Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.
Digital Object Identifier XX.XXXX/YYY.20XX.ZZZZZZZ

Gender and Age Recognition using CNN and ResNet

SUMIT KUMAR[1] DR. VISHNUVARTHANAN
[1] *VIT Bhopal University, India (E-mail: info@vitbhopal.ac.in)*

ABSTRACT This study explores the fascinating realm of predicting age and gender through facial images. Imagine a computer looking at a picture of your face and making an educated guess about how old you are and whether you're male or female.
This research focuses on predicting age and gender using advanced deep learning models, namely ResNet and Convolutional Neural Networks (CNNs). These models are trained to analyze facial images and extract features that help determine a person's age and gender. The study uses a dataset with diverse facial images, each labeled with age and gender information.
Results from this investigation not only enhance our understanding of the intricate relationship between features and demographic information but also have practical applications in various domains, including human-computer interaction, content recommendation systems, and demographic analysis.

INDEX TERMS Gender Recognition, Age Classification, Deep Learning, Resnet50, Convolutional Neural Network Model, Image classification

I. INTRODUCTION

The human face holds a lot of information about a person, like who they are, how they feel, their gender, and age. As people get older, their faces change, and these changes are crucial in how we communicate without words. Figuring out someone's age and gender just by looking at their face is a crucial job for computers and machines.
Guessing someone's age and gender using computers has become a big deal lately. It's super useful in real life for things like showing ads to specific groups of people, solving crimes, watching over places, finding stuff online, and making devices that understand us better.

i) Age Recognition: Age recognition, also known as age estimation uses deep learning techniques to get the relevant features from images or videos and make predictions about the individual's age.
The deep learning technique analyzes attributes like wrinkles, facial lines, skin texture, and sometimes hair color to estimate the person's age.

ii) Gender Recognition: Gender recognition is the process of identifying and categorizing the individual's gender based on facial images. In some cases, it may recognize the non-binary or transgender faces, but this system classifies the individuals as either male or female. Gender recognition algorithms use facial features like facial shape, hair length, and sometimes clothing cues to make gender predictions.

A. WHY IS THIS PROJECT REQUIRED?

The main aim of this research paper is to create a smart computer vision tool that can look at a picture of a person and tell the age and gender accurately based on various features like eyes, lips, beard, and hair. This might sound like a fun idea, but it can be super useful in many areas of our lives. [7]

B. WHERE CAN WE USE IT?

1) **Voting Eligibility:** The age detection component aims to prevent individuals below the legal voting age from participating in elections. This could help maintain the integrity of the electoral process by ensuring that only those who meet the age requirements can cast their votes.
2) **Security Systems:** Imagine security cameras in public places like malls or airports. This deep learning tool can help these cameras figure out how old people are and whether they are male or female.
3) **Entertainment:** For various streaming service, like Netflix and Amazon Prime. It could help in suggesting movies or shows that match your age and gender. So, you get recommendations that suit you better.
   In video games, the tool could make the game adapt to your age and gender, making it more enjoyable and tailored just for you.
4) **Social Media:** On platforms like Instagram or Facebook, This phenomenon could look at users' profiles

and estimate their age and gender. Companies could use this info to understand who is using their platform and create ads that appeal to specific age groups and genders.

5) **Healthcare:** Analyzing the age and gender of people might help spot health issues or risks in hospitals. For instance, it could notify a nurse if an older person seems to be in trouble.[7]

