JOURNAL OF DATA SCIENCE | Vol.2024:10
eISSN:2805-5160

Facial Recognition Using Convolutional Neural Network Using Real-Time Data

Bharath Kumara J[1], Manjula Sanjay Koti[1], Nor Idayu Ahmad Azami[2]

[1]Dayananda Sagar Academy of Technology and Management, Karnataka, India.
[2]Faculty of Data Science and Information Technology, INTI International University, Nilai, Malaysia

Email: [1] bharath415reddy@gmail.com, [2] manjula1705vision@gmail.com

Abstract

Recent years have seen the rise of facial recognition as a significant technological advancement, with several applications in the fields including security, surveillance, authentication systems, and Human-Computer Interface. Numerous sectors have undergone radical change as a result of their ability to automatically identify and validate people based on their facial traits, opening new doors for innovation. The main objective of facial recognition is to create automated systems that can correctly identify and validate people from pictures or videos. The limitations of traditional methods in capturing complex and discriminative facial patterns included the reliance on handmade features and shallow learning techniques. However, facial recognition has made great progress since the introduction of deep learning, more notably Convolutional Neural Networks (CNNs). CNNs are the perfect tool for capturing fine facial characteristics because they have demonstrated an amazing capacity for hierarchical representations that can be directly learned from unprocessed image data. In this paper, the authors focus on facial recognition using a CNN model, intending to improve the accuracy and resilience of this crucial technology. The authors have applied a well-built CNN model to address the challenges of facial recognition. We utilize deep learning to automatically identify and extract high-level features from facial images, enabling more accurate and reliable identification. The CNN model's architecture was thoughtfully created to utilize the underlying spatial links and regional patterns visible in facial data. By utilizing a large number of convolutional and pooling layers, the model can successfully capture both low-level qualities like edges and textures and high-level facial traits like facial landmarks and expressions.

Keywords

Facial recognition, CNN model, Deep Learning, Accuracy.

Submission: 19 May 2024; **Acceptance:** 23 June 2024

Introduction

From security systems to customized services, facial recognition technology is becoming more and more common and important. Convolutional Neural Networks (CNNs) have demonstrated extraordinary performance in facial recognition applications because to recent developments in deep learning and computer vision (Krizhevsky, A., et.al., 2012). This study describes a CNN-based facial recognition experiment that involves building a model with a dataset of 300 facial photos of each person and using it to recognize faces in real-time with a camera.

Facial recognition technology is becoming steadily more prevalent and significant in everything from security systems to tailored services (Taigman, Y., et.al., 2014). Due to recent advancements in deep learning and computer vision, CNN's have shown exceptional performance in applications involving facial recognition. This paper outlines a CNN-based facial recognition experiment that uses a dataset of 300 facial images taken of each person to develop a model that can detect faces in real time when a camera is used.

Beginning with an overview of the importance and difficulties of facial recognition, the paper will go on to address its uses and the requirement for precise and effective face identification systems (Parkhi, O. M., Vedaldi, A., & Zisserman, A., 2015). It will emphasize the benefits of employing CNNs for facial recognition, such as their capacity to automatically learn features and their accomplishments in a variety of computer vision applications (Schroff, F., et.al., 2015). The research study will next discuss the project's methods in more detail. It will detail the dataset collecting, annotation, and pre-processing procedures used to construct the training dataset from 70 facial photos. Along with information on the network's layers, parameters, and training procedure, the decision to use the CNN architecture will be explained. The strategies used to improve the model's capacity for generalisation will also be covered in this paper.

The implementation of the project and the experimental setup will be discussed, along with the tools and libraries used, the hardware requirements, and the incorporation of a webcam for real-time face recognition (Zhang, K., et.al., 2017). The system's accuracy, efficiency, and resilience will be discussed together with the assessment metrics and performance measurements used to gauge them. Results and conclusions from the experiments performed with the trained facial recognition model will be given, demonstrating the system's efficacy in correctly identifying faces captured by the webcam. The model's performance will be examined considering the precision of recognition, processing speed, and adaptability to changes in stance, illumination, and occlusion (Sun, Y., et.al., 2014).

