RESEARCH
O&G Forum 2024; 34 – 3s: 2633-2644

GENDER & AGE DETECTION USING GEOMETRIC-BASED & APPEARANCE-BASED DEEP LEARNING APPROACH ON REAL-TIME VIDEO CROWD

Dimpal A. Kapgate[1], Prof. Abhimanyu Dutonde[2], Prof. Pragati Patil[3]
[1] Student, Department of Computer Science Engineering, Tulsiramji Gaikwad Patil Collage of Engineering & Technology, Nagpur, India. dimpalkapgate2000@gmail.com
[2] Professor, Department of Computer Science Engineering, Tulsiramji Gaikwad Patil Collage of Engineering & Technology, Nagpur, India. abhimanyudutonde@gmail.com
[3] Professor, Department of Computer Science Engineering, Tulsiramji Gaikwad Patil Collage of Engineering & Technology, Nagpur, India. pragatimit@gmail.com

Abstract
Gender Classification method using biometric features, we access from the face of a human. That means the face of the human being is an important biometric feature. As well as Age Estimation is also a challenging task in object recognition. Due to differences in head attitude, scale fluctuation in face photos, varying lighting conditions, occluded faces, and noisy face images, identification of faces and classification according to gender have become extremely challenging tasks. Two feature extraction techniques have been used: appearance-based and geometric-based techniques. Variations in lighting circumstances, head posture, facial expressions, partial occlusion by hats or spectacles, and camera quality can all affect how well the gender classification algorithm performs in terms of categorization rate. Therefore, it is ideal to have an algorithm that can withstand changes in illumination, position, occlusion, and emotion. The study paper's primary goal is to examine an automatic gender detection technique employing Deep Convolution Neural Networks (D-CNNS) on facial photos. The study paper's second goal is to evaluate an algorithm for age detection utilizing a novel geometric-based approach and appearance-based method based on the image, taking into account factors such as significant illumination variance. This research paper analyzes gender categorization for computer vision applications and presents improvements in gender classification & age detection accuracy. It has been discovered that texture descriptors perform a better job of classifying gender & Age detection than edge descriptors. The classification of gender that combines more characteristics provides greater accuracy than the other methods covered in this study. The intensity, shape, and texture of the face image are the features that are employed to get the best accuracy. As a result, the research project offers a thorough analysis of gender classification & Age Detection with improved accuracy when compared to the body of existing literature. Future studies in this area may involve developing a novel method for classifying gender in addition to identifying the Transgender ethnicity of voice captured in unrestricted environments.
Keywords: Deep Convolutional Neural Networks (D-CNNs), deep learning, Artificial intelligence (AI), Computer Vision, Gender classification, Age detection, Geometric-based model, Appearance-based model.

LITERATURE

[1] A.A. Vogan et al, the research discussed in the paper focuses on the issues that developing countries deal with as a result of an aging population and age-related cognitive decline. A study that examines the present and future applications of Artificial Intelligence (AI) & Human-Robot Interaction (HRI) as intervention strategies for cognitive training finds that human-like & pet robots show promise in improving cognition & emotional health markers, especially when equipped with AI & Deep Learning capabilities. ***[2] S.T.Rahman et al***, positive outcomes are also obtained by testing on other databases. This paper addresses the increased demand for online platform integration and offers a thorough overview of AI-based methods for Facial Expression Recognition (FER), spanning a variety of age groups. ***[3] V. Carletti et al,*** this study summarizes the progress made in the last six years in estimating age from faces, with a special emphasis on deep learning techniques. The analysis encompasses multiple facets, such as network design, datasets, preprocessing, and the incorporation of supplementary data like gender, race, and facial expressions. The influence on system performance is highlighted in the conclusion, along with the continued difficulties in the field.

[4] M. Rajput et al, the study presented a method for estimating age and gender from iris images using CNNs that were trained for age recognition and gender estimation using modified Deep Convolutional Neural Networks (D-CNNs) models such as GoogLeNet & AlexNet. Gender classification outperformed age prediction because there were more subjects in gender classes. The suggested system used AlexNet to classify gender with an accuracy of 89.34% and to approximate a person's age using iris features. ***[5] Fu, G. Guo et al***, learning parameters and weights for the input, hidden, and output network layer (loss through

BPNN) is comparatively easy. They have the same number of hidden layers between the input and output layers but fewer parameters than fully linked multi-layer perceptron modeling neural networks. Consequently, Deep Convolutional Neural Networks (D-CNNs) have shown remarkable performance in a wide range of applications, such as face recognition (FR), traffic signal recognition (TSR), optical character recognition (OCR), object, article, and people surveillance systems, tracking individuals in crowds (also referred to as human tracking systems or HTS), and many more. In computer vision applications, Deep Convolutional Neural Networks (D-CNNs) are also frequently utilized.

[6] E. R. Enbar et al, feature selection is a strategy used in neural network training, particularly with VGGNet, to assist the network in determining the weights used in feature extraction. Deep Convolutional Neural Networks (D-CNNs) may extract a variety of properties from an unprocessed input image with little or no pre-processing required. Deep Convolutional Neural Networks (D-CNNs) give halfway resistance and boosts for geometric deformation, transformation, and two-dimensional modifications in shapes and figures. Because existing feature extractors on the market exhibit static behavior, Deep Convolutional Neural Networks (D-CNNs) are specifically intended to address this issue. The goal of taking this method was to reduce extraneous information and simplify the task. Down-sampled pictures from high-resolution shots were specifically used to build discriminative patches. ***[7] Rodríguez et al,*** offered an alternative model with an unconstrained dataset validation stage. Their model was built on an attention mechanism and a feedforward Convolutional Neural Network (CNN) pipeline. The attention mechanism was utilized to extract more complete features of the face to learn more about certain portions of the face.

[8] Islam T. UI et al, the study illustrated the benefits of deep learning-based strategies by providing a complete overview of the age estimation from gait challenge. More research was needed to overcome the problem, expand the Gait databases, and increase accuracy by processing video data directly with deep learning models. ***[9] R. Yamashita et al***, CNNs (Convolutional Neural Networks) are one of the most common deep learning algorithms, and they can train to do classification tasks from photos. This approach, which treats gender and age range prediction as a classification problem with two classes for gender (male and female) and many classes for age ranges, will be quite useful in our case. In this study, we attempt to provide a form validator based on automatic gender classification and age identification using Convolutional Neural Networks (CNNs) and evaluate it using a dataset of real people's images to validate a gender and age range reflected in user photos. ***[10] Haing Htake, & Khaung Tin***, predicted the age of facial photographs using Principal Component Analysis (PCA). Nagesh Singh Chauhan uses CNN and OpenCV to reliably identify the gender and age of every human face in a video.

[11]* Haibin Liao et al,** the author evaluated the age of facial photographs using Convolutional Neural Networks (CNNs) using the divide-and-rule method. Convolutional Neural Networks (CNNs) were used to extract the robust properties in the photographs, followed by a suggested divide-and-rule face age estimator, and an age-based and sequential analysis of rank-based age estimation learning methods. The study emphasized the importance of accurately assessing consumer age and gender for organizational objectives. It proposed a novel method for estimating age and validating gender from user photographs using Deep Learning—specifically, Convolutional Neural Networks—and developed a web application for validation. ***[12] A. Khan et al, facial expression recognition, scene classification, and visual detection are among the many computer vision applications. After experimenting with several strategies for FER, a thorough review of the literature on emotion recognition suggests that Convolutional Neural Networks (CNNs) are an effective tool for facial expression recognition.

