Deep Face Recognition Using Deep Learning for Fractional Visible Faces

Amruta Nagesh Chitari[1], Dr. Sharanabasava Inamadar[2], Dr. Pradip Salve[3]
1Research Scholar, Ajeenkya DY Patil University Pune, India
[2]Associate Professor, Ajeenkya DY Patil University Pune, India
[3]Assistant Professor, Ajeenkya DY Patil University Pune, India

ABSTRACT
In today's fully digitized society, there is a necessity for a primary, dependable, and sophisticated mechanism similar to computer-based facial recognition systems utilized in privacy management contexts. Numerous face recognition (FR) systems are dependable for "ideal frontal facial images/data." In some instances, a portion of the face is obscured, rendering the entire frontal visage potentially invisible. Consequently, the issue of face identification utilizing fractional facial data (occluded face data) is an emerging study domain. The implementation of deep learning solutions in this domain has been prevalent recently. However, numerous challenges persist, such as incomplete face data. So in this work, topic of face-recognition with fractional facial data is addressed. Experiments supporting the architecture of high-tech Convolutional Neural Networks (Hereinafter referred as CNN)within deep learning, together with the VGG-Face prototype/model, have been conducted. The VGG-Face prototype is a pre-trained CNNemployed for characteristicabstraction in mechanismknowledge. We performed research to assess the efficacy of the VGG prototype on fractional face images, along with other modifications including zooming and rotation of face images. Before using the image as input, face detection and alignment are executed using multi-task cascaded MTCNN.
We performed experiments on the publicly accessible Brazilian FEI dataset. Our findings indicate that the created model also identifies faces in RGB images that are partially visible in the input image.
Keywords: Face Recognition, Deep Learning, Convolution neural network, Cosine similarity.

1. Introduction

1.1 Basic concepts

Humans ownaintelligenceareaexplicitlydevoted to the dispensation and acknowledgement of faces. It is quite inevitable that humans possess a remarkable ability to recognize faces. The human brain recognizes the face by remembering its fundamental details, viz., shapes and vital features like the nose, forehead, cheeks, and eyes. Moreover, our brain can accommodate significant differences in lighting and facial expressions. Nonetheless, face recognition in computers is directly influenced by variations in facial appearance.

Despite this, the vast amount of available data enables a computer to adapt to and function in conjunction with the superior processing capabilities of machines. It is suggested that certain domains for assessment, such as a CNN, can attain real-time efficacy while performing facial recognition tasks.The paper seeks to investigate the compelling question of how a machine performs facial recognition using incomplete and partial facial information. This study seeks to investigate the impact of distant facial perception—specifically by means of enlarging out and rotating facial images—on the face identification process.

Fig. 1 Samples of fractional/covered face images

In Fig. 1, we present several unsatisfactory fractional face photos that were used in our methodology. Researchers [2-4] examined recent methodologies for face classification and identification. The prototype implies that the entire face is not attainable; instead, only particular parts such as the mouth, eyes, nose, or cheeks can be accessed as significant content.The judicious application of this work demonstrates the need to practically engage in programmed facial recognition tasks. Similar to other biometric verification systems, such as facial recognition, facial observation, and fingerprint analysis, these methods are commonly employed. Also, steadfast face processing [10, 11] and face recognition [12-15] technologies, which can exploit facial information, for example, photographs that originate from conventional CCTV cameras, are gaining in importance.This illustration demonstrates our face recognition framework, that utilizes partial faces for testing purposes.

