ISSN 2063-5346

ANALYSIS OF FACE RECOGNITIONAND SMART ADVERTISEMENT USING DEEP LEARNING

Anie Selva Jothi[1], Ragavendra N[2], Rakesh P[3], Richard Jacob Saju[4]

Article History: Received: 01.02.2023 Revised: 07.03.2023 Accepted: 10.04.2023

Abstract

Face Facial recognition is a process that involves analysing a person's face to confirm their identity. It is a type of technology that uses a system to compare a digital image of a person's face, which can be from a photograph or a video frame, with a database of stored faces. This process can be done in real time, and it can identify individuals captured in either photographs or video footage. The most sophisticated face recognition technique locates and measures facial features from a picture, and it is also used for identifying individuals through ID verification services. Since the 1960s, automated facial recognition made its debut. Face recognition uses a variety of algorithms. Identification of the algorithm that is best suited for facial recognition is necessary. The algorithm must have characteristics like low memory consumption, high accuracy, and high prediction %. To determine the optimal approach for creating a facial recognition system, it is necessary to analyse various face recognition algorithms. By evaluating different methods, it is possible to identify the most effective techniques and technologies for developinga facial recognition system.

***Key terms**: Convolution Neural Network, Deep Learning,Haar Feature Extraction.*

[1]Assistant ProfessorSri Krishna College ofEngineering & Technology, India
niranbme@gmail.com

[2]U.G. Scholar Sri Krishna College of Engineering & Technology, India
19euec110@skcet.ac.in

[3]U.G. Scholar Sri Krishna College of Engineering & Technology, India
19euec113@skcet.ac.in

[4]U.G. Scholar SriKrishna College ofEngineering &Technology, India
19euec115@skcet.ac.in

DOI:10.31838/ecb/2023.12.s1-B.268

I INTRODUCTION

Facial recognition is a technology that uses algorithms to identify and verify individuals based on their facial features. It has various applications, including in marketing and advertising. Brands can use facial recognition technology to improve customer experiences by capturing unfiltered and unreserved emotional reactions from customers. This allows for more accurate and scalable analysis compared to traditional focus groups and surveys. Facial recognition advertising offers several advantages, such as the ability to combine digital technology with in- store purchasing, providing real-time customer relationship management information, and notifying stores when products are running low on stock. It also allows for dynamic ads that adjust based on the customer's preferences and may display customized promotions and commercials. Brands can use facial recognition to sell lesser-known products and increase sales by effectively grabbing customers' attention while they browse. One example of facial recognition advertising is Cooler Screens, which use proximity sensors and facial recognition technology to display products that are likely to interest customers. It also monitors the items customers select and displays advertisements based on their selection. With its unique features, facial recognition advertising offers a promising way for brands to develop insightful relationships with customers and differentiate themselves from the competition. Modern marketing requires the ability to target products to clients in real- time, which can be achieved by utilizing face recognition technology. This enables advertisers to make educated assumptions about a person's age and gender, allowing for more effective message targeting. To determine the most appropriate face recognition method for such applications, various factors such as computational complexity, memory usage, accuracy, and prediction percentage were evaluated. The analysed methods included the Local Binary Patterns Histogram (LBPH), Convolutional Neural Network (CNN), and Haar Cascade Classifier utilizing OpenCV. The optimal facial recognition algorithm for a smart advertising system should possess high accuracy, low computing complexity, high prediction percentage, and minimal memory consumption.

II LITERATURE SURVEY

In recent times, the demand for personalized AI- powered services has increased significantly to cater to individual users' specific needs. Samsung introduced "Future Home," an AI-powered companion robot avatar at CES 2022, acting as a personalized avatar life assistant. Moreover, MINDs Lab has created lifelike AI virtual individuals that businesses can use as announcers, curators, and counsellors to convey information, showcasing the fast expansion of intelligent agents in daily life.

To provide tailored services, an intelligent service powered by digital humans is proposed. The system includes adaptive data wrangling, digital human control, and personal identification approaches. Face recognition is initiated by detecting a face through image-based deep learning for individualized recognition. Identifying personal identity information (PID) is crucial for personalized service. Different methods can identify a person's face from a photograph, such as compressing the key elements of a face image into vectors and matching them to the input and target image vectors, as proposed by H. J. Mun et al. Another method is based on eigenface, using explicit feature definitions obtained via principal component analysis (PCA).

Moreover, digital humans or virtual characters are used in digital signage and kiosk devices to display digital information and deliver efficient services through natural conversation. User-friendly controls and animation can maximize the benefits of digital humans to communicate information through non-verbal expression. Digital human services can also be used in applications requiring greater intuitiveness, such as the health and medical industries, where virtual agents have been proposed to support people living with physical and mental comorbidities. Avatar services integrated with various virtual spaces and AI technologies are also being investigated.

