

深層学習による高性能インテリジェントカメラの開発

—小型・高性能画像認識システム Nano Deep—

1. 背景

近年、画像認識技術の進展は目覚ましく、特に深層学習 (Deep Learning) と呼ばれる技術の登場によって、コンピュータの画像認識能力は飛躍的に向上した。しかし、現状の Deep Learning は計算資源が潤沢なサーバでのみ実用的な時間内で実行が可能なものとなっている。これは Deep Learning がその他の既存手法に比べて膨大な計算を必要とするからである。

一方で、近年 IoT (Internet of Things) への関心が高まっており、近い将来、大量のカメラデバイスが画像や動画を撮影し、インターネットを介して通信するようになると予想される。しかし、末端のデバイスが撮影したデータを全てサーバに送信する現状のモデルでは、サーバの処理能力やネットワーク帯域に大量の負荷がかかる問題が残る。そこで、末端のデバイスでデータをある程度処理してから、最小限のデータをサーバへと送信して処理をする、エッジコンピューティングというモデルが必要になる。

Deep Learning をエッジコンピューティングに利用するためには、Deep Learning の膨大な計算を処理できる、携帯可能な小型デバイスを開発しなければならない。現在、サーバにおける Deep Learning の計算には、大規模な並列演算が可能な GPU が用いられているが、多くの GPU は大きさ、消費電力、コストの面で小型デバイスへの搭載には適さない。

2. 目的

本プロジェクトでは FPGA を利用することで、並列演算による高い処理能力を維持しつつ、小型、低消費電力、低コストの Convolutional Neural Networks (CNN) の計算に特化したハードウェアを開発する。また、FPGA を利用して作成した CNN 計算用ハードウェアをカメラと組み合わせ、エッジコンピューティングが可能なカメラデバイスを作成する。さらに、デベロッパがこのカメラデバイスを簡単に扱えるように、高度に抽象化された API を提供し、包括的な開発環境を構築する。

3. 開発の内容

本プロジェクトで開発したシステム Nano Deep は、CNN 計算モジュールやカメラ、その他周辺回路を含むハードウェアと、CPU からそれらを制御するソフトウェア、及び CPU 上でのプログラミングを簡素化するライブラリからなる包括的な環境である。図 1 にハードウェアの外観を示す。本デバイスは CPU と FPGA がワンチップに載った Xilinx 社 Zynq SoC を採用し、FPGA 上には独自ハードウェアが実装され、それは CPU から制御されるシステムとした(図 2)。

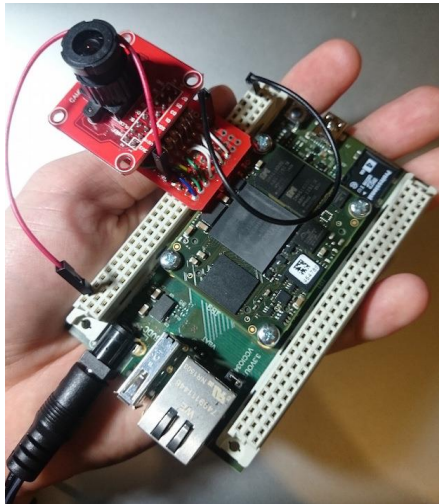


図 1 Nano Deep 外観

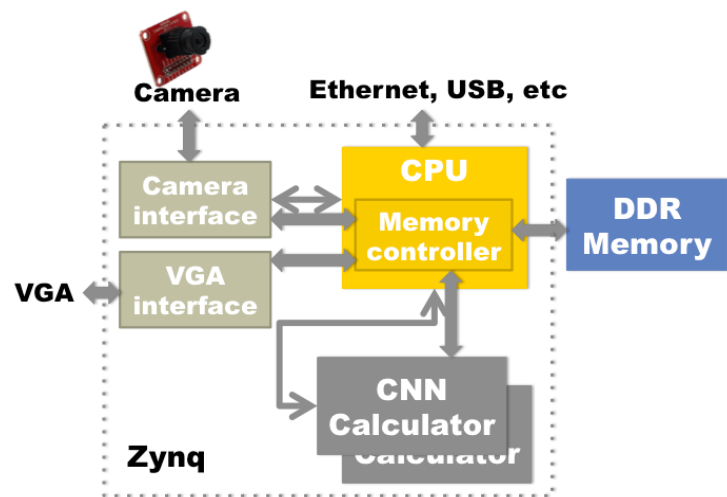


図 2 Nano Deep システム構成図

CNN 計算モジュールは FPGA 上に構築された独自回路で、CNN の計算を効率よく行うことを目的に設計されている。演算の固定小数点化、並列化、I/O の最適化など様々なチューニングによって、高いコスト対性能比を実現した。同一 SoC 内の CPU と比較した場合、100-180 倍の処理の高速化が認められた。

CPU 上では Linux(Ubuntu)が動作し、デバイスドライバを介して CNN 計算モジュールやカメラ等のハードウェアを制御できる。さらに、ハードウェア制御などの処理を抽象化して扱いやすくした Python のライブラリを整備した。これによって本デバイスを使ったシステムを開発するデベロッパは、低レイヤの処理を意識することなく、通常の Python プログラミングの延長として Nano Deep を活用するプログラムを書くことができる。

本プロジェクトでは、この Nano Deep 環境を利用してサンプルアプリケーションを開発した。このサンプルアプリケーションはカメラで撮影した画像を Web ストリーミングしつつ、その画像をリアルタイムで認識するものである(図 3, 図 4)。認識結果は保存され、認識されたカテゴリ毎に整理され、検索することができる(図 5)。本アプリケーションは上記の Python ライブラリを利用して開発されている。Python 用インタフェースが提供されていることで、Nano Deep のハードウェア処理と Web アプリケーションのような高レベルの処理とを組み合わせたアプリケーションが簡単に作成できる(図 6)。

