IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Comparative study of deep learning models on face recognition

[1]Aanchal Singh ,[2] Vivek Kumar Sinha

[1]Student,[2]Assistant Professor

[1]Department of Computer Science and Engineering,

[1]Raipur Institute of technology, Raipur,India

Abstract: In the fields of image analysis and computer vision, face identification poses a difficult problem. This work, Face Recognition, analyses these problems and uses them to identify or verify a person using deep learning techniques. A face in an image is automatically found via facial recognition. 40 attributes and datasets with a maximum of 200,000 faces were used in this study on celebrity faces. This dataset features a variety of angles, backgrounds, and facial expressions. Using VGG16, AlexNet, and GoogleNet, the technique recorded test accuracy of 96.15 percent, 96.81 percent, and 97.73 percent.

***Index Terms* - Transfer Learning, Face Recognition, VGG16, Alexnet, GoogleNet, Classification Model, celebrity face dataset.**

I. Introduction

Face detection technique is utilised extensively in the field of vision because it has a distinctive stability and flexibility compared to the identification of fingerprints and other technologies [1]. Humans perform face recognition on a daily basis, but developing the same computer system and increasing precision is an ongoing effort. The study of face recognition dates back to the 1950s [2]. Visual pattern recognition struggles with face identification. When viewing visual information, people frequently notice visual patterns. The brain interprets this information as logical thinking. A computer is a pixel in a single matrix, whether it be for video or images [3].

In the fields of image analysis and computer vision, face identification poses a difficult problem. This work, Face Recognition, analyses these problems and uses them to to identify or verify a person using deep learning techniques. A face in an image is automatically found via facial recognition. 40 attributes and datasets with a maximum of 200,000 faces were used in this study on celebrity faces. This collection of facial expressions, perspectives and backgrounds.

The machine must identify what a particular bit of data in the data means. This is a challenging diagnostic issue for the identification of visual models. Knowing which information relates to which component and component with regard to facial recognition How can we confirm or establish the identification of the person in the photo given the provided face photo and celebrity photo site? This is a significant issue with facial recognition[3].

Complex face traits can be detected using deep learning-based methods [4]-[5]. The in-depth learning-based strategy achieves notable success in resolving issues that have for years hampered the best efforts of the intelligence community is a factor that pertains to many commercial and government research fields and has shown to have a complicated architecture in high-quality data. Convolutional Neural Network (CNN), Stacked Auto encoder [6], and Deep Belief Network (DBN) [7] are a few in-depth learning algorithms that can be utilised to tackle this problem. CNN is the most used picture and visual algorithm [8].

II. RELATED WORK

For facial recognition categorization, a few have utilized a variety of Machine Learning, DeepLearning, and Hybrid algorithms A couple of the papers are listed below :

Di Wang et al, have created a convolutional neural network-based algorithm for facial recognition. Every data set is gathered locally. We must personally review some of the inappropriate images when gathering data sets online, choose three Chinese characters as our research subjects, gather 200 images for each character, and then gather the recognition accuracy of our improved model, which ranges from 68.85 percent to 79.41 percent after sets checking[1].

Edy Winarno et al., has introduced an attendance system based on a face recognition system using the CNN PCA method and real-time camera. In order to detect and identify faces as a person's identity and to store them in a face database, the attendance system uses faces as objects. Up to 98 percent accuracy in facial recognition can be achieved using the suggested technique [3].

Yong Li et al, introduced a neural network testing-based facial recognition system. The suggested activity fireliably locates and measures faces in the image using the AdaBoost algorithm. Next, facial features are isolated and removed using a deep convolution neural network. CNN training database is tested using CMU PIE. The recognition results will then be displayed visually. Preliminary testing using current laboratory employees as test subjects revealed that the system properly detects facial recognition function and has improved visual accuracy[4].

Sudha Sharma et alintroduced the machine learning algorithm used in the face recognition system. attempts utilising Naive Bayes, a vector support machine, direct discrimination analysis, and multi-category perception. On the ORL website, tests are taken while exercising. A variety of facial images that were captured in the lab between April 1992 and April 1994 are included in the ORL Face Database. On the ORL website, there were 400 images drawn from 40 distinct lessons. The face data collection for 40 persons consists of 10 photos in the 92 x grey PGM format with a size of 112 pixels [5]. Through the application of PCA and linear discriminate analysis, the authors achieved 97 and 100% recognition accuracy[6].