Intelligent agents are being deployed in several domains, including the mobile market and the home Internet of Things field. Notable examples include Apple's Siri and Samsung's Bixby in the mobile market and Amazon's Echo speaker, SK's NUGU candle, and Google's Home AI speaker performing integrated functions inside the house. To create intelligent agents optimized for specific domains, it is crucial to go beyond scenario-based knowledge and incorporate speech recognition and natural language processing (NLP) technology.

III EXISTING SYSTEM

A proposed approach for face identification using a multi-task Convolutional Neural Network (CNN) involves using identity classification as the primary task and Pose Illumination, and Expression (PIE) predictions as auxiliary tasks. The side tasks act as regularization to separate PIE variants from learned identification traits. Multitask learning (MTL) is used to learn multiple tasks concurrently and has been successful in various applications. The suggested method utilizes divide-and-conquer to improve CNN learning, with the CNN using a dynamic-weighting approach to learn loss weights for each side activity. In addition, a straightforward and efficient descriptor called Extended Local Binary Patterns on Three Orthogonal Planes (ELBPTOP) is proposed for recognizing facial micro-expressions. ELBPTOP consists of three complimentary binary descriptors that examine local second-order information in ME video sequences: LBPTOP, Radial Difference LBPTOP (RDPLAPTOP), and Angular Difference LBPTOP (ADLBPTOP). This approach is highly effective for ME recognition and computationally efficient.

A machine learning-based facial recognition system that uses support vector machines (SVM) is presented, which utilizes the Viola-Jones algorithm for face detection, and histogram of oriented gradients (HOG) for feature extraction. The system trains and classifies facial databases using multi-class SVM, where each face represents a class. The study explores the drawbacks of a global approach to feature extraction, but it is still accurate and easier to use when there is less diversity in facial alignment.

Facial recognition was initially built using Hidden Markov Models (HMM) and SVM, and it has since been advanced using neural networks and Ada boost. The study primarily focuses on facial recognition as it pertains to recording student attendance.

IV PROPOSED SYSTEM

This study aims to accurately classify unrestricted real-world facial photos into predetermined age and gender categories using age and gender predictions on unfiltered faces. While significant advancements have been made in this area, established approaches have demonstrated their inability to handle significant degrees of variation in such unrestricted photos. To address this issue, Convolutional Neural Networks (CNNs) have recently been widely used for the classification problem due to their superior performance in facial analysis. The two-level CNN architecture includes feature extraction and classification processes. A strong image pre-processing technique is used to prepare and process real-world faces before feeding them into the CNN model, specifically addressing the significant variances in unfiltered real-world faces. The feature extraction component of the model utilizes the WideResNet architecture and Convolutional Neural Networks to learn facial features ranging from edges and corners to more abstract features such as eyes and mouth.

Smart Advertising:

Marketing is evolving to provide real-time product targeting to customers, who expect prompt service. One approach to achieve this is by using face recognition to infer customers' age and gender, leading to more targeted advertising. Effective face recognition algorithms should have low computational complexity, high accuracy, and low memory usage, among other characteristic.

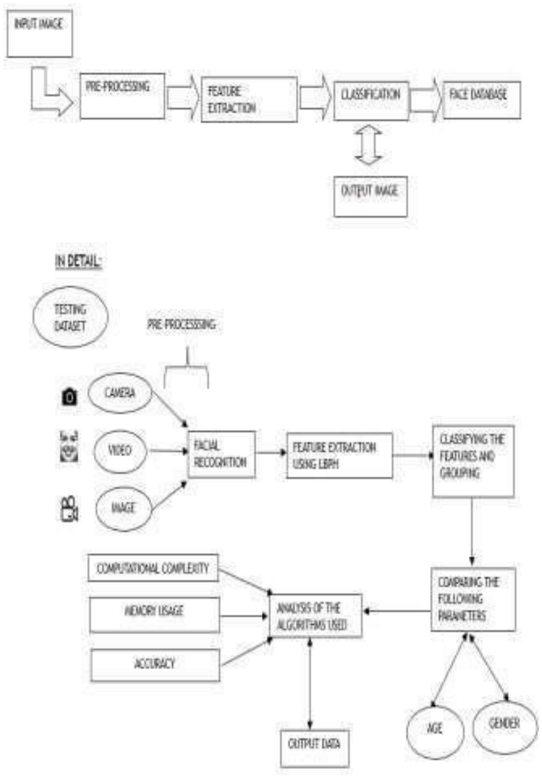


Fig.1. Block diagram

METHODOLOGY

Image Acquisition:

The process of acquiring an input image for the automated detection of normal and abnormal blood cell images through digital image processing is commonly referred to as image acquisition.

Pre-processing:

The process of pre-processing involves a series of operations that are performed on intensity images, which represent the lowest level of abstraction. The primary objective of pre-processing is to enhance image data by reducing unwanted distortions or highlighting essential elements that are necessary for subsequent processing.