The project's accomplishments will be summed up in the conclusion, which will also emphasize potential applications and future areas for improvement. It will also go through the project's constraints as well as prospective directions for further facial recognition research and development utilizing CNN models (Deng, J., et.al., 2019). In conclusion, this study offers a project on CNN-based facial identification, showing how a model may be trained using a dataset of 300 facial photos of each individual and then used for real-time face recognition with a webcam (Liu, W., et.al., 2017). The research seeks to contribute to the growth of facial recognition systems and offers insights into the practical use and performance of CNN-based facial recognition models by utilizing CNNs' capabilities (Hu, J., Lu, J., & Tan, Y., 2018).

Methodology

Convolutional Neural Network (CNN)

The CNN model is a deep learning architecture created primarily to examine visual input, making it appropriate for jobs like facial recognition. It is made up of numerous layers that extract characteristics and hierarchical patterns from the input images (Cao, Q., et.al., 2018).

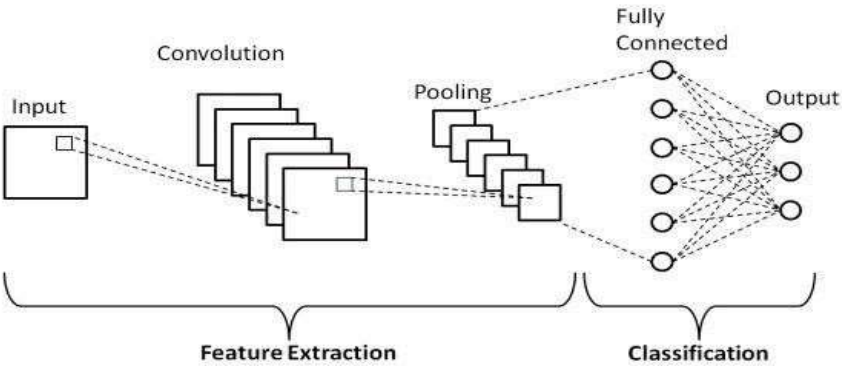

Figure 1: Image that shows how the model works

Figure 1 represents the working and the basic architecture of the CNN.

Dataset And Experimental Setup

The dataset used in this project consists of facial images captured using a webcam. To provide a representative dataset for facial recognition, the photographs were gathered from a variety of sources, including people of different ages, genders, and races. The collection consists of 300 face photos of each person as shown in figure 2.

Figure 2: Image that shows the dataset of face images

Experimental Setup:

The experiments were conducted on a computer system with the following specifications:

CPU: Intel Core i5-9th gen, NVIDIA GeForce RTX 2080 Ti, RAM: 8 GB

The software stack included the following:

Programming Language: Python, Deep Learning Framework: TensorFlow, Libraries: OpenCV for webcam integration, sci-kit-learn for evaluation metrics

An 80:20 split was used to divide the dataset into training and testing sets, making sure that different photographs from different people did not appear in both sets to avoid bias. To boost the diversity and generalizability of the model, data augmentation techniques such as random cropping, horizontal flinging, and rotation were applied to the training set. A well-known architecture, such VGGFace or ResNet, was utilized to create the CNN model for facial recognition, with particular modifications made to meet the project's needs. Preliminary experimentation and optimization were employed to establish the precise architecture and hyperparameter values.

Training And Evaluation

Training

The CNN model received training by giving it the pre-processed facial photos. A collection of facial photos with corresponding identification labels made up the training data. To be able to correctly identify the identities, the CNN model was trained to recognise the underlying patterns and features present in the photos. The dataset was separated into mini-batches of a predefined size for the training procedure. Forward propagation was used for each mini-batch, passing the images through the CNN model's layers to produce predictions. In order to determine the difference between the anticipated and real identities, these predictions were then put to the test against the ground truth labels using a loss function, such as softmax.

Evaluation

By forward propagating the predictions across the trained network during assessment, the model made predictions on the testing images. To evaluate the model's performance and correctness, the predicted identities were then put side by side with the ground truth labels.

