

実験研究 注目AIコンパイラで 広がる組み込み人工知能の世界

ご購入はこちら

今井 健男, 山田 貴登

人工知能開発環境の現状

● フレームワークの世界はデファクト・スタンダードTensorFlow + α に収束

2015年頃から始まったディープ・ラーニング・フレームワークの開発競争は、2017年でほぼ収束した



写真1 推論/エッジ側で注目の人工知能コンパイラTVMを試す

ようです。世界的にデファクト・スタンダードとなったのはTensorFlowです。

そこに加えて、より簡単に使えるKeras、プロトタイプ開発に向いているPyTorch、そして国内では根強い人気のChainer、といった辺りでユーザも定着した感があります。

● 仁義なき戦い…AIコンパイラが絶賛競争中

代わりに始まったのはディープ・ラーニング用コンパイラの開発競争です。ここでいうコンパイラ(「グラフ・コンパイラ」、「バックエンド」などとも呼ばれる)とは、TensorFlowのようなフレームワークを使って学習済みのディープ・ラーニング・モデルを最適化し、さまざまな実行環境・デバイス向けに展開できるようにする、といったものです。代表的な人工知能コンパイラを表1に示します。

● 今回紹介する人工知能コンパイラ「TVM」& FPGA 演算

本稿ではこれらのうち、比較的古参で、また最近人気が出てきているコンパイラ「TVM」を紹介します。

表1 注目ポイント…AIコンパイラの開発競争が激しさを増している

| コンパイラ名 | 開発元 | 特徴 | 公式リリース |
|---------------------------------|-------------------|---|----------|
| Nervana Graph Compiler (nGraph) | インテル | Nervana Systems社(インテルが買収)が自社製プロセッサ向けに開発していたもの。今はインテルのCPUやNVIDIAのGPUなどにも対応。 | 2016年11月 |
| TVM | ワシントン大学 | さまざまなフレームワークとデバイスに対応。現在オープンソースでは1番人気。 | 2017年8月 |
| PraidML | Vertex.ai | 各ベンダのGPUに対応。2018年8月に開発元のVertex.aiがインテルに買収され、10月にnGraphに組み込まれた。 | 2017年10月 |
| TACO | マサチューセッツ工科大学(MIT) | 汎用コンパイラの手法を転用するというコンセプトで開発。疎行列計算の最適化を組み込んでいる。 | 2017年10月 |
| DLVM | イリノイ大学 | Swift言語で使えてアップルの機械学習ライブラリCoreMLフォーマットをターゲットにするiOS用コンパイラ。2018年8月に開発終了。 | 2017年11月 |
| Tensor Comprehensions | Facebook | パラメータの自動チューニングを遺伝的アルゴリズムで行う。しかし9月頃に開発メンバがこぞって抜けてしまったといううわさで、現在開発はほぼ停止。 | 2018年2月 |
| TIRAMISU | マサチューセッツ工科大学(MIT) | x86マルチコアCPU、NVIDIAのGPU、ザイリンクスのFPGAのそれぞれに対応。ティラミスのようにたくさんの層(4層)に重なったスタックで構成される。 | 2018年4月 |
| Glow | Facebook | PyTorch/Caffe2の派生プロジェクトで、さまざまなエッジ・デバイスに対応予定。インテル、Maxwell、Qualcommが公式サポート表明。 | 2018年5月 |
| ONNC | Skymizer/BITMAIN | ONNXを入力としてさまざまなプロセッサ向けにコンパイルできる。現在オープンソースで公開されているのはAIプロセッサのSophon(BITMAIN)が主なターゲット。 | 2018年8月 |