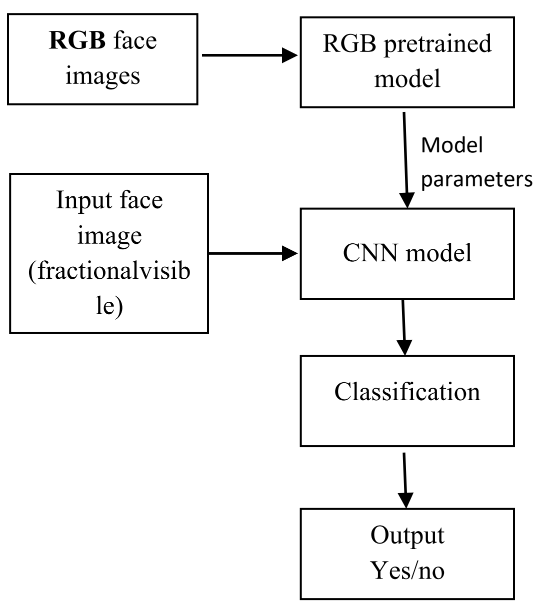


Fig. 2: Overview of face recognition with fractional visible face as input

2. Related Work

The literature regarding computer-aided facial recognition, particularly in relation to incomplete or obstructed facial images, appears to be relatively inadequate and inconsistent.

Savvvides et al. [16] proposed a technique utilizing kernel interaction filtration on monochrome pictures to diminish picture complexity alongside the extraction of features. Subsequently, support vector machines (SVM) were utilized to distinguish between different face features, focusing primarily on the eyes, nose, and mouth. The results of their experiments demonstrate a superior confirmation rate for the ocular region relative to other areas, including the mouth and nose. Simultaneously, L. He, H. Li, et al. [18] demonstrated rudimentary facial recognition algorithms referred to as Dynamic Feature Matching (DFM). Their suggested approach was based on a combination of complete CN and sparse representations. The VGG-Face model has been utilized for feature extraction. These qualities were employed in the FCN. This procedure produces enhanced precision in classification relative to other methods.

Proposition of an innovative model by H. Li and Y. Ching [21] for face recognition derives dynamical subdomains using images to capture distinctive traits of every human being. They mostly concentrated on frontal perspectives with differing illumination and occlusion.

In any event, when confronting significant obstacles, the efficacy of contemporary methods diminishes markedly. Many previous studies have shown that familiarity is crucial in facial recognition. The rate of commonality is evidently affected when the target facial image is occluded, partially obscured, or modified in relation to the individual's age [22, 23].

Then again, human face recognition is challenged by the computing capacity of machine learning algorithms to employ a vast quantity of input data as training data to produce outputs by using numerical analysis. Moreover, it can be articulated that machine learning algorithms have the potential to yield superior recognition rates in instances of partial representations or in the most adverse scenarios.

Despite ongoing research in machine-based facial identification, none of the studies have addressed how machine learning consistently utilizes fractional faces in this domain. Consequently, our objective in this analysis is to bridge that gap. We focus on how different facial features contribute to recognition. Our procedurepractices a CNN-based construction in concurrence with the pre-trained VGG-Face model to excerptcharacteristics. Subsequently, we employ a classifier, specifically cosine similarity (CS), to evaluate accuracy.

The subsequent paragraphs of the study provide a concise overview of the VGG-Face model, detailing its CNN architecture and cosine similarity classification, followed by experiments conducted on the FEI dataset using partial face inputs. In conclusion, we finalize the article.

3. Proposed method.

Currently, the most prominent instances of machine learning are derived from deep learning methodologies. One of the most used deep learning architectures is the Convolutional Neural Network (CNN). The field of Visual Computing is extensively investigated through the use of Convolutional Neural Networks (CNN). The non-linear, multilevel design of CNNs is adept at learning intricate features [24]. Convolutional Neural Networks (CNNs), a supervised machine learning methodology, derive profound insights from data through intensive example-based training. State-of-the-art deep learning models, derived from CNN architectures, are employed across practically all visual computing areas, including image processing, image recognition, image classification, and information retrieval.

A variety of pretrained models for facial perception tasks are accessible through CNN. These models can be used for feature extraction eg. VGGFace, VGG16, VGG19[29] — created by Oxford Visual Geometry Group [29]. We have utilized VGG-FACE model for feature extraction [30] mentioned below. This is followed by Cosine similarity, CS [30]. The VGG model was trained on 2.6 million facial photographs of over 2,600 individuals. VGG-Face employs a total of 38 layers, encompassing the input as well as the output surfaces. The input image is an RGB image with dimensions 224 x 224 x 3. A mean is calculated from the input image throughout the preparatory stage. VGG-Face consists of 13 convolutional layers, each featuring a collection of composite variables.

