Computers and Electrical Engineering 68 (2018) 484–489

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

# More discriminative convolutional neural network with inter-class constraint for classification☆

Jianchao Fei[a], Ting Rui[*,a,b], Xiaona Song[a], You Zhou[c], Sai Zhang[a]

[a] Army Engineering University of PLA, Nanjing, China
[b] State Key Lab. For Novel Software Technology, Nanjing University, Nanjing, China
[c] Jiangsu Institute of Commerce, Nanjing, China

ARTICLE INFO

*Keywords:*
Convolutional neural network
Classification
Softmax loss
Inter-class constraint

ABSTRACT

Recently, convolutional neural network (CNN) has achieved impressive results in object classification tasks. Through various modifications and careful design of the inner structure of CNN, its performance has already become human-competitive, according to recent reports. To some extent, the testing accuracy depends largely on the decision boundaries produced by classifiers, which classify different objects into specific feature spaces. Hence, the relationships among samples in the feature space are critical. However, the softmax loss function that is used in most CNN models, does not directly contain the relationship information. In this paper, we propose a novel loss function, named inter-class constraint loss function, that maximizes the distance between different classes. Together with softmax loss, we can obtain larger inter-class distances and smaller intra-class distances in CNN, thus significantly improving the accuracy in classification. We achieve substantial improvements for the SVHN, CIFAR-10 and CIFAR-100 datasets using our proposed loss function.

## 1. Introduction

Deep learning has made surprising improvements in various fields of computer vision. As one of the most widely used deep learning algorithms, convolutional neural network (CNN) has achieved great success on visual tasks due to its special design for images derived from human visual mechanisms [1,2], such as object detection [3,4], classification [5,6] and segmentation [7,8]. Inspired by the excellent recent performance of CNN, more efforts are being devoted to improving its inner structure to achieve higher accuracy and effectiveness. Some researchers focus on the hyper-parameter optimization, such as Xavier [9]. Other improvements focus on more specially on aspects of the training process, such as dropout [10]. Further, increasing the width and depth of network is also an important trend, such as [11].

In this paper, we propose a novel loss function for classification, named inter-class constraint loss function, that accounts for the relationships among different classes during optimization. The traditional softmax loss function is conceptually identical to a softmax layer followed by a multinomial logistic loss layer to offer a numerically stable gradient. However, as a loss function for classification, it does not pay attention to the relationship between different classes directly. Our main goal is to directly maximize the inter-class distance which is critical for classification performance. During each training epoch, we compute the difference square distances between each sample belonging to one class and all the other class centers, then maximize the sum of these difference square

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. Huimin Lu.

* Corresponding author at: Army Engineering University of PLA, Nanjing, China.
*E-mail addresses:* 349318312@qq.com (J. Fei), rtinguu@sohu.com (T. Rui), songxiaona1@163.com (X. Song), 354442511@qq.com (Y. Zhou), 466908114@qq.com (S. Zhang).

https://doi.org/10.1016/j.compeleceng.2018.05.002
Received 30 September 2017; Received in revised form 30 April 2018; Accepted 2 May 2018
0045-7906/ © 2018 Elsevier Ltd. All rights reserved.

distances. With the joint supervision of inter-class constraint loss and softmax loss, a smaller intra-inter class distance ratio can be obtained, thus significantly increasing the probability of correct classification.

The rest of this paper is organized as follows. Section 2 reviews the related work of loss function. Section 3 introduces our proposed methods. In Section 4, we conduct a series of experiments. A summary is shown in Section 5.

## 2. Related work

Loss function aims to estimate how much the prediction deviates from its corresponding true value. Parameters can be optimized by minimizing (or maximizing) the loss function. We usually define our task as a specific form of loss function. For classification, the main goal of the loss function is to produce best decision boundaries by which every sample can be classified correctly such as Support Vector Machine (SVM) [12], softmax loss function [13].

As we know, representing a categorical distribution is the main use of the softmax loss function [14]. However, we tend to realize other functions simultaneously in some cases. Combining the softmax (or logistic) loss function with other specific loss functions has been demonstrated to be an efficient way to supervise the training of a CNN and can achieve certain design goals. In [15], a temporal sparsity term is added to logistic loss function to ensure that each branching component focuses on a subset of fine categories. Shen et al. [16] propose a positive-sharing loss function that considers the loss caused by incorrect estimation between zero and nonzero labels. During our research, the center loss integrated with softmax loss, minimizing the intra-class variations, is explored in concurrent work by Yandong Wen [17]. In that study, center loss applies a constraint on intra-class distance, whereas our proposed inter-class constraint loss function pays attention to both the intra-class distance and the inter-class distance. As demonstrated by our experiments, inter-class constraint can obtain lower an intra-inter class distance ratio and achieve higher accuracy for the SVHN, CIFAR-10 and CIFAR-100 datasets.

## 3. CNN with inter-class constraint

### 3.1. Inter-class constraint

The inter-class constraint loss function that we propose can be viewed as an indicator that shows the square distance between a sample of one class and other class centers. Therefore, we define inter-class constraint using Eq. (1).

$$J_{\text{inter}} = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{m}\|x_i^j - M_k\|_2^2,\ i \neq k,$$

$$M_k = \frac{\sum_p^q x_k^p}{q}. \tag{1}$$

where $n$ is training batch size, $m$ is the number of classes, $q$ is the number of samples in class $k$, $x_i^j$ is the $i$th deep feature (with the same definition of softmax loss used in Eq. (3)) belonging to sample $j$ and $M_k$ is the center of class $k$.

When training models, we concatenate inter-class constraint loss and softmax loss instead of using softmax loss alone. Therefore, we get the whole loss function as follows:

$$J_{loss} = J_{soft\ \max} + \alpha\frac{1}{J_{\text{inter}}}, \tag{2}$$

in which

$$J_{soft\ \max} = -\sum_{i=1}^{m}\sum_{j=1}^{n}\log(P_i^j),$$

$$P_i^j = \frac{\exp(x_i^j)}{\sum_l^m \exp(x_l^j)}. \tag{3}$$

where $J_{softmax}$ is the softmax loss function and $\partial$ is a scalar to balance softmax loss and inter-class constraint loss. In order to minimum $J_{loss}$, we use the reciprocal value of $J_{inter}$. We notice that Linear Discriminant Analysis (LDA) can also obtain larger intra-inter class distance in classification. LDA not only maximizes the inter-class distance but also minimizes the intra-class distance. We test the LDA model in our experiments (model C).

### 3.2. Discussion

To present the discussion more clearly, a simplified CNN model is designed for the experiments of this section on SVHN, similar to Wen et al. [17]. Fig. 1 shows the overall architecture.

The softmax loss function can provide an efficient and stable gradient with excellent performance for multi-class classification. Fig. 2(a) shows the distributions of 2-D features obtained via softmax loss function supervision. Softmax loss function enforces a constraint on inter-class and intra-class distance. However, two problems still exist:

Download English Version:

https://daneshyari.com/en/article/6883415

Download Persian Version:

https://daneshyari.com/article/6883415

Daneshyari.com