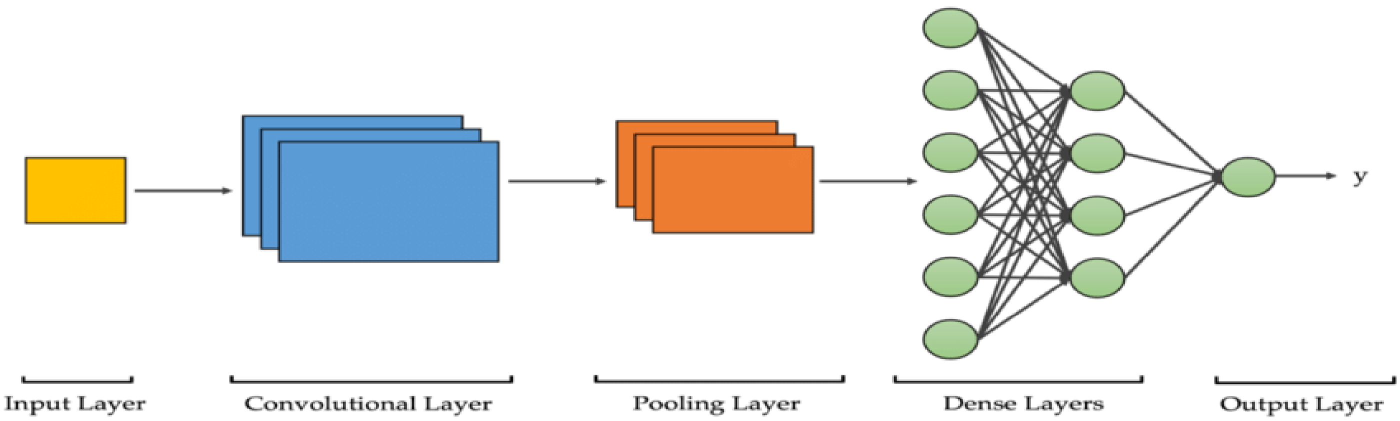


Figure1.1 Convolution neural network architecture.

Convolutional Neural Networks (CNNs) outperformed RNNs and Deep Learning. Position shifts and scale variations are handled using autoencoders, Multilayer Perceptron (MLP), and DBN. CNNs are often made up of convolutional, pooling, and fully linked layers. Convolutional Neural Networks have two features: local connectivity and weight sharing, which result in fewer network parameters, faster training, and an impact on regularization. Figure 1 depicts a Convolutional Neural Network (CNN)-based FER approach.

[13] H. Zhu et al, to estimate age, a deep learning system was created that included global convolutional neural networks (CNNs), three local convolutional neural networks, and an ordinal distribution regression model. The three local Convolutional Neural Networks (CNNs) are trained on three separate GEI components: the head, body, and feet. The full GEI dataset is utilized to train global CNNs (Convolutional Neural Networks). The model was trained and evaluated using data from the OULP Age dataset. It yielded a MEA of 5.24 and a CS (k = 5) of 69.95%. Deep learning techniques, such as Convolutional Neural Networks (CNNs), may recognize and learn from the gait description's distinctive patterns, eliminating the need to explicitly extract characteristics. Sakata et al. did another recent study.

[14] A. Sakata et al, the author proposed a cutting-edge Convolutional Neural Network-based technique for predicting gender and age. They used sequential Convolutional Neural Networks (CNNs) to perform age regression and estimate gender and age group. The GEI would initially travel through a CNN (Convolutional Neural Network) that predicted its gender, followed by two further CNNs that predicted its age and age group, respectively. The OULP Age dataset was utilized to train and evaluate the model, and the results were encouraging, with the approach reaching an average age of 5.84 years. The framework predicts less accurately for older people. ***[15] Zakariya Q et al,*** a deep neural network. The authors trained the VGG-face model using the LFW database. A connected layer was replaced with four brand-new totally connected layers. The first layer is 4096 pixels in size, followed by three layers of 5000 pixels each. Their output layer includes a total of eight age classes. 80.57% one-off accuracy was achieved. The symbol I1 off accuracy represents a poor outcome class caused by one neighboring age symbol on the left-right. They also trained a pre-trained GoogLeNet network with modified totally linked layers, attaining an age estimation accuracy of 45.07%.

[16] Murat Berksan, this trend is caused by the low number of participants in the senior age category accessible in the OULP database. After analyzing multiple Convolutional Neural Networks (CNNs) systems over gait, an outline mean approximation was used to estimate gender with an accuracy of 79.45% and age with a MEA of 5.74 years. ***[17] J. Tapia & C. Aravena,*** an unsupervised method for pre-training a Deep Belief Network using a large number of unlabeled iris photos has been proposed. The Deep Multilayered System was then applied to classify gender. They achieved 74.66% accuracy utilizing data augmentation and 83.00% using CNN-2 to identify genders. ***[18] A. Bansal et al,*** the authors classified the iris picture into one of three categories. 1. Young people (60). They conducted their research using a Bio-Secure Multimodal database. This database's photos were captured with the LG Iris Access EOU3000 system. The database contains 200 subjects, aged 18 to 73. Only five geometrical features were retrieved from the iris. The classification methods used included KNN, SVM, MLP, and multi-classifier (fusion and negotiation-based). They achieved 75% accuracy using the multi-classifier negotiation strategy.

[19] O. Agbo-Ajala et al, suggested a new Convolutional Neural Networks (CNNs) model for classifying and extracting characteristics from unconstrained real-life facial photographs based on gender and age. They achieved 74.8% accuracy in age group classification and 79.7% in gender classification. The study emphasized the importance of accurately assessing consumer age and gender for organizational objectives. ***[20] Wang et al,*** the paper examines the increasing use of biometric techniques' supplementary data. The photographs were divided into age groups using the Furthest Nearest Neighbor (FNN) algorithm. ***[21] M. Singh & S. Nagpal,*** used the ND-GFI iris dataset. They proposed employing Deep class encoding in conjunction with two classifiers, RDF and NNet, in their research. They claimed an accuracy rate of 73.17%.

INTRODUCTION

One of the most potent forms of artificial intelligence that focuses on copying or reproducing the human visual system is computer vision. The computer recognizes and then processes the items that are present in the pictures and videos with the use of computer vision. As soon as the strategy is put into practice, we may begin evaluating its correctness. Because it is a training model, our application's accuracy will rise with continued use. With the advent of machine learning and deep learning, computer vision performance was restricted, requiring manual coding. But Main Task we are done in our project is that Gender Classification & Age Detection system using facial features of the object. Human face is the main object in our project.

Deep learning is recognized as the branch of machine learning that performs machine learning at a multiple layer level. With the advent of deep learning, machines are now able to handle data in a nonlinear manner as opposed to the old method's linear analysis.

The following are a few of the numerous uses for deep learning:

- **Recognition of images**
- **Identification and detection of objects**
- Fraud detection
- Healthcare
- The processing of natural language

To address this difficulty, corporations have implemented AI systems to automate support. Machine learning is one of the most prevalent AI approaches for processing large amounts of data. At first, distinguishing between related computer vision tasks might be tricky. It may be difficult to distinguish between object detection, object localization, & picture categorization. This is because all three concepts are related to face recognition. Image classification assigns a class name to an image, whereas object localization creates a bounding box around its objects(face). Face detection is a harder task than the previous two as it includes both bounding a box around an image and applying a class label. When looking at an object using a video camera, inter- and intra-frame differences are used for object detection. But in order to recognize things, thresholding or object detection algorithms take into account neighbouring pixels when making both global and local decisions. Any shift in the observer's environment is seen as ego-motion. In actuality, the picture theory is erroneous due to the ego-motion reality. This issue can be resolved by calculating the camera motion, which can be accomplished by creating an altered sequence.

Convolutional neural networks (CNNs) are widely used deep learning models that are utilized for image classification. Its primary focus is on visual imagery analysis and picture classification. The face on webcam photos needs to be removed before we can apply the remedy. The OpenCV library in Python is used to achieve this. An efficient artificial intelligence method for face identification is the use of Hair feature-based **cascade classifiers** to identify faces. Many, many positive and negative photos will be used in the training process of the machine learning algorithm. After then, it's used to identify faces in other pictures. A schematic representation of CNN's architecture.