Fig. 3: VGG-Face model architecture

Every cluster of convolution layers has max pooling layers and also rectified linear units, or ReLUs. There are fifteen ReLUs. Subsequent to these layers are three completely connected layers: FC6, FC7, and FC8. The final layer is the softmax classifier, which categorizes the input face. Figure 3 depicts the framework of the VGGF model.

Feature extraction and classification: Convolutional Neural Networks (CNNs) utilize various layers to learn features, subsequently employing this knowledge to categorize images. Each convolutional layer is designed to learn distinct features. For instance, if an initial layer extracts features such as colors and edges of a picture, subsequent deeper layers acquire more complicated features. Convolutional layers produce channels and multiple 2D arrays. In VGG-FACE, thirteen levels are convolutional layers, while the remainder consists of a combination of pooling, ReLU, and fully connected layers and concludes with a softmax layer. After applying input to convolutional layers with filters, features are extracted for classification purposes. The primary characteristics are derived by analyzing the activation of the layer.

Classification occurs subsequent to feature extraction. Classification is a supervised machine learning method. The objective of a classification model is to assign class labels based on expected features. Our classification studies calculate the cosine similarity (CS) to determine the smallest distance using the following equation:

$CS_{min} = \min(dist(im\_test, im^n\_training))$

whereim_test is test image and $im^n$is the training image and n is total number of images in training set

4. Experiments and Implementation

Researchers performed experiments on facial recognition using diverse facial characteristics. For this job, researchers employed the renowned FEI facial dataset.[31] The FEI face dataset is a Brazilian resource that contains a compilation of facial pictures. Each one of the 200 individuals has 14 photographs, culminating in a total of 2800 images. All images are vibrant and captured over a uniform white background in an upright frontal orientation with a profile rotation of approximately 180 degrees. The scale may vary by around 10%, and the primary measurements for each image are 640 by 480 pixels. The faces predominantly consist of students and personnel at FEI, aged between 19 and 40, characterized by diverse appearances, hairstyles, and adornments.

The quantity of male and female subjects is identical, totaling 100 each. Figure 1 presents some instances of image changes from the FEI face database. [31]. We utilized the Python programming language together with essential development tools and various deep learning/machine learning packages, including NumPy, Keras, TensorFlow, OpenCV, and Matplotlib. Researcher deployedKeras deep learning outline to develop model. Our output layer differs from the ImageNet version. Preprocessing is conducted before utilizing the image as input. Face detection and alignment are accomplished using multi-task cascaded convolutional networks (MTCNN).[58]. In [32], the authors (2016) proposed that "MTCNN employs a cascaded architecture comprising three stages of meticulously designed deep convolutional networks that predict facial and landmark locations in a coarse-to-fine approach." The image is entered into the VGG model, followed by classification by cosine similarity. The model's output primarily fluctuates based on the threshold level of the distance matrices. This model uses a threshold value of 0.40.

