Grenze International Journal of Engineering and Technology, June Issue

Age, Gender and Emotion Detection using Convolutional Neural Networks (CNNs)

Kevin Varghese[1] and Ramchand Hablani[2]
[1]Student of Master of Technology, Department of Computer Science & Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, India
Email: varghesekv@rknec.edu
[2]Associate Professor, Department of Computer Science & Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, India
Email: hablanir@rknec.edu

***Abstract*—Convolutional Neural Networks (CNNs) are used in this study to determine age, gender, and emotion in face analysis. Using a variety of datasets for training, the age diagnosis component accurately predicts age categories by identifying small changes in face features associated with ageing using CNNs' hierarchical feature learning. Large datasets were used to train the gender detection module, which helps facial recognition and marketing applications by automatically extracting gender-specific traits for accurate gender identification. Sentiment analysis and human-computer interaction benefit from the emotion detection framework's improved emotion classification, which is achieved by capturing temporal and spatial correlations in facial expressions. CNNs are effective for face analysis, as evidenced by the model's adaptability and robustness across a range of demographics, as well as by testing and training on a variety of datasets. This paper provides a thorough framework for age, gender, and emotion recognition based on CNN, which holds promise for breakthroughs in a number of sectors depending on facial analysis.**

***Index Terms*— Convolutional neural networks (CNN), deep learning, age estimation, gender classification, emotion recognition.**

I. INTRODUCTION

A key deep learning approach called Convolutional Neural Networks (CNNs) has transformed computer vision by dramatically increasing accuracy and efficiency. Facial analysis presents interesting issues and is an area of great interest, especially when it comes to age, gender, and emotion detection. Understanding complex facial expressions is important for a variety of fields, including the social sciences, marketing, healthcare, emotional intelligence, and human-computer interaction. It also has implications for demographic profiling. In order to clarify the intricacies and practical uses of these interrelated facets of facial analysis, this paper explores the complicated field of age, gender, and emotion recognition using CNNs. With advances in technology comes an increasing need for intelligent systems that can comprehend and react to human expressions. Taking into account the complex interactions between age development, gender differences, and the range of human emotions requires advanced computer models that can identify subtle face traits.

Grenze ID: 01.GIJET.10.2.624_1

II. LITERATURE REVIEWS

AM Abu Nada, et al[1]. This research suggests a CNN-based approach for predicting age and gender. It probably explains the CNN model's architecture, the dataset that was used for validation and training, and the evaluation criteria that were employed to gauge the model's effectiveness. Insha, Rafique[2], and others CNN-based age and gender prediction is also covered in this work. It could offer an alternative solution to the issue, perhaps involving modifications to the CNN architecture or training process. It probably contains comparisons with other approaches and experimental outcomes. Lim, Sooyeon[3] - In this study, they employ a CNN-based facial recognition method to estimate age and gender. It probably includes information on the particular CNN architecture that was employed, any picture pre-processing methods that were utilized, and the model's overall accuracy and computational efficiency. Overall, by examining CNNs' ability to correctly identify age and gender from single user photos, these articles advance the science of computer vision.Ueda, M., et al.[4] - This study, which focuses on the use of sophisticated imaging methods for age prediction, proposes an age estimation approach employing 3D-CNN from brain MRI images. Shanthi, N., et al[5] - Talks about age and gender detection with deep CNNs, perhaps introducing new techniques or increasing accuracy. Research into identifying emotions from facial expressions is indicated by Jaiswal, Akriti, et al.[6] exploration of facial emotion recognition using deep learning. Mehendale, N[7]. - Advances the field of affective computing by describing the use of CNNs for face emotion identification. Acheampong, Francisca Adoma, et al[8]. - Talks about text-based emotion recognition, probably emphasizing the difficulties and possibilities involved. In the context of pandemic-related applications, Golwalkar, Rucha, and Ninad Dileep Mehendale's [9] presentation on age identification with face mask using deep learning may be of interest. A emphasis on age estimate in social media situations is indicated by Alghieth, Manal, et al.'s[10] proposal for smart age detection for social media. Jia Guo[11] - Talks about a deep learning method for text analysis to identify human emotions from large amounts of data, proposing improvements in text analysis for emotion identification.The first portion of our work involves age detection with CNNs, which is challenging since ageing is characterized by subtle modifications to skin tone, face structure, and distinctive features. These complex designs are often hard for traditional technologies to effectively portray. CNNs, on the other hand, are exceptionally good at detecting age-associated variations in pictures of faces because of their hierarchy feature learning capabilities. The age detection model wants to offer a more precise and trustworthy framework for age identification by surpassing conventional methods by training on a variety of datasets encompassing a range of age groups.