The combination of machine learning and computer vision that is also called geometric-based approach used to perform a virtual task through the machine. This method gathers geometric features from photos and uses effective machine learning techniques to study them in order to create a set of representative features of geometric shape to describe a human face. Humans use their ability to extract perceptual information from what they perceive to perform visual problems and respond quickly to their surroundings. To address computer vision challenges, researchers model human object recognition skills. Techniques for extracting geometric features Connected-component labeling, Corner detection, Curve fitting, Edge detection, Global structure extraction, Feature histograms, Line detection, Image texture, and Motion estimation etc. An object's corners are a very basic yet important aspect. In particular, complex objects typically have distinct corner characteristics from one another. Corner detection allows for the extraction of an object's corners. the separation and angle of two segments of a straight line. By defining features as a parameterized combination of several components, this is a novel approach. An image's edges are its one-dimensional structural elements. They stand in for the edges of several picture sections. By applying the edge detection approach to locate the edge, one can readily discover a face outline. Using the blob detection approach, portions of a picture can be identified as blobs.

METHEDOGOLY

Artificial Intelligence

Artificial intelligence (AI) is a collection of technology that is the computer and machines worked like a human being. It allows allowing to perform tasks like image analysis & speech recognition. Because it has a problem-solving capability. It is the theory of the development of computer system or machine are capable to perform tasks that typically require. The main aim of artificial intelligence is to build a machine that are capable to perform human like tasks with their own understanding. Machine can perform a task with iteratively improve themselves based on.

The use of AI in healthcare industry, security workplace, public surveillance, identification & verification etc. Ther are some applications of artificial intelligence like natural language processing, deep learning, expert system, machine vision, & speech recognition etc. Artificial intelligence also has a multiple advanced functions like ability to see, translate spoken, written language, understand language, data analyzation, recommendation and so on(fig.1.2).

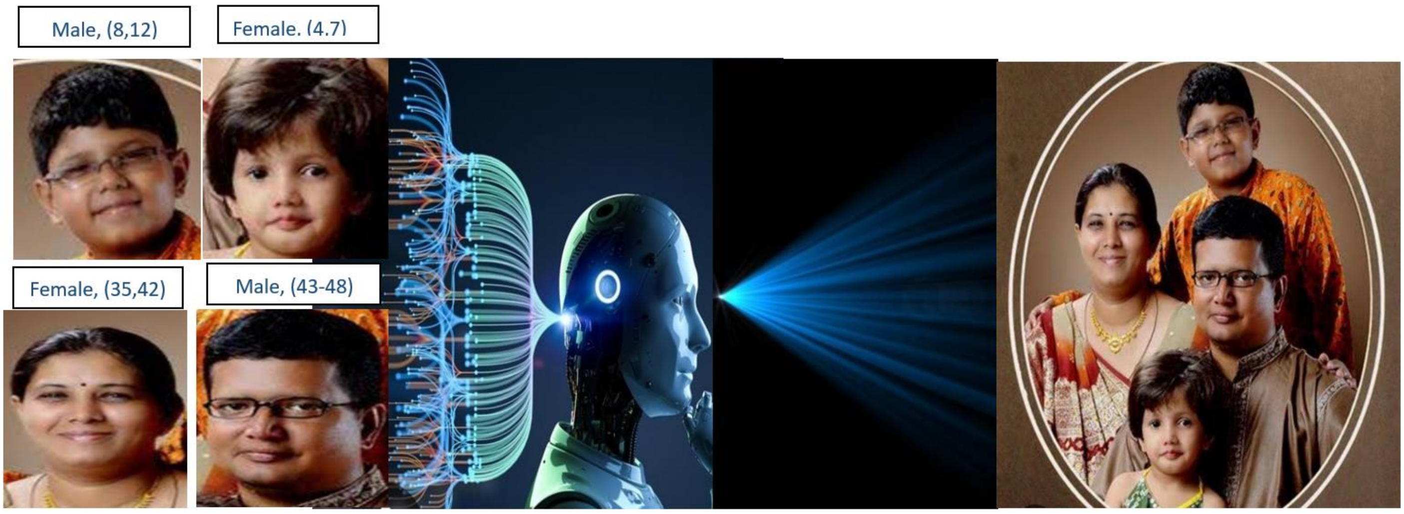


fig1.2 Artificial intelligence (Sample)

Deep Learning:

Deep learning is a tools of artificial intelligence (AI) technology that learn the machines/computers to process data in a way that comes from human beings. By using a deep learning model, we can recognize complex patterns like video, picture, sound, text, and other data to produce accurate results & predictions. Deep learning is a sophisticated artificial intelligence technique that enables computers/machines to learn from data. It's like learning a computer/machine to detect patterns on its own, which is comparable to how human brains operate when we learn something new.

Deep learning can grasp and analyze complicated data including images, audio, and text by employing layers of interconnected nodes known as neural networks. This innovation has resulted in achievements in fields such as image recognition, language translation, and even self-driving automobiles, making computers more intelligent and accomplished in handling real-world problems as shown in the below fig.1.3.

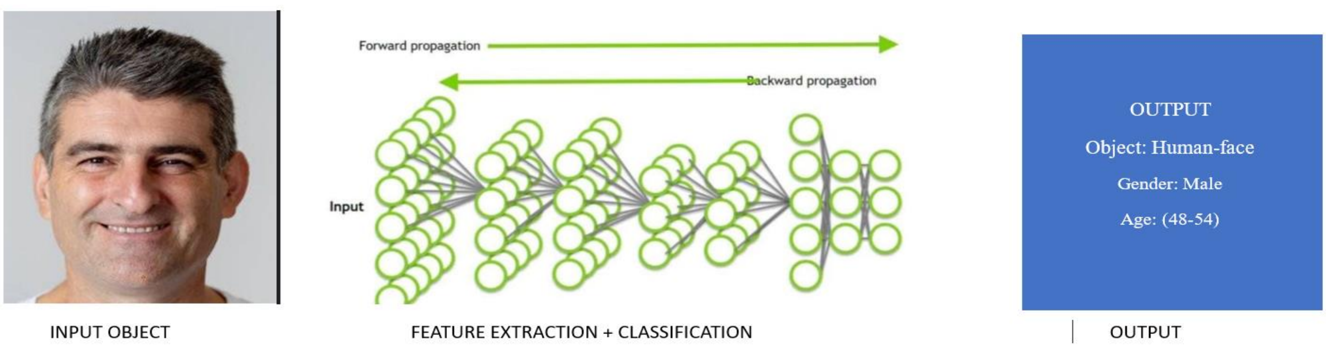


fig1.3 Deep Learning (Sample)

The properly recognizing gender, robots can provide more personalized and relevant support for clients by adapting to their nature. Human-computer interaction is when machines, computers, or robots are designed to interact with individuals. If a robot can detect a person's gender, it will deal with them correctly. Age prediction methodology is most typically employed in an e-photo compilation, where individuals may manage and retrieve their photos by specifying the needed age, reducing query time, and yielding more accurate results. Age prediction is utilized to retrieve and categorize images according to their age through collections.

Computer Vision:
Computer vision is a subfield of artificial intelligence that allows systems to analyze and interpret visual information from the environment, such as images and videos. this method is about making computers "see" and understand what they perceive. In computer vision, there are some key areas like image/video processing, Feature Extraction, Object Detection & Recognition, Image Segmentation, Motion Analysis, **&** 3D Reconstruction, etc**.**
The goal of computer vision is to enable systems to evaluate and comprehend data that is visually appealing, having applications ranging from healthcare to autonomous driving. It has some hardware & software including OpenCV used to process the data(image/video). And Hardware are sensors, cameras, GPUs, they capture the data for processing. The frameworks of computer vision are PyTorch & TensorFlow this software is used to create complex models for analyzing visual data. Ensuring that image analysis is faster than other tools and technology for real-time applications.