II. LITERATURE SURVEY

1) "Gender Recognition and Age Approximation using Deep Learning Techniques"[1]. In their research paper, Shubham Patil, Bhagyashree Patil, and Ganesh Tatkare present a study on gender recognition and age approximation using deep learning techniques.
   The authors aim to leverage Haar Cascade for gender prediction and Caffe-based deep learning for age estimation from facial images. The authors categorize age into eight groups: (0–2, 4–6, 8–13, 15–20, 25–32, 38–43, 48–53, 60+). This granularity allows for a more detailed understanding of the age distribution. The research concludes that while the results obtained are not perfect, they exhibit promise for future research. The authors acknowledge the imperfections in the current model and outline plans for improvement.
2) "HUMAN AGE AND GENDER PREDICTION USING DEEP MULTI-TASK CONVOLUTIONAL NEURAL NETWORK"[2]. This research utilizes the Adience Benchmark face dataset, consisting of 17,603 images with variations in age and gender.
   Age classes are categorized into 10 groups, and gender is classified into two types (male and female). The CNN is trained using two learning methods: Single-Task Learning (STL) and Deep Multi-Task Learning (DMTL). The proposed CNN using STL achieves 80.11
3) "Age and Gender Detection"[3]. This study employs the caffenet CNN arcitecture, which is trained on 1250 images (650 females and 600 males). The developed system accurately detects faces and predicts gender and age ranges based on the implemented model.
4) "Age and Gender Classification - A Proposed System".[4] In this proposed system, David Pereira da Silva introduces an innovative approach to age and gender classification using advanced neural network architectures. The primary methodologies employed in the system are GoogLeNet for face detection and Wide Residual Networks for age and gender classification. In summary, David Pereira da Silva's proposed system combines the strengths of GoogLeNet and Wide Residual Networks to create a comprehensive approach to age and gender classification, with a particular emphasis on achieving high accuracy in face detection using OpenCV's GoogLeNet.

III. DATASET DESCRIPTION

The dataset is the fundamental component for training and testing models, and having various age and gender recognition datasets is incredibly helpful for researchers. Each dataset comes with unique features.

This study uses the UTKFace dataset which comprises 16,000 facial images for the training phase of age and gender recognition. The dataset covers the ages of persons ranging from 1 to 99. It also includes images of both males and females.

Each file name in the dataset is encoded with the following two labels:

- Age, and
- Gender.

Overall, the UTKFace dataset serves as a valuable resource for advancing research in the fields of computer vision, machine learning, and facial analysis. Its comprehensive annotations make it suitable for training models to predict age and gender accurately.

IV. METHODOLOGY

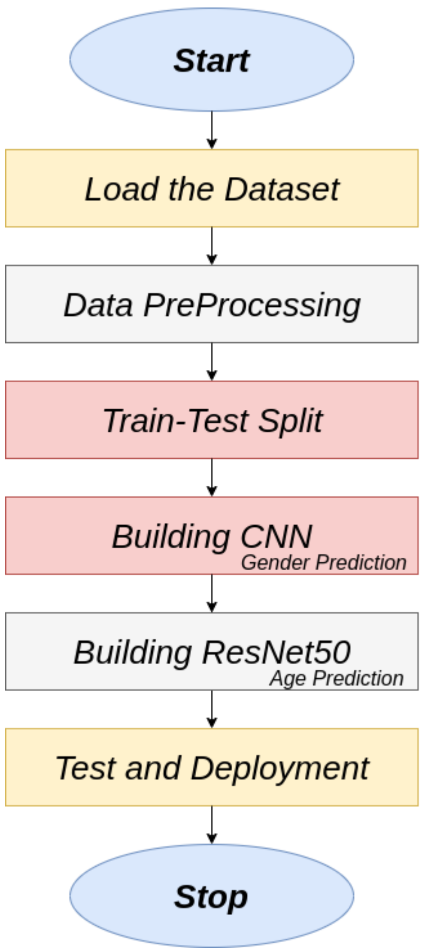

FIGURE 1. Flowchart of

In this section, we are going to explain the methods used in the research and outline the steps taken to achieve the objective.

The methodology used for age and gender recognition using deep learning consists of the following various steps:
1. Load the Dataset
2. Data Preprocessing
3. Splitting the Dataset
4. Training the models
5. Test and Deployment

A. LOAD THE DATASET:

- For this research study, it is necessary to load the image files from the UTKface dataset.
- The loading of the dataset is done by importing the path function from the *pathlib* library. The path() function is mainly used to access the directory containing the face image files. And, the path.glob function retrieves all the filenames from the directory that matches the *.jpg.chip.jpg* extension.

B. DATA PREPROCESSING:

Data Preprocessing is the second phase whose goal is to clean and organize the collected data so that it can be effectively used for analysis and training of the models.

The below steps delineate the data preprocessing phase of the research study:

1) Extracting Information:

As specified in the dataset description, the age and gender information is encoded in the filenames. So, the split() function is utilized which separates the age and gender labels from each filename.
Subsequently, the new dataframe is created to store the image path, age labels, and gender labels.