We investigate gender identification, a crucial component of face analysis. Because of its intrinsic complexity, traditional gender differentiation based on facial traits is difficult. CNNs provide an efficient way to classify gender by automatically identifying discriminative characteristics. Our model has been extensively trained on a variety of datasets with varying face traits in order to reach its goals of high accuracy and wide applicability.

We investigate emotion detection in our work, which complicates face analysis. Traditional approaches have issues due to the vast range of facial expressions used by people to indicate their emotions, which are context-dependent and dynamic. We provide a CNN-based framework that takes into account both spatial and temporal correlations for detecting emotions in facial expressions. With potential applications in sentiment analysis, mental health monitoring, and human-computer interaction, our approach seeks to provide a flexible and reliable solution for emotion identification. Training on annotated datasets that cover a wide variety of emotional states allows for this.

The reliance on age, gender, and emotion detection using CNNs is revealed. In order to offer an integrated structure for detailed face analysis, this paper offers an extensive framework that combines these elements. Age, gender, and emotion are all interconnected, which emphasizes the necessity for an integrated approach to fully understand the intricate interactions between these factors and improve our understanding of human expressions.

Besides theory, our study has applications. Personalized recommendations for content based on age demographics, improved human-computer interaction, and the development of intelligent systems that respond to users' emotions are among the many real-world applications for this technology. While gender detection has consequences for security and surveillance, emotion recognition can help studies on mental health and well-being.

In an effort to improve facial analysis and further computer vision and artificial intelligence research, this study investigates age, gender, and emotion identification using CNNs. The integrated framework demonstrates how well CNNs perform in intricate face analysis and provides opportunities for cutting-edge technical solutions that take into account the subtleties of human emotion. We encourage readers to investigate how facial analysis is developing and influencing many industries as we dive into these elements.

III. Research Methods

A. Proposed Methodology Overview

While there are other models that can be used to predict age, gender, or emotion alone, our model is one of the few that can accurately predict age, gender, and emotions using CNN. We intend to use the Cohn-Kanade Dataset and the UTK Face Dataset, two distinct datasets, to accomplish this. These datasets were selected due to their extensive coverage of age, gender, and emotion annotations on facial images. These datasets will be used to train a reliable CNN model that can perform precise and complex facial analysis.

B. Dataset Selection

Our method is based on the use of UTK Face and Cohn-Kanade databases. Because the UTK Face Dataset contains a wide variety of facial photos from various age groups, it is perfect for training age detection algorithms. Similar to this, the Cohn-Kanade Dataset's large variety of facial expressions makes it ideal for training emotion recognition algorithms. We guarantee a thorough training environment that tackles the challenges of age, gender, and emotion categorization by merging these datasets.

C. Labelling Classes

For our methods to work, the classes for Age, Gender, and Emotion must be accurately labelled. Precise age, gender, and emotion labels will be applied to every face image in the collection. The CNN model is trained using this annotated data as the ground truth, which enables it to discover the complex correlations and patterns unique to each categorization job.

D. Processing Pipeline

Our process flow is explained by a pipeline of processing stages. Firstly, picture inputs are obtained from pre-selected databases. Preprocessing is applied to these inputs in order to guarantee consistency and improve the model's capacity to recognize pertinent characteristics. Grayscale picture conversion is an important part of the preprocessing stage. This decision is based on the fact that computations are more efficient when using greyscale pictures since they have a lower resolution than colourful ones. The requirement to lessen the computing weight of ensuing analysis justifies this change.

E. Analysis and Result Presentation

The CNN model conducts an in-depth analysis of the greyscale pictures following preprocessing. The method, which was trained on labelled data sets, produces precise forecasts about the age, gender, and emotion of each input image based on its collected properties. The analysis that ensues shows the way the model can classify data.

