2023 International Conference for Advancement in Technology (ICONAT)
Goa, India. Jan 24-26, 2023

User Recognition Via Facial Parameters With Occlusion Using CNN

1st Mahadeo D Narlawar
Datta Meghe College of Engineering Airoli Navi Mumbai
University of Mumbai
India, narlawaratul@gmail.com

2nd DR. D. J. Pete
Datta Meghe College of Engineering Airoli Navi Mumbai
University of Mumbai
India, dnyandeo.pete@dmce.ac.in

2023 International Conference for Advancement in Technology (ICONAT) | 978-1-6654-7517-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICONAT57137.2023.10080688

***Abstract*—The prime essence of occlusion based face detection has been derived across the world with a simple agenda to tackle the rise of frauds that security reinforcement or surveillance techniques provide concrete solution with bio metric database. Though the prime solution is by bio metric authentication via smart phones through facial recognition of users; still the misuse is a challenge as the user's face is often partially covered or occluded. Therefore, face detection or rather authentication even when occluded has become very important to prevent cyber frauds. Traditional approaches in terms of facial bio metric authentication typically have comprised of image processing or machine learning steps: feature extraction, pooling, segmentation, flattening and recognition. The researchers of this paper propose a facial occlusion detection network utilizing Convolutional Neural Networks (CNN). The result showed that accuracy in our developed model is in the range of 88 to 91 percent which in competence with the reported algorithms till occlusion of face in terms of percentage is restricted up to 50 percent of the focus area and 40 percent of the complete face.**

***Index Terms*—Convolutional Neural Network (CNN), Automated Teller Machines (ATMs), Machine Learning, Epochs (Iterations)**

I. Introduction

In a persons life a human can not only remember but identify various faces without difficulty in life; yet when we talk about modern machines made to do human job faces tremendous difficulties especially in the case of human face detection atomization. The limitation of any machine is evident instantaneously as the task done by a human takes hardly fraction of seconds whereas for a machine algorithm time constraints are enormous based on resource, area and cost limitations. The complexities are just not restricted to resources but on the matter even the surrounding environment conditions like that of light, or angle of vision for a human and in that context angle of capture of image for a machine show machines lacking behind human by a huge gap. Facial recognition has been under research since 1960's but the success of algorithms is tested recently with the advent of mask usage because of pandemic. In current research taking information from surveillance systems is not just restricted to gender identification or human traffic monitoring rather it is required especially for bio metric authentication [1-5]. In bio metric authentication systems facial recognition is an important category that extracts the pinpoints of a face under test by performing certain image processing mathematical operations required during training which have a unique id or authentication key associated with it. The algorithm for testing on the other hand is required to compare information extracted from a live feed with the ones of pre-trained network so as to bring about identification success. Smartphones have made testing of such computerized systems a tremendous easy but necessary option for any login requirements on platforms like that of Aadhar, interstate border crossing, commercial shopping, banking etc. One such application that was restricted in primary phases was that of app purchases made by apple users where the platform was restricted by only the quality index users that could afford apple smartphones. With the open source boom android users also could feel the ease of life when Google Cloud Vision API came as a savior for android users. This machine learning technology is slowly getting adapted by various domain like that of its usage in amazon go shopping marts to reduce the billing constraints or rather the sports industry where third umpire decisions are automated [6-8].

II. Facial Recognition

A biometric system meant for security uses the facial characteristic's of the person which comes in the range of the input image capturing system. Not only these captured images are utilised for login of various devices but they are also a tremendous help when we talk about system module meant for such applications like ATM, Traffic signals and many more. Facial recognition systems are broadly categorised based on the methods that are utilised for face processing namely feature analysis, neural network, eigen faces, and automatic face processing. One of the most cited and utilised algorithm is "Principal Component Analysis based on eigen faces. The flow of that algorithm can be visualised as in figure 1. The steps on which the algorithm evaluates are based on the Training Phase and the testing phase.

- Training
  1) Evaluate the eigen faces from input images per user.
  2) Feature extraction based on weights to form the weight matrix.
  3) Repeat step two for all the training images.
- Testing:
  1) Evaluate the eigen faces from test image under test.
  2) Calculate the average distance to compare it with weights in the weight matrix.

3) Compare the weight with the threshold value and classify empirically based on statistical analysis that is Principal component analysis.