OpenCV: - The OpenCV library provides the tools to do computer vision tasks. Some functions of this library that is face detection, video processing & image processing, etc.

The OpenCV library import into the program in the following way: -
import cv2import

CONVOLUTION NEURAL NETWORK:
Convolutional neural networks, or CNNs for short, are a class of neural networks composed of computers that are frequently used to identify and classify images or objects. A CNN is used in the deep learning technique to identify objects in a picture. The three layers of a traditional neural network are input, hidden, and output.The CNN Architecture Inspired by the Brain. Similar to how a neuron in the real brain functions and transmits information between cells, the artificial neurons, or nodes in CNNs, gather inputs, process them, and output the results. Images of people's faces are a valuable source of input data. CNNs can have many hidden layers, each of which uses computations to extract features from a human face.
Convolution is the first layer that extracts characteristics from an input image. The fully linked layer in the output layer is responsible for classifying and identifying the object. The most crucial part of CNNs is the convolutional layer. A computer procedure called convolution combines two different kinds of data. Gender prediction from a social perception storage; images that can access subjects' non-public personal information without their consent, like their precise birthdate; and the standard technique, which consists of gathering additional information about an individual and using that information as the foundation for manual gender identification.
For this reason, we employ D-CNN, which processes images quickly and aids in precise gender classification. Usually, such behaviour is indicative of a minor problem. The intricacy is seen when Deep learning-based methods are applied to a set of images that contains so few face pictures.

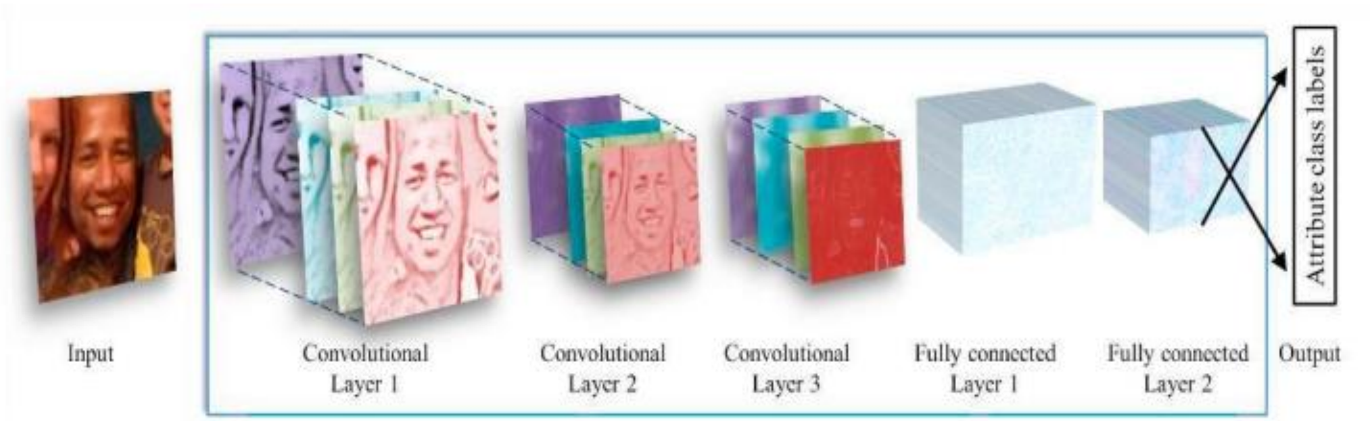


Fig 1.4 CNN Diagram

GEOMETRIC-BASED APPROACH:
The later group is based on the intensity distribution of the entire face image, whereas the former category uses the geometrical separation among various facial features that are included in the image. Every one of these strategies has advantages and disadvantages of its own. Local features are another name for geometric-based characteristics. It specifies the features of the face, including the length, width, and thickness of the brows, as well as the spaces between the mouth, nose, and eyes. Angles and spaces between face landmarks are used. As indicated in Figure 1.5, correct facial point positions yield excellent geometrical features.

❖ Limitations of Geometric based features:
1. It requires high-resolution and quality face photos to locate the facial critical spots.
2. Locating facial key points is influenced by the partial obscuring of faces by eyeglasses or other objects.
3. Geometrical feature-based approaches overlook the textural qualities of the facial image, which are an essential indicator of gender classification.

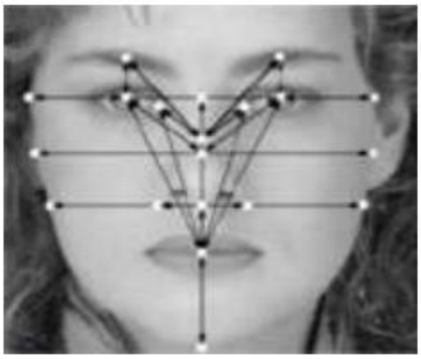


Fig. 1.5 Geometric Features in face

Four essential components are integral pictures for feature computation, Adaboost for feature selection, attention cascade for effective computational resource allocation, and Haar features to identify the existence of features in the input image.

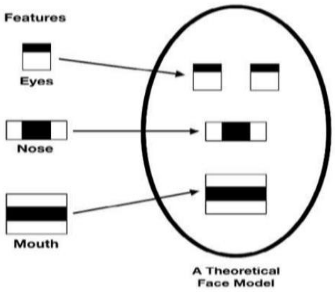


Figure 1.6 Facial Haar Features using Rectangular Filter

Appearance-Based Approach:
The appearance-based approach extracts a feature vector from the entire face image or a subset of the image. The method has a benefit over the geometric-based type, which extracts data from all areas of the face. Appearance-based features, also known as universal features, use basic information about facial image portions based on brightness patterns. The appearance-based method can be called holistic, as the feature vector is taken from the entire face image. Figure 1.6 displays a sample image of a face's appearance-based pattern.
Texture is another key aspect of image analysis. Self-similar structures are repeated repeatedly to create the texture. Texture analysis is commonly used in industrial inspection, biomedical image analysis, remote sensing picture evaluation, and content-based retrieval from image databases.

Fig. 1.7 Appearance based face pattern

An image's textured area is identified by its uneven distribution of spatial intensities. Anything made up of a set of pixels & pixels that are mutually related is referred to as a texture. The texture's exact structure is determined by the face topography and reflectivity, the surface lighting, the position of the camera, & response of the frequency. In numerous biometric applications, a variety of strategies for defining photo texture have been presented.
Data may be retrieved from all parts of the face, which is an advantage over the geometric-based technique. One disadvantage of the holistic appearance-based representation is its susceptibility to local appearance fluctuations, which can be induced by variations in position, emotion, and illumination.
The most common appearance-related characteristics include:

- Various texture properties, such as the local binary pattern (LBP).
- Local Directional Patterns (LDP)
- Histogram of gradients (HOG)
- Scale-invariant feature transform (SIFT)
- The coefficients of an image wavelet transformation, such as the Gabor or Haar wavelet.

Feature Extraction for Age Progression:
The age estimate is used to retrieve and index images based on their age from databases. This kind of technology is most frequently used in electronic photo albums, where users may manage and retrieve their photos by specifying an age requirement. This shortens search times and yields more targeted results. The age estimation system consists of two modules: age estimation algorithms and age image representation.

