Deep Learning Approach for Gender and Age Prediction Using CNN

Daram Vinay
vinaydaram@my.unt.edu
University of North Texas
Denton, TX, USA

Tejaswini Kanukunta
Tejaswinikanukunta@my.unt.edu
University of North Texas
Denton, TX, USA

Asish Inuganti
AsishInuganti@my.unt.edu
University of North Texas
Denton, TX, USA

ABSTRACT

The creation of a Convolutional Neural Network (CNN) model for precisely estimating age and gender from face photos is the main goal of this work. The model seeks to provide a strong response to gender and age detection problems by utilizing deep learning techniques. Its ramifications span a number of several industries, including advertising, healthcare, and security. Prior studies in this field have looked at deep learning techniques using CNNs as well as conventional techniques based on human feature extraction. By using a multi-output CNN architecture, this study improves on earlier research by enabling the simultaneous prediction of age and gender. Data collection, preprocessing, feature extraction, model architecture creation, and training using the Adam optimizer are critical elements in the technique. Early findings show promising performance with low mean absolute error in age prediction and good gender prediction accuracy. Notwithstanding, possible constraints including bias in the dataset and external influences on the model's performance are recognized. Prospective improvements might encompass mitigating dataset bias, enhancing resilience to environmental fluctuations, and integrating a wider range of datasets. The suggested CNN model has the potential to greatly progress facial recognition technology and its practical applications by surmounting these obstacles. This study emphasizes how crucial it is to forecast gender and age accurately in a variety of industries, such as marketing, healthcare, and surveillance. The study intends to aid in the creation of dependable and scalable solutions for these important activities by utilizing CNNs.

KEYWORDS

Convolutional Neural Networks (CNNs), real-world applications, dataset bias, multi-output architecture, deep learning, face image analysis, gender and age prediction, and facial recognition technology

ACM Reference Format:

Daram Vinay, Tejaswini Kanukunta, and Asish Inuganti. 2024. Deep Learning Approach for Gender and Age Prediction Using CNN. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Convolutional Neural Networks (CNNs) are the main tool used in this project to determine gender and age from face image data accurately. This work is important for many fields, such as targeted advertising, demographic analysis, and facial recognition systems. Gender and age recognition from photos is essential for many applications, including improving marketing tactics, enabling customized healthcare services, and strengthening security protocols. Thus, creating a workable answer to this issue is essential for handling practical issues and enhancing decision-making procedures.

This issue is significant since it has broad ramifications across several industries. For example, precise gender and age prediction from facial photos can help identify people in surveillance films, improve public space security, and support law enforcement in their efforts. Furthermore, advertising campaigns can be optimized in marketing by knowing the demographics of target consumers through gender and age prediction, which improves customer engagement and boosts sales. Furthermore, based on facial image analysis, gender, and age prediction in healthcare can improve patient outcomes, allow for individualized treatment strategies, and enable early diagnosis of age-related health conditions.

This project's contribution is the creation of a CNN model that can correctly identify age and gender from facial image data. This model attempts to offer a strong and dependable answer to the problem of gender and age recognition by utilizing deep learning techniques. The study aims to show how well the suggested CNN model performs in attaining high levels of accuracy and generalization across a variety of datasets and real-world settings through rigorous experimentation and evaluation. In the end, this model's effective use could have far-reaching effects across various sectors and domains, advancing both technology and society at large.

1.1 Importance of the Problem

Enabling personalized healthcare services, enhancing security measures, and improving marketing techniques all depend on the ability to recognize gender and age from facial pictures. The accuracy of gender and age prediction from face photos may help with several applications, such as helping law enforcement with their investigations, identifying people in surveillance films, and improving public space security. Furthermore, precise demographic research via age and gender prediction helps to improve consumer interaction, customize advertising campaigns, and eventually increase revenue. Furthermore, early detection of age-related health issues, tailored treatment plans, and better patient outcomes can all result from the use of facial image analysis for gender and age prediction in the healthcare industry.

