Chapter 6
Transfer Learning in 2.5D Face Image for Occlusion Presence and Gender Classification

Sahil Sharma
Thapar Institute of Engineering and Technology, India

Vijay Kumar
Thapar Institute of Engineering and Technology, India

ABSTRACT

Face depth image has been used for occlusion presence and gender prediction by transfer learning. This chapter discusses about the overfitting problem and how augmentation helps overcoming it. Preprocessing of the dataset includes converting a 3D object image into depth image for further processing. Five state-of-the-art 2D deep learning models (e.g., AlexNet, VGG16, DenseNet121, ResNet18, and SqueezeNet) have been discussed along with their architecture. The effect of increasing the number of epochs on the top-1 error rate has been presented in the experimental section. The result section consists of error rates in comparison of with and without augmentation on the datasets.

INTRODUCTION

Face recognition is the oldest problem which has been worked in the field of computer vision. Many researchers have contributed to build robust face recognition systems since more than a three decade now. Transfer learning has come handy when comparing the efficiency of the face recognition systems. Using Labelled Faces in Wild face benchmark dataset multiple deep learning-based models have given human level accuracy. Some of the methods are DeepFace (Taigman et al., 2014), DeepID3 (Sun et al., 2015), FaceNet (Schroff et al., 2015), Baidu (Liu et al., 2015), VGGface (Parkhi et al., 2015), light-CNN (Wu et al., 2018), SphereFace (Liu et al., 2017), Cosface (Wang et al., 2018), Arcface (Deng et al., 2018) and Ring loss (Zheng et al. 2018). All the mentioned face recognition methods are done on two dimensional faces. No work has been done with deep learning for depth images of faces.

DOI: 10.4018/978-1-5225-7862-8.ch006

There are various ways to implement face recognition using facial modalities, 2D face images (Liu et al., 2017), 2.5D depth face images (Lee et al., 2016), and 3D face images (Liu et al., 2018). The capturing of depth images is the key for the next level problems in the field of computer vision. Plethora of work has been done in the field of 2D face recognition (Xu et al., 2017; Ranjan et al., 2017; Grm et al., 2017; Hu et al., 2015). There are various challenges that need to be worked for building a robust face recognition system which can deal with real-world problems of occlusion, low-lightening, makeup, augmented reality and low-resolution image etc. To handle these kinds of challenges, there is need for a robust face detection algorithm that has to be implemented in spatial region. Considering the challenges in depth face images, this becomes more tough. Considering the most famous viola-jones face detection algorithm (Viola et al., 2003) to be state-of-the-art for two dimensional algorithms, face detection can be extended for depth images to handle real world scenarios. The lack of depth image face datasets makes the recognition problem challenging. Augmentation can be used to counter the challenge of overfitting because of less number of images. Augmentation is the concept of improving the number of images by variations of the same image (Perez et al., 2017).

The motivation of the current work is the enormous amount of work being done in the field of deep learning for two dimensional images, but no work has been done on depth face images. Pose and expression normalization has been done on depth images by Feng, & Zhao, 2018 for the purpose of face recognition but not for occlusion presence and gender recognition. Occlusion and gender prediction of depth facial images has not been implemented using transfer learning. In real time, quadcopters having depth sensors coupled with these models can be used to track head count of male and female for attendance purposes or in public gatherings. Counting the number of people in crowds by correctly segmenting different people can be done if presence of occlusion can be identified. Detection of faces in depth images is a challenging task.

In current work, the problem of occlusion classification as well as gender recognition is discussed in detail. The face image can have different types of occlusion viz. glasses, hand, paper, cap, cigarette, bottle, and cropping etc. Deep learning is used for classifying the face image into occluded image or the non-occluded image. The gender binary classification into male and female is presented on the two datasets. The transfer learning is implemented for the task of binary classification in both type of problems.

The following sections are as follows: Section 2 discusses the background of the work. Section 3 presents the types of transfer learning as well as architecture of five deep neural networks. Section 4 presents the experimentation and results section. The last Section 5 presents the conclusion.

BACKGROUND

This section discusses the basics of two-dimensional deep learning.

Occlusion and Gender Prediction From Face Depth Image

The phases of face recognition with depth image for the classification of presence of occlusion and gender can be seen in Figure 1.

It can be seen in Figure 1 that the initial dataset images are converted from the 3D face object images to pre-processed depth images of size 224x224. The size is set to 224x224 because the input layer of the deep learning models is set to default 224x224 in AlexNet [20] and other models viz. VGG16 (Simonyan

Related Content