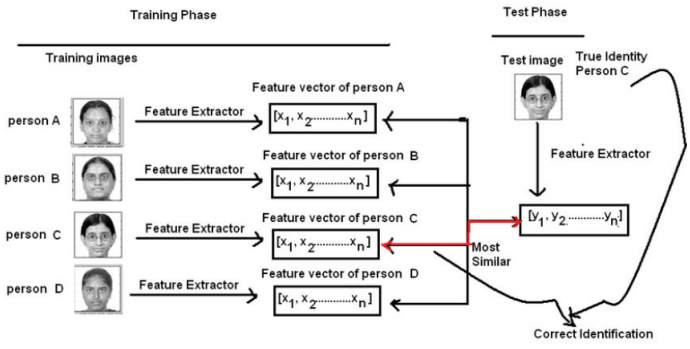

Fig. 1. Facial Recognition Based on PCA [9]

With the advent of Covid-19 this successful algorithm of facial recognition is facing setbacks because of facial occlusion which in this current scenario is not just restricted to environmental changes like that of light, or prop addition like that of turban, sunglasses and many more; rather the most primary obstruction or occlusion is because of facial masks that cover almost 25/40 percent of the face. This advent of occlusion has become a great challenge in regards to user authentication which is covered in the next section of facial occlusion.

III. FACIAL OCCLUSION

Any hindrance to existing solutions be it any domain provides challenges that no developer could have forecast during the development phases of research. One such domain of facial occlusion comes about with many such problems where wearing masks reduces the area of feature validation and user identity authentication. From the view point of practicality creating enormous updates to databases that could provide user authentication in facial recognition domain is a impossible task hence researches are trying out to find solutions that a machine can work with partial facial yet occluded features to provide sustainable results. Facial recognition not only considers wearing a specific size masks in fact there are various considerations that one can consider as occlusion. They primarily can be use of glasses, sunglasses, scarfs, masks, helmets and many more which could block the view of capturing a image in front of the camera. The accuracy of determination or identification varies tremendously which one can say is not just restricted to camera environment considerations but rather it could be even regards to geographic locations and human behaviors. Example of occlusion can be better understood by the figure 2 below illustrating various possibilities of occlusion. Just categorizing such occlusions is not self sufficient providing solutions to such challenges requires development of algorithms that provide a fast breakthrough [10-13].

So basic requirement of any machine/deep learning algorithm is computational requirements making sure GPU is optimised to provide faster solutions. Various solutions

Fig. 2. Occlusion of Face [12]

have been reported but each has one or the other limitation associated with it:

1) The major disadvantage of frontal face detection is the difficulty of incorporating a automatic feature detection in regards to decision making which is a key aspect in machine learning algorithms.
2) In face detection based on illumination approach the major setback is that of local contrast which irrespective of filter configurations like HAAR, Wavelet etc. still lack the maturity of enhancement of retain the quality of imaging for factorial feature extraction.
3) Further more when we talk about expression recognition in face recognition this technique completely fails when facial occlusion especially greater than that of 50 percent occurs.
4) Finally pose recognition is also a translated part of above technique but because of video processing multiple sampling of data under test have provided better opportunities of feature extraction.

Overcoming these enlisted limitation of facial recognition systems utilizing CNN machine learning approach tailored with the vision of human/user authentication in both occluded and non-occluded scenarios we have developed an algorithm as covered in next section.

IV. ALGORITHM

A typical CNN algorithm that utilises ALEXNET/RESNET is utilised for image prepossessing or feature extraction. The modification to the algorithm that makes the algorithm hybrid is based on ”auto focus” concept. In auto focus the detection is first made to check weather the required object is present in the field of test or not. If the object is in the field then the object is enhanced further known as sharp focus. Typically the area of the object is filtered out so as to achieve the best possible image enhancement. Hence in our algorithm this focusing is performed in way that the typical facial extracted features are further focused to extract only the features of the upper face mainly comprising of the eyebrows and the eyes. This tweaking is based on the concept of distance weights that are supplied for the support vector machine based interpolation. This modified or hybrid algorithm can be visualised from figure 3 and 4 respectively.