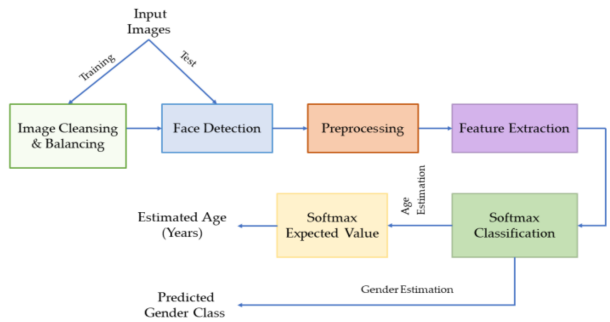

Figure 1: Schematic diagram of the proposed age and gender estimation method.

1.2 Project Contribution

The main contribution of this study is the creation of a CNN model that can reliably determine gender and age from facial picture data. The model seeks to offer a strong and dependable solution to the gender and age recognition issue by employing deep learning techniques. The work aims to show how well the suggested CNN model performs in attaining high levels of accuracy and generalization across a variety of datasets and real-world settings through rigorous experimentation and assessment. In the end, this model's effective use might have significant and far-reaching consequences on a variety of industries and fields, advancing both society and technology.

2 RELATED WORKS

Numerous machine learning algorithms have been used to study gender and age prediction from facial photos in great detail. Conventional methods frequently entail the manual extraction of features, which is followed by the use of classifiers like Random Forests or Support Vector Machines (SVMs). To represent facial traits, these methods rely on handmade features such as color histograms, texture descriptors, or facial landmarks.

On the other hand, deep learning models—in particular, Convolutional Neural Networks, or CNNs—have become extremely effective instruments for automatically extracting features from unprocessed data. CNNs can efficiently extract hierarchical characteristics from images by pooling and convolutioning many layers of data. By doing away with the necessity for manual feature extraction, this end-to-end learning process can produce predictions that are more reliable and accurate.

Earlier research in this field investigated several CNN designs for problems predicting age and gender. Single-output CNN models have been proposed by some researchers, in which each prediction task (e.g., gender or age) is trained on a different network. Some have embraced multi-output CNN topologies, in which a single network simultaneously predicts age and gender. Enhanced computing efficiency and shared feature representations are two benefits of these multi-output models.

By using a multi-output CNN architecture for simultaneous gender and age prediction, this study expands on earlier findings. The model can potentially achieve higher accuracy than traditional methods by autonomously learning discriminative characteristics from facial photos through the use of deep learning techniques. The model's capacity to generalize to previously unseen data is further enhanced by the end-to-end learning process, which allows it to recognize intricate patterns and variations in face traits.

Overall, this study seeks to establish a strong and accurate model for gender and age prediction from facial photos by utilizing deep learning skills and building upon the foundation laid by prior studies in this field.

3 METHODS

To create and train the CNN model for gender and age prediction, this study used the following crucial procedures as part of its methodology:

3.1 Data Collection and Preprocessing

Age and gender labels are appended to each facial image, which is mostly sourced from the UTKFace collection. Preprocessing involves processing photos to extract gender and age labels from filenames, making it simple to associate labels with images. To guarantee uniformity in input size for the CNN model, photos are additionally scaled to a common dimension of 128x128 pixels.

3.2 Feature extraction

Preprocessing stages include grayscale conversion and scaling the photos to the appropriate proportions before putting them into the CNN model. This conversion from grayscale preserves important face information while decreasing the dimensionality of the input. Next, convolutional layers of the CNN architecture are used for feature extraction. By using filters, these convolutional layers can extract hierarchical features from the pictures and capture important structures and patterns.

3.3 Model Architecture

The CNN model architecture is made to analyze face photos efficiently and produce age and gender predictions. The max-pooling layers come after a sequence of convolutional layers that work to extract and downsample features from the input pictures. The model is then able to learn intricate correlations between the collected characteristics and the target variables (gender and age) by using fully linked layers for additional processing. Dropout regularization is used to the fully connected layers to help in generalization to previously unknown data and prevent overfitting. The gender and age forecasts make up the model's final output.