Xiujie Qu • et al., the Fast Facial Recognition Program based on In-Depth Learning was unveiled. There are two sections of the road. First, network training is finished and network parameters are obtained using the PC terminal. The second is an FPGA-based face recognition system. Key component analysis (PCA), the LBP algorithm, and the in-depth learning method are the main topics of the study. The University of Carnegie Mellon's CMU-PIE face site was used for the testWebsite size for CMU-PIE is 32 x 32. Each individual contains 170 images. On the CMU-PIE facial site, 17 individuals were chosen, 120 training shots were randomly chosen from the individual face website, and an additional 50 images were utilised as test sets. After software simulation and board rating, the authors were able to determine the accuracy of the recognition. The face recognition system speed is 400FPS, the recognition rate is 99.25 percent, and it also has good endurance, allowing it to finish alert work in column lighting settings [7].

S. Sharma et al Tan-Triggs pre-screening for 96 x 96 and 64 x 64 pixel images was first introduced by S. Sharma et al. The (FRGC) database was used for the experiments. This analysis employed the Face recognition grand challenge (FRGC) database containing 11,2B4 2D previews for 286 distinct labels. This examines eight distinct strategies for modifying the orientation, size, and pre-processing of faces. According to the findings, facial alignment with CLM offers the best facial recognition precision, and the best facial verification technique uses the advance processing method. When employing distant photos for training, Dlib alignment also offers good accuracy. On matching pathways, it has generated results with accuracy ranging from 90 to 98.30 percent[8].

Wael AbdAImageeda et al.,Utilizing numerous pose-aware deep learning models, we have presented our approach and face recognition system. The IJB-A data set and the CASlA-WebFace [15] test set were used. The CASlA-WebFace [15] website is the largest known public face recognition website, and the QB-A website was instantly recognised by the scientific community. There are 494,414 photos in all across 10,575 studies in the CASIA WebFace data. In both verification and acquisition activities, Ournovel representation outperforms the CS2 implementation of IARPA and NIST'sUB-A[16].

III. METHODOLOGY

We've picked a deep learning architecture called Deep Convolution Neural Networks. The weights of the deep learning model are employed and the deep learning model is adjusted to construct a novel architecture for performing classification in face recognition on the Celebrity face dataset, which consists of millions of photos dispersed among thousands of classes:

A. Data Collection

For training, a celebrity face dataset with a maximum of 200,000 faces and 40 attributes was used [9]. The 40 features in this dataset include various facial expressions, viewpoints, and backgrounds. The sample properties included in this dataset are shown in Figure 1. Gender is one of the attributes that can be found in the databases [9].

Figure 1 :- Celebrity face classified by gender[9] .

B. Training datasets

In this system training and testing data have been separated from the dataset. In Proposed dataset, data have been split into 7:3 of training to testing data

C. Transfer Learning:

In the field of deep learning, there is a concept known as transfer learning. The extent of its operation is unknown to us, though. The proposed research looks at the transferability of neural networks. Efficiency is significantly impacted by transfer learning when high layer characteristics are transferred. It is still better than creating a network from scratch, though. When compared to using random weights, modest tweaking yields significantly better outcomes. This illustrates how transfer learning is far more effective than learning from scratch. [10][11].

D. Applied Model

VGG16: A deep convolutional neural network, which is frequently employed in the field of photo recognition, scored 92.7 percent top-5 test accuracy in the ImageNet dataset, which consists of 14 million photos belonging to 1000 classes. The network consists of 41 layers in total, 16 of which have learnable weights, including 13 convolutional layers and 3 fully connected layers. It can handle images with dimensions of (224, 224, 3). (224, 224, 3). The VGG-16 design has 138,357,544 parameters in total. [12].

Alexnet: In this In deep CNN architecture, AlexNet is highly regarded[11]. It has made ground-breaking advancements in the classification and image recognition sectors. The final two layers of AlexNet, the softmax and output layers, make complete its twenty-six layers[12]. Krizhevesky et al.[11], who created numerous parameter optimization techniques and expanded the CNN's depth, initially proposed the AlexNet. Figure15[13] depicts the AlexNet architecture.