F. Benefits of Greyscale Conversion

The strategic choice to convert input images to greyscale greatly enhances the efficacy of our suggested methodology. By lowering the dimensionality of the data, greyscale images reduce the computational burden and free up the CNN model to concentrate on important features without being overtaken by colourful, high-resolution data. This decrease in complexity speeds up training and inference procedures while also improving computational efficiency.

In conclusion, our suggested approach improves upon the drawbacks of the current model for the detection of Age, Gender, and Emotion using CNNs. We seek to improve the accuracy and precision of facial analysis by combining two different datasets, carefully labelling classes, and a streamlined processing pipeline that includes greyscale conversion. Using CNNs that have been trained on these annotated datasets allows our methodology to produce more accurate and nuanced results.

Future directions for this research include training the model iteratively on more diverse datasets in order to continuously refine it. In addition, it is imperative to take into account practical implementations, moral ramifications, and possible partialities within the dataset to guarantee the conscientious and impartial implementation of the created model. Our suggested methodology provides the groundwork for advanced facial analysis systems in the future, which can be useful for a variety of applications ranging from personalized content recommendations to human-computer interaction.

Flow Chart of the Process is shown in figure 1 below: -

Dataset used for Age Detection: -

UTKFace Dataset with Facial Age

Image Resolutions: 200*200

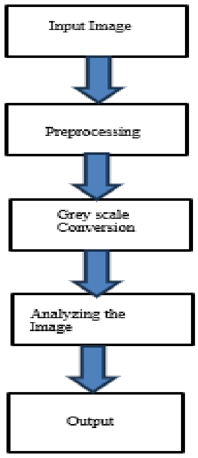

Figure: 1

Figure 1: Flowchart of the Process

For Age detection we have used the following Class labels: -
UTKFace Dataset with Facial Age Image Resolutions: 200*200 For Age detection we have used the following Class labels, and it is specified in the Table I.

TABLE I. AGE RANGES AND CLASS LABEL

Class Label	Age Ranges
0	1-2
1	3-9
2	10-20
3	21-27
4	28-45
5	46-65

G. Process of Age Detection

Here's a methodical explanation of the process for age detection:
Data Augmentation: - Augment the dataset by rotating the photos at various angles. This artificially expands the training dataset, aiding in the learning of robust features and improving model generalization.
Dataset Size After Augmentation:
After augmentation, the training dataset contains 234,000 images.
Preprocessing: - Convert the images to greyscale to remove RGB color information. Greyscale images reduce resolution and improve computational efficiency, simplifying computational procedures and reducing analysis complexity.
Age Classification Schema: - Define classes based on age ranges for age classification.
CNN Architecture Design:
Input Layer**:** Include a Conv2D layer with 32 filters paired with an AveragePooling2D layer for initial feature extraction.
Intermediate Convolutional Blocks: Use three pairs of Conv2D and AveragePooling2D layers with increasing filter counts (64, 128, 256) to capture complex features progressively.
Global Average Pooling Layer: Add a GlobalAveragePooling2D layer to reduce spatial information.
Dense Layers: Include two dense layers, the first with 132 nodes, to learn complex patterns.
Output Layer: Use an output dense layer with seven nodes for the defined age-range classes, with SoftMax activation for age prediction.
Training Process:
Train the model over 60 epochs to allow it to adjust parameters and pick up complex patterns for accurate age range prediction.
Conclusion:
The methodology, including the CNN architecture, data augmentation, greyscale conversion, and extended training, aims to achieve high accuracy in age range prediction.

