

Journal Pre-proof

Learning Cascade Attention for fine-grained image classification

Youxiang Zhu, Ruochen Li, Yin Yang, Ning Ye

PII: S0893-6080(19)30331-4
DOI: <https://doi.org/10.1016/j.neunet.2019.10.009>
Reference: NN 4300

To appear in: *Neural Networks*

Received date : 1 March 2019
Revised date : 15 August 2019
Accepted date : 11 October 2019



Please cite this article as: Y. Zhu, R. Li, Y. Yang et al., Learning Cascade Attention for fine-grained image classification. *Neural Networks* (2019), doi: <https://doi.org/10.1016/j.neunet.2019.10.009>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. 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.

Published by Elsevier Ltd.

Learning Cascade Attention for Fine-grained Image Classification

Youxiang Zhu^a, Ruochen Li^a, Yin Yang^b, Ning Ye^{a,*}

^a*College of Information Science and Technology, Nanjing Forestry University, No.159 Longpan Road, Nanjing, 210037, PR China*

^b*Department of Electrical and Computer Engineering, The University of New Mexico, Albuquerque, NM 87131, USA*

Abstract

Fine-grained image classification is a challenging task due to the large inter-class difference and small intra-class difference. In this paper, we propose a novel Cascade Attention Model using the Deep Convolutional Neural Network to address this problem. Our method first leverages the Spatial Confusion Attention to identify ambiguous areas of the input image. Two constraint loss functions are proposed: the Spatial Mask loss and the Spatial And loss. Second, the Cross-network Attention, applying different pre-train parameters to the two stream architecture. Also, two novel loss functions called Cross-network Similarity loss and Satisfied Rank loss are proposed to make the two-stream networks reinforce each other and get better results. Finally, the Network Fusion Attention, fusions the middle stage outputs by the novel entropy add strategy and get the final predictions. All of these modules can work together and can be trained end to end. Besides, different from previous works, our model is fully weak-supervised and fully paralleled, which leads to easier generalization and faster computation. We obtain the state-of-art performance on three challenge benchmark datasets (CUB-200-2011, FGVC-Aircraft and Flower 102) with results of 90.8%, 92.1%, and 98.5% respectively. The model will be publicly available at <https://github.com/billzyx/LCA-CNN>.

*Corresponding author

Email address: yening@njfu.edu.cn (Ning Ye)

Download English Version:

<https://daneshyari.com/en/article/13429970>

Download Persian Version:

<https://daneshyari.com/article/13429970>

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