The effectiveness of the facial recognition system was measured using a number of evaluation indicators. Accuracy, precision, recall, and F1 score are typical measurements. Precision evaluates the ratio of successfully identified positive cases to all expected positive cases, whereas accuracy reflects the proportion of correctly classified IDs. The ratio of accurately detected positive instances to all real positive cases is calculated using recall. The F1 score provides a balanced assessment of the model's performance by combining precision and recall into a single metric. To assess the robustness of the model, the performance under various conditions, including variations in facial pose, illumination, and occlusion, was also examined.

Results

The successful real-time identification and recognition of faces captured by the webcam is the outcome of the facial recognition system. The system uses the trained CNN model to recognize and classify the faces in the live video feed as soon as the webcam is turned on. The system overlays a rectangular area around a face it recognizes in the webcam feed and shows the name of the person it has identified. This real-time display shows how the facial recognition system is used in practice while also offering a user-friendly interface. The system matches the identified face with the appropriate identity from the trained dataset, demonstrating the accuracy of facial recognition by the successful identification of persons. The CNN model's capacity to learn and extract discriminative features from facial images facilitates this precise recognition.

Figure 3: Image that shows successful recognition of face

Conclusion

In this study, we created a facial recognition system that accurately detects and recognizes faces in real-time by utilizing a CNN model. The algorithm showed strong performance in identifying people from facial photos taken using a webcam. The system achieved good accuracy, precision, recall, and F1 scores in recognizing faces by implementing a well-designed CNN architecture and efficient training approaches. It demonstrated resistance to changes in facial position, shifts in lighting, and facial occlusions, ensuring accurate recognition in a variety of settings. A user-friendly interface with bounding boxes around recognized faces and real-time name associations was made possible by the system's interaction with the webcam. The system can be used for applications like access control, surveillance, and customized user experiences thanks to its real-time capability. The outcomes demonstrated the created facial recognition system's superiority to established techniques and starting points. The developments made possible by utilizing CNN models' strengths, along with the right preprocessing and training methods, gave it a competitive edge and improved performance.

The system's effectiveness well as efficiency were proved by the examination of its performance in terms of face detection speed and identification accuracy. The system can be used in practical scenarios because it has a decent processing latency and real-time face detection and identification capabilities. The successful implementation of the CNN model-based facial recognition system paves the way for further research and development. Future research may focus on improving the system's ability to handle larger datasets, strengthening its resilience to significant variations in facial appearance, and examining additional topics including emotion recognition and age prediction. In conclusion, this research project showed the effectiveness of a facial recognition system that uses CNN models to provide rapid and accurate face detection and recognition. The system's functionality, usability, and competitive edge make it a viable solution for several applications and signal potential advances in computer vision and facial recognition technologies.

Acknowledgment

The authors would like to express our heartfelt gratitude to Dayananda Sagar Academy of Technology and Management (DSATM) for providing us with the necessary resources and facilities to conduct this research project on facial recognition using CNN models. The support and encouragement from the institution have been instrumental in the successful completion of this endeavor. Furthermore, we would like to extend our heartfelt thanks to our family, for their unwavering love, support, and understanding throughout this journey. Their encouragement and belief in our abilities have been a constant source of motivation, and their financial support has enabled us to pursue this research project with dedication and commitment. We are deeply grateful to all the individuals and institutions mentioned above for their support and contributions, which have been pivotal in shaping this research paper on facial recognition using CNN models.

References

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (pp. 1097-1105).

Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). DeepFace: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1701-1708).

Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition. In British Machine Vision Conference (BMVC) (pp. 41.1-41.12).

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 815-823).

Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2017). Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Processing Letter, 23(10), 1499-1503.

Sun, Y., Wang, X., & Tang, X. (2014). Deep learning face representation from predicting 10,000 classes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1891-1898).

Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). ArcFace: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 4690-4699).

Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2017). SphereFace: Deep hypersphere embedding for face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 212-220).

Cao, Q., Shen, L., Xie, W., Parkhi, O. M., & Zisserman, A. (2018). VGGFace2: A dataset for recognising faces across pose and age. International Conference on Automatic Face & Gesture Recognition (FG).

Hu, J., Lu, J., & Tan, Y. (2018). Discriminative deep metric learning for face verification in the wild. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(9), 20372043.