Local Binary Pattern Histogram:

It is a widely used texture operator that identifies pixels in an image by analyzing the surrounding area of each pixel and encoding the result as a binary number. LBPH has proven to be a highly effective approach for texture categorization and has gained popularity since its development in 1994. Studies have shown that combining LBPH with histograms of oriented gradients (HOG) descriptors can significantly improve detection performance on specific datasets. The LBPH method utilizes four parameters to encode the texture characteristics of the image.

Radius:

LBPH relies on the radius parameter to specify the neighborhood size around the central pixel and generate the circular local binary pattern. Typically, a radius value of 1 is frequently employed to encode the image's texture characteristics, as it determines the region surrounding the pixel that will be taken into account during the encoding process.

Neighbors:

The circular local binary pattern is created using a certain number of sample points. Setting a higher number of sample points can increase the computational cost of the process. However, the usual number of sample points used in LBPH is 8, as it has been found to provide a good balance between computational efficiency and accuracy in texture characterization.

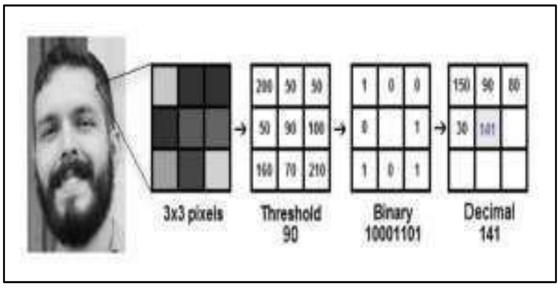
Grid X:

The dimensionality of the feature vector is determined by the number of cells in the horizontal direction, with a higher number of cells leading to a more detailed grid and a higher-dimensional feature vector. It is common practice to set the number of cells in the horizontal direction to 8.

Grid Y:

The amount of vertically oriented cells. The resulting feature vector has a higher dimensionality the more cells there are and the finer the grid is. Usually, it is setto 8.

The initial computational phase of the LBPH is to produce an intermediate image that, by emphasizing the face features, more accurately describes the original image. The algorithm does this by utilising a sliding window notion depending on the radius and neighbors of the parameter.



Fi g.2. Computation in LBPH

Haar Cascade Classifier Algorithm:

Haar features are digital image properties utilized for detecting objects in images, especially faces. These features are named after Haar wavelets due to their similar appearance. The Haar cascade classifier is a well-known computer vision technique that uses machine learning to accurately and swiftly detect objects in images. The Integral Image concept is employed by this technique to enable efficient computation of the detector's feature set. Furthermore, the AdaBoost learning algorithm is employe to select a small subset of vital features from a large feature set, resulting in highly effective classifiers. The Integral Image method entails using the corner values of a rectangle to calculate its total pixel count, lowering the computational expense of feature calculation.

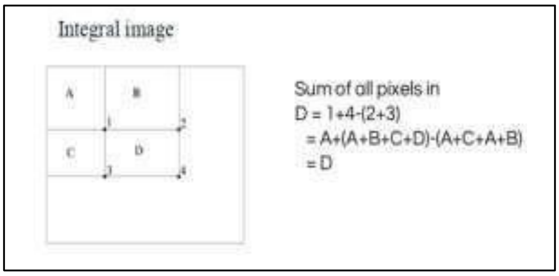


Fig.3. Haar Feature Extraction

Convolutional Neural Network:

A deep cascaded multi-task framework is used to enhance performance by leveraging the connection between detection and alignment. The framework uses a cascaded architecture with three levels of deep convolutional networks to predict face and landmark placement in a coarse-to-fine manner. In the first step, a shallow CNN quickly generates candidate windows. The windows are then refined by a more intricate CNN to reject non-face windows. Finally, a more powerful CNN refines the results and produces the positions of the face landmarks.

The cascade model (PNet, R-Net, and O-Net) is a multi-task network as each of the three models is trained on three tasks, including face classification, bounding box regression, and localization of facial landmarks. The outcomes of the previous stage are given as input to the next one, allowing additional data processing between phases. To illustrate, the candidate bounding boxes produced by the first-stage P-Net are filtered using non-maximum suppression (NMS) before being passed to the second-stage R-Net model.

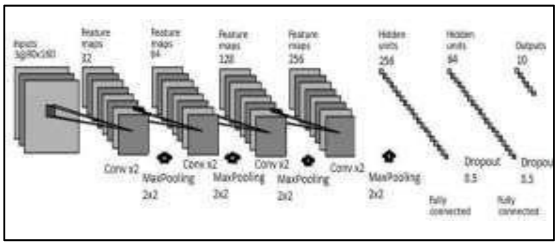


Fig .4. CNN Architecture

Implementing the MTCNN architecture takes some effort, but there are already trained models and open- source implementations of the architecture available for face identification on new datasets.

