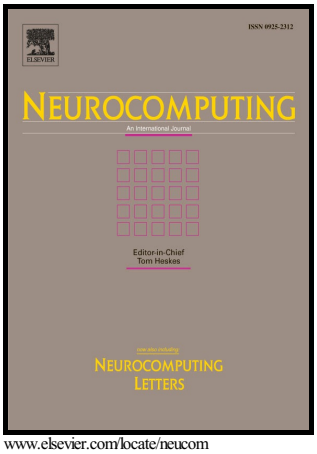

www.elsevier.com/locate/neucom

Data Augmentation for Unbalanced Face Recognition Training Sets

Biao Leng, Kai Yu, Yu Liu, Jingyan QIN

PII: S0925-2312(16)31488-6
DOI: http://dx.doi.org/10.1016/j.neucom.2016.12.013
Reference: NEUCOM17830

To appear in: *Neurocomputing*

Received date: 24 December 2014
Revised date: 29 October 2015
Accepted date: 6 December 2016

Cite this article as: Biao Leng, Kai Yu, Yu Liu and Jingyan QIN, Data Augmentation for Unbalanced Face Recognition Training Sets, *Neurocomputing* http://dx.doi.org/10.1016/j.neucom.2016.12.013

Data Augmentation for Unbalanced Face Recognition Training Sets

Biao Leng[a], Kai Yu[a], Yu Liu[a], Jingyan QIN[b,*]

[a]School of Computer Science , Beihang University, Beijing, 100191, P.R. China
[b]School of Computer and Communication Engineering, University of Science & Technology Beijing, Beijing, 100191, P.R. China

Abstract

Face recognition is gaining great significance. The last step of face recognition can be divided into face verification and face classification. Face verification algorithms have been well-developed in recent years, but face classification researches are still facing problems. For example, when training face classification models like Support Vector Machines and Softmax, a problem exists that some subjects have only one or a few samples, while others have abundant samples, and directly trained models might show unsatisfying performance on classifying these small-sample-size subjects, while applying single-sample algorithms might waste information. To address this problem, we propose in this work a novel method to generate reasonable virtual samples, so as to prevent imbalance classification results. Our method is based on Joint Bayesian Face Analysis, and we also develop an algorithm to boost the whole procedure. We conduct experiments based on high dimensional LBP features and features extracted by a shallow Convolutional Neural Network, and succeed to verify the effectiveness of this method, using image data from benchmark dataset LFW.

Keywords: Data Augmentation, Face Recognition, Convolutional Neural Network, Support Vector Machine

1. Introduction

Face recognition is an integrated task to locate the face, analyze it and distinguish to whom the face belongs to. The last sub-task can be categorized into two classes: verification and classification. Verification models take two images as input, and determine whether they belong to a person, using tools like Bayesian Face Recognition [1]. For practical use, a database of photos of candidates is stored, and the target face image if compared to every photo in that database. As shown in a benchmark face dataset LFW [2], verification algorithms have reached accuracy as high as human. However, the speed of these algorithms is not yet satisfying, and the time cost of the whole system grows linearly as the candidate amount set grows. It still needs great effort to build a face recognition system for practical use under this framework.

On the other hand, classification models take a single picture as input, then directly output the class label. Representatives are multi-class Support Vector Machine (SVM) and Softmax. Once trained, the candidate set cannot be changed, but as long as the feature extracted from the face (in the analyzing step) is discriminative enough, these models can deal with tasks of large candidate set size. However, to train classification models, we usually require that each target subject should have as many samples as possible. But in real world situations, this requisite is hard to satisfy. For example, in some surveillance scenarios, we have only accidently taken one picture of a subversives profile, and we need our system to automatically recognize him. Directly training SVMs or Softmax with this would cause the system to be only able to recognize him when he is showing the camera almost exactly the same pose and dressing as he was formerly shot by our camera.

To solve this problem, many algorithms focus on single sample training conditions [3, 4, 5]. They indeed achieve higher accuracy than traditional algorithms on this specified extreme condition. However, directly applying them on situations where subjects with only one sample and others with multiple samples both exist in the datasets (see LFW [2]), called the Unbalanced Training Dataset Condition (UTDC), might waste the rich information laying in the multi-sample classes. Also, under the UTDC, common classifiers such as neural networks and SVMs often perform unbalanced classification, that is, they more possibly classify test samples into classes with more training samples.

In this work, we present a novel method to deal with face recognition task under UTDC at a relatively high speed. Our algorithm augments data at feature level, and we will discuss about its superiority over other sample augmentation algorithms. We also design a shallow Convolutional Neural Network (CNN) to extract features, compared to a high dimensional Local Binary Patterns (LBP) feature extraction proposed in [6], then perform sample augmentation on feature level for classes with insufficient samples. We estimate the improvement brought by the augmentation to the training process of SVM on the LFW dataset.

This paper is organized as follows. In section 2, we give a brief review of some related works on face verification and face identification. In section 3, we introduce the detailed implementation of our algorithm. We show our experiments in section 4, and make a conclusion in section 5.

*Corresponding author
Email address: qinjingyanking@foxmail.com (Jingyan QIN)