Gender Detection: -
Dataset used: - UTKFace
Resolution: 200 * 200
Class Labels: - Male or Female
Gender Detection Process: -
Here's a methodical breakdown of the described process:
Extract Gender Labels and Import Datasets:
Extract gender labels from the file names of the UTK Face Dataset.
Import the Gender Dataset from the specified Drive location.
Combine Datasets and Begin Preprocessing:
Combine the UTK Face Dataset and the Gender Dataset.
Start preprocessing the images by converting them to grayscale.
CNN Architecture Design:
Input Layer: Use a 32-filter-equipped Conv2D layer with a paired MaxPooling2D layer for initial feature extraction.
Intermediate Convolutional Blocks: Include three pairs of Conv2D and MaxPooling2D layers, gradually increasing the filter count (64, 128, 256) to capture complex features.
Dense Layer: Add one Dense layer with 128 nodes to identify and combine crucial patterns.
Output Layer: Use an output Dense layer with three nodes corresponding to the three gender classes, with SoftMax activation for probability distribution.
Training the Model:
Train the model over a specified number of epochs (denoted as "X") to learn discriminative features and optimize parameters for accurate gender classification.
Model Evaluation:
Evaluate the trained model's performance on a separate validation dataset to assess its accuracy and adjust parameters if needed.
Conclusion:
The combination of gender-related datasets, thoughtful CNN architecture design, and grayscale image conversion enhances the model's ability to accurately predict gender based on facial features.
This methodical approach ensures a systematic and clear process for implementing the gender classification task using CNNs.
Emotion Detection: -
Dataset: - CK+
Here's a clear and concise explanation of the methodical approach for emotion detection:
Dataset Preparation:
Use the CK+ dataset, which contains various facial expressions labelled with numerical codes representing emotions.
Preprocess the images by converting them to grayscale for consistency and reduced computational complexity.
Emotion Labelling:
Define emotion labels based on task requirements, such as 'Negative' (anger, contempt, disgust, fear, sadness) and 'Positive' (happy, surprise).
Alternatively, use a finer granularity if needed, treating each emotion category separately.
CNN Architecture Design
Input Layer: Start with a 32-filter Conv2D layer followed by a MaxPooling2D layer for feature extraction.
Intermediate Convolutional Blocks: Include three pairs of Conv2D and MaxPooling2D layers, progressively increasing filter counts (64, 128, 256) to capture complex features related to facial expressions.
Dense Layer: Add a Dense layer with 128 nodes to identify and combine crucial patterns.
Output Layer: Use an output Dense layer with three nodes for the defined emotion categories, with SoftMax activation for probability distributions.
Training the Model:
Train the model over a specified number of epochs to learn discriminative features and optimize parameters for accurate emotion classification.
Conclusion: The methodology ensures a systematic approach to emotion detection, leveraging a well-designed CNN architecture, grayscale image conversion, and appropriate emotion labelling for effective analysis of facial expressions in the CK+ dataset.

IV. Result and Discussion

The The age, gender, and emotion models are incorporated after the pretraining the stage, and the resulting files are kept in the HDF5 (h5) file format. In this phase, the precision of the models in predicting age, gender, and emotions is evaluated. An example of this would be to take an image of Kartik Aryan and evaluate its efficiency using the integrated model.

Figure 2: Input Image for testing

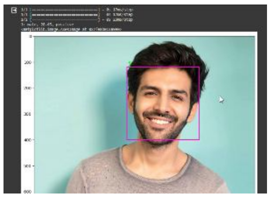

Figure 3: Output of the detection

The combined model correctly evaluates Kartik Aryan's picture in the final prediction phase, yielding the following results: a positive emotion, a male gender, and an age range of 28 to 45. This illustrates how well the model processes face characteristics and generates accurate multi-dimensional predictions.

V. Conclusions

In conclusion, Convolutional Neural Networks (CNNs) have revolutionized computer vision, particularly in the complex realm of facial analysis. This study delves into age, gender, and emotion detection, highlighting the intricate interplay between these aspects. CNNs prove adept at identifying age-related nuances, gender differences, and dynamic emotional expressions, offering a robust framework for facial analysis. The integration of age, gender, and emotion detection opens avenues for diverse applications, from personalized content recommendations to improved human-computer interaction and mental health monitoring. By combining these elements into a single system, we gain deeper insights into human expressions and behavior. Looking ahead, the potential for CNNs in facial analysis is vast, promising advancements in various fields. Continued research in this area will likely lead to further innovations, enhancing our understanding and utilization of facial analysis across industries.

References

[1] Nada, AM Abu, et al. "Age and gender prediction and validation through single user images using CNN." Int. J. Acad. Eng. Res. (IJAER) 4 (2020): 21-24.

[2] Rafique, Insha, Awais Hamid, Sheraz Naseer, Muhammad Asad, Muhammad Awais, and Talha Yasir. "Age and gender prediction using deep convolutional neural networks." In 2019 International conference on innovative computing (ICIC), pp. 1-6. IEEE, 2019.

[3] Lim, Sooyeon. "Estimation of gender and age using CNN-based face recognition algorithm." International journal of advanced smart convergence 9.2 (2020): 203-211.