Stage 1 of the MTCNN framework uses the Proposal Network (PNet), which is a fully convolutional network that produces potential face windows and their bounding box regression vectors. These candidates are then calibrated based on the calculated regression vectors, and non-maximum suppression (NMS) is used to merge candidates with significant overlap.

In Stage 2, all candidates are sent to the Refine Network (R-Net), another CNN that performs bounding box regression calibration, NMS, and further rejects a large number of incorrect candidates.

In Stage 3, the network identifies facial regions more precisely, and in particular, produces five different facial landmarks.

Compared to other picture classification techniques, CNN-based face recognition requires minimal pre- processing, meaning the network learns filters instead of relying on manually constructed filters as in traditional techniques. This independence from prior knowledge and human effort in feature design is asignificant advantage.

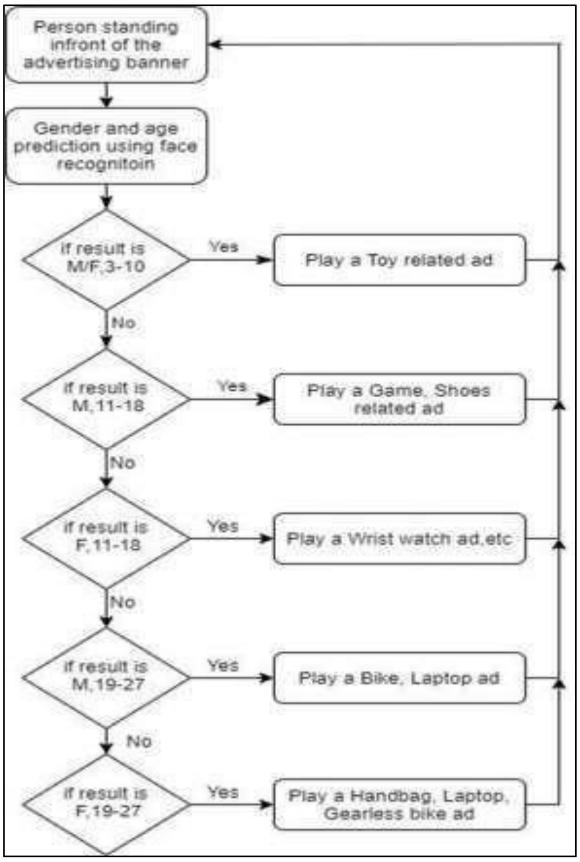


Fig.5. Flowchart of Smart Advertising

FACE RECOGNITION WAS DETECTED BYDEEP LEARNING ALGORITHM

We use a library file 'OpenCV'

1. OpenCV is a free to use software library for computer vision and machine learning. To hasten the inclusion of artificial intelligence into goods, OpenCV was used to develop a common infrastructure for computer vision applications. To carry out LBPH, I created the lbph.py module, which offers the required features. Face detection has been accomplished using the viola jones haar cascade classifier and face identification has been accomplished using local binary pattern histograms (lbph).

2. Before entering their name and password, the user must first start facerecognition.py and select 1 to train the algorithm. The software generates 64 images of the user and keeps them in a database folder. Once the algorithm has

performed LBPH on the images, the trained classifier is then saved in trainedRec.py. The user can then select recognise to check his authorisation status. The threshold provided by the condition if (prediction [1])100 should be changed to change the sensitivity of face recognition.

3. Local binary pattern histograms:Consider a grayscale facial image that can be segmented into a 3x3 matrix holding pixel intensities ranging from 0 to 255. The matrix's central value is considered the threshold, following which, the eight neighboring pixels are assigned new binary values. The binary values of 0 and 1 correspond to the pixels with intensities below and above the threshold, respectively. Performing this procedure on all the pixels (except the central one) produces a binary matrix. The concatenation of binary values from each location in the matrix results in a single binary value (e.g., 10001101), which replaces the original central pixel value after converting it to decimal. The process, known as LBP, generates a new image that characterizes the original image more precisely. Subsequently, the generated image can be divided into various grid sizes for further examination. It should be noted that although people concatenate binary values differently, the output remains the same. Histograms can be extracted from the grayscale image by dividing it into grids and calculating the histogram for each grid. The histogram will contain 256 positions (0-255) to represent the occurrence of each pixel intensity. All histograms can then be combined into a single, larger histogram, which will have 16,384 positions (8x8x256) for an 8x8 grid. This final histogram provides insight into the characteristics of the original image. However, it is worth noting that the accuracy of the histogram and the performance of the face recognition system can be influenced by the chosen grid size.

4. Age and Gender prediction:

We can use deep learning to predict a person's gender and age from a single image of their face. Tal Hassner and Gil Levi have developed models that can accomplish this task. The models can predict the gender as either male or female, and the age can be predicted within one of eight ranges, including (0-2), (4-6), (8-12), (15-20), (25-32), (38-43), (48-53), and (60-100). These predictions are based on the final soft max layer of the trained models. It is important to note that accurate predictions depend on the quality of the input image and the performance of the deep-learning models. It is difficult to accurately determine a person's age from a single image due to various factors such as makeup, lighting, facial expressions, and obstacles in the image. To tackle this challenge, the problem can be reframed as a classification task instead of a regression task.