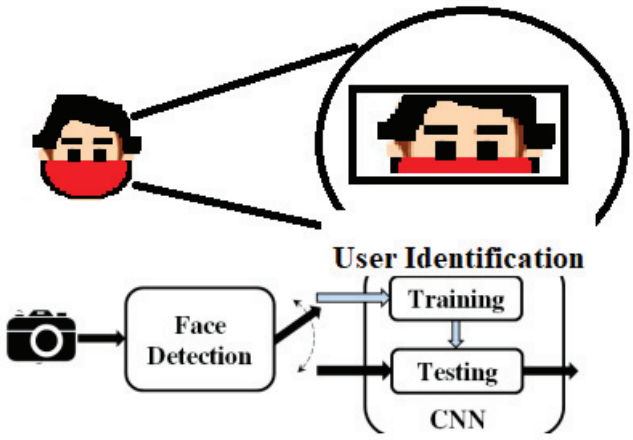

Fig. 3. Algorithm Flow

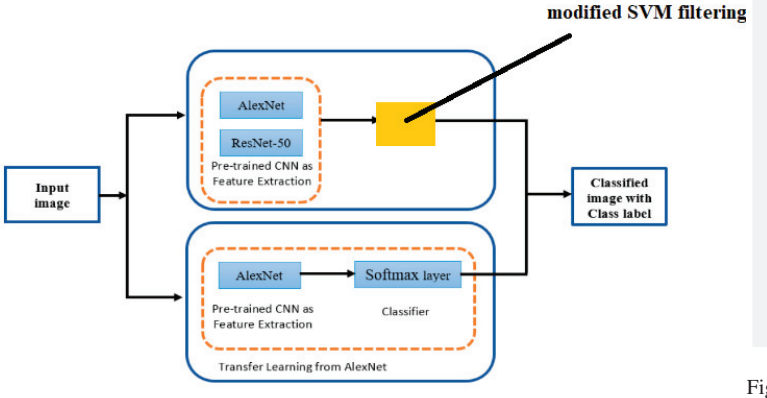

Fig. 4. Modified CNN Algorithm Block Diagram

V. IMPLEMENTATION

So steps that can strengthen the research in regards to the hybrid algorithm as mentioned above is development on MATLAB. The steps that are performed on MATLAB can be listed below:

- Database expansion than the reported in regards to image numbers in terms of face coverage (10 to 40 percent)
- Feature Extraction - CNN based algorithm to train the database so that it can be compared to the previous citing's or reported literature's.
- Classification – based on image input analysis to detect the person in regards to varying level of occlusion as per the database development.

A. Training

Example screenshots in terms of training are shown in figure 5 and 6 respectively. Accuracy in terms of training as in figure 5 is at very preliminary level which is tuned to 100 percent as shown in figure 6.

B. Testing

We tested the impact of accuracy in terms of occlusion as shown for user identification for different case scenarios as highlighted via figure 7, 8 and 9 for a single user for example purposes.

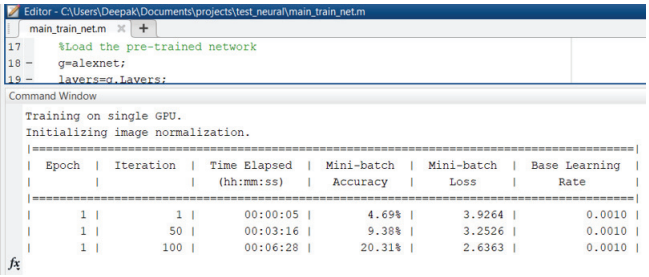

Fig. 5. Screenshot of Training at initial Level where accuracy building starts

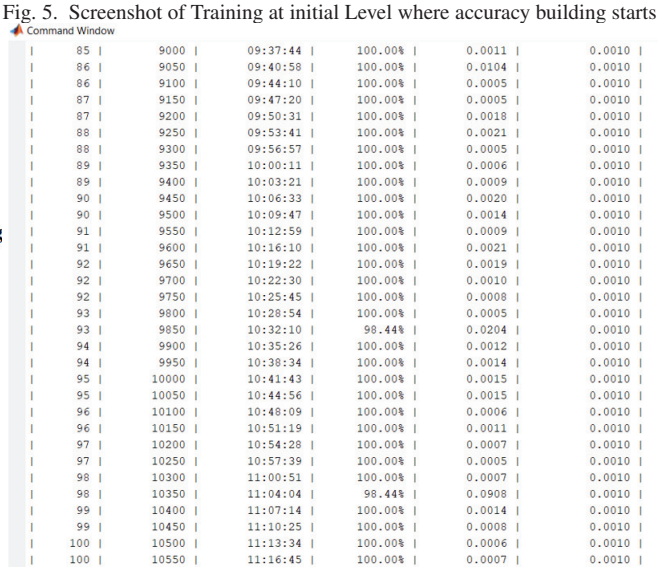

Fig. 6. Screenshot of Training at Final Level where accuracy achieved is 100 Percent

