JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# CLASSIFICATION OF OCCLUDED FACE IMAGES THROUGH DEEP LEARNING TECHNIQUES

**Deepak Jangid**
Lecturer (Computer Engineering)
Govt. Polytechnic College, Jodhpur

***Abstract:***

Face occlusion generally describes circumstances in which a person's face is partially veiled or obscured, frequently by another bodily part or item. Face occlusion, in the context of computer vision and augmented reality, is a major difficulty that entails precisely identifying and handling scenarios in which objects in a scene conceal parts of a person's face. In this paper, two well-known convolutional neural network architectures MobileNet and InceptionNet, are used to effectively classify face occlusions, even when such faces are partially covered by different kinds of occlusions.

***Keywords:*** Raspberry Pi, Face occlusion, Deep learning, Transfer learning

## I. INTRODUCTION

Face occlusion is a critical obstacle in computer vision and augmented reality systems because it necessitates the identification and handling of cases in which entities within the frame obstruct particular parts of the face. For instance, an occlusion situation is a scenario when the face's wrist partly covers the face or the subject wears glasses. Deep learning technologies such as MobileNet and InceptionNet can be used to face images classification in obscured images. The mentioned two examples are CNN configurations that are employed for accurate image classification. InceptionNet is legitimately acclaimed for image recognition capabilities, whereas MobileNet is more specialized for mobile and embedded devices.

## III. LITERATURE SURVEY

The single shot multibox detector deep learning technique is used in this work [1] to accurately classify and locate face occlusions. A self-constructed dataset with seven different types of common facial occlusion contributed to the mean average precision reaching 95.46%.
The authors of research [2] provide a dependable and effective method for precisely identifying facial occlusion utilizing convolutional neural networks and multi-task learning. Accurately predicting the coverage of many facial regions, such as the nose, mouth, and both eyes, is possible using the multitask CNN.
In order to tackle the issue of face de-occlusion, this paper [3] divides the problem into three steps: face parsing, occlusion detection, and face reconstruction. The last stage reconstructs the face using the information from the previous two stages. As a result, the model performs well on actual occlusion data, which was not achievable with previous methods.

The paper [4] introduces a sophisticated convolutional neural network that uses facial features for face detection. Convolutional neural networks trained to discern features from uncropped images of faces can be used to learn the characteristics without the need for explicit supervision of each individual component.
In this study [5], a convolutional neural network with an attention mechanism is presented. Its purpose is to identify occlusion regions and to recognize and focus on the most informative un-occluded portions of the face. Real and artificial occlusions are used to evaluate the ACNNs under consideration.

## III. MOBILE NET AND INCEPTION NET

MobileNet[6] is developed for mobile and embedded devices. It is a lightweight convolutional neural network architecture. Depthwise separable convolutions are used to keep performance high while lowering computational complexity. Because MobileNet is parameter-efficient, it can be used on platforms with limited resources.
InceptionNet[7] is known for its inception modules, also referred to as GoogLeNet, is an architecture of deep convolutional neural networks. These modules efficiently collect characteristics at multiple scales by using parallel convolutional layers with varied filter thicknesses

## IV. DATASET COLLECTION & PREPROCESSING

To develop the face occlusion classification datasets, a wide range of images that describe numerous real-life conditions must be assembled. To develop a model, we gathered a collection of images of faces that had been partially occluded. Various sorts of occlusions were displayed in all images, including those caused by individuals ' hands, masks goggles, and glasses. All of the images are labeled with the type of occlusion. To develop training and testing model, these images were gathered from various sources. The image dataset shown in Table 1 contain a total of 7514 images depicting partially occluded faces.

Table 1: Dataset for face occlusion [8]

| Occlusion Type | Number of Images |
|---|---|
| Glasses | 2011 |
| Mask | 1412 |
| Sunglasses | 2185 |
| Hand | 1906 |

The dataset contains four types of occlusion: mask, glasses, hand, and sunglasses.

## V. EXPERIMENTAL SETUP AND RESULTS

The models for Inception net and MobileNetV2 were trained using the collected dataset. The classification of facial occlusion is done methodically. After the collection and preparation of images from many sources, a deep learning model was developed to categorize facial occlusion. The proposed models were developed and trained using a high-performance computing system equipped with an HP Z6 workstation's Quadro P5000 Graphics Processing Unit, which has 2560 NVIDIA CUDA cores, 16 gigabytes of GDDR5X GPU memory, and 8.9 Teraflops of floating processing capability.

The accuracy curve using MobileNet is shown in figure 1 and the precision and recall for MobileNet is shown in Table 2.

