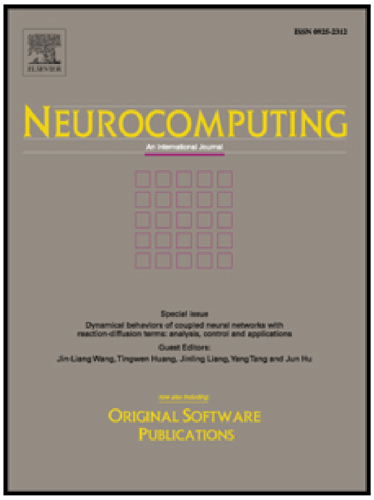
Communicated by Prof. R. Capobianco Guido

Accepted Manuscript



Fuzzy Qualitative Deep Compression Network

Wei Ren Tan, Chee Seng Chan, Hernán E. Aguirre, Kiyoshi Tanaka

PII: S0925-2312(17)30693-8
DOI: 10.1016/j.neucom.2017.04.023
Reference: NEUCOM 18356

To appear in: *Neurocomputing*

Received date: 26 April 2016
Revised date: 4 April 2017
Accepted date: 4 April 2017

Please cite this article as: Wei Ren Tan, Chee Seng Chan, Hernán E. Aguirre, Kiyoshi Tanaka, Fuzzy Qualitative Deep Compression Network, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.04.023

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
NEUROCOMPUTING 00 (2017) 1–23
Neurocomputing

Fuzzy Qualitative Deep Compression Network

Wei Ren Tan[a], Chee Seng Chan[b,*], Hernán E. Aguirre[a], Kiyoshi Tanaka[a]

[a]*Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano 380-8553 Japan*
[b]*Center of Image and Signal Processing, Faculty of Computer Science & Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia*

Abstract

Recently, Convolutional Neural Networks (CNNs) have become a popular choice to tackle image classification tasks. Despite that, it is almost infeasible to embed the CNNs into resource limited hardware (e.g. mobile devices) due to its extremely high memory requirement. To address this problem, several methods were proposed to reduce the CNN memory requirement with a minimum compensation on the classification accuracy. In this paper, we propose a novel one-shot deep compression method based on the fuzzy quantity space to remove redundant CNN weights. Experiments in three public datasets (i.e. MNIST, CIFAR-10 and ImageNet) showed that our proposed approach is able to compress the CNN up to 14× with a minimal loss of classification accuracy. Also, we present the first attempt to train an end-to-end fuzzy qualitative deep compression model in the fine-art paintings classification problem. We argue that the classification of fine-art collections is a more challenging problem in comparison to objects classification. This is because some of the artworks are neither non-representational nor figurative, and might even require imagination to recognize them. Hence, a question may arise as to whether a machine is able to capture imagination in paintings. One way to find out is by training a deep model and then visualize the low-level to high-level features learnt. Extensive experiments have been conducted on the recently publicly available Wikiart paintings dataset that consists of more than 80,000 paintings and our solution achieve state-of-the-art results (68%) in overall performance. The source code and models are available at: `https://github.com/cs-chan/fuzzyDCN`.

Keywords: Convolutional neural network, fuzzy deep learning, image classification, deep compression, painting classification

1. Introduction

For the past decade, deep learning has become the state-of-the-art technique for many computer vision, natural language processing and speech recognition tasks [1]. The convolutional neural network (CNN), in particular, has been the first choice method for image classification after Krizhevsky et al. [2] successfully implemented AlexNet to achieve state-of-the-art result in the ImageNet competition [3] in 2012. To date, the top performers for image classification tasks are all CNN-based models [4, 5, 6, 7].

As the performance of the CNNs become more and more promising, it is desirable to deploy them on embedded systems for practical use. However, CNN requires intensive memory and computational cost, where many embedded systems may not be able to cope with, specifically mobile devices. For example, the well-known CNN architecture AlexNet [2] has 61 million parameters, and it requires more than 200MB memory space. Meanwhile, the VGG-16 model [4] has 138 million parameters and requires over 500MB memory. Downloading any of these networks into the

*Corresponding author
Email addresses: `14st203c@shinshu-u.ac.jp` (Wei Ren Tan), `cs.chan@um.edu.my` (Chee Seng Chan), `ahernan@shinshu-u.ac.jp` (Hernán E. Aguirre), `ktanaka@shinshu-u.ac.jp` (Kiyoshi Tanaka)