2) Loading, Resizing, and Converting the Image into a NumPy array:

The $load\_img$ function loads the image path from the dataframe and the img.resize() function of the PIL library resizes each image into (128X128) pixels. Then, the resizing image is converted into a NumPy array and appended in X variable which is used as an input feature for training the model. Conversion of an Image into a numPy array is a common step when working with images in machine learning for numerical operations.

3) Converting Age and Gender Label into Array:

The np.array function creates a numpy array for the age and gender values from the dataframe and stores them in age and gender labels respectively.

C. SPLITTING THE DATASET:

The way of splitting the image dataset in training and testing is the most promising factor in deep learning. After data preprocessing, it is the next phase where the data is divided into training and testing sets. The training set is that part of our data that helps to train the models, while the testing set is mainly used to evaluate or validate the performance of the models.

In this research paper, the $train\_test\_split()$ function splits the dataset into **80:20** respectively.

After splitting, the trained parameters (X_train and Y_train) to train the models on training data and evaluate the performance on the testing data for both age and gender recognition separately.

D. TRAIN THE MODELS

Training a model for age and gender detection involves creating and optimizing the deep learning model to predict the age and gender of individuals based on input data, especially images.

1) CNN or Convolutional Neural Network:

Convolutional Neural Network (CNN) is a class of deep learning algorithms that is specifically designed for processing and analyzing visual (images and videos) data. CNN model is used for gender prediction as it is effective for image-related projects because it automatically learns the hierarchical features and spatial relationships within the images.

The following three convolutional 2D layer is used to detect the accurate gender of a person:

1) **Convolutional Layer:** Convolutional layers are translant-invariant which is crucial for facial recognition because the facial features (such as eyes, nose, and mouth) may appear in different positions across images. In this research study, convolution layers apply the filters to the input image, enhancing the model efficiency and minimizing the risk of overfitting.
a. Conv2D(32, (3, 3), activation='relu', input_shape=(128, 128, 3)):
The first convolutional 2D layer uses 32 filters or kernels, each with a size of 3x3. Each filter is responsible for detecting patterns and learning the diverse set of features from the input 128x128 image.
The use of Relu (Rectified Linear unit) is important for this study because the images introduce non-linearity to the model which helps in capturing more complex patterns.
b. Conv2D(64, (3, 3), activation='relu'):
The second convolutional 2D layer with 64 filters preprocesses the sophisticated features returned by the first layer. This layer also uses the Relu activation function for defining non-linearity transformations.
c. Conv2D(128, (3, 3), activation='relu'):
The last and third convolutional 2D layer with 128 filters, each of size 3x3, is used to extract the high-level features from the features learned by the second convolutional layer.
The use of adding a second and third convolutional layer with a high number of filters enhances the

model's capacity to detect intricate patterns within the image data.

2) **Pooling Layer:** The pooling layer is used to reduce the spatial dimensions of the feature map while selecting the necessary information from the images. MaxPoolin, AveragePooling, and MinPooling are the three types of pooling layer.
This research study uses the three MaxPooling 2D layers with a 2x2 pool size. Each pooling layer reduces the computational load by maintaining the highest value within each 2x2 square region of the feature map. The Relu activation function is connected with each pooling layer to learn the complex connections for capturing the facial characteristics associated with the gender of a person.

3) **Fully Connected Layer:** This layer is typically used in the final part of the network for making predictions. They connect all the neurons from the previous layer to the current layer and are responsible for the final classification or regression output.
a. Dense(128, activation='relu'):
This is the first fully connected layer in the model because it applies the Relu (rectified linear unit) activation function with 128 neurons to capture the various features extracted by the last convolutional layer.
$$ReLU(x) = max(0, x)$$
b. Dense(1, activation='sigmoid'):
This final output layer uses the single neuron with the sigmoid function which is used for binary classification problems such as gender prediction (Male as 1 or Female as 0).
$$Sigmoid(x) = 1/(1 - e.pow(x))$$
where, x is the sum of inputs to the neuron, and e is the base of the natural algorithm.
The sigmoid function transforms the neuron output into the probability score. A probability score close to 1 indicates a high probability, while a probability score close to 0 indicates a low probability. The high probability score predicts the positive instance i.e., male, and the low probability score predicts the negative instance i.e., female.