3.4 Training

To optimize model parameters, the model is trained using the Adam optimizer, a stochastic gradient descent variation that dynamically modifies the learning rate. The loss function for age prediction is mean absolute error (MAE), whereas the loss function for gender prediction is binary cross-entropy loss. To track model performance and avoid overfitting, the dataset is divided for validation once the model has been trained over several epochs. As it goes through training, the model becomes better at predicting age and gender and learns to minimize the loss function.

By using this technique, the study hopes to expand facial recognition technology and associated applications by creating a reliable

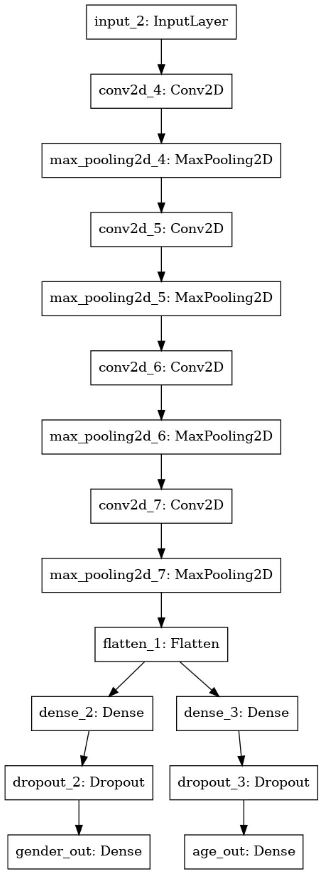

Figure 2: Model Summary

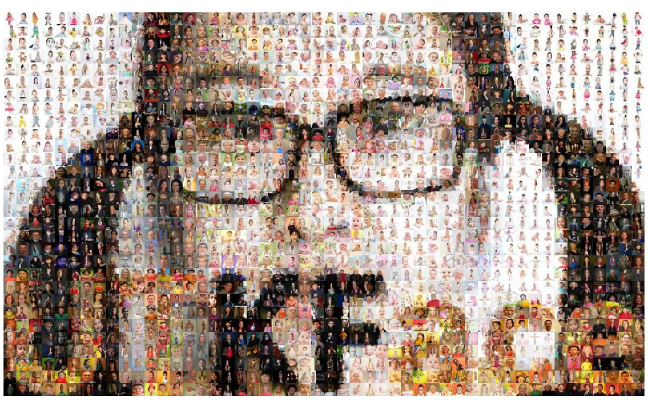

Figure 3: You can download the UTKFace dataset from Kaggle

CNN model that can correctly identify gender and age from face photographs.

4 PLOTTING GRAPHS AND RESULTS

23,708 face photos that have been resized to 128 by 128 pixels make up the dataset. The dataset's gender balance and age distribution are revealed by preliminary exploratory data analysis. With a 90

