Deep Learning Model for Face Identification using Localization of Faces

Dnyanesh Walwadkar
dept. EECS, Queen Mary University of London
London, United Kingdom

Abstract:- An unknown face in a face dataset can be identified using a deep learning model and a wide range of issues can be addressed based on it. Identifying celebrities' faces for tagging in photos or other deep learning functions is one example. It can be used to identify missing people, identify criminals, verify government documents, collect images, and search for images.

To address this issue. (1) The paper provides a comparison between the challenges and solutions proposed by the researchers and (2) the use of VGG Face models with localized faces to solve organizational face comparison problems using 113 different classes and 8664 images with different ages, angles, and Spectacles. Analyzes traditional and deep learning-based models. Good accuracy was achieved by identifying the face from the image and performing the identification on the identified face rather than the entire image.

Keywords:- *Deep Learning, VGG, Face Identification, Face Localisation.*

I. INTRODUCTION

Facial recognition has received a lot of attention due to numerous applications in access control, law enforcement, security, surveillance, internet communications, and computer entertainment. Although significant advances have been made, state-of-the-art facial recognition systems provide satisfactory performance only in controlled scenarios and are significantly degraded in the face of real-world scenarios. Real-world scenarios have unlimited conditions such as lighting and pose variations, occlusions, and facial expressions. So there are still many challenges and opportunities in front of us.

Face comparison has always been a daunting task for researchers, as the appearance of a person's face can change over a short period of time. For example, the same person looks quite different due to differences in head poses, facial expressions, lighting, aging, and the presence of accessories. Therefore, it is very difficult to compare two facial images of the same person. Comparing faces over age (comparing two facial images of people taken at different ages) is important, but it is rarely dealt with by computer vision. Previous studies have focused on changes in facial poses, lighting, and facial expressions (PIEs). However, less research has been done by researchers on the issue of aging. Face comparisons can be used in many applications. Search for missing persons, check government documents and criminal identification.[1]

However, in reality, facial recognition technology faces many challenges, such as changing lighting conditions, facial expressions, and facial obstructions caused by obstacles such as sunglasses, masks, and scarves.

Face recognition algorithms require intensive calculations and are thought to make face recognition difficult to perform. Most face recognition processes are implemented on one face at a time. Recognizing a single face can take a short time, but for many faces/people, using a single face recognition can take a long time. Therefore, in order to speed up the recognition process, it is necessary to develop a system that recognizes multiple faces at once.[6] This paper implements research on multiple face detection using the hybrid method of Haar Cascade. This study aims to improve the performance of the face identification process using the Haar Cascade.

Verification of a photo of the same person or of two different people is done using facial features such as facial expressions, beauty ratings, facial emotion, and more. The app will display "Photos match, person is XYZ" if the photo matches an existing record. Otherwise, the app returns "No match".

II. CHALLENGES IN FACE IDENTIFICATION

A. Illumination Changes

A human face appears differently depending on factors such as illumination (spectrum, light source distribution and intensity) and camera characteristics (sensor response and lens). Lighting changes can also influence this due to the reflective properties of the skin and internal camera control.

B. Angle of View

Face images are affected by the angle of the camera (front, 45°, profile, upside down) and partial or complete concealment of some facial features can be observed. It may have been. In fact, postural changes affect the cognitive process by introducing projection deformation and self-occlusion.

C. Skin Anging and Wrinkles

Aging can be natural (due to aging) or artificial (using cosmetic tools). In both cases, aging and wrinkles can significantly impair the performance of facial recognition methods. In general, the effects of age variation and age factors are not widely considered in facial recognition studies.

III. GENERAL TERMS IN FACE RECOGNITION

A. Face Identification

This means searching for an unknown face image from the face dataset and returning its name. This is shown in Figure 1. The problem is simplified with a 1: n mapping face dataset.

