iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XI **Month of publication:** November 2023

DOI: https://doi.org/10.22214/ijraset.2023.56680

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

CNN Based Facial Recognition with Age Invariance

Mr. Abhilash L Bhat[1], N Nithesh Kumar[2], Poojitha Y[3], Siripireddy Thulasi[4], V Arvind[5]
Department of Computer Science and Engineering, K S Institute of Technology, Bengaluru, India

Abstract : ***As the world has seen exponential expansion over the previous decade, there has been an unusual increase in the crime rate as well as an increase in the number of criminals/missing persons. Face recognition can extract the individualistic characteristics of the human face. A straightforward and adaptable biometric technology is face recognition. The technology used to recognize and identify faces in images or videos is called face detection and recognition. The process of removing facial features has gotten easier as technology has advanced. This study describes the use of an automated security camera for real-time face recognition. With this system, we can instantly identify and detect the faces of the criminals in a live video feed captured by a camera. Criminal records typically include the offender's picture and personal information. Thus, we are able to use these photos in addition to his information. The security camera's recorded footage is transformed into frames. After a face is identified in a frame, it undergoes pre-processing and feature extraction. The characteristics of the real-time image processing are compared to those of the images kept in the criminal database.***

Keywords: CNN (Convolution Neural Networks), Deep learning, Face net, MTCNN, Image Processing, Python, Keras.

I. INTRODUCTION

The main objective of the age-invariant facial recognition project using convolutional neural networks (CNN) is to develop a system that can reliably and accurately identify people of any age. This means creating a strong facial recognition model that works well for a variety of age groups, guaranteeing and authentication capabilities. One of the most important and pervasive problems in our society is crime/missing of a person, and it is our duty to prevent it. In any society, there are many different kinds of crimes and the protection and safety of its citizens must be taken into full consideration. These are important factors that directly affect the citizens' quality of life. A person's life can be disturbed and stressed out by some criminal incidents, like theft, identity theft, or even pickpocketing.

The increasing concerns about crime and its threat to security and safety have led to the widespread adoption of closed-circuit television (CCTV) systems in both public and private settings. Since a deep learning-based approach outperforms existing methods in terms of performance and speed, it is used to provide real-time data that police forces can use to operate more effectively. In the context of the convolutional neural network (CNN) based age-invariant facial recognition project, the main goal is to build an advanced system that can accurately identify and authenticate people of any age. This entails creating and putting into practice a strong CNN-based model that can manage the various facial traits that people of various ages have, improving the facial recognition system's overall accuracy and inclusivity. The project's main goal is to create a CNN-powered facial recognition system that works for people of all ages

II. RELATED WORK

Face detection and alignment are necessary for many facial applications, including face recognition and facial expression analysis. Large visual fluctuations of faces, such as occlusions, significant pose fluctuations, and severe lightings, present major challenges for these tasks in real-world applications. A cascade face detector with good performance and real-time efficiency is provided by Viola and Jones. It trains cascaded classifiers using AdaBoost and Haar-Like features. Nevertheless, a number of studies demonstrate that in real-world applications where there are greater visual variations in human faces, this kind of detector may perform significantly worse, even with more sophisticated features and classifiers. Apart from the cascade structure, Mathias offered face detection deformable component models that worked quite well.

They may, however, require expensive annotation during the training phase and are computationally demanding. Recently, convolutional neural networks (CNNs) have achieved notable progress in various computer vision applications such as face recognition and image classification. In order to obtain a high response in face regions and identify candidate face windows, deep neural networks are trained to identify facial attributes. However, because of its intricate CNN structure, this technique takes a lot of time to implement in practice.

©IJRASET: All Rights are Reserved

The inherent connection between bounding box regression and facial landmark localization is overlooked when using cascaded CNNs for face detection, which results in additional processing costs for bounding box calibration from face detection. Convolutional neural networks (CNNs) have shown remarkable success in image classification and face recognition, and recent advances in this field have shown promise for revolutionizing a variety of computer vision applications.