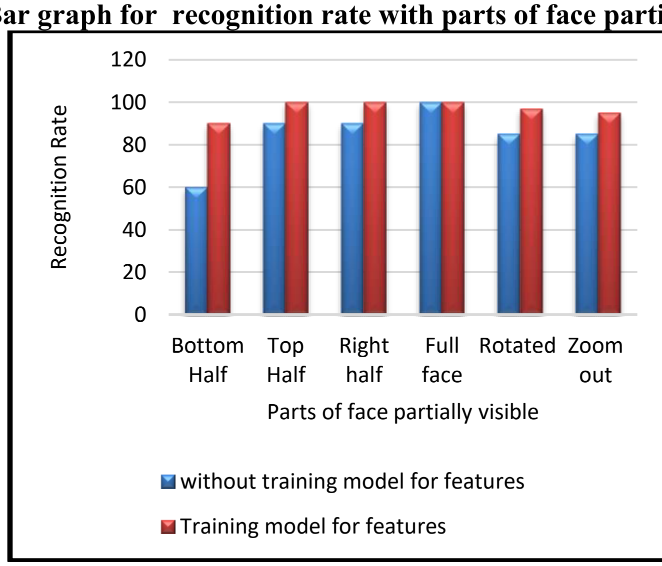
Result analysis

The test sets corresponded to different portions of the face. The model was trained on various test sets corresponding to different facial regions. The subsequent results are produced with and without training the VGG face model for various facial segments, including the bottom half, top half, right half, left half, entire face, zoomed-out face (up to 50%), and rotated face (up to 50%). The data in tables 1 and 2 indicate that the recognition rate improves when the model is trained on various facial regions.

Table 1: Face recognition rate for parts of face partially visible without training VGG model and with training model for parts of face

Parts of face partially visible	without training model for parts of face	Training model for parts of face
Bottom Half	60	90
Top Half	90	100
Right Half	90	100
Full Face	100	100
Rotated	85	97
Zoom out	85	95

Table 2 Bar graph for recognition rate with parts of face partially visible



6 CONCLUSION:

The efficacy of contemporary machine-based facial recognition methods in operating effectively using incomplete facial data, including obscured characteristics or flipped faces, presents a considerable difficulty in the field of artificial intelligence and visual computation. The present report outlines various methods for addressing these challenges. We have investigated the potential of deep learning for face identification utilizing partial or fractional faces. We outline the extraction of features utilizing the CNN VGG-Face model and the application of cosine similarity and classifiers in face recognition tasks. This study demonstrates the utilization of RGB characteristics extracted from the CNN VGG-Face model for the detection of obscured faces. This method can be further extended to near-infrared pictures for facial identification. Additionally, other prevalent deep learning methodologies, such as generative adversarial networks, can be included, necessitating minimal datasets for model training.

Acknowledgements

We would like to thanks to the Doctoral committeeandallthefacultymembersat Ajeenkya DY Patil University and also to Ajeenkya D.Y.Patil Schoolof Engineering, Lohegaon, Pune to support this work. Specifically, we would like to acknowledge the use of “FaceSurvdataset”andthesupportprovidedbytheIIITDelhi,Indiateamtoextractthedata

7 REFERENCES

[1] S.M. Andersen, C.A. Carlson, M.A. Carlson, S.D. Gronlund, Individual differences predict eyewitness identification performance, Pers. Individ. Differ. 60 (2014) 36–40, http://dx.doi.org/10.1016/j.paid.2013.12.011

[2] Z. Li, J. Liu, J. Tang, H. Lu, Robust structured subspace learning for data representation, IEEE Trans. Pattern Anal. Mach. Intell. 37 (10) (2015) 2085–2098, http://dx.doi.org/10.1109/TPAMI.2015.2400461.

[3] C. Ding, D. Tao, Pose-invariant face recognition with homography-based normalization, Pattern Recognit. 66 (2017)144–152, http://dx.doi.org/10.1016/j.patcog.2016.11.024.

[4] C. Ding, D. Tao, Robust face recognition via multimodal deep face representation, IEEE Trans. Multimed. 17 (11) (2015)
2049–2058, http://dx.doi.org/10.1109/TMM.2015.2477042.

[5] T. Zhang, J. Li, W. Jia, J. Sun, H. Yang, Fast and robust occluded face detection in ATM surveillance, Pattern Recognit. Lett. 107 (2018) 33–40, Video Surveillance-oriented Biometrics. http://dx.doi.org/10.1016/j.patrec.2017.09.011.

[6] D. Marˆcetić, S. Ribarić, Deformable part-based robust face detection under occlusion by using face decomposition into face components, in: 2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2016, pp. 1365–1370. http://dx.doi.org/10.1109/MIPRO.2016.7522352.