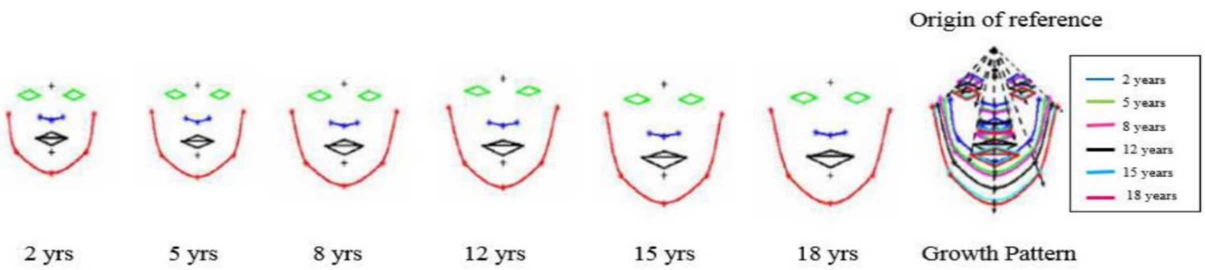


Fig. 1.8 Craniofacial growth model for age progression

There are two stages of face growth and aging in humans: the childhood phase and the aging phase. From birth to adulthood, there is a childhood phase. Figure 1.3 illustrates how the individual's craniofacial growth has significantly changed throughout this phase (Seung & Lee 2000). The size and thickness of the face increase with age as a result of craniofacial growth. The forehead drops back, the chin and cheeks enlarge, and the areas of the eyes, nose, ears, and mouth widen to fill the newly formed interstitial space. The aging process starts when a person reaches old age. The skin's texture is where people notice the most change. Skin gets thinner, darker, leatherier, and less elastic. Also, with time, modifications such as sagging cheeks, wrinkles, bags under the eyes, weakness in the muscles, and freckles develop.

Gender Discriminant Features in Face:
Due to the small variations in facial features, it is challenging to compare the features of men and women. There aren't many facial traits that can make a female portrait appear more feminine or a male portrait appear more manly.

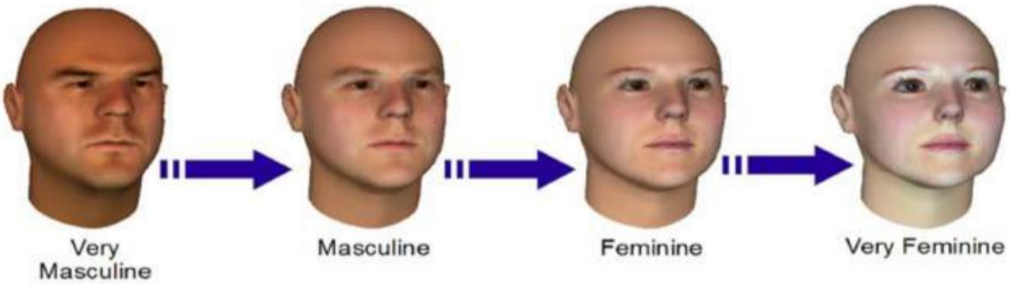


Fig. 1.9 Appearance of male & female faces with minor differences

Figure 1.9 depicts a sampling of the male & female face models. In some circumstances, the primary distinction between male & female face photos is the change in image contrast & shape. The image on the left represents a man, while the image on the right represents a woman.

Table 1.1 Discriminant features in face

ELEMENTS	FEMALE	MALE
DIFFERENT TEXTURE IN FACE REGIONS		
EYES	Smaller and ovel shape	Larger and rectangular shape
EYEBROW	More arched and thin. Space between eyes and eyebrow is large.	Straight and thick. Space between eyes and eyebrow is small
NOSE	Smaller and shorter	Mid-length, Angular shape and sharp edges
MOUTH AND LIPS	Distance between the nose's base and the lip's top is small. Thick and curve upper lip	Distance between the nose's base and the lip's top is large. Very thin upper lip.
CHIN	Shorter and rounded	More square with flat base

Table 1.1 compares the geometrical information of the face's gender-discriminating features. Practically speaking, gender classification of digital facial photos is more challenging because of differences in lighting, posture, expression, and occlusion. The majority of automatic and systematic frameworks in today's society are built with gender classification in mind. Human body forms, gait cycle, and iris are used in gender recognition algorithms. But most research employed facial features to identify gender. In computer vision, facial image analysis is one of the most difficult problems. One of the best traits that set a person apart is their face, and facial attractiveness is essential to human recognition.

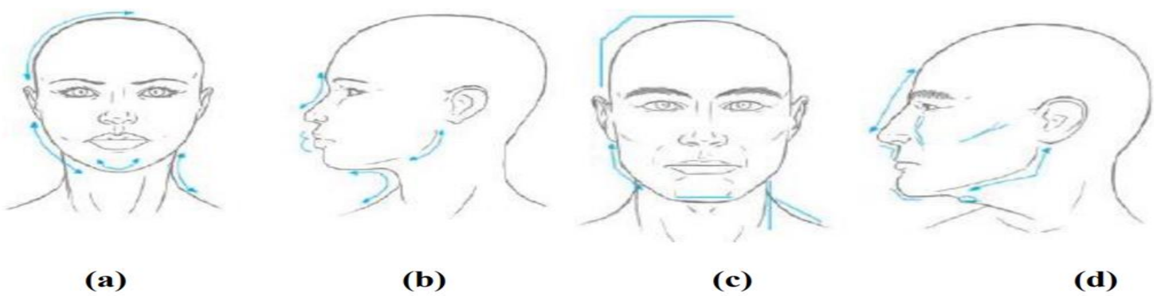
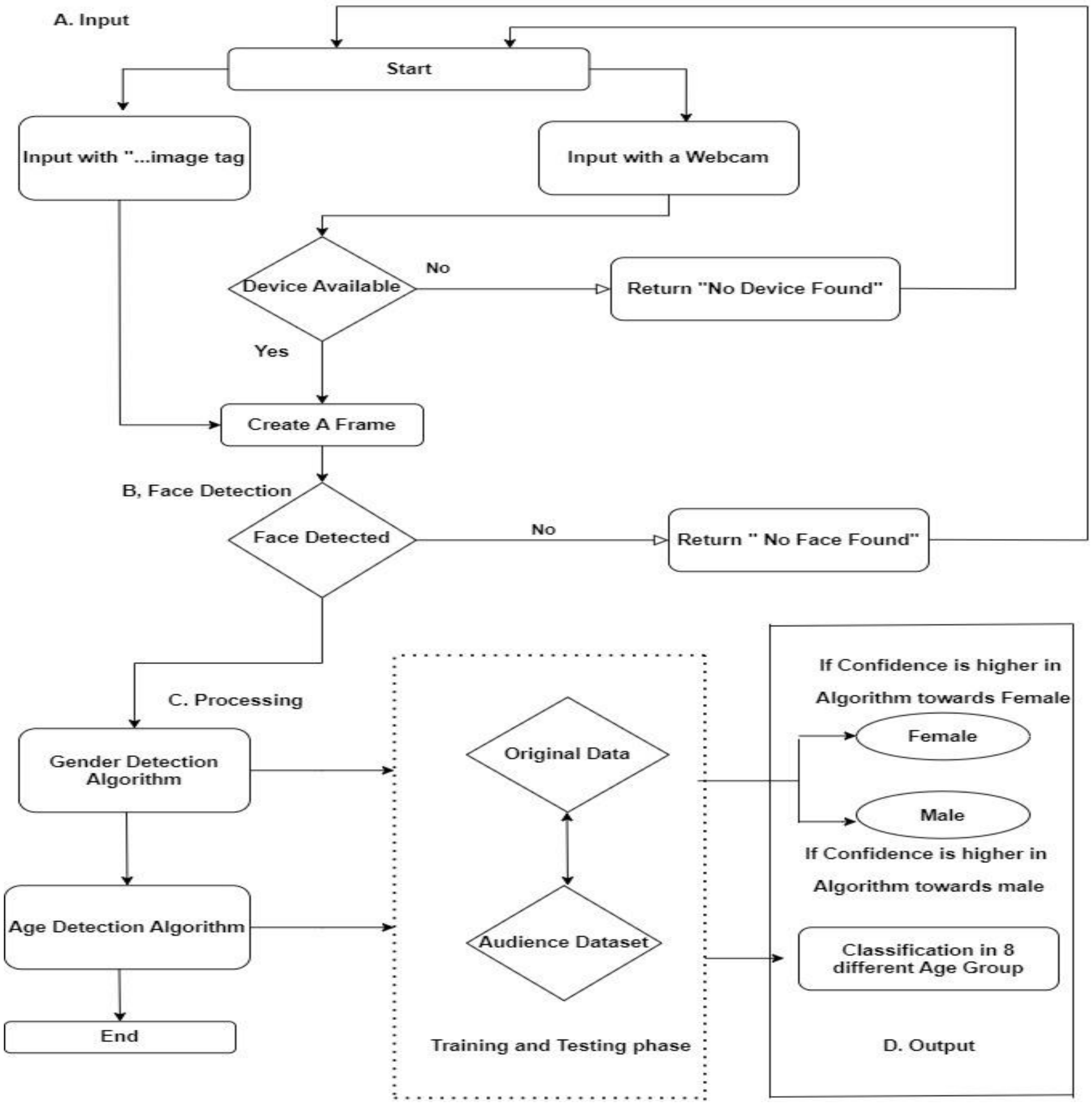