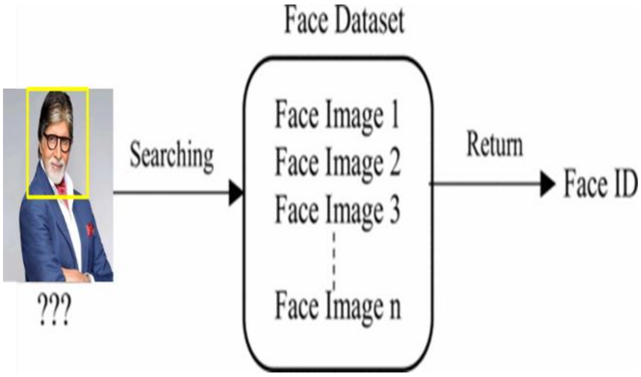

Fig. 1: Face Identification

B. Face Verification

This refers to comparing one face to another to see if they match. The problem is simplified to a 1: 1 mapping. Figure 2. Shows a model for face recognition.

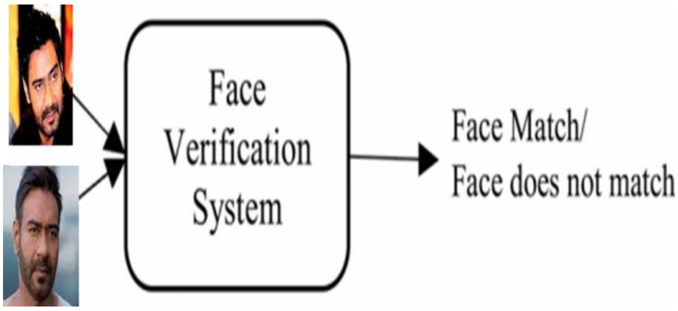

Fig. 2: Face Verification

C. Face Comparison

This refers to comparing one face image with another and comparing the similarity score between two input face images. It has also been simplified to a 1: 1 mapping. Figure 3 shows a face comparison model.

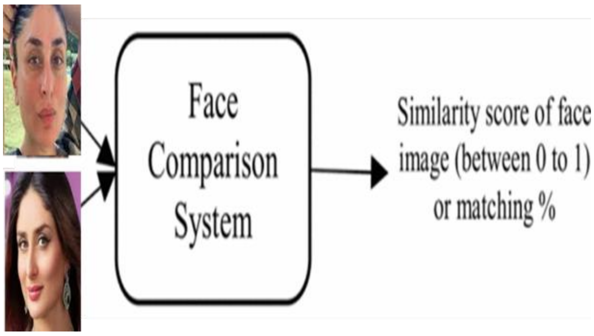

Fig. 3: Face Comparison

IV. ARCHITECTURE FOR FACE IDENTIIFCATION SYSTEM

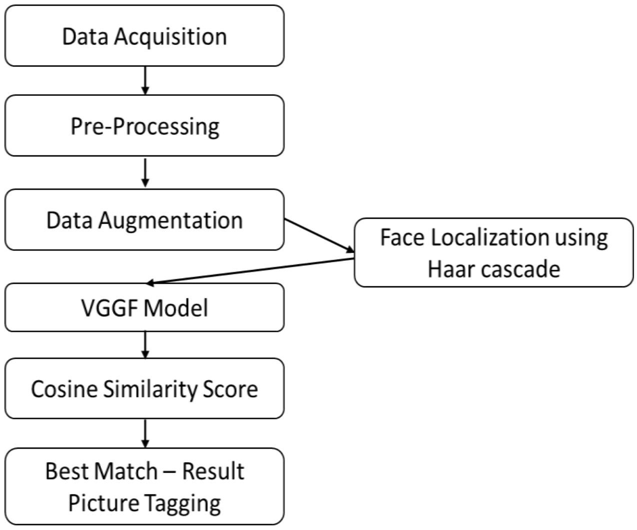

Fig. 4: Architecture

A. Data Acquisition / Data Collection

For FI systems, you need a dataset that contains images of the same subject of different ages. Many datasets are available, including images such as pose images, wild images, frontal face images, illumination difference images, and age difference images.[1]