Deep CNNs, specifically designed for facial attribute identification, show a tendency to produce large responsive face regions, which results in candidate face windows. However, because of CNNs' complex structure, this technique requires a lot of resources to implement in practice. The model can concentrate on particular areas of the input image thanks to attention mechanisms, which is especially useful when dealing with difficult situations like occlusions and changing lighting. By utilising convolutional neural networks' potent representation learning capabilities in conjunction with attention mechanisms, these methods seek to increase the robustness and accuracy of face detection systems in practical applications.

III. METHODOLOGY

A. Mtcnn

The Multi Task Cascade Neural Network, or MTCNN, is a neural network that can identify faces and facial landmarks in photos and videos. There are three stages to the MTCNN concept, the third of which entails simultaneously identifying facial landmarks and facial detection. These stages are made up of different CNNs with different levels of complexity. In the first step, the MTCNN generates a large number of frames that scan the entire image from the top left corner to the bottom right corner. P-Net, or proposal net, is the name of the information retrieval process. It is a shallow, fully connected CNN. In the subsequent phase, the entire dataset from PNet serves as input for the R-Net (Refinement Network), a complex, fully connected CNN that rejects most of the frames that don't have faces. The third and last step involves a more sophisticated and potent CNN called O-Net (Output Network), which, as its name implies, outputs the facial landmark position when a face is detected in the provided image or video.

Facenet:

Facial features are encoded into embedding vectors during the FaceNet algorithm's creation of face embedding. This results in a representation where vectors from two different photos of the same person are closer together and those from different people are farther apart. The distance between the face encodings produced by the Encoder network (Inception- ResNet-v1) is the metric used to compare two faces. The Triplet Loss is used to train the Encoder network, which requires skilled Triplet Mining. Triplet mining is the process of choosing triplets of face images (anchor, positive, and negative) in which the anchor represents the face of one person, the positive represents the image of a different person, and the negative represents the image of a different person. Facial features are encoded into embedding vectors during the FaceNet algorithm's creation of face embedding same person are closer together and those from different people are farther apart. The distance between the face encodings produced by the Encoder network (Inception-ResNet-v1) is the metric used to compare two faces.

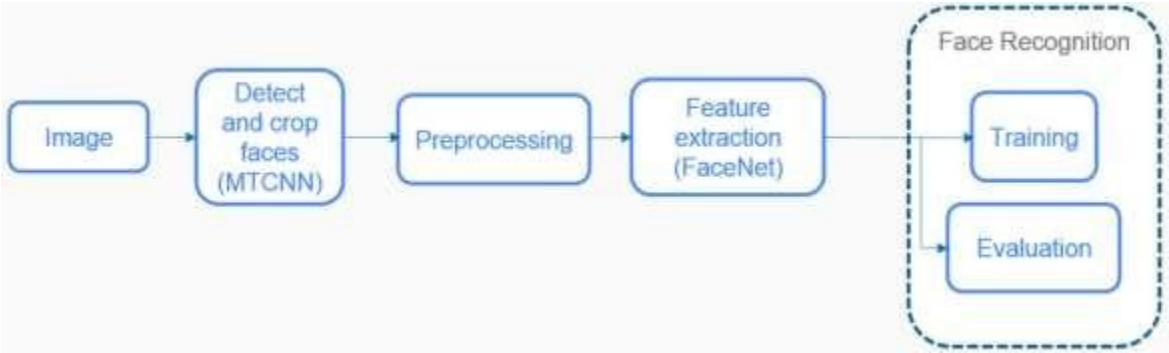


Fig. 1 Flow diagram Flowchart for face detection and face recognition

Refining the face embeddings for improved facial recognition performance involves minimizing the distance between the anchor and positive and maximizing the distance between the anchor and negative. In order to maximize learning efficiency and guarantee that the encoder network generates distinct and well-separated face embeddings, efficient triplet mining is essential. By iteratively selecting informative triplets, this rigorous Triplet Mining process helps the Encoder network learn robust facial representations, leading to improved face recognition accuracy in the resulting embeddings.

By leveraging effective Triplet Mining in the training of the Encoder network, the FaceNet algorithm ensures that the embedded facial features are finely tuned for discerning similarities and differences, enhancing the overall precision of face recognition systems.

