

発表概要

深層学習フレームワークにおける CPUとGPUの性能解析および最適化

樋口 兼一^{1,a)} 田浦 健次朗¹

2018年1月15日発表

今日では深層ニューラルネットワーク技術がAI分野に大きく貢献している。その学習プロセスは非常に時間がかかるため、計算環境に合わせた最適化が必要である。この発表では計算環境としてCPUとGPUを考え両者間の深層学習性能を比較する。広く使われている大半の深層学習フレームワークではGPU用の最適化は十分に行われているが、CPU用の最適化は積極的に行われていない。そのようなフレームワークの1つであるChainerも同様であり、内部の関数はCPU向けには十分に最適化されていないといえる。たとえばGPU上で走るほぼすべての関数はcuDNNやCuPyといった深層学習に特化したライブラリを用いて高速化されているが、CPU上で走る部分はNumPyでしか高速化されていない。NumPyは汎用的な科学計算ライブラリであるために、深層学習のような特殊な用途では十分にCPUを活用できていないといえず、元のChainerを使った性能比較ではCPUとGPU間のプロセッサレベルの差を反映できない。そこで正確な性能差を得るために、CPU上で走る関数内のNumPyで記述された部分をC言語で書き換えてOpenMPとIntel MKLで高速化を行った。その結果として得られる計測結果を用いて、CPUとGPUの深層学習における性能差とその特性を詳細に示す。

Performance Analysis and Optimization of CPU and GPU in Deep Learning Framework

TOMOKAZU HIGUCHI^{1,a)} KENJIRO TAURA¹

Presented: January 15, 2018

Nowadays, deep neural network technology has made a significant contribution to AI field. Its learning process is very time-consuming and needs to be optimized for a computing environment. In this work, we will show a compare of deep learning performance between CPUs and GPUs. Major deep learning frameworks are fully optimized for GPU, but not for CPU. Specifically, Chainer, a widely-used deep learning framework, does not use CPU-specialized kernels for deep learning. Almost all kernels performing well on GPU are optimized with cuDNN and CuPy, libraries specialized for deep learning. In contrast to that, kernels running on CPU are optimized only with NumPy. NumPy, general scientific computing library, is difficult to get high performance in deep learning. So the performance comparison between CPUs and GPUs with the original Chainer implementation does not reflect the performance difference of the processors. To deal with this problem, we rewrite existing kernels running on CPU with C language and optimize them with OpenMP and Intel MKL. We analyze the performance with detailed profiling and discuss characteristics of performance.

¹ 東京大学工学部電子情報工学科
Department of Information and Communication Engineering,
The University of Tokyo, Bunkyo, Tokyo 113-8656,
Japan

^{a)} thiguchi@eidous.ic.i.u-tokyo.ac.jp