B. Preprocessing and Data Expansion

Pretreatment methods can be divided into three main categories: 1) lighting normalization, 2) face detection, and 3) pose normalization. In the deep learning approach, we recommend that you grow the data after preprocessing and apply various operations such as rotation, scaling, shearing, and mirroring to increase the size of the training dataset. [1]

C. Model Architecture

a) Face Detection:

In real-time scenarios, the face image captured by the camera includes the background and face image. To classify face images, this article uses the Haar Cascade Classifier to detect faces from images. For face recognition in a variety of lighting conditions, we used a fast and sufficiently good Haar Cascade classifier. The Haar Classifiers method classifies organs using the statistical properties of facial features . we implemented HaarCascadeClassifier using Open Computer Vision Library (OpenCV)[5]. Results of face detection from original images are as follows:

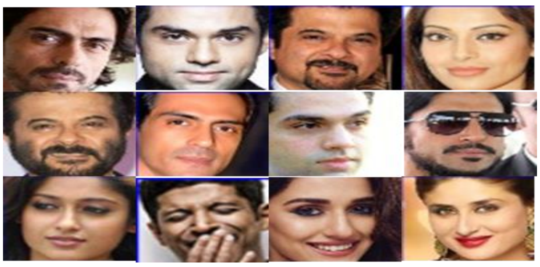

Fig. 5: Face Detection

b) VGGF Model
Our study relied on this model because it shows its efficiency in face recognition problems. The Visual Geometry Group created the VGF Face model and trained the model with a vast set of facial data containing 2. 6 million pain images from over 26,000 different people. The VGG16 model has 16 layers and the VGG Face model has 38 layers, so it has the same neural network structure as the VGG16 model, but with a different number of layers that can be trained.[2] In our VGGF Architecture Total trainable params are 145,002,878.

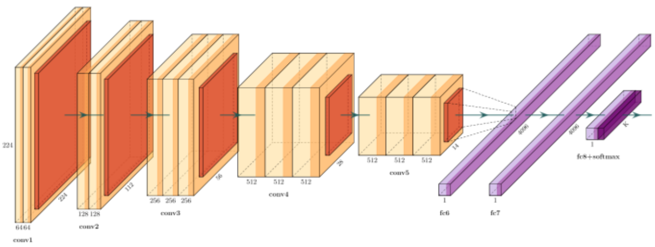

Fig. 6: VGG16 Model

The model captures an image in the format of 224 * 224. The application takes a picture from the user and changes its shape to 224 * 224 according to the requirements of the model. Next, cosine-similarity images containing different faces are classified into different classes.

To reduce computational effort and false-positive rates, the rough areas of interest (ROIs) of the eyes, nose, lips, and eyebrows were selected based on the geometric position of the face. All parts were individually detected using a haar classifier trained for each part. The hair classifier returns the vertices of the detected rectangular area of the face.

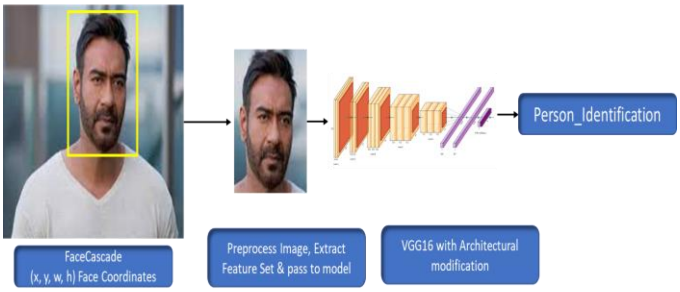

Fig. 7: Pipeline

c) Analysis of Datasets
This dataset carries the localized face of one hundred celebrities. each magnificence has samples among eighty to a hundred and fifty of length 64 X 64 pixels. samples include wild situations which includes distinct orientations, illuminations, age transitions, etc The dataset contains a total of 8664 images with 113 different classes (celebrities), different age groups up to 5 to 6 years duration, and different wearables such as eyeglasses. The dataset contains images of men and women. The validation dataset contains 440 images.