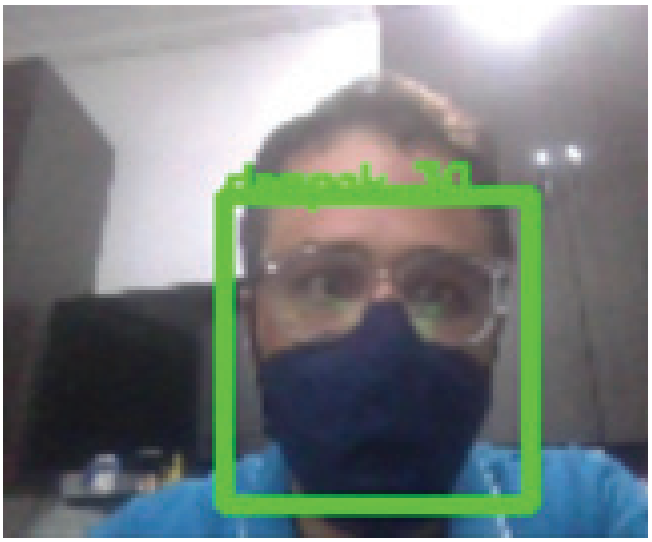

Fig. 7. Scenario 1: Occlusion with mask and Glasses

As shown in the example figure 7 above we conducted experiments to evaluate the performance of our CNN network performing the facial recognition as classification task with presence of glasses as well as face mask. Second scenario comprises of the same user but the position of the glasses obstructing in face has been taken into consideration. Figure 8 highlights that the user is identified easily as the amount of focus area is not altered in terms of coverage/occlusion. Third case scenario is where in the same user increases the

occlusion by covering the focus area with one hand as shown in figure 9. This figure shows that the user is still identified but the amount of accuracy takes a toll. As during the capture of frame if the position or angle of face is not in favour the user will not be identified which brings about limitation to user identification algorithm that we have devolved. This limitation is highlighted by graph as in figure 10.

As shown in the above example figures we conducted experiments to evaluate the performance of our CNN network performing the facial recognition as classification task. We developed a network using AlexNet where the training was done on images for 0 to 40 percent occlusion class problems. We tested the impact of accuracy in terms of occlusion as shown in the figure 10.

VI. CONCLUSION AND FUTURE WORK

In this paper we have tried to highlight the need of face recognition for occluded face coverage in this real world of covid where in the existing and the past work has been covered to provide a breakthrough for this research. Our developed algorithm provides 88 to 91 percent accuracy when we talk about facial occlusion till the focus occlusion level is restricted to 50 percent as beyond that occlusion accuracy percentage drops drastically. In this current situation our algorithm framework based on the use of CNN for face recognition as a milestone our research has to compete with higher accuracy rates that is over 94 percent but in presence of occlusion is still under development. Hence tremendous level of database testing and development is desired in near future to justify that our algorithm has achieved a breakthrough in facial recognition in occlusion presence.

REFERENCES

[1] Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A, "Face recognition: A literature survey", ACM Computing Surveys, 2003, Volume 35, page no: 399–458.

[2] Peter N. Belhumeur, Joao P. Hespanha, and David J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 7, JULY 1997.

[3] Yan, J., Zhang, X., Lei, Z. and Li, S. Z. Real-time high performance deformable model for face detection in the wild. IEEE International Conference on Biometrics (ICB), 2013, 1-6.

[4] Leung, T.K., Burl, M.C. and Perona, P. Finding faces in cluttered scenes using random labeled graph matching. Fifth International Conference on Computer Vision, 1995, 637-644.