5. The library functions are used in this program:

- OpenCV pip install OpenCV-python
- argparse pip install argparse

6. The Data Sets:

You may find the public domain version of the Adience dataset, which we had used, here. This dataset is used as a benchmark for face pictures and contains a variety of real-world imaging situations, such as noise, illumination, pose, and look. The images have a Creative Commons (CC) licence and were obtained from Flickr galleries. It has a total of 26,580 photos with 2,284 subjects in eight age categories, and it is roughly 1GB in size. These are the models that we utilised after learning them from this dataset.

AGE AND GENDER PREDICTION USING CNN

Using age and gender predictions on unfiltered faces, unfiltered real-world facial images are categorised into predefined age and gender groups. This research area has advanced significantly because of its importance in sophisticated practical uses. The unfiltered benchmarks, however, reveal that the established methods are unable to manage appreciable levels of variation in such unrestricted images. In this study, we propose a novel end-to-end CNN method for accurate age group and gender classification of unfiltered real-world faces. The two-level CNN design includes the feature extraction and classification procedure in its entirety. The feature extraction pulls out traits that match age and gender while the classification assigns the face photos to the proper age group and gender. Strong picture pre-processing algorithms allow for Due to their superior performance in facial analysis, Convolutional Neural Networks (CNNs)-based methods have recently gained popularity for the classification task. Before giving the CNN model the unfiltered real-world faces, we explicitly handle the significant variations in those faces. The neural network component that pulls features employs wide residual networks, also known as wide residual networks. It uses convolutional neural networks, or ConvNets for brief, to learn the facial features. From eyes and mouth, which are more abstract features,to edges and corners, which are less abstract qualities

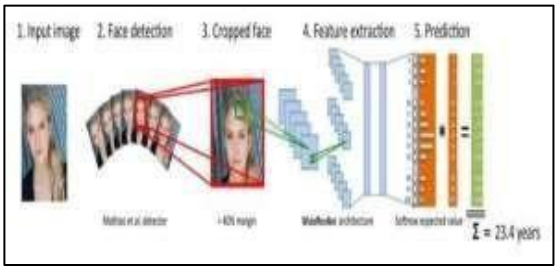


Fig.6. Stages of Age and Gender Prediction using CNN

V EXPERIMENTAL RESULTS

Because it works better in accuracy, good light, medium light, and low light conditions than the Haar Cascade Classifier and LBH algorithm, the CNN-based face recognition algorithm has been used for Smart Advertising.

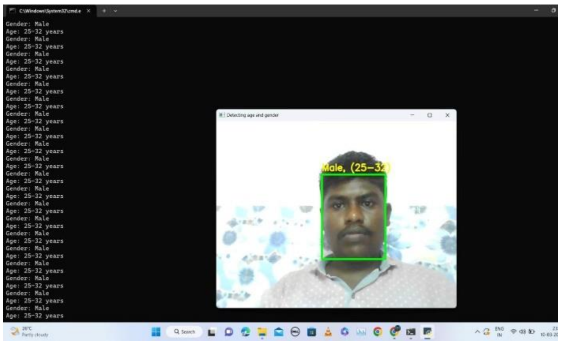


Fig.6.1 Age Prediction using CNN

- When there is no one in front of the banner,no advertisement is played.

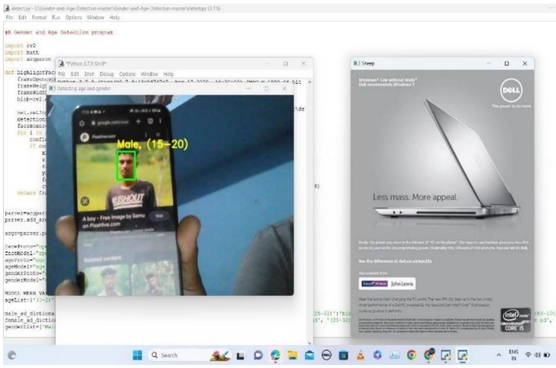


Fig.6.2 A person (Male, Age-20) in front of the camera

- When there is a person (Male, Age-20) in front of the banner, a laptop advertisement is played.

