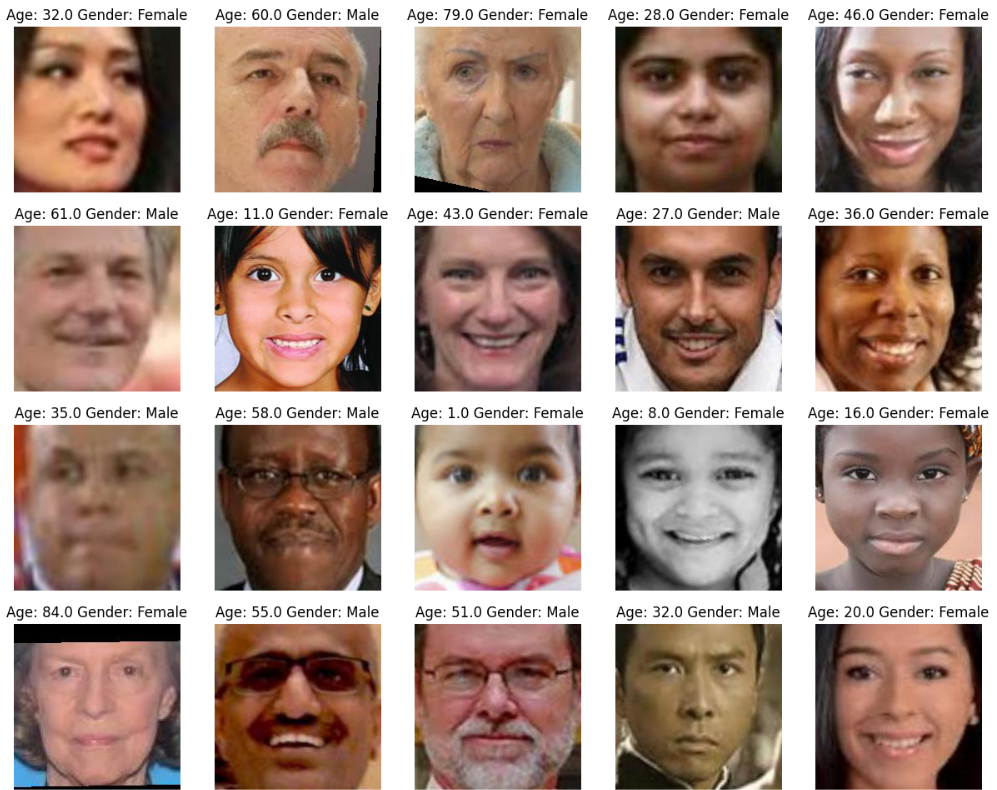
# **4.0 Empirical Studies**

# **4.1 Description of dataset**

The dataset [10] is a huge dataset containing over 20,000 pictures of faces with an age span from 0 to 116 years old. The dataset also contains annotations of age, gender, and finally ethnicity, which is irrelevant in the scope of this project. There is a wide range of poses, facial expressions, lighting, occlusion, resolution, etc. in the photographs. Several tasks, including face identification, age estimation, age progression and regression, landmark localization, etc., might be performed using this dataset.

The sheer size of the dataset presented a challenge in training the model as did the nature of the project, leaving us with the conclusion that the model will need a lot of time to train.

As seen below in Figure 1, the images are labeled by age and gender. These are some examples taken from the dataset.



*Figure 1- Images Example*

The dataset contains different numbers of images for different ages, as shown in Figure 2. A picture containing diagram, screenshot, plot

Description automatically generated

*Figure 2- Image Number by Age*

The dataset is near balanced, as we have more male pictures than female pictures. The males are represented with a (0) while the females are represented as a (1) as shown in the figure.

A picture containing text, screenshot, rectangle, diagram

Description automatically generated

*Figure 3- Number of Images by Gender*

# **4.2 Data Pre-Processing**

Pre-Processing is a critical phase in the making of a model as it dictates if a model will train on relatively clean data or not. The initial step was to read the age and gender values from each image name, with the first number being the age and the second being the gender with 2 possibilities either (0) for male or (1) for female, and then they are saved as new attributes for in the DataFrame. Then splitting the data in a 80:20 ratio, with 80% being the training size, using the ‘sklearn.model\_selection.train\_test\_split’. The training size was that of 18966 images. Then after the splitting of the data, the feature extraction phase began by converting the images into NumPy arrays and resizing them so that we have a uniform format across the dataset. Lastly, before creating the model, a normalization technique was applied for the color scale to be the standard 255. That was achieved by dividing the training set by 255.

**4.3 Experimental Setup**



The study was conducted using Google Colab, and a Convolutional Neural Network (CNN) on the image dataset. Also, some Skip Connections, which are typically used in Residual Neural Networks (RNNs), were used to save time. Skip Connections are considered “shortcuts” of sorts, they connect the output of a certain layer to the input of another that is not close to it. [11]

As mentioned, the age and gender are read from the file names and stored as new columns in the DataFrame. Then the dataset was split into an 80% training and 20% testing ratio. The CNN model was built in TensorFlow (TF) as a Functional model and it included Conv2D, MaxPooling2D, a batch normalization layer, and fully connected layers with 2 activation layers, Sigmoid and ReLU, for the 2 dropout layers included. The Adam optimizer was used with two loss functions, Binary\_CrossEntropy and MAE, and Accuracy as the metric. The model was originally trained for 20 epochs with a batch size of 36, but the was tuned after trial and error to be trained for 100 epochs with a batch size of 128.