[4] Ueda, M., Ito, K., Wu, K., Sato, K., Taki, Y., Fukuda, H. and Aoki, T., 2019, April. An age estimation method using 3D-CNN from brain MRI images. In 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019) (pp. 380-383). IEEE.

[5] Shanthi, N., P. Yuvasri, S. Vaishnavi, and P. Vidhya. "Gender and Age Detection using Deep Convolutional Neural Networks." In 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 951-956. IEEE, 2022.

[6] Jaiswal, Akriti, A. Krishnama Raju, and Suman Deb. "Facial emotion detection using deep learning." In 2020 international conference for emerging technology (INCET), pp. 1-5. IEEE, 2020.

[7] Mehendale, N., 2020. Facial emotion recognition using convolutional neural networks (FERC). SN Applied Sciences, 2(3), p.446.

[8] Acheampong, Francisca Adoma, Chen Wenyu, and Henry Nunoo-Mensah. "Text-based emotion detection: Advances, challenges, and opportunities." Engineering Reports 2, no. 7 (2020): e12189.

[9] Golwalkar, Rucha and Ninad Dileep Mehendale. “Age Detection with Face Mask Using Deep Learning and FaceMaskNet-9.” CompSciRN: Other Computing Methodology (Topic) (2020): n. pag.

[10] Alghieth, Manal, Jawaher Alhuthail, Kholod Aldhubiay and Rotan Alshowaye. “Smart Age Detection for Social Media.” International Journal of Advanced Computer Science and Applications (2019): n. pag.

[11] Guo, Jia. “Deep learning approach to text analysis for human emotion detection from big data.” Journal of Intelligent Systems 31 (2022): 113 - 126.

[12] Nair, Sidharth Sukumaran, Dipesh Nair, Gautam Nair, Anoop Pillai and Prof Sujith Tilak. "Detection of Gender, Age and Emotion of a Human Image using Facial Features." (2020).
[13] Sharma, Dibya Jyoti, S. Dutta and Dibya Jyoti Bora. "REGA: Real-Time Emotion, Gender, Age Detection Using CNN - A Review." International Conference on Research in Management & Technovation (2020).
[14] Kancharlapalli, Tara Pranay and Pulkit Dwivedi. "A Novel Approach for Age and Gender Detection using Deep Convolution Neural Network." 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom) (2023): 873-878.
[15] Rajeshwari, Menugu et al. "Voice Based Gender Detection." 2023 International Conference on Computational Intelligence, Communication Technology and Networking (CICTN) (2023): 288-291.
[16] Alamri, Hayat Rasheed Hamzah et al. "Face Recognition and Gender Detection Using SIFT Feature Extraction, LBPH, and SVM." Engineering, Technology & Applied Science Research (2022): n. pag.
[17] Mao, Rui et al. "The Biases of Pre-Trained Language Models: An Empirical Study on Prompt-Based Sentiment Analysis and Emotion Detection." IEEE Transactions on Affective Computing 14 (2023): 1743-1753.
[18] Guo, Jia. "Deep learning approach to text analysis for human emotion detection from big data." Journal of Intelligent Systems 31 (2022): 113 - 126.
[19] Nandwani, Pansy and Rupali Verma. "A review on sentiment analysis and emotion detection from text." Social Network Analysis and Mining 11 (2021): n. pag.
[20] Vasavi, S. et al. "Age Detection in a Surveillance Video Using Deep Learning Technique." SN Computer Science 2 (2021): n. pag.
[21] Pascucci, Antonio et al. "Computational Stylometry and Machine Learning for Gender and Age Detection in Cyberbullying Texts." 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW) (2019): 1-6.
[22] Chowdary, Boyilla Sushma et al. "Age and Gender Detection to Detect the Manipulated Images using CNN." 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) (2023): 1187-1192.
[23] Kumar, Sandeep et al. "Face Spoofing, Age, Gender and Facial Expression Recognition Using Advance Neural Network Architecture-Based Biometric System." Sensors (Basel, Switzerland) 22 (2022): n. pag.
[24] Saxena, Aryan et al. "Gender and Age detection using Deep Learning." 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (2021): 719-724.
[25] Meinedo, Hugo and Isabel Trancoso. "Age and gender detection in the I-DASH project." ACM Trans. Speech Lang. Process. 7 (2011): 13:1-13:16.