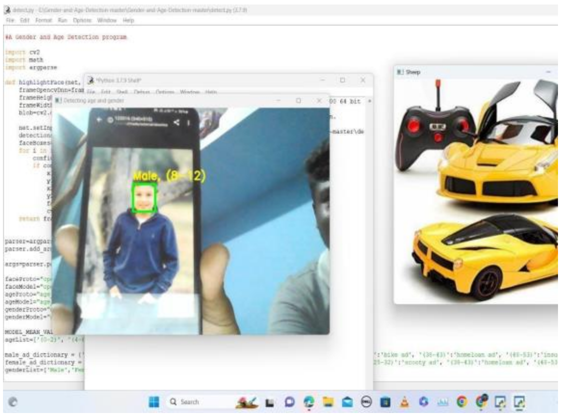


Fig.6.3 A child (Male, Age-10) in front of the camera

- When there is a child (Male, age 10) in front ofthe banner, a Toy car advertisement is played.

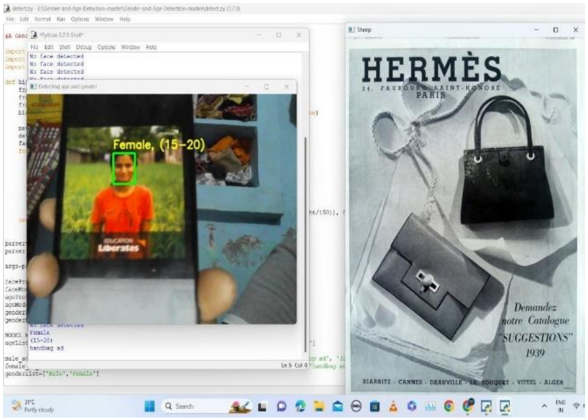


Fig.6.4 A person (Female, Age-20) in front of the camera

- When there is a person (female, Age-20) in front of the banner, a ladies bag advertisementis played.

