

# Bi-Real Net: Enhancing the Performance of 1-bit CNNs With Improved Representational Capability and Advanced Training Algorithm

Zechun Liu<sup>1</sup>, Baoyuan Wu<sup>2</sup>, Wenhan Luo<sup>2</sup>, Xin Yang<sup>3</sup>, Wei Liu<sup>2</sup>, and Kwang-Ting Cheng<sup>1</sup>

<sup>1</sup> Hong Kong University of Science and Technology

<sup>2</sup> Tencent AI lab

<sup>3</sup> Huazhong University of Science and Technology

zliubq@connect.ust.hk, {wubaoyuan1987, whluo.china}@gmail.com,  
xi nyang2014@hust.edu.cn, wliu@ee.columbia.edu, timcheng@ust.hk

**Abstract.** In this work, we study the 1-bit convolutional neural networks (CNNs), of which both the weights and activations are binary. While being efficient, the classification accuracy of the current 1-bit CNNs is much worse compared to their counterpart real-valued CNN models on the large-scale dataset, like ImageNet. To minimize the performance gap between the 1-bit and real-valued CNN models, we propose a novel model, dubbed Bi-Real net, which connects the real activations (after the 1-bit convolution and/or BatchNorm layer, before the sign function) to activations of the consecutive block, through an identity shortcut. Consequently, compared to the standard 1-bit CNN, the representational capability of the Bi-Real net is significantly enhanced and the additional cost on computation is negligible. Moreover, we develop a specific training algorithm including three technical novelties for 1-bit CNNs. Firstly, we derive a tight approximation to the derivative of the non-differentiable sign function with respect to activation. Secondly, we propose a magnitude-aware gradient with respect to the weight for updating the weight parameters. Thirdly, we pre-train the real-valued CNN model with a clip function, rather than the ReLU function, to better initialize the Bi-Real net. Experiments on ImageNet show that the Bi-Real net with the proposed training algorithm achieves 56.4% and 62.2% top-1 accuracy with 18 layers and 34 layers, respectively. Compared to the state-of-the-arts (e.g., XNOR Net), Bi-Real net achieves up to 10% higher top-1 accuracy with more memory saving and lower computational cost.

## 1 Introduction

Deep Convolutional Neural Networks (CNNs) have achieved substantial advances in a wide range of vision tasks, such as object detection and recognition [12, 23, 25, 5, 3, 20], depth perception [2, 16], visual relation detection [29, 30], face

---

Corresponding author.