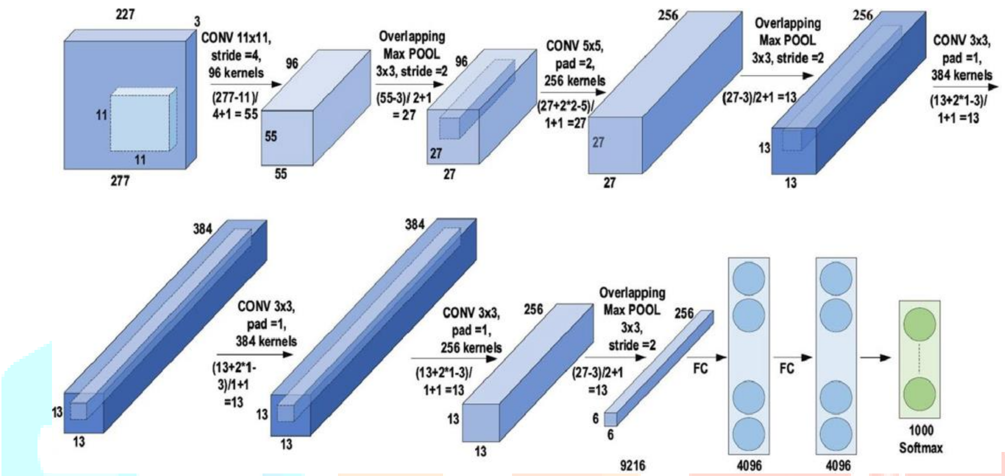

Figure 2 :- The architecture of AlexNet[13].

GoogleNet: GoogleNet (also known as Inception-V1) was declared the winner of the 2014-ILSVRC competition [14]. High-level accuracy at a lower computing cost is the primary goal of the GoogleNet architecture[13]. GoogleNet had a top-5 error rate of 6.67 percent, according to "Deep Sparse Rectifier Networks." The 14th International Conference on Statistics and Artificial Intelligence [15]. After only a few days of training, the human expert (Andrej Karpathy) was able to achieve a top-5 error rate of 5.1 percent for single models and 3.6 percent for ensembles. RMSprop, image distortions, and batch normalisation were all used.. This module uses numerous, extraordinarily tiny convolutions to significantly reduce the number of parameters. There are 22 levels in GoogleNet's design, yet there are only 4 million parameters as opposed to 60 million in AlexNet. Figure 4 depicts the GoogLeNet [15] architecture. In order to take clustering and network within network into consideration, GoogleNet has nine inception modules..During the inception modules, the module range is established and fully connected layers are eliminated. The number of factors used in the inception modules is reduced by pooling, meanwhile. To enhance the outcomes, a shadow network and an auxiliary classifier have also been used. GoogLeNet has additional layers thanks to the nine inception modules that repeat the convolutional, pooling, softmax, and concat procedures [15],[16].

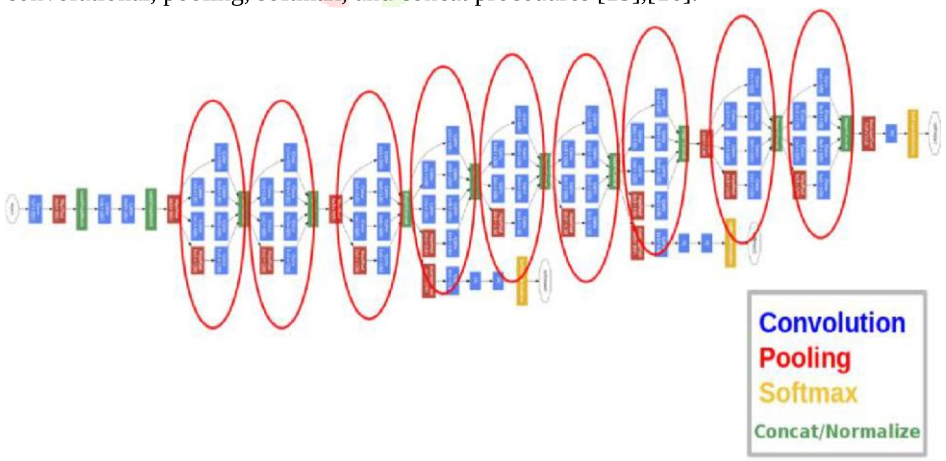

Figure : Image of GoogLeNet architecture [15].

In this section, the performance measures of the specified model will be discussed. The quality metric used for facial recognition is accuracy. In three models ko VGG-16, AlexNet, and GoogleNet have been applied to the celebrity face dataset. Then vgg16 achieved 96.15%, AlexNet achieved 96.81% and GoogleNet 97.73% achieved accuracy, which has been achieved on the Celebrity dataset.

Table 1.

Training Models			
	VGG-16	AlexNet	GoogleNet
Test Accuracy	96.15%	96.81%	97.73%

IV. CONCLUSION AND FUTURE SCOPE

Face detection is a method of identifying or confirming who you are using your face. Face recognition systems can identify people in photos, videos, or in real time. these technologies have made significant strides. The next generation of facial recognition technologies will be widely used in the identification of a person's face which includes focusing on specific unique features, such as jaw, cheeks, facial expressions, cheers, sadness, drowsiness, surprise or blinking and so on. In smart environments, where robots and computers will act as assistants, the next generation of face recognition technologies will be widely used. Due to the outbreak of COVID-19, face recognition systems are also being deployed. In this paper, I have used three deep learning models where the Resnet model gained the highest accuracy.

