Proceedings of the OAGM Workshop 2018 DOI: 10.3217/978-3-85125-603-1-01

Image Retrieval with BIER: Boosting Independent Embeddings Robustly

Michael Opitz, Georg Waltner, Horst Possegger and Horst Bischof

***Abstract*— Deep metric learning methods embed an image into a high dimensional feature space in which similar images are close to each other and dissimilar images are far apart from each other. However, state-of-the-art deep metric learning approaches typically yield highly correlated embeddings. To address this issue, we propose a method called Boosting Independent Embeddings Robustly (BIER) which divides the last embedding layer of a metric CNN into several smaller embeddings. We train these embeddings with online gradient boosting to make the learners more diverse from each other. During training, each learner receives a reweighted training sample from the previous learner. Additionally, we use an auxiliary loss function to increase the diversity between learners. In our experiments we show that BIER significantly reduces correlation in the embedding layer and consequently improves accuracy. We evaluate BIER on several image retrieval datasets and show that it significantly outperforms the state-of-the-art.**

I. INTRODUCTION

Deep Convolutional Neural Network (CNN) based metric learning approaches learn a distance function between images. This function maps semantically similar images close to each other and dissimilar images far apart from each other.

State-of-the-art approaches in metric learning typically saturate or decline due to over-fitting, especially when large embeddings are used [4]. To address this issue, we proposed a learning approach, called Boosting Independent Embeddings Robustly (BIER) [5], [6], which leverages large embedding sizes more effectively. Rather than using a single large embedding, BIER divides the last embedding layer of a CNN into multiple non-overlapping groups (see Fig. 1). Each group is a separate metric learning network on top of a shared feature extractor. To make learners diverse from each other we train our learners with online gradient boosting [6], and use auxiliary loss functions between pairs of learners [5].

We demonstrate the effectiveness of our metric on several image retrieval datasets [4], [7] and show that we can significantly outperform state-of-the-art approaches.

II. BIER

To train our network, we adapt an online gradient boosting algorithm [1]. During forward propagation we sample an mini-batch and compute the loss function for the first learner. The learner then reweights the training set according to the negative gradient of the loss function for the successive learner. After the last learner computes the loss, the gradients are backpropagated to the hidden layers of the CNN, as illustrated in Fig. 1.

*This work was supported by the Austrian Research Promotion Agency (FFG) Project MANGO (836488) and DARKNET (85891).

Graz, University of Technology, michael.opitz@icg.tugraz.at

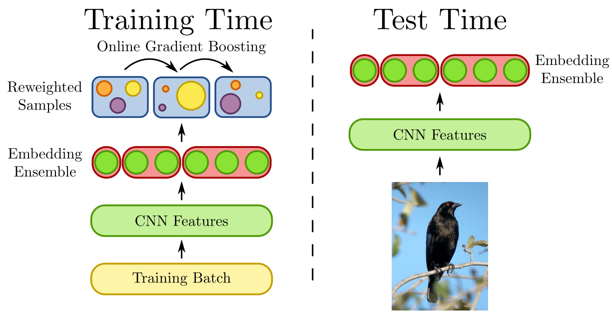

Fig. 1. During training time BIER uses online gradient boosting to train the individual learners. During test time we simply concatenate the predictions of all our learners to a single feature vector.

To further increase diversity in our method, we propose a novel auxiliary loss function [5]. We add adversarial regressors on pairs of learners. These regressors try to map one embedding to an other embedding, maximizing their similarity. Since we are using a gradient reversal layer [2], our hidden layers minimize the similarity w.r.t. to these regressors, making the embedding more diverse.

III. RESULTS

In our experiments we observe that BIER significantly reduces correlation of the embedding on the CUB dataset [7] by about 47.8%. We also compare our method and baseline to the state-of-the-art in Table I. BIER significantly improves performance and outperforms state-of-the-art methods.

TABLE I

EVALUATION OF BIER ON CUB [7] AND STFD. ONLINE PRODUCTS [4].

	CUB (R@1)	Stanford Online Products (R@1)
Proxy NCA [3]	49.2	73.7
Baseline	51.8	66.2
BIER	**57.5**	**74.2**

REFERENCES

[1] A. Beygelzimer, S. Kale, and H. Luo, "Optimal and Adaptive Algorithms for Online Boosting." in *Proc. ICML*, 2015.
[2] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-Adversarial Training of Neural Networks," *JMLR*, vol. 17, no. 59, pp. 1–35, 2016.
[3] Y. Movshovitz-Attias, A. Toshev, T. K. Leung, S. Ioffe, and S. Singh, "No Fuss Distance Metric Learning Using Proxies," in *Proc. ICCV*, 2017.
[4] H. Oh Song, Y. Xiang, S. Jegelka, and S. Savarese, "Deep Metric Learning via Lifted Structured Feature Embedding," in *Proc. CVPR*, 2016.
[5] M. Opitz, G. Waltner, H. Possegger, and H. Bischof, "Deep Metric Learning with BIER: Boosting Independent Embeddings Robustly," *arXiv:cs/1801.04815, submitted to TPAMI*, 2018.
[6] ——, "BIER: Boosting Independent Embeddings Robustly," in *Proc. ICCV*, 2017.
[7] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie, "The Caltech-UCSD Birds-200-2011 Dataset," California Institute of Technology, Tech. Rep. CNS-TR-2011-001, 2011.