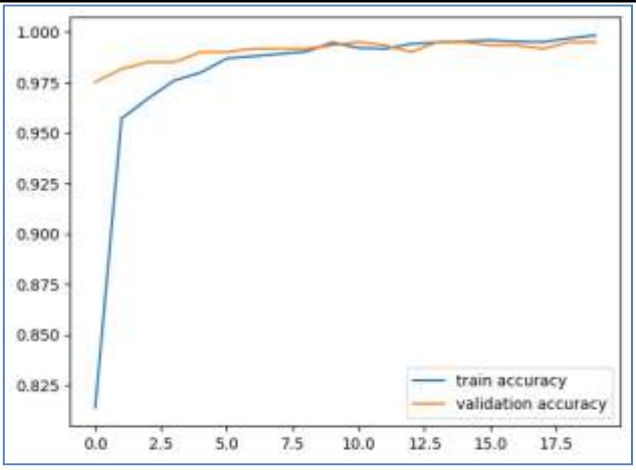


Figure 1: Accuracy using MobileNet

Classifying occluded face images with the MobileNet model worked well. All classes are well predicted with a high precision. Class precisions are 97% glass, 99% hands, 99% mask, and 99% sunglass. The recall is also high for the model, varying between 97 and 99% over the occlusion classes. These numbers show the capability of the MobileNet model to predict face occlusion properly, even in difficult cases.

Table 2: Precision and Recall for MobileNet

| Class | Precision | Recall |
|---|---|---|
| Glasses | 0.97 | 0.97 |
| Hand | 0.99 | 0.99 |
| Sunglasses | 0.99 | 0.98 |
| Mask | 0.99 | 0.99 |

The accuracy curve using InceptionNet is shown in figure 2 and the precision and recall for InceptionNet is shown in Table 3.

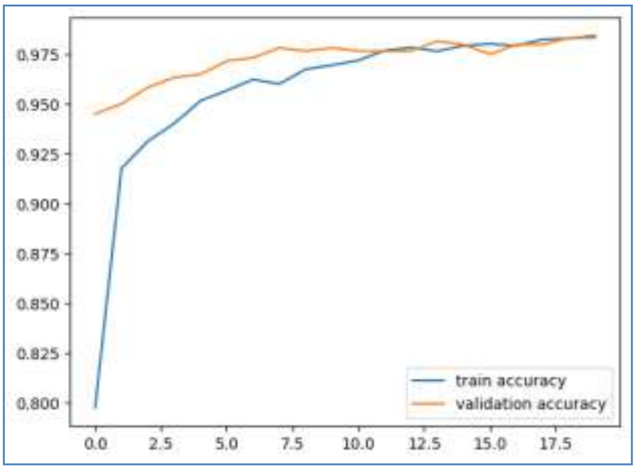


Figure 2: Accuracy using InceptionNet

The performances of InceptionNet model are good. It gives the high accuracy rates of 95% for hands, 96% for glasses, 97% for sunglasses, and 99% for masks - it is a perfect prediction. Moreover, the model gives the recall rates of 97% for glasses, 98% for hands, 96% for sunglasses, 97% for masks. It is illustrated in Table 3 below.

Table 3: Precision and Recall for InceptionNet

| **Class** | **Precision** | **Recall** |
|---|---|---|
| Glasses | 0.95 | 0.97 |
| Hand | 0.96 | 0.98 |
| Sunglasses | 0.97 | 0.96 |
| Mask | 0.99 | 0.97 |

These findings demonstrate InceptionNet's ability in classifying face occlusion in a variety of occlusion environments..

## V. CONCLUSION

In the present work, two popular convolutional network models, MobileNet and InceptionNet, are used to efficiently classify faces that may be only partially visible due to diverse types of occlusions. While the MobileNet model is good to accurately classify occluded facial images with several different types of occlusions, InceptionNet is able to consistently and accurately classify a face across a diversity of different occlusion types.

## REFERENCES


1. X. Ziwei et al., “Face occlusion detection based on SSD algorithm,” IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, vol. 2020, 2020, pp. 362-365, doi:10.1109/ICEIEC49280.2020.9152335
2. Yizhang Xia et al., “Face occlusion detection based on multi-task convolution neural network,” 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Zhangjiajie, 2015, 2015, pp. 375-379, doi:10.1109/FSKD.2015.7381971.
3. N. Zhang et al., “Face de-occlusion with Deep Cascade guidance learning” in *IEEE Trans. Multimedia*, vol. 25, pp. 3217-3229, 2023, doi:10.1109/TMM.2022.3157036.
4. S. Yang et al., “Faceness-net: Face detection through deep facial part responses” in *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 8, pp. 1845-1859, Aug. 1 2018, doi:10.1109/TPAMI.2017.2738644.
5. Y. Li et al., “Occlusion aware facial expression recognition using CNN with attention mechanism” in *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2439-2450, 2018, doi:10.1109/TIP.2018.2886767
6. M. Sandler, et al., “Mobilenetv2: Inverted residuals and linear bottlenecks” in Proc. IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, Jun. 18-23 2018, pp. 4510-4520, doi:10.1109/CVPR.2018.00474.
7. Pierre et al., “Szegedy, Christian & Liu, Wei & Jia, Yangqing & Sermanet,” Going Deeper with Convolutions. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
8. Datasets: Face-Obstructions, Glasses or No Glasses, MAFA_data, Face Mask Detection, Covid Mask, Sunglasses / No Sunglasses, Glasses and Coverings, With/Without Mask, Face Mask Detection, Masked Face Detection in the Wild Dataset.
   [Online]. Available at: https://www.kaggle.com/