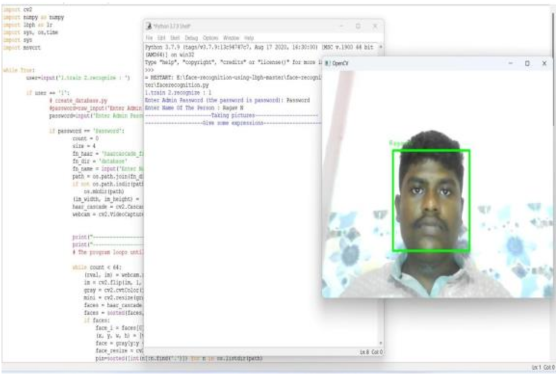


fig.6.5 Compilation in the software

HAAR CASCADE CLASSIFIER OUTPUT

Fig.6.6 Haar cascade in low light

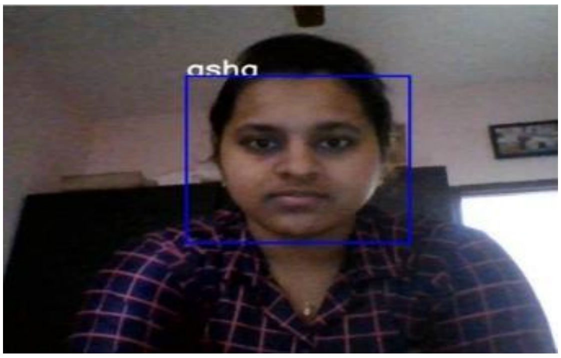


Fig.6.7 Haar cascade in Medium Light

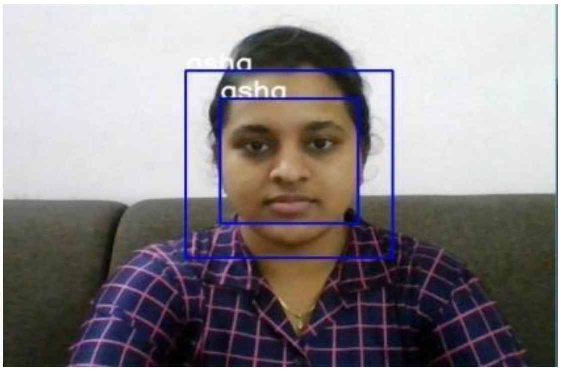


Fig.6.8 Haar cascade in outdoor light

LOCAL BINARY PATTERN HISTOGRAM OUTPUT

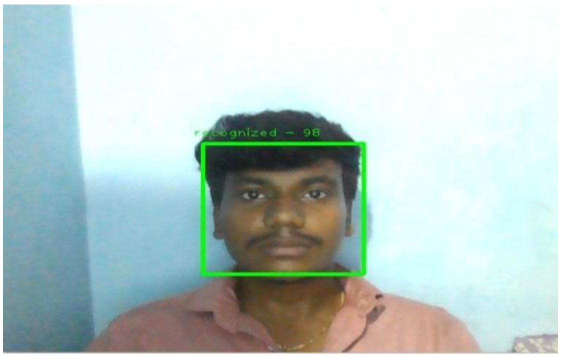


Fig.6.9 LBPH in low light

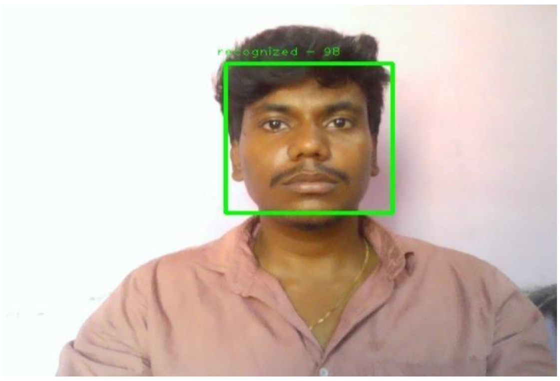


Fig.6.10 LBPH in medium light

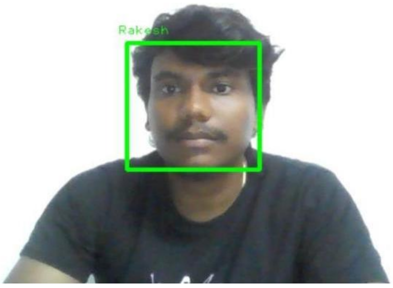


Fig.6.11 LBPH in outdoor light

CONVOLUTIONAL NEURAL NETWORK OUTPUT

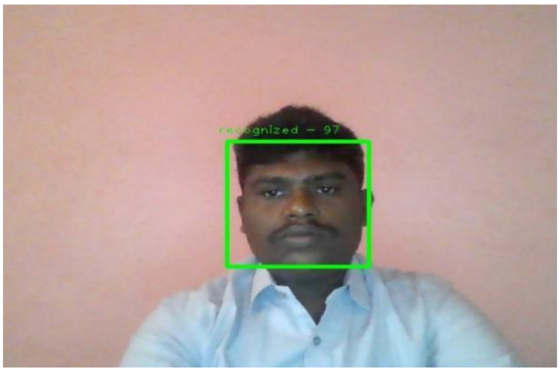


Fig.6.12 CNN in Low Light

Fig.6.13 CNN in Medium Light

Fig.6.14 CNN in Outdoor Light

ANALYSIS ON THE ALGORITHMS COMPUTATIONAL COMPLEXITY

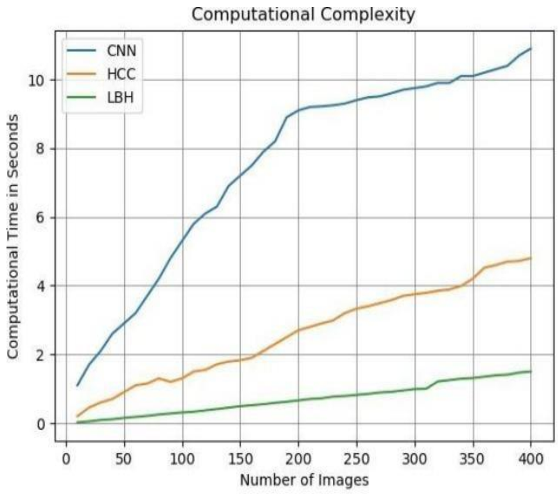


Fig.6.15 Computational ComplexityGraph

When the number of images increases, CNN algorithm takes more time but Haar Cascade and LBH takes lesser time compared to CNN. But it depends on the hardware used in the machine.

MEMORY USAGE

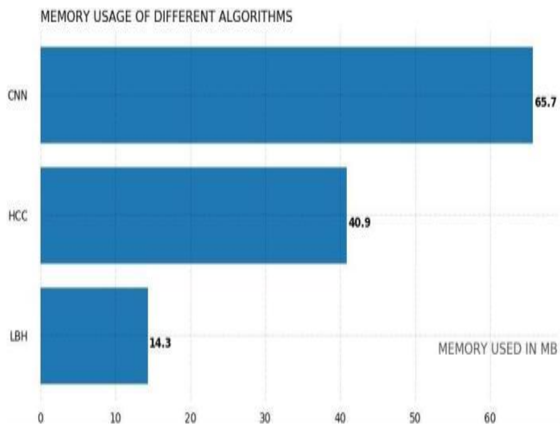


Fig.6.16 Memory Usage Graph

CNN algorithm takes more memory when compared to Haar Cascade Classifier and LBH algorithm.

ACCURACY

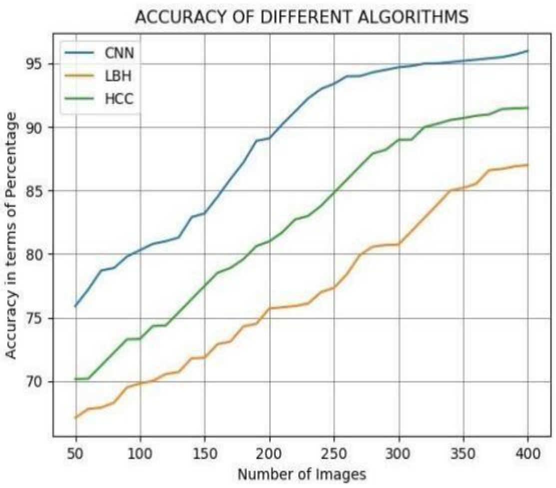


Fig.6.17 Accuracy Graph

In terms of Accuracy, CNN algorithm performs well when compared to Haar Cascade Classifier and LBH algorithm.