Fig.1.10 Difference in face size & shape of female and male a, b female image & c, d male image

As seen in the aforementioned figures (a), (b), (c), and (d), distinctive characteristics like gender, age, race, and identity are in fact essential for sophisticated real-world analysis. Such might prove quite useful in identifying users and gathering user-related data for market field analyses. The growing corporate demand for gender classification in digital photographs has prompted study in this area. It is simple for humans to distinguish between males and females based on facial features, but it is more challenging for computers. Machines require sufficient data from attributes to conduct identification. Machines can identify a person's gender from facial picture data by utilizing certain qualities that set male and female looks apart. Although the approach for classifying genders appears straightforward, the automatic gender classification method remains a difficult pattern recognition challenge because of changes in lighting, position, expression, and other factors.

Data Flow Diagram for Gender & Age Detection:



A. Input

To expedite the process, there are multiple ways to feed data into the algorithm. To start, the user can quickly take data using the webcam on the machine or any webcam-equipped digital device. Simplifying and speeding up a complete system is the main goal of this research.

B. Face Detection

Software that matches a human face in a frame of a digital image or video with a database of faces is called a facial recognition system. Charles Bisson, Helen Chan Wolf, and Woody Bledsoe were some of the pioneers in developing facial recognition technology. In 1964 and 1965, Bledsoe, Wolf, and Bisson started experimenting with using computers to detect faces. Certain digital (noise, interference) and natural (posing angles, facial labeling) changes are applied when a face is discovered in a frame. It is more challenging to identify a human face when using two features of a human face as a template: (1) There are an enormous and almost infinite number of templates, or faces, to classify. (2) Nearly all patterns have the same appearance. We can use a range of crowd recordings to solve this problem and improve the algorithm's efficiency. The audience set is used as a reference in neural networks to classify gender.

C. Processing

If a face is discovered when using the facial detection method. A Deep convolutional neural network, or CNN, may be used to begin processing. It is a particular kind of deep neural network that is primarily employed in image processing. Following a training phase, CNN produces an estimation range. This kind of deep neural network is widely used for image and natural language processing. The actual training phase will be handled by CNN, when a variety of projections will be made. It is possible to predict two genders: male and female. The task of estimating age involves multiple classes, with the eras being divided into groups. Since different age groups have diverse face features, it is difficult to find accurate data. We divided the population into age categories to speed up the procedure. The age estimate may fall into one of eight categories: 0–2, 4-6, 8–12, 15–20, 25–32, 38–43, 48–53, and 60–100.

D. Output

After launching the project using the Command Prompt, the Login form will show up as the initial screen. The project window displays after the credentials have been entered correctly. The procedure then checks to see if anything is in front of the webcam and, if so, what kind of object it is. Here are some examples taken from our research.

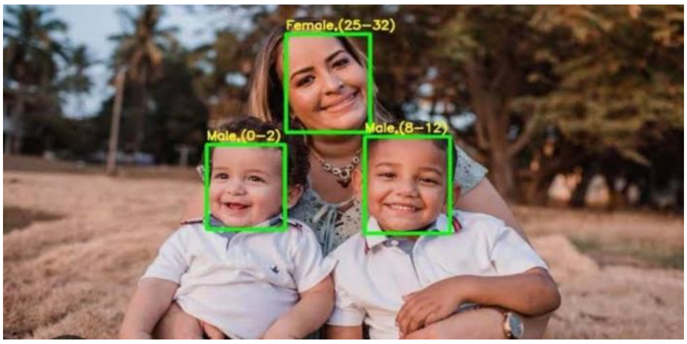


Fig 1.11 Example of the Result

IMPLEMENTATION:

❖ **Define faceBox function:**

Code:
```
def faceBox(faceNet, frame):
    frameHeight = frame.shape[0]
    frameWidth = frame.shape[1]
    blob = cv2.dnn.blobFromImage(frame, 1.0, (300,300), [104,117,123], swapRB=False)
    faceNet.setInput(blob)
    detection = faceNet.forward()
    bboxs = []
    for i in range(detection.shape[2]):
        confidence = detection[0,0,i,2]
        if confidence > 0.7:
            x1 = int(detection[0,0,i,3] * frameWidth)
            y1 = int(detection[0,0,i,4] * frameHeight)
            x2 = int(detection[0,0,i,5] * frameWidth)
            y2 = int(detection[0,0,i,6] * frameHeight)
            bboxs.append([x1, y1, x2, y2])
            cv2.rectangle(frame, (x1, y1), (x2, y2), (0, 255, 0), 1)
    return frame, bboxs
```

Explanation:

- def faceBox(faceNet, frame): This defines a function faceBox which takes a neural network faceNet and an image frame as input.
- frameHeight = frame.shape[0]: Gets the height of the frame (image).
- frameWidth = frame.shape[1]: Gets the width of the frame.
- blob = cv2.dnn.blobFromImage(frame, 1.0, (300,300), [104,117,123], swapRB=False): Converts the image to a blob, a format required by the neural network. Here, it resizes the image to 300x300 pixels and normalizes it using the mean values [104,117,123].
- faceNet.setInput(blob): Sets the input to the face detection network.
- detection = faceNet.forward(): Performs forward pass to get the face detections.
- bboxs = []: Initializes an empty list to store bounding box coordinates of detected faces.
- for i in range(detection.shape[2]): Iterates over each detection.
- confidence = detection[0,0,i,2]: Extracts the confidence level of the detection.
- if confidence > 0.7: Checks if the confidence is above 70%.
- x1, y1, x2, y2: Calculates the bounding box coordinates by scaling the normalized coordinates by the width and height of the frame.
- bboxs.append([x1, y1, x2, y2]): Adds the bounding box coordinates to the list.
- cv2.rectangle(frame, (x1, y1), (x2, y2), (0, 255, 0), 1): Draws a rectangle around the detected face on the frame.
- return frame, bboxs: Returns the frame with rectangles drawn and the list of bounding boxes.

❖ **Load Pre-trained Models:**

Code:
```
faceProto = "opencv_face_detector.pbtxt"
faceModel = "opencv_face_detector_uint8.pb"
ageProto = "age_deploy.prototxt"
ageModel = "age_net.caffemodel"
genderProto = "gender_deploy.prototxt"
genderModel = "gender_net.caffemodel"
faceNet = cv2.dnn.readNet(faceModel, faceProto)
ageNet = cv2.dnn.readNet(ageModel, ageProto)
genderNet = cv2.dnn.readNet(genderModel, genderProto)
```

Explaination:

- faceProto, faceModel, ageProto, ageModel, genderProto, genderModel: These are paths to the configuration and weights files for the face detection, age prediction, and gender prediction models respectively.
- faceNet = cv2.dnn.readNet(faceModel, faceProto): Loads the face detection model.
- ageNet = cv2.dnn.readNet(ageModel, ageProto): Loads the age prediction model.
- genderNet = cv2.dnn.readNet(genderModel, genderProto): Loads the gender prediction model.