V. REFERENCES

1. Anjum D. Wang, H. Yu, D. Wang, and G. Li, "Face recognition system based on CNN," *Proc. - 2020 Int. Conf. Comput. Inf. Big Data Appl. CIBDA 2020*, pp. 470–473, 2020, doi: 10.1109/CIBDA50819.2020.00111..
2. M. Coskun, A. Ucar, O. Yildirim, and Y. Demir, Face recognition based on convolutional neural network," *Proc. Int. Conf. Mod. Electr. Energy Syst. MEES 2017*, vol. 2018-Janua, no. January 2018, pp. 376–379, 2017, doi: 10.1109/MEES.2017.8248937
3. E. Winarno, I. H. Al Amin, H. Februariyanti, P. W. Adi, W. Hadikurniawati, and M. T. Anwar, "Attendance System Based on Face Recognition System Using CNN-PCA Method and Real-Time Camera," *2019 2nd Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2019*, pp. 301–304, 2019, doi: 10.1109/ISRITI48646.2019.9034596.
4. Y. Li, Z. Wang, Y. Li, X. Zhao, and H. Huang, "Design of face recognition system based on CNN," *J. Phys. Conf. Ser.*, vol. 1601, no. 5, 2020, doi: 10.1088/1742-6596/1601/5/052011.
5. P. Dinkova, P. Georgieva, A. Manolova, and M. Milanova, "Face recognition based on subject dependent Hidden Markov Models," *2016 IEEE Int. Black Sea Conf. Commun. Networking, BlackSeaCom 2016*, 2017, doi: 10.1109/BlackSeaCom.2016.7901570.
6. S. Sharma, M. Bhatt, and P. Sharma, "Face recognition system using machine learning algorithm," *Proc. 5th Int. Conf. Commun. Electron. Syst. ICCES 2020*, no. Icces, pp. 1162–1168, 2020, 10.1109/ICCES48766.2020.09137850.X. Qu, T. Wei, C. Peng, and P. Du, "A Fast Face Recognition System Based on Deep Learning," *Proc. - 2018 11th Int. Symp. Comput. Intell. Des. Isc. 2018*, vol. 1, pp. 289–292, 2018, doi: 10.1109/ISCID.2018.00072.
7. S. Sharma, A. Kalyanam, and S. Shaik, *Technical Analysis of CNN-Based Face Recognition System—A Study*, vol. 711. Springer Singapore, 2019. doi: 10.1007/978-981-10-8055-5_49.
8. S. Yang, P. Luo, C. C. Loy, and X. Tang, "From facial parts responses to face detection: A deep learning approach," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2015 Inter, no. 3, pp. 3676–3684, 2015, doi: 10.1109/ICCV.2015.419.
9. J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" *Adv. Neural Inf. Process. Syst.*, vol. 4, no. January, pp. 3320–3328, 2014.
10. S. A. P. Raja Sekaran, C. Poo Lee, and K. M. Lim, "Facial Emotion Recognition Using Transfer Learning of AlexNet," *2021 9th Int. Conf. Inf. Commun. Technol. ICoICT 2021*, pp. 170–174, 2021, doi: 10.1109/ICoICT52021.2021.9527512.
11. V. Mnih and G. Hinton, "Learning to label aerial images from noisy data," *Proc. 29th Int. Conf. Mach. Learn. ICML 2012*, vol. 1, pp. 567–574, 2012.
12. L. Alzubaidi *et al.*, *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions*, vol. 8, no. 1. Springer International Publishing, 2021. doi: 10.1186/s40537-021-00444-8.
13. J. A. Gliner, G. A. Morgan, N. L. Leech, J. A. Gliner, and G. A. Morgan, "Measurement Reliability and Validity," *Res. Methods Appl. Settings*, pp. 319–338, 2021, doi: 10.4324/9781410605337-29.
14. H. Li, J. Wang, M. Tang, and X. Li, "Polarization-dependent effects of an Airy beam due to the spin-orbit coupling," *J. Opt. Soc. Am. A Opt. Image Sci. Vis.*, vol. 34, no. 7, pp. 1114–1118, 2017, doi: 10.1002/ecs2.1832.
15. S. Arora, A. Bhaskara, R. Ge, and T. Ma, "Provable bounds for learning some deep representations," *31st Int. Conf. Mach. Learn. ICML 2014*, vol. 1, pp. 883–891, 2014.