[5] Takatsuka, H., Tanaka, M. and Okutomi, M. Distribution-based face detection using calibrated boosted cascade classifier. IEEE 14th International Conference on Image Analysis and Processing, 2007, 351-356.

[6] Huang, L.L., Shimizu, A., Hagihara, Y. and Kobatake, H. Face detection from cluttered images using a polynomial neural network. Neuro computing 51 (2003) 197-211.

[7] Romdhani, S., Torr, P., Schölkopf, B. and Blake, A. Efficient face detection by a cascaded support–vector machine expansion. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 460 (2051) (2004) 3283-3297.

[8] Yang, M.H., Kriegman, D.J. and Ahuja, N. Detecting faces in images: A survey. IEEE Transactions on pattern analysis and machine intelligence 24 (1) (2002) 34-58.

[9] M. Meenakshi, "Real-Time Facial Recognition System—Design, Implementation and Validation",Journal of Signal Processing Theory and Applications (2013) 1: 1-18, doi:10.7726/jspta.2013.1001.

[10] Jiang, N., Yu, W., Tang, S. and Goto, S. A cascade detector for rapid face detection. IEEE 7th International Colloquium on Signal Processing and its Applications (CSPA), 2011, 155-158.

[11] F. Zohra, Md Wasiur Rahman, M. Gavrilova, "Occlusion Detection and Localization from Kinect Depth Images", International Conference on Cyberworlds (CW), Published 1 September 2016.

[12] F. S. Samaria and A. C. Harter, "Parameterization of a stochastic model for human face identification," Proceedings of 1994 IEEE Workshop on Applications of Computer Vision, 1994, pp. 138-142, doi: 10.1109/ACV.1994.341300.

[13] B. Huang et al., "When Face Recognition Meets Occlusion: A New Benchmark," ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021, pp. 4240-4244, doi: 10.1109/ICASSP3972

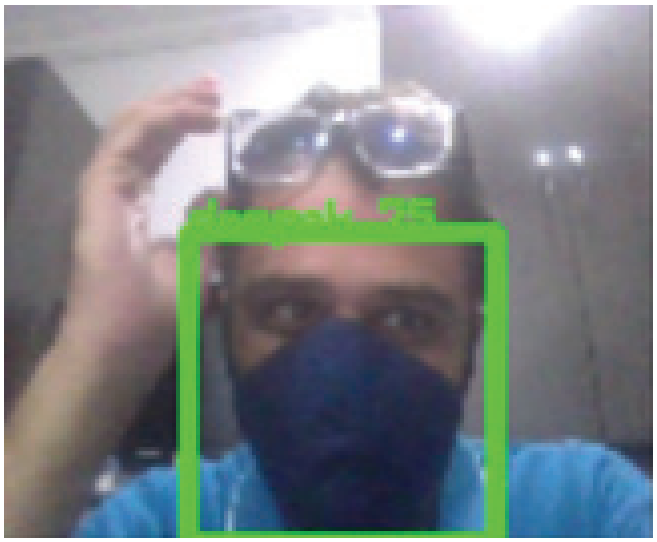

Fig. 8. Scenario 2: Glasses not in focus area

Fig. 9. Scenario 3: 50 percent Focus area occlusion

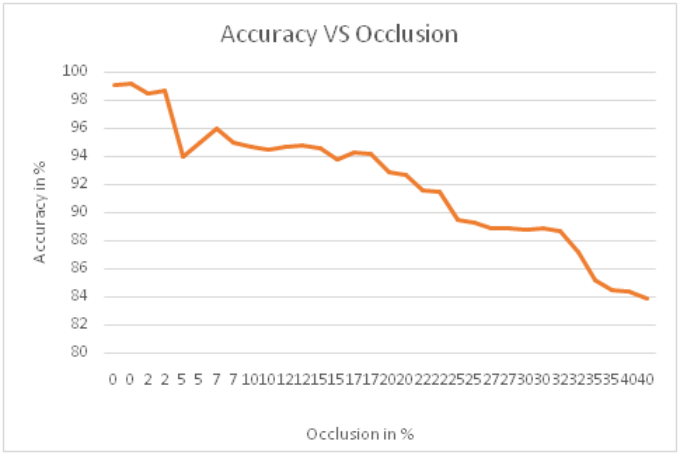

Fig. 10. Accuracy in terms of Increase in Occlusion percentage of face.