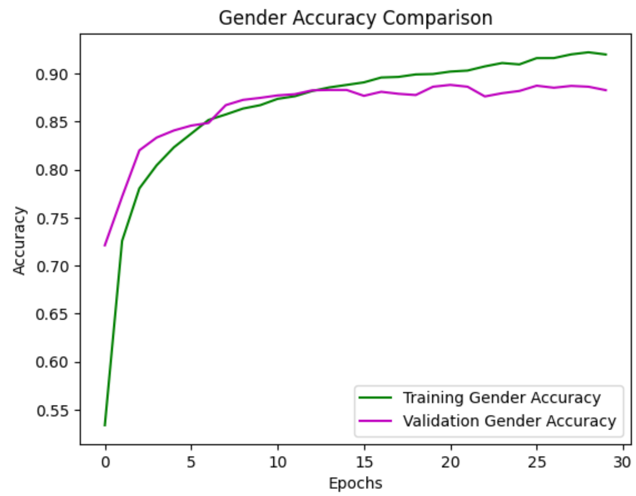

Figure 4: Gender Accuracy Graph

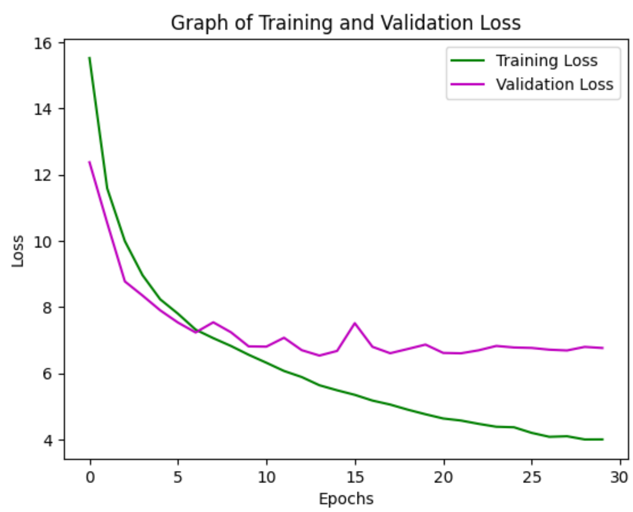

Figure 5: Gender Loss Graph

percent gender prediction accuracy and an age prediction mean absolute error of 6.5, the CNN model shows promise. Training and validation loss curves show how well the model learns and generalizes across epochs.

4.1 Gender Accuracy Graph

This charts a machine learning model's accuracy in predicting gender throughout training epochs. In order to visualize the performance, a line plot is created using the accuracy data that is extracted for the training and validation sets. Epochs are represented by the x-axis, while accuracy is shown by the y-axis. Training accuracy is indicated by the green line, while validation accuracy is indicated by the magenta line. With labeled axes, the figure is captioned "Gender Accuracy Comparison". To distinguish between validation

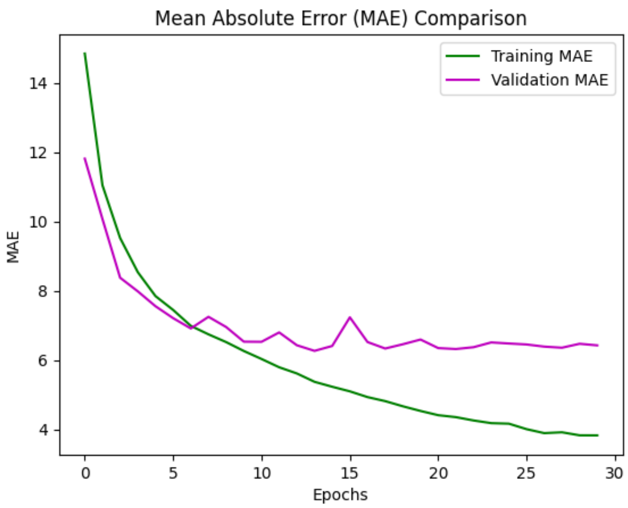

Figure 6: MAE Graph for Age

and training accuracy, a legend is provided. In general, it aids in evaluating the model's learning curve and room for development.

4.2 Gender Loss Graph

This illustrates a machine learning model's training and validation loss over training epochs. For both sets, it obtains loss data and presents it on a graph. Epochs are represented by the x-axis, and loss is represented by the y-axis. Validation loss is represented by the magenta line, and training loss is indicated by the green line. With labeled axes, the graphic is called "Graph of Training and Validation Loss". To differentiate between validation loss and training loss, a legend is used. All in all, it sheds light on how the model's loss varies during training.

4.3 MAE Graph for Age

This graphs a machine learning model's age prediction's Mean Absolute Error (MAE) throughout training epochs. It generates a line plot after extracting MAE data from the training and validation sets. Epochs are represented by the x-axis, and MAE is represented by the y-axis. Validation MAE is represented by the magenta line, and training MAE is shown by the green line. With labeled axes, the plot is titled "Mean Absolute Error (MAE) Comparison". To distinguish between training and validation MAE, a legend is supplied. All in all, it sheds light on how the model's MAE varies during training.

