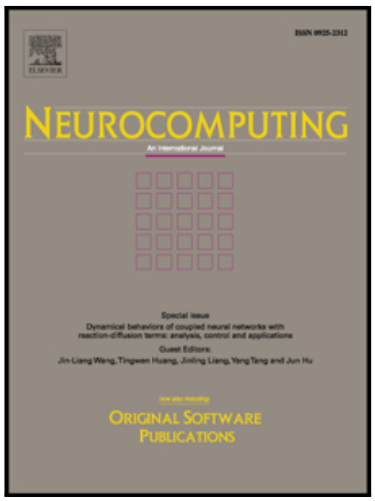

Roles of Pre-training in Deep Neural Networks from Information Theoretical Perspective

Yasutaka Furusho, Takatomi Kubo, Kazushi Ikeda

PII: S0925-2312(17)30443-5
DOI: 10.1016/j.neucom.2016.12.083
Reference: NEUCOM 18193

To appear in: *Neurocomputing*

Received date: 28 June 2016
Revised date: 11 November 2016
Accepted date: 17 December 2016

Please cite this article as: Yasutaka Furusho, Takatomi Kubo, Kazushi Ikeda, Roles of Pre-training in Deep Neural Networks from Information Theoretical Perspective, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2016.12.083

Roles of Pre-training in Deep Neural Networks from Information Theoretical Perspective

Yasutaka Furusho[a], Takatomi Kubo[a], Kazushi Ikeda[a]

[a] *Graduate School of Information Science, Nara Institute of Science and Technology, Ikoma, Nara 630-0192 Japan*

Abstract

Although deep learning shows high performance in pattern recognition and machine learning, the reasons remain unclarified. To tackle this problem, we calculated the information theoretical variables of the representations in the hidden layers and analyzed their relationship to the performance. We found that entropy and mutual information, both of which decrease in a different way as the layer deepens, are related to the generalization errors after fine-tuning. This suggests that the information theoretical variables might be a criterion for determining the number of layers in deep learning without fine-tuning that requires high computational loads.

Keywords: Deep Neural Networks, Pre-Training, Information Theory

1. Introduction

Deep learning, which is a multi-layered neural network, has been changing the history of pattern recognition and machine learning in accuracy performance [1] and can be applied to computer vision, automatic speech recognition and translation, and so on [2, 3]. However, the reasons for its high performance remain relatively unclarified since layered models have singular points that are difficult to treat statistically [4, 5]. In addition, the performance improved by combining several heuristics, such as drop-out [10] and pre-training [1, 6, 7].

Pre-training is an unsupervised learning algorithm that improves the classification performance. How pre-training affects classification performance through fine-tuning has been studied in several ways [7, 8, 9], suggesting that it works as a kind of regularization that resembles manifold learning.