A picture containing text, diagram, receipt

Description automatically generated

Figure 4- The Architecture of the CNN Model (Python Generated)

# **4.4 Performance Measures**

We used a variety of performance indicators, taking into account the values of the False Positive (FP), which is a projected positive but the observation is actually negative, the True Positive (TP), which is a positive for both observations and predictions, the False Negative (FN), which is a positive but predicted to be a negative, and the True Negative (TN), which is a negative for both observations and predictions and is present in a Confusion Matrix. The accuracy metric, which measures the percentage of accurate forecasts over all predictions, is widely used as the foundational metric for model assessment. It may be used to evaluate a Classification model's performance. Precision is a measure of how many of the positive predictions produced by the classifier were correct (true positives), and Recall is a measurement of how many of all the positive instances in the data that the classifier properly predicted. A measurement that combines recall and accuracy is the F1-Score. The Harmonic Mean of the two is the term used to describe it. The harmonic mean is another way to compute an "average" of numbers, which is typically seen as more suited for ratios than the conventional arithmetic mean (such as recall and accuracy). [12]

Also, Binary Cross Entropy is a model measure that records inaccurate data class labeling by a model, punishing the model if variations in probability occur when categorizing the labels. It is sometimes referred to as logarithmic loss or log loss. High accuracy values are equivalent to low log loss values. [13]

Finally, the average of the absolute deviations between the actual value and the value predicted serves as the MAE loss function. The second most popular regression loss function is this one. Without taking their directions into account, it calculates the average size of mistakes in a group of forecasts. [14]

The equations for the performance measures are as follows:

**Accuracy:**

**Recall:**

**Precision:**

**F1-Score: 2\***

**Binary-Cross-Entropy:**

**MAE:**

# **4.5 Optimization strategy**

In our model, we experimented with the Adam optimizer.

Finally, we selected Adam as the best optimizer for our needs. Adam optimizer's job is to update the network weights based on the gradients of the loss function with respect to the weights.

# **5.0 Result and Discussion**

Here we will discuss the result of Adam and compare the Age and Gender attributes.

Table 1- Results, Classifiers. & Loss Functions

|  |  |  |  |
| --- | --- | --- | --- |
| Quality measures | Gender | | Age |
| Validation loss | 0.1106 | | 2.5250 |
| Accuracy | 96.59% | | - |
| Optimizer | Adam | | |
| Classifier | CNN | | |
| Loss Function | Binary Cross-Entropy | Mean Absolute Error (MAE) | |

The table above shows the results of our experimental setup. The Age accuracy score is unavailable as it falls under Regression, and the entire model deals with a Classification issue, so it produces bizarre accuracy scores. Seeing as both the Gender and Age are trained with the same model at the same time, they both had Adam as an optimizer and CNN as a classifier. The Gender on one hand used the Binary Cross-Entropy loss function as it is a classification issue, meaning it is either a Male or a Female, whereas the Age is a regression issue, so we used MAE as a loss function.

In figures 5 and 6 we can observe the loss values of the Gender and Age with respect to the number of epochs (iterations).

A picture containing text, diagram, line, plot

Description automatically generated

Figure 5- Gender Loss Graph

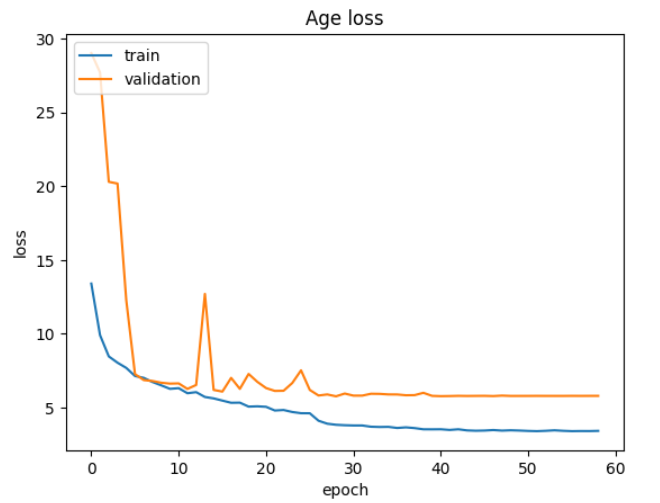


Figure 6- Age Loss Graph

In figure 7, we can observe the Accuracy Score of the Gender with respect to the number of epochs.

