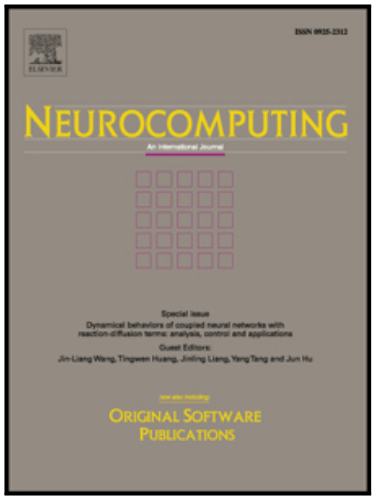
Communicated by Dr Jiwen Lu

Accepted Manuscript



Improving Deep Neural Network with Multiple Parametric Exponential Linear Units

Yang Li, Chunxiao Fan, Yong Li, Qiong Wu, Yue Ming

PII: S0925-2312(18)30125-5
DOI: 10.1016/j.neucom.2018.01.084
Reference: NEUCOM 19286

To appear in: *Neurocomputing*

Received date: 11 August 2016
Revised date: 14 January 2017
Accepted date: 31 January 2018

Please cite this article as: Yang Li, Chunxiao Fan, Yong Li, Qiong Wu, Yue Ming, Improving Deep Neural Network with Multiple Parametric Exponential Linear Units, *Neurocomputing* (2018), doi: 10.1016/j.neucom.2018.01.084

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.

ACCEPTED MANUSCRIPT

Improving Deep Neural Network with Multiple Parametric Exponential Linear Units

Yang Li[a], Chunxiao Fan[a,*], Yong Li[a,*], Qiong Wu[a], Yue Ming[a]

[a]*Beijing University of Posts and Telecommunications, School of Electronic Engineering, No.10 Xitucheng Road, Beijing, P.R.China, 100876*

Abstract

Activation function is crucial to the recent successes of deep neural networks. In this paper, we first propose a new activation function, Multiple Parametric Exponential Linear Units (MPELU), aiming to generalize and unify the rectified and exponential linear units. As the generalized form, MPELU shares the advantages of Parametric Rectified Linear Unit (PReLU) and Exponential Linear Unit (ELU), leading to better classification performance and convergence property. In addition, weight initialization is very important to train very deep networks. The existing methods laid a solid foundation for networks using rectified linear units but not for exponential linear units. This paper complements the current theory and extends it to the wider range. Specifically, we put forward a way of initialization, enabling training of very deep networks using exponential linear units. Experiments demonstrate that the proposed initialization not only helps the training process but leads to better generalization performance. Finally, utilizing the proposed activation function and initialization, we present a deep MPELU residual architecture that achieves state-of-the-art performance on the CIFAR-10/100 datasets. The code is available at https://github.com/Coldmooon/Code-for-MPELU.

Keywords: Deep learning, activation function, weight initialization

*Corresponding author

Email addresses: yli@bupt.edu.cn (Yang Li), cxfan@bupt.edu.cn (Chunxiao Fan), liy@bupt.edu.cn (Yong Li), qiongwu@bupt.edu.cn (Qiong Wu), yming@bupt.edu.cn (Yue Ming)

ACCEPTED MANUSCRIPT

Download English Version:

https://daneshyari.com/en/article/6863861

Download Persian Version:

https://daneshyari.com/article/6863861

Daneshyari.com