

Accepted Manuscript

Structured Deep Fisher Pruning for Efficient Facial Trait Classification

Qing Tian, Tal Arbel, James J. Clark



PII: S0262-8856(18)30104-5
DOI: [doi:10.1016/j.imavis.2018.06.008](https://doi.org/10.1016/j.imavis.2018.06.008)
Reference: IMAVIS 3701
To appear in: *Image and Vision Computing*
Received date: 16 September 2017
Revised date: 4 June 2018
Accepted date: 21 June 2018

Please cite this article as: Qing Tian, Tal Arbel, James J. Clark , Structured Deep Fisher Pruning for Efficient Facial Trait Classification. *Imavis* (2018), doi:[10.1016/j.imavis.2018.06.008](https://doi.org/10.1016/j.imavis.2018.06.008)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Structured Deep Fisher Pruning for Efficient Facial Trait Classification

Qing Tian^{a,*}, Tal Arbel^a, James J. Clark^a

^a*Centre for Intelligent Machines & ECE Department, McGill University
3480 University Street, Montreal, QC, Canada*

Abstract

High efficiency is desirable for many interactive biometrics tasks, including facial trait recognition. Although deep convolutional nets are effective for a multitude of classification tasks, their high space and time demands make them impractical for PCs and mobile devices without a powerful GPU. In this paper, we propose a structured filter-level pruning approach based on Fisher LDA [1], which boosts efficiency while maintaining accuracy for facial trait classification. It starts from the last convolutional layer where we find filter activations are less correlated. Through Fisher's LDA, we show that this decorrelation makes it safe to discard directly filters with high within-class variance and low between-class variance. The pruning goes on by tracing deconvolution based dependency over layers. Combined with light classifiers, the reduced CNNs can achieve comparable accuracies on example facial traits from the LFWA(+) and CelebA datasets, but with great reductions in model size (96%-98% for VGG-16, 81% for GoogLeNet) and computation (as high as 80% for VGG-16, 61% for GoogLeNet).

Keywords: neural network pruning, Fisher LDA, facial trait classification

1. Introduction

Deep artificial neural networks have revolutionized many computer vision areas, in large part due to effective algorithms for large network training [2, 3,

*Corresponding author

Email addresses: qing.tian@mail.mcgill.ca (Qing Tian), arbel@cim.mcgill.ca (Tal Arbel), clark@cim.mcgill.ca (James J. Clark)

Download English Version:

<https://daneshyari.com/en/article/6937688>

Download Persian Version:

<https://daneshyari.com/article/6937688>

[Daneshyari.com](https://daneshyari.com)