B. CNN

Our approach makes use of transfer learning by automatically extracting features from face images using a multilayer convolutional neural network (CNN) that has been pre-trained. The convolutional neural network's multilayer structure makes it possible for the extracted features to be extremely discriminative and interoperable across ageing variation. Compared to handcrafted features, this method of feature extraction is more resistant to intraperitoneal variability. Because of this, our strategy is better suited for implementation in security systems operating in uncontrolled environments. Artificial neural networks with both fully connected and locally connected layers, referred to as convolutional layers, are called convolutional neural networks. The process of cnn is represented in Figure 2.

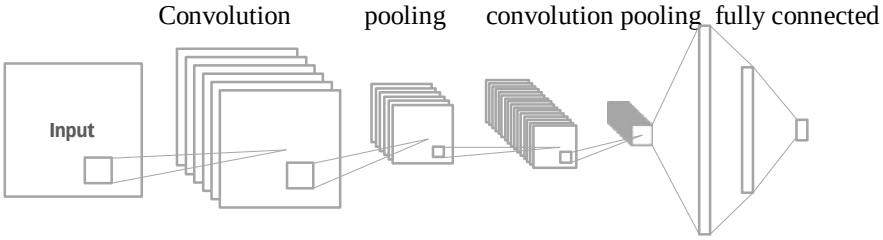


Fig. 2. Cnn process

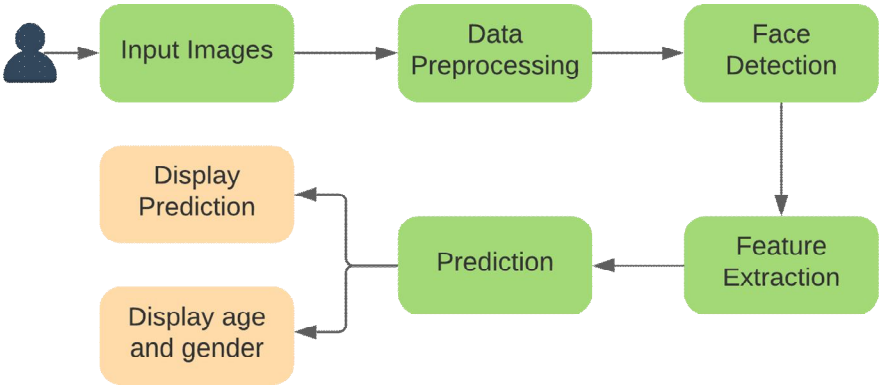


Fig. 3. System Flow

The block diagram in Figure 3 consists of:

1) Input Image

Pictures are entered into the system as input. To guarantee consistent input, images are first obtained and preprocessed, which may include resizing, normalisation, and possible alignment. These images, which have been accurately labelled with age information, are part of a diverse dataset that is subjected to augmentation, which introduces variations for better model generalisation. Using this pre-prepared dataset as training data, the CNN extracts features that allow the system to recognise faces with robustness and to take age-related variations into account.

2) Data Pre-processing

To enhance the efficiency of the face detection and prediction algorithms, the input images undergo pre-processing. This could entail operations like colour normalisation, image resizing, and noise reduction. The input images go through crucial pre-processing procedures in order to maximise the effectiveness of face detection and prediction algorithms.

To standardise the colour distribution and help the model recognise facial features consistently, colour normalisation is used. Furthermore, the algorithm employs image resizing and noise reduction techniques to ensure uniformity in input dimensions and improve its accuracy in face detection and prediction under a variety of conditions.

3) Face Detection

The input images contain faces that the system recognises. Convolutional neural networks (CNNs) and other machine learning algorithms are commonly used for this. Recognition systems use Convolutional Neural Networks (CNNs) and other machine learning algorithms to process the input images, which contain faces.

Accurate face recognition is made possible by these algorithms, which are trained to recognise distinctive patterns and features on faces. Specifically, CNNs are very good at learning hierarchical representations, which enables reliable and accurate face detection and recognition in the input images.