Fig. 8: Tagging

Fig. 8: Tagging

This dataset is used to test generative and facial recognition models. It can also be used to reconstruct image patterns using generative models such as variational autoencoders and generative hostile networks.

Database Name	Image Size	Images	Subjects	Colour	Imaging Conditions
FERT	256*384	14126	1564	colour	Controlled
FRGC	1200*2272	12776	688	colour	uncontrolled
Our Used Dataset	64*64	8664	113	colour	uncontrolled

V. RESULTS & EXPERIMENTS

Face detection and identification algorithms work well with data for the entire face, but for partially occluded faces, these algorithms are works partially. After using the VGGF model architecture, the average score improved by 10-15% for each category.

VGG16, ResNet50, etc. are deep architectures of convolutional neural networks for images.

Such architectures are typically trained to classify images into one of 1000 possible categories. Such an architecture typically "consumes" an RGB image with a size of 224x224 and uses a convolution layer to extract visual features on five different scales. At the end of the network, the width and height are 1/32 of the original image.

The model can recognize facial images with the same facial expressions and characteristics. You can use the same model to identify faces that are similar to a particular face.

VI. CONCLUSIONS

Good accuracy was achieved by identifying the face from the image and performing the identification on the identified face rather than the entire image. We were able to achieve average verification accuracy of 86.8%, 83.47%, and 79.34% for VGG16, ResNet50, and SENet50, respectively. Besides accuracy, other performance matrices used in this task are precision and recall.

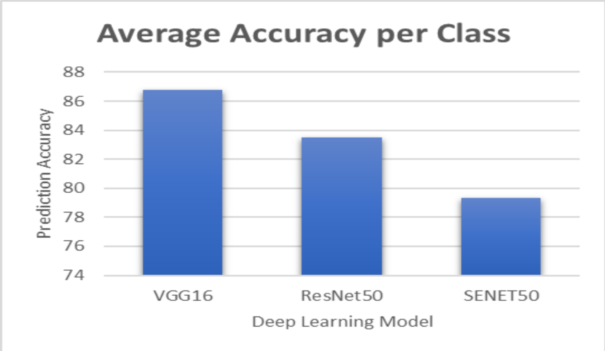

Fig. 9: results comparison

Our assessment based on these performance matrices shows that the VGG16 is the most accurate and provides accurate face detection from all three networks after localizing faces from low-medium quality images using Haar Cascade.

REFERENCES

[1.] S. D. Ginoya, H. B. Prajapati and V. K. Dabhi, "Survey on Cross-Age Face Comparison," *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2020, pp. 445-451, doi: 10.1109/ICACCS48705.2020.9074389.

[2.] S. Alfattama, P. Kanungo and S. K. Bisoy, "Face Recognition from Partial Face Data," 2021 International Conference in Advances in Power, Signal, and Information Technology (APSIT), 2021, pp. 1-5, doi: 10.1109/APSIT52773.2021.9641286.

[3.] D. Gyawali, P. Pokharel, A. Chauhan and S. C. Shakya, "Age Range Estimation Using MTCNN and VGG-Face Model," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020, pp. 1-6, doi: 10.1109/ICCCNT49239.2020.9225443.

[4.] P. Goel and S. Agarwal, "Hybrid Approach of Haar Cascade Classifiers and Geometrical Properties of Facial Features Applied to Illumination Invariant Gender Classification System," 2012 International Conference on Computing Sciences, 2012, pp. 132-136, doi: 10.1109/ICCS.2012.40.

[5.] Open Computer Vision Library http://sourceforge.net/projects/ opencvlibrary/

[6.] T. Mantoro, M. A. Ayu and Suhendi, "Multi-Faces Recognition Process Using Haar Cascades and Eigenface Methods," 2018 6th International Conference on Multimedia Computing and Systems (ICMCS), 2018, pp. 1-5, doi: 10.1109/ICMCS.2018.8525935.