[7] I. Choi, D. Kim, Facial fraud discrimination using detection and classification, in: G. Bebis, R. Boyle, B. Parvin, D. Koracin, R. Chung, R. Hammound, M. Hussain, T. Kar-Han, R. Crawfis, D. Thalmann, D. Kao, L. Avila (Eds.), Advances in Visual Computing, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 199–208, http://dx.doi.org/10.1007/978-3-642-17277-9_21.

[8] M.A. Alsmirat, F. Al-Alem, M. Al-Ayyoub, Y. Jararweh, B. Gupta, Impact of digital fingerprint image quality on the fingerprint recognition accuracy, Multimedia Tools Appl. 78 (3) (2019) 3649–3688, http://dx.doi.org/10. 1007/s11042-017-5537-5.

[9] S. Atawneh, A. Almomani, H. Al Bazar, P. Sumari, B. Gupta, Secure and imperceptible digital image steganographic algorithm based on diamond encoding in DWT domain, Multimedia Tools Appl. 76 (18) (2017) 18451–18472, http://dx.doi.org/10.1007/s11042-016-3930-0.

[10] A.M. Bukar, H. Ugail, Automatic age estimation from facial profile view, IET Comput. Vis. 11 (8) (2017) 650–655, http://dx.doi.org/10.1049/ietcvi.2016.0486.

[11] H. Ugail, A. Al-Dahoud, Is gender encoded in the smile? A computational framework for the analysis of the smile driven dynamic face for gender recognition, Vis. Comput. 34 (9) (2018) 1243–1254, http://dx.doi.org/10.1007/s00371-018-1494-x.

[12] B. Olk, A.M. Garay-Vado, Attention to faces: Effects of face inversion, Vis.Res. 51 (14) (2011) 1659–1666, http://dx.doi.org/10.1016/j.visres.2011.05.007.

[13] J. Taubert, D. Apthorp, D. Aagten-Murphy, D. Alais, The role of holistic processing in face perception: Evidence from the face inversion effect, Vis.Res. 51 (11) (2011) 1273–1278, http://dx.doi.org/10.1016/j.visres.2011.04.002.

[14] J.J. Richler, M.L. Mack, T.J. Palmeri, I. Gauthier, Inverted faces are (eventually)processed holistically, Vis. Res. 51 (3) (2011) 333–342, http://dx.doi.org/10.1016/j.visres.2010.11.014.

[15] H. De, Adélaïde, R. Bruno, M. Daphne, Developmental changes in facerecognition during childhood: Evidence from upright and inverted faces, Vis. Res. 27 (1) (2011) 17–27, http://dx.doi.org/10.1016/j.cogdev.2011.07.001.

[16] M. Savvides, A. Ramzi, H. Jingu, S. Park, X. Chunyan, B.V.K. Vijayakumar, Partial and holistic face recognition on FRGC-II data using support vector machine kernel correlation feature analysis, in: Computer Vision and Pattern Recognition Workshop, New York, USA, 2006. http://dx.doi.org/10.1109/CVPRW.2006.153.

[17] A. Sleit, R. Abu-Hurra, W. AlMobaideen, Lower-quarter-based face verification using correlation filter, J. Imaging Sci. 59 (1) (2011) 41–48, http: //dx.doi.org/10.1179/136821910X12863757400286.

[18] L. He, H. Li, Q. Zhang, Z. Sun, Dynamic feature learning for partial face recognition, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 2018. http://dx.doi.org/10.1109/CVPR.2018.00737.

[19] L. He, H. Li, Q. Zhang, Z. Sun, Dynamic feature learning for partial face recognition, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 2018. http://dx.doi.org/10.1109/CVPR.2018.00737.

[20] P. Omkar, V. Andrea, Z. Andrew, Deep face recognition, in: X. Xianghua, J. Mark, T. Gary (Eds.), Proceedings of the British Machine Vision Conference (BMVC), BMVA Press, 2015, pp. 41.1–41.12, http://dx.doi.org/10.5244/C.29.41.