4) Feature Extraction

Following facial detection, features are taken from the images. These characteristics can be used to recognise the faces and forecast characteristics like emotion, gender, and age. After facial detection is finished, features are taken out of the pictures to capture unique qualities that are unique to every face.

The system can recognise faces and predict other attributes like emotion, gender, and age by using these features. This method makes use of the acquired representations to offer a thorough comprehension of facial data, allowing for more complex forecasts than just identification.

5) Prediction

The system forecasts the characteristics of the faces based on the features that were extracted. Usually, a machine learning algorithm like a support vector machine (SVM) is used for this. 6. Prediction Displayed: The predicted features of the faces are shown by the system. Numerous methods are available for accomplishing this, including displaying a table of results or superimposing text on the input images.

The block diagram is summarized as follows in points: To enhance the efficiency of the face detection and prediction algorithms, input images are pre-processed. A CNN-style machine learning algorithm is used to identify faces in the input images. The identified faces are used to extract features.

The predicted attributes of the faces are displayed. A machine learning algorithm, such as an SVM, is used to predict the attributes of the faces using the extracted features

IV. CONCLUSION AND FUTURE SCOPE

In summary, this paper uses deep learning and similarity distances to effectively address the problems associated with face recognition in the ageing context. The study demonstrates superior performance in one-to-one verification and one-to-many identification tasks by utilising a pre-trained convolutional neural network (CNN) to extract discriminative feature descriptors. This highlights the superiority of set distances over singletons. The results, especially the ability to distinguish older subjects from younger ones, have practical implications, especially in situations involving cosmetic surgery and plastic surgery. The study highlights the promise of CNN-based facial recognition with age invariance and makes recommendations for future directions in the field, especially with regard to investigating the use of similarity set distances when confronted with deception and denial difficulties.

Moreover, the employment of minimum distances and minimum modified Hausdorff distances is recognised as a crucial factor in attaining optimal overall efficiency.

The study's emphasis on dealing with deception and denial issues, such as situations involving cosmetic procedures and plastic surgery, emphasises how applicable the suggested approach is to dealing with difficulties that arise in the real world. The ease with which older subjects were recognised in comparison to younger ones provides important new information about the subtleties of age-related facial recognition inside the CNN-based framework.

All things considered, the paper offers a thorough investigation of deep learning methods for facial recognition, with a focus on age invariance, and lays the groundwork for future developments in managing challenging real-world situations.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] Approaches on Partial Face Recognition: A Literature Review Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2021) IEEE Xplore Part Number: CFP19J32- ART; ISBN: 978-1-5386- 9439-8 Reshma MR, Kannan. BAlmajai, S. Cox, R. Harvey, and Y. Lan. Improved speaker independent lip-reading using speaker adaptive training and deep neural networks. In IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 2722–2726, 2016.

[2] Design and Implementation of Remote DeepFace Model Face Recognition System Based on sbRIO FPGA Platform and NB- IOT Module Lu Peng, Zhou Xin, Gan Ping 2019 2nd International Conference on Safety Produce Informatization (IICSPI) Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al.: TensorFlow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467 (2016).

[3] Wasserstein CNN: Learning Invariant Features for NIR-VIS Face Recognition Ran He, Senior Member, IEEE, Xiang Wu, Zhenan Sun, Member, IEEE, and Tieniu Tan, Fellow, IEEE JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST2017

[4] R. He, B. C. Lovell, R. Chellappa, A. K. Jain,and Z. Sun, “Editorial: Special issue on biometrics,” Pattern Recognition, vol. 66,pp. 1–3, 2017