❖ **Define Mean Values and Labels:**

Code:
```
MODEL_MEAN_VALUES = (78.4263377603, 87.7689143744, 114.895847746)
ageList = ['(0-2)', '(4-6)', '(8-12)', '(15-20)', '(25-32)', '(38-43)', '(48-53)', '(60-100)']
genderList = ['Male', 'Female']
```

Explanation:

- MODEL_MEAN_VALUES: Mean values for the dataset the model was trained on. Used for normalization.
- ageList: List of age ranges corresponding to the output labels of the age prediction model.
- genderList: List of genders corresponding to the output labels of the gender prediction model.

❖ **Capture Video from Webcam:**

Code:
```
video = cv2.VideoCapture(0)
padding = 20
```

Explanation:

- video = cv2.VideoCapture(0): Captures video from the webcam. 0 is the default camera.
- padding = 20: Padding around the detected face for better prediction accuracy.

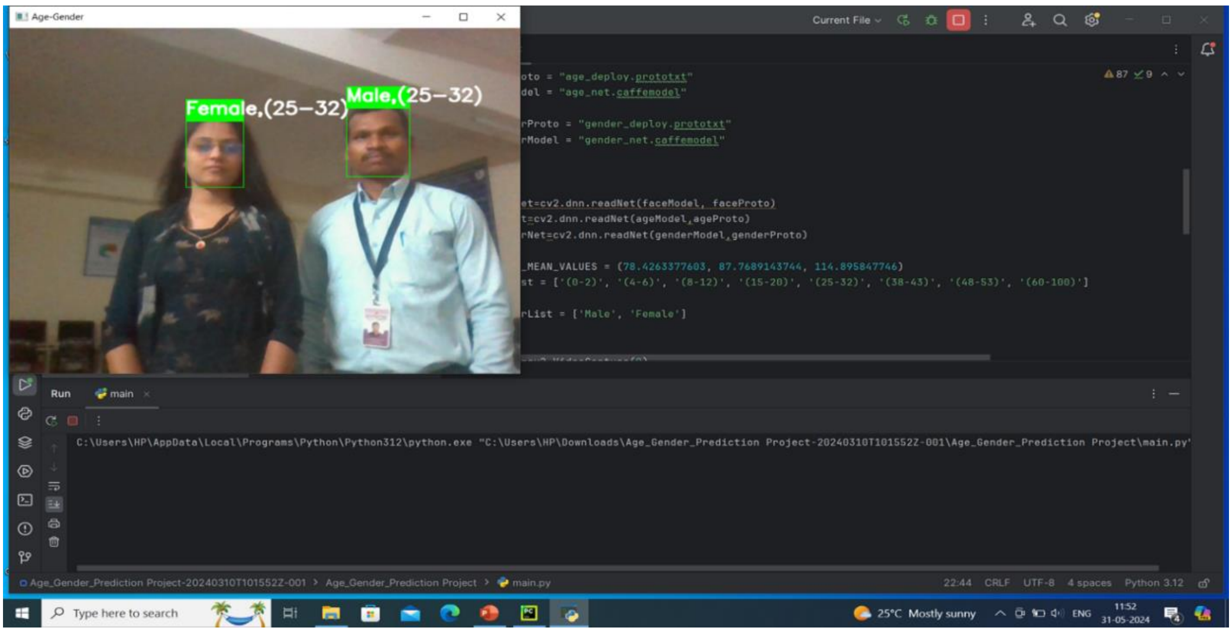
❖ **Real-time Age and Gender Detection:**

Code:
```
while True:
    ret, frame = video.read()
    frame, bboxs = faceBox(faceNet, frame)
    for bbox in bboxs:
        face = frame[max(0, bbox[1]-padding):min(bbox[3]+padding, frame.shape[0]-1), max(0, bbox[0]-padding):min(bbox[2]+padding, frame.shape[1]-1)]
        blob = cv2.dnn.blobFromImage(face, 1.0, (227,227), MODEL_MEAN_VALUES, swapRB=False)
        genderNet.setInput(blob)
        genderPred = genderNet.forward()
        gender = genderList[genderPred[0].argmax()]
        ageNet.setInput(blob)
        agePred = ageNet.forward()
        age = ageList[agePred[0].argmax()]
        label = "{},{}".format(gender, age)
        cv2.rectangle(frame, (bbox[0], bbox[1]-30), (bbox[2], bbox[1]), (0,255,0), -1)
        cv2.putText(frame, label, (bbox[0], bbox[1]-10), cv2.FONT_HERSHEY_SIMPLEX, 0.8, (255,255,255), 2, cv2.LINE_AA)
    cv2.imshow("Age-Gender", frame)
    k = cv2.waitKey(1)
    if k == ord('q'):
        break
video.release()
cv2.destroyAllWindows()
```

Explanation:

- while True: Infinite loop to process video frames in real-time.
- ret, frame = video.read(): Reads a frame from the video capture.
- frame, bboxs = faceBox(faceNet, frame): Detects faces and gets bounding boxes.
- for bbox in bboxs: Iterates over each detected face bounding box.
- face = frame[max(0, bbox[1]-padding):min(bbox[3]+padding, frame.shape[0]-1), max(0, bbox[0]-padding):min(bbox[2]+padding, frame.shape[1]-1)]: Extracts the face region with padding.
- blob = cv2.dnn.blobFromImage(face, 1.0, (227,227), MODEL_MEAN_VALUES, swapRB=False): Converts the face region to a blob.
- genderNet.setInput(blob): Sets the blob as input to the gender prediction network.
- genderPred = genderNet.forward(): Gets gender prediction.
- gender = genderList[genderPred[0].argmax()]: Converts the prediction to a human-readable label.
- ageNet.setInput(blob): Sets the blob as input to the age prediction network.
- agePred = ageNet.forward(): Gets age prediction.
- age = ageList[agePred[0].argmax()]: Converts the prediction to a human-readable label.
- label = "{},{}".format(gender, age): Combines gender and age into a label.
- cv2.rectangle(frame, (bbox[0], bbox[1]-30), (bbox[2], bbox[1]), (0,255,0), -1): Draws a filled rectangle for the label background.
- cv2.putText(frame, label, (bbox[0], bbox[1]-10), cv2.FONT_HERSHEY_SIMPLEX, 0.8, (255,255,255), 2, cv2.LINE_AA): Puts the label text on the frame.
- cv2.imshow("Age-Gender", frame): Displays the frame.
- k = cv2.waitKey(1): Waits for 1 ms for a key press.
- if k == ord('q'): Checks if the 'q' key is pressed.
- break: Breaks the loop if 'q' is pressed.
- video.release(): Releases the video capture object.
- `cv2.destroyAllWindows()

Once the credentials have been input correctly, the project window appears, and the algorithm proceeds to determine whether there is an item in front of the webcam, and if so, the gender type. Results from our study are provided below.



Accuracy Testing:

As soon as the strategy is put into practice, we may begin evaluating its correctness. Because it is a training model, our application's accuracy will rise with continued use. A typical method for determining accuracy is:

- Enter the input data as face.
- Create a frame.
- Detect the face image.
- Perform processing of the face image.
- Classify the gender as either male or female.
- Detect the age group in range.
- Provide the result in an image.
- result of the image to the considering position.