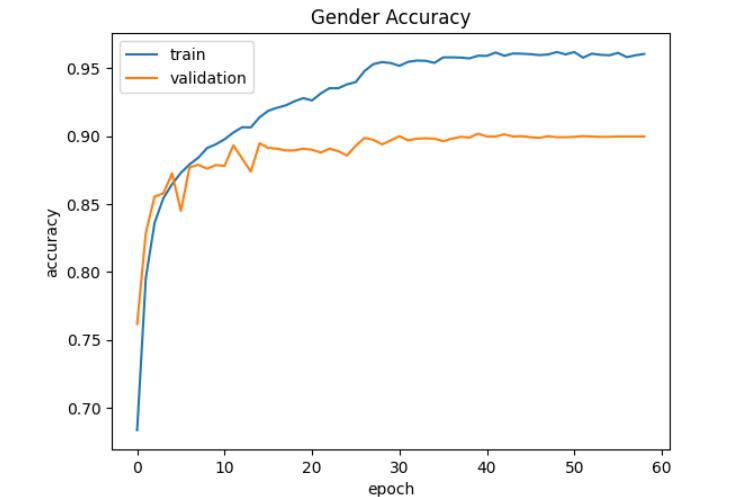


Figure 7- Gender Accuracy Score

# **6.0 Conclusion and recommendation**

In conclusion, this study aimed to improve the existing age and gender detection models by applying the latest advances in technology. We achieved a relatively high accuracy score for the Gender attribute after the epoch number and batch size tuning as we had a previous accuracy of 90% but after the tuning it was bumped to 96.59%. This study has the potential for more improvements as the technology used in it is ever evolving and any breakthrough made in this area could be considered the standard for age and gender detection and could be extended to real-time detection. In the future, the accuracy score can be improved, and more image data can be gathered to increase the training data for a better outcome. Overall, the sky is the limit for the applications and improvements possible for the gender and age detection models currently existing and this study is just, and our study is an improvement on the current existing ones.

# **Acknowledgement**

Dr. Aadam is gratefully acknowledged by the authors for his excellent advice and support throughout the study process. His expertise and perspective have been crucial to the project's success. The authors also acknowledge the invaluable assistance and resources provided by their respective universities. Without their support, this research would not have been feasible.

Finally, the authors appreciate the time and cooperation of all the study participants.

# **References**

[1] K. D. Kaur and P. Rai, “An Analysis on Gender Classification and Age Estimation Approaches,” *Int J Comput Appl*, vol. 171, no. 10, pp. 975–8887, 2017.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Commun ACM*, vol. 60, no. 6, 2017, doi: 10.1145/3065386.

[3] J. Lin *et al.*, “Exploring Explicit Domain Supervision for Latent Space Disentanglement in Unpaired Image-to-Image Translation”, Accessed: Jun. 05, 2023. [Online]. Available: https://github.com/

[4] “(PDF) AGE AND GENDER DETECTION.” https://www.researchgate.net/publication/357448747\_AGE\_AND\_GENDER\_DETECTION (accessed May 21, 2023).

[5] M. R. Dileep and A. Danti, “Human Age and Gender Prediction Based on Neural Networks and Three Sigma Control Limits,” *https://doi.org/10.1080/08839514.2018.1451217*, vol. 32, no. 3, pp. 281–292, Mar. 2018, doi: 10.1080/08839514.2018.1451217.

[6] H. Kaur, “Age and Gender Detection,” *International Journal of Computer Science Trends and Technology (IJCST)*, vol. 5, 2013, Accessed: May 21, 2023. [Online]. Available: www.ijcstjournal.org

[7] Y. Lee, C. L. Grady, C. Habak, H. R. Wilson, and M. Moscovitch, “Face Processing Changes in Normal Aging Revealed by fMRI Adaptation,” *J Cogn Neurosci*, vol. 23, no. 11, pp. 3433–3447, Nov. 2011, doi: 10.1162/JOCN\_A\_00026.

[8] C. Yin and X. Geng, “Facial age estimation by conditional probability neural network,” *Communications in Computer and Information Science*, vol. 321 CCIS, pp. 243–250, 2012, doi: 10.1007/978-3-642-33506-8\_31/COVER.

[9] “What are Convolutional Neural Networks? | IBM.” https://www.ibm.com/topics/convolutional-neural-networks (accessed Jun. 05, 2023).

[10] “UTKFace | Kaggle.” https://www.kaggle.com/datasets/jangedoo/utkface-new/code (accessed Jun. 01, 2023).

[11] “What are the benefits and drawbacks of using skip connections in CNNs?” https://www.linkedin.com/advice/0/what-benefits-drawbacks-using-skip-connections (accessed Jun. 02, 2023).

[12] “A Look at Precision, Recall, and F1-Score | by Teemu Kanstrén | Towards Data Science.” https://towardsdatascience.com/a-look-at-precision-recall-and-f1-score-36b5fd0dd3ec (accessed Jun. 02, 2023).

[13] “Binary Cross Entropy: Where To Use Log Loss In Model Monitoring - Arize AI.” https://arize.com/blog-course/binary-cross-entropy-log-loss/ (accessed Jun. 02, 2023).

[14] “Common Loss Functions in Machine Learning | Built In.” https://builtin.com/machine-learning/common-loss-functions (accessed Jun. 02, 2023).