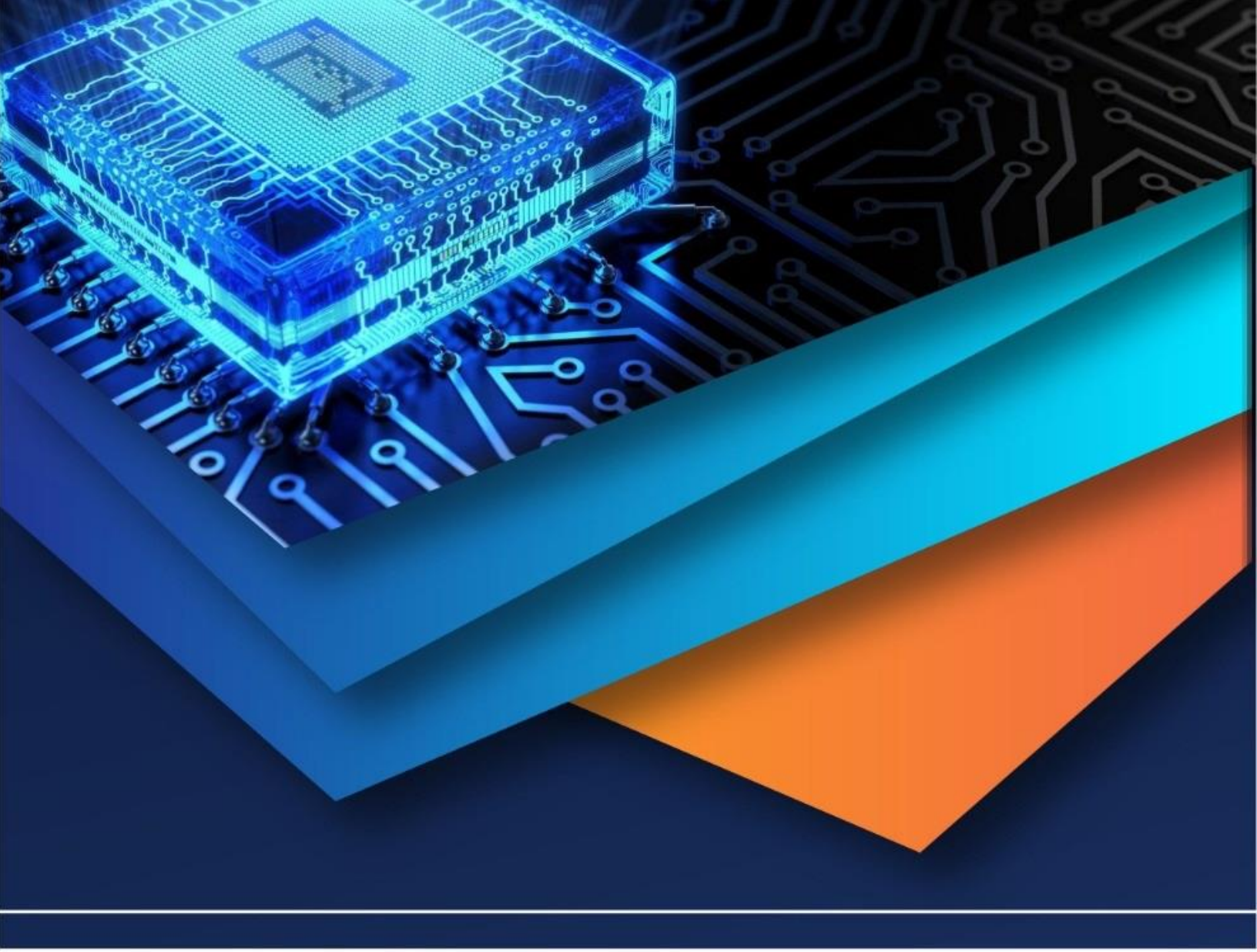
[5] S. Ouyang, T. Hospedales, Y.-Z. Song, X. Li, C. C. Loy, and X. Wang, “A surveyon face recognition: Sketch, infra┐ red, 3d and low-resolution,” Image and Vision Computing, vol. 56, 2016.

[6] Florian Schroff, Dmitry Kalenichenko:“Face Net: A Unified Embedding for Face Recognition and Clustering”, 2015; arXiv:1503.03832.DOI:10.1109/CVPR.2015.7298682.

[7] FaceNet: A unified embedding for face recognition and clustering”, in 2015 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp. 815-823, 2015.

[8] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, “Joint face detection and alignment using multi-task cascaded convolutional networks”

[9] F. Schroff, D. Kalenichenko and J.Philbin, “ FaceNet: A unified embedding for face recognition and clustering,” in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 815-823, 2015. segmentation. The final result is Basedon extraction of border from the detected images



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)