LIMITATIONS

Skin color segmentation poses a significant obstacle to face identification. The distance from the camera, lighting, noise, and position of the object all have an impact on the accuracy of facial segmentation. The following are the different forms of barriers that may develop during detection:

- Alignment;
- Expression;
- Age;
- Dimensions of the face;
- Distinctive features of the face;
- Lighting

CONCLUSION

A suggested deep learning approach utilizes deep convolutional neural networks, this is meant to be able to classify and detect the different types of faces in a video with high accuracy. The paper begins with an explanation of the architectures needed for Gender Classification & Age detection, along with a thorough analysis of deep learning algorithms. Because of the quick learning ability, in recent years, deep learning-based face detectors have been a popular area of study. For real-time recognition of the face, particularly for Makeup face evaluation, the current models did not produce findings that were as exact as possible. When the clarity of a camera is good, then its more helpful to archived accurate result. The Real-Time Gender & Age detection is so slow because it is a training model when we use more & more time it become gives more accurate results and time-consuming also. In this system, the age detection shows in the form of ranges like 0–2, 4-6, 8–12, 15–20, 25–32, 38–43, 48–53, and 60–100. To get around this problem, a model that can reliably identify gender & age from a video sequence is offered. There are three primary modules in our proposed system: -

- Proposing an architecture that can accurately detect and recognize the static and the moving people's Gender & Age from a video sequence.
- Applying the proposed model to work well with a face detection application including more features in real-time application.
- Also, here we are not working on the transgender type of gender detection so in the proposed system it is also possible to develop.

The proposed model preserves the full detail feature of the Makeup face by extracting the images to do more accurate results. When we talk about transgender, we use an object as well as voice recognition include. Then, we may like to predict the result also for transgender detection. If the face is detected as a female and the voice are predicted like as a man then it is called transgender and vice versa for males. This condition may work for the proposed system.

References

1. *A. A. Vogan, F. Alnojiar, M. Gochoo, S.A. Khalid. "Robots, AI & Cognitive Training in an Era of mass Age- Related Cognitive Decline: A Systematic Review", IEEE Access, DOI: 10.1109/Access 2020.2966819, Vol: 8,2020, pp:18284-18304.*
2. *S. T. Rahman, A. Arefen, S.S.Mridea, A.L.Khan, S.Subrina, " Human Gender & age Estimation using Facial Image Processing", IEEE Access, Region 10 Symposium (TENSYMP), DOI: 10.1109/Access 2020.1102986, Vol: 6, pp: 6-28, 2020.*
3. *V. Carletti, A. Greco, G. Percannells, M. Vento. "Age from Faces in the Deep Learning Revolution", IEEE Translations on pattern Analysis & Machine Intelligence-2019.*
4. *M.R.Rajput, G.S. Sable, "Deep Learning Based Gender & Age Estimation from Human Iris", Proceeding of ICAEEC-2019 vol:06, pp:8-10.*
5. *Fu, G. Guo, & T. S. Huang(Nov 2010)"Age synthesis & estimation via faces:A Survey", IEEE Transactions on Pattern Analysis & Machine Intelligence,32(11):19551976.*
6. *Eidinger, R. Enbar, & T. Hassner (Dec 2014) "Gender & age estimation of unfiltered faces"IEEE Transactions on Information Forensics & Security,9(12):21702179.*
7. *P. Rodríguez, G. Cucurull, J. M. Gonfaus, F. X. Roca, & J. González, "Gender & age recognition in the wild with deep attention," Pattern Recognit., vol. 72, pp. 563–571, Dec. 2017.*
8. *T.UI Islam, L.K.Awasthi, V. Garg, " Gender & Age Estimation from gait: A Review", IEEE International conference on Analysis of Image , Social Network & Texts-2021, pp-6-15.*
9. *R. Yamashita, M. Nishio, R. K. G. Do, & K. Togashi, ―Convolutional neural networks: an overview & application in radiology, Insights into Imaging, vol. 9, no. 4. Springer Verlag, pp. 611–629, Aug. 01, 2018, doi: 10.1007/s13244-018-0639-9.*
10. *N. Chauhan, "Predict Gender & age using Convolutional Neural Network & openCV", Medium, 2020. [Online]. Available: https://towardsdatascience.com/predict-age-&-gender-using convolutional-neural- network-&-opencv-fd90390e3ce6. [Accessed: 05- Jun- 2020].*
11. *S. Lagree, K. W. Bowyer.: Predicting ethnicity & gender from iris texture. In: IEEE International Conference on Technologies for Homel& Security (HST), (2011).*
12. *A. Khan, A. Sohail, U. Zahoora, & A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," Artif. Intell. Rev., vol. 53, no. 8, pp. 5455–5516, Apr. 2020, doi: 10.1007/ S10462-020- 09825-6.*
13. *H. Zhu, Y. Zhang, G. Li, J. Zhang, H. Shan, Ordinal Distribution Regression for Gait-based Age Estimation (2019). arXiv:1905.11005.*
14. *A. Sakata, N. Takemura, Y. Yagi, Gait-based age estimation using multi-stage convolutional neural network. IPSJ Trans. Comput. Vis. Appl. 11(1), 4 (2019).*
15. *Zakariya Q., Arafat A., M., B. D. Barkana.: Deep Convolutional Neural Network for Age Estimation based on VGG-Face Model. In: CoRR (2017).*
16. *M. Berksan, Gender recognition & age estimation based on human gait (Master's thesis, Ba¸skent Üniversitesi Fen Bilimleri Enstitüsü) (2019).*

17. *Tapia J., Aravena C.: Gender Classification from NIR Iris Images Using Deep Learning. In: Bhanu B., Kumar A. (eds) Deep Learning for Biometrics. Advances in Computer Vision & Pattern Recognition. Springer, Cham. (2017).*
18. *A. Bansal, R. Agarwal, R.K. Sharma.: SVM based gender classification using iris images. In: Fourth International Conference on Computational Intelligence & Communication Networks pp. 425–429. (CICN) (2012).*
19. *O. Agbo-Ajala & S. Viriri, "Face-Based Gender & age Classification Using Deep Learning Model", Image & Video Technology, pp. 125-137, 2020.*
20. *J. Wang, E. Sung, & W. Yau, "Active learning for solving the incomplete data problem in facial age classification by the furthest nearest-neighbor criterion," IEEE Trans. Image Process., vol. 20, no. 7, pp. 2049–2062, Jul. 2011.*
21. *M. Singh, S. Nagpal.: Gender & Ethnicity Classification of Iris Images using Deep Class-Encoder. In: IEEE(2011).*
22. *Prince, SJ & Aghajanian, J 2009, 'Gender classification in uncontrolled settings using additive logistic models', In 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 2557-2560.*
23. *Mousavi, SMH & Lyashenko, V 2019, 'Age Estimation in Make-up (Cosmetics) Using Color & Depth Images', 3rd International Conference on Soft Computing.*
24. *Kim, HC, Kim, D, Ghahramani, Z & Bang, SY 2006, 'Appearance1based gender classification with Gaussian processes', Pattern Recognition Letters, vol. 27, no. 6, pp. 618-626.*
25. *Fu, Y, Xu, Y & Huang, TS 2007, 'Estimating human age by manifold analysis of face pictures & regression on aging features', In IEEE International Conference on Multimedia & Expo, pp. 1383-1386.*
26. *Chen, C, Dantcheva, A & Ross, A 2013, 'Automatic Facial Makeup Detection with Application in Face Recognition', In 2013 international conference on biometrics (ICB), IEEE, pp. 1-8.*
27. *Choi, SE, Lee, YJ, Lee, SJ, Park, KR & Kim, J 2011, 'Age estimation using a hierarchical classifier based on global & local facial features', Pattern Recogn., vol. 44, no. 6, pp. 1262-1281.*
28. *Iga, R, Izumi, K, Hayashi, H, Fukano, G & Ohtani, T 2003, „A gender & age estimation system from face images", SICE Annual Conference in Fukui, Japan, vol. 1, pp. 756-761.*
29. *A. Dantcheva, F. Bremond, "Gender Estimation Based on Smile-Dynamics", IEEE Conference on Computer Vision & Pattern Recognition, Vol:4, pp: 146-158, 2018.*