[21] H. Li, Y. Ching, Robust face recognition based on dynamic rank representation, Pattern Recognit. 60 (2016) 13–24, http://dx.doi.org/10.1016/j.patcog.2016.05.014

[22]J. Murphy, R. Cook, Revealing the mechanisms of human face perception using dynamic apertures, Cognition 169 (2017) 25–35, http://dx.doi.org/10.1016/j.cognition.2017.08.001

[23] M.G. Calvo, A. Fernández-Martín, L. Nummenmaa, Facial expression recognition in peripheral versus central vision: Role of the eyes and the mouth, Psychol. Res. 78 (2) (2014) 180–195, http://dx.doi.org/10.1007/s00426-013-0492-x..

[24] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444, http://dx.doi.org/10.1038/nature14539.

[25] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (3) (2015) 211–252, http://dx.doi.org/10.1007/s11263-015-0816-y.

[26] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, http://dx.doi.org/10.1109/CVPR.2016.90.

[27] G. Huang, Z. Liu, L. van der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017. http://dx.doi.org/10.1109/CVPR.2017.243.

[28] Y. Zhong, R. Arandjelović, A. Zisserman, Faces in places: Compound query retrieval, in: BMVC - 27th British Machine Vision Conference, York, United Kingdom, 2016. http://dx.doi.org/10.5244/C.30.56

[29] P. Omkar, V. Andrea, Z. Andrew, Deep face recognition, in: X. Xianghua, J. Mark, T. Gary (Eds.), Proceedings of the British Machine Vision Conference (BMVC), BMVA Press, 2015, pp. 41.1–41.12, http://dx.doi.org/10.5244/C.29.41.

[30] G. Sidorov, A. Gelbukh, H. Gmez-Adorno, D. Pinto, Soft similarity and soft cosine measure: Similarity of features in vector space model, Comput. Sist. 18 (2014) 491–504, http://dx.doi.org/10.13053/CyS-18-3-2043.

[31] C.E.Thomaz, G.A.Giraldi, FEI Face database, 2010,
https://fei.edu.br/~cet/facedatabase.html.

[32] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, Yu Qiao,**Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks,** IEEE Signal Processing Letters (Volume: 23 , Issue: 10 , Oct. 2016)https://arxiv.org/abs/1604.02878

[33] Pan, Qiaoying, Liaoying Zhao, Shuhan Chen, and Xiaorun Li. Fusion of low-quality visible and infrared images based onmulti-level latent low-rank representation joint with Retinexenhancement and multi-visual weight information." IEEEAccess 10 (2021): 2140-2153.

[34] Yang, Yiming, Weipeng Hu, and Haifeng Hu. "Neutral face learning and progressive fusion synthesis network for NIR-VIS face recognition." IEEE Transactions on Circuits and Systems for Video Technology (2023).

[35]Mei, Lin, and Cheolkon Jung. "Deep fusion of rgb and nir paired images using convolutional neural networks." In 2020 25th International Conference on Pattern Recognition (ICPR), pp. 6802-6803. IEEE, 2021.

[36] Yang, Yiming, Weipeng Hu, Haiqi Lin, and Haifeng Hu. "Robust cross-domain Pseudo-labeling and contrastive learning for unsupervised domain adaptation NIR-VIS face recognition." IEEE Transactions on Image Processing (2023).

[37]Ran He, Jie Cao, Lingxiao Song, Zhenan Sun, and TieniuTan. Adversarial cross-spectral face completion for nir-visface recognition. IEEE transactions on pattern analysis andmachine intelligence, 42(5):1025–1037, 2019.

[38]Chunlei Peng, Nannan Wang, Jie Li, and Xinbo Gao. Re-ranking high-dimensional deep local representation for nir-visface recognition. IEEE Transactions on Image Processing,28(9):4553–4565, 2019.

[39] Kim, Jeyeon, Moonsoo Ra, and Whoi-Yul Kim. "A dcnn-based fast nir face recognition system robust to reflected light from eyeglasses." IEEE Access 8 (2020): 80948-80963.).
[40] WeipengHu,WenjunYan,andHaifengHua.Dualfacealignmentlearning network for nir-vis face recognition. *IEEE TransactionsonCircuitsandSystemsforVideoTechnology*,2021.