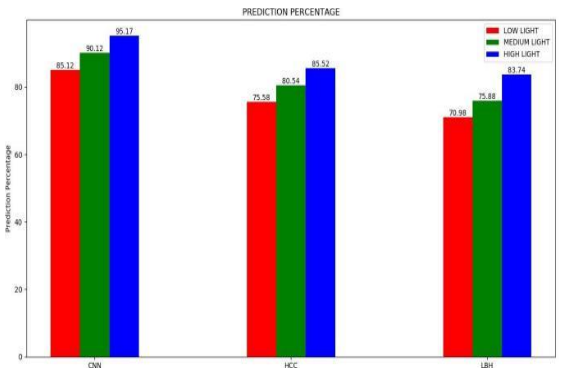


Fig.6.18 Prediction Percentage Graph

The CNN based algorithm performs well in Good Light, Medium Light and also in Low light conditions.

VI CONCLUSION

Face detection has many applications today. The accuracy of face recognition differs depending on the algorithm used. Three face recognition algorithms are looked at in this research. Convolutional Neural Network (CNN)-based face recognition worked well in all of the scenarios considered, including low light, medium light, and good light. Despite having a higher prediction rate than other algorithms, the CNN-based algorithm requires more memory and has a higher level of computational complexity. Because of its greater accuracy, CNN is used to forecast age and gender. The age and gender data will be used to show an advertisement film. By comparing the algorithms, you can choose the one that will make predictions more correctly. By playing advertisements videos based on age and gender, it will be possible to reach the appropriate group of people with the products in a larger quantity. Banners placed in crowded public spaces like bus stops, metro stations, and shopping malls can use this sort of clever advertising. The only drawback is that it requires appropriate controllers and processors in addition to cameras in the banner, raising the price of the advertising banner. This can be improved further by simplifying and reducing memory-intensive calculations in the CNN algorithm. The project's long-term objective is to reduce working time and memory usage.

FUTURE SCOPE

By making the CNN algorithm's computations simpler and less memory-intensive, this can be further

enhanced. It is possible to decrease memory usage and computational time, which is the project's future goal. threatens privacy, restricts personal freedom, infringes on individual rights, and creates data risks.

VII REFERENCES

[1] Kim, G.; Choi, I.; Li, Q.Kim, J. “A CNN-Based Advertisement Recommendation through Real-Time User Face Recognition”. Appl. Sci. 2021

[2] Xi Yin, Xiaoming Liu, “Multi-Task Convolutional Neural Network for Pose Invariant Face Recognition”, IEEE Transactions on Image

Processing, 2017.

[3] Chengyu Guo, Jingyun Liang, Geng Zhan, Zhong Liu, "Extended Local Binary Patterns for Efficient and Robust Spontaneous Facial Micro- Expression Recognition", IEEE Access, 2017.

[4] Mellor, C. (2016). "Cloudian clobbers car drivers with targeted ads • The Register". Retrieved August 24, 2017, from https://www.theregister.co.uk/2016/06/21/cloudian_could_clobber_car_drives_with_targeted_ads/

[5] Chen, D., Ren, S., Wei, Y., Cao, X., & Sun, J. (2014). "Joint cascade face detection and alignment". Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8694 LNCS(PART 6), 109–122. https://doi.org/10.1007/978-3-31910599-4_8.

[6] Ferreira, a J., & Figueiredo, M. a T. (2012). "Boosting Algorithms: A Review of Methods, Theory, and Applications". Ensemble Machine Learning: Methods and Applications, 35. https://doi.org/10.1007/9781441993267_2

[7] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., … Adam, H. (2017). "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications". https://doi.org/arXiv:1704.04861.

[8] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., … Zitnick, C. L. (2014). "Microsoft COCO: Common objects in context". Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8693 LNCS (PART 5), 740–755. https://doi.org/10.1007/978-3-31910602-1_48. 31

[9] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). "SSD: Single shot multibox detector". Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9905 LNCS, 21–37. https://doi.org/10.1007/978-3-319-46448-0_2.

[10] Ren, S., He, K., Girshick, R., Zhang, X., & Sun, J. (2015). "Object Detection Networks on Convolutional Feature Maps", 1–8. https://doi.org/10.1109/TPAMI.2016.2601099.

[11] Richter, S. R., Vineet, V., Roth, S., & Koltun, V. (2016). "Playing for data: Ground truth from computer games". Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9906 LNCS, 102– 118. https://doi.org/10.1007/978-3319-46475- 6_7.