V. RESULTS

In this section, we present the outcomes of our study on age and gender recognition experiments using facial images. The evaluation was conducted on 16000 images of the UTKface dataset, which compromises the male and female facial images from 0 to

Gender Recognition Results: For Gender prediction, the model achieved an accuracy of 86.7 .

Test Accuracy = Number of Correct Predictions / Total number of samples

The table specifies precision, recall, and F1 score for males and females. The model has a relatively high accuracy, indicating that it performs well on the test set. The Figure2 shows the confusion matrix which provides a more detailed view and showing the balance between correct and incorrect predictions.

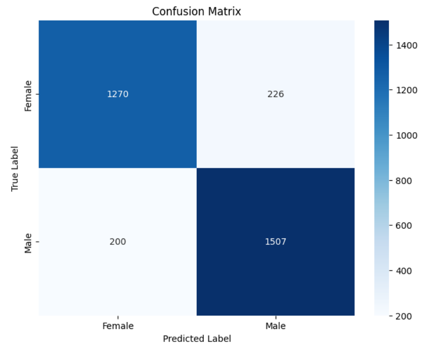

FIGURE 2. Confusion matrix for Gender prediction

Age Recognition Results:
For Age prediction, this study specifies the MAE (Mean Absolute Error) function for evaluation which is utilized to assess the average magnitude of errors between predicted and actual age values in the model.

$$\text{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

In this MAE equation for Age prediction, Yi is the Predicted age, and xi the Actual age.
The Resnet50 model yields a Mean Absolute Error (MAE) of 4.5, which suggest that the lower MAE value indicates that the model's predictions are closer to the true age values.

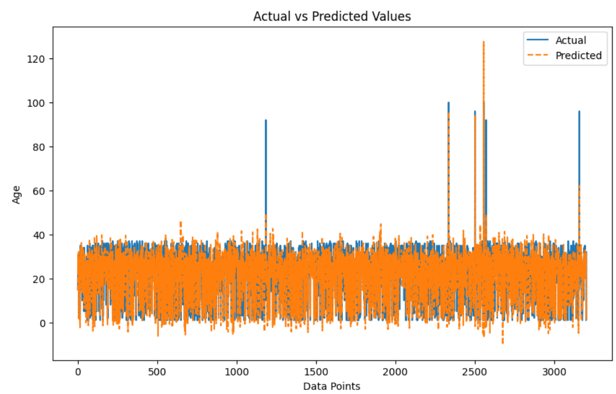

FIGURE 3. Line graph for Age prediction: It shows the variation between actual and predicted values.

VI. CONCLUSION

In conclusion, this project creates a smart computer vision tool which is capable to determine the age and gender of an individual based on the facial features such as eyes, lips,

FIGURE 4. Enter Caption

FIGURE 5. Enter Caption

beard, and hair.
The methodolgy specifically explains the systematic approach by covering the key stages such as data loading, data preprocessing, dataset splitting, and model training. The convolutional layers, pooling layers, and fully connected layers in CNN model play a crucial role in accurately predicting gender. Overall, this study not only enhance the scientific understanding of age and gender prediction but also have practical implementations with real-world applications across diverse domains.

VII. FUTURE ENHANCEMENT

We will look into a more complex CNN architecture and a more reliable image processing approach for estimating exact ages for future work. We can use this project for voting eligibility.

VIII. REFERENCES

1) Patil, Shubham, Bhagyashree Patil, and Ganesh Tatkare. "Gender Recognition and Age Approximation using Deep Learning Techniques." y (IJERT) 9.4 (2020).
2) Al-Azzawi, D. S. (2019). Human Age and Gender Prediction Using Deep Multi-Task Convolutional Neural Network. Journal of Southwest Jiaotong University, 54(4).
3) Wadkar, Parth Anand, Varad Wangikar, and Pramod Kanjalkar. "Age and Gender Detection."
4) Silva, David Pereira da. Age and gender classification: A proposed system. MS thesis. 2019.
5) Mansour, A.I. and Abu-Naser, S.S., 2022. Classification of Age and Gender Using ResNet-Deep Learning.
6) 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.
7) Afnan, Syed Mohammed, N. K. Shreyas, K. C. Prajwal, Gagan Kumar, and K. Swarnalatha. "AGE AND GENDER DETECTION."