4.4 Results

The gender prediction accuracy in the final report is 90 percent, which is encouraging. The age prediction has a mean absolute error (MAE) of 6.5 and a standard deviation of 4, which indicates reasonable accuracy. The model continuously increases in age prediction MAE and gender prediction accuracy over training epochs, indicating efficient feature learning from facial photos. To further improve the model's performance, more optimization and investigation of different strategies are advised.

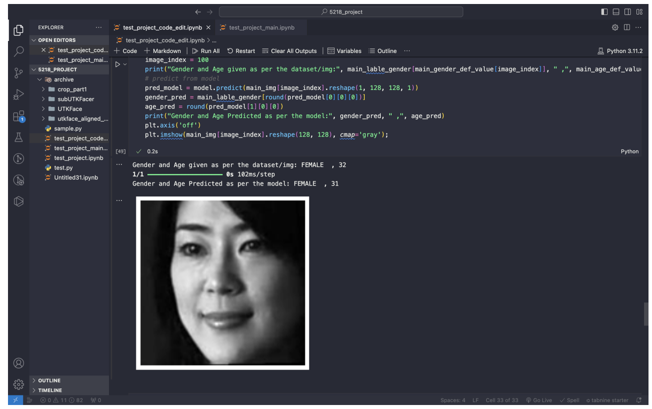

Figure 7: Final result overview

5 DISCUSSION

The CNN model performs admirably in identifying gender and age from facial images, as shown by its low mean absolute error in age prediction and high accuracy in gender classification. Nonetheless, it is imperative to recognize specific constraints that are intrinsic to the model and dataset. Bias in the dataset, which could result from things like the underrepresentation of particular age groups or gender identities, is one possible drawback. This bias may have an impact on the model's performance, possibly producing skewed or inaccurate predictions, especially for people who are underrepresented in the dataset. Ensuring the fairness and generalizability of the model across varied populations requires addressing dataset bias.

Moreover, even though the CNN model shows promise in identifying gender and age from facial images, variations in facial expressions, lighting, and other environmental factors may have an impact on the model's performance. For example, variations in lighting and face orientation between photos may add noise and impair the model's capacity to identify significant features. Subsequent investigations may concentrate on augmenting the model's resistance to these fluctuations using methods such as data augmentation, which entails faking variations in training data to enhance model generalization. Furthermore, adding more varied datasets with a wider range of demographic features might improve the model's capacity to generalize to various populations and lessen the negative effects of dataset bias on model performance. Future revisions of the CNN model can enhance its accuracy and practicality in real-world circumstances by tackling these constraints.

ACKNOWLEDGMENTS

We sincerely thank all the people and institutions whose support helped to make this research feasible. We are especially grateful to the authors of the above papers for their insightful contributions and pioneering work in the fields of deep learning and facial image analysis. We also thank the UTKFace collection for supplying the study's dataset. We also value the assistance and direction our mentors and colleagues gave us during the study process. Lastly, we would like to sincerely thank our family and friends for their continuous support and comprehension.

REFERENCES

(1) Levi, Gil, and Tal Hassner. "Age and gender classification using convolutional neural networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2015.

(2) Rothe, Rasmus, Radu Timofte, and Luc Van Gool. "Deep expectation of real and apparent age from a single image without facial landmarks." International Journal of Computer Vision 126.2-4 (2018): 144-157.

(3) Zhang, Kaipeng, et al. "Joint face detection and alignment using multitask cascaded convolutional networks." IEEE Signal Processing Letters 23.10 (2016): 1499-1503.

(4) Wang, Shiqi, et al. "Deeply Learned Face Age Estimation from Partially Labelled Data." Proceedings of the IEEE International Conference on Computer Vision. 2019.

(5) Escalera, Sergio, et al. "ChaLearn Looking at People 2016: Apparent Age and Cultural Event Recognition Databases and Results." IEEE Transactions on Pattern Analysis and Machine Intelligence 40.8 (2018): 1912-1918.

(6) Antipov, Grigory, et al. "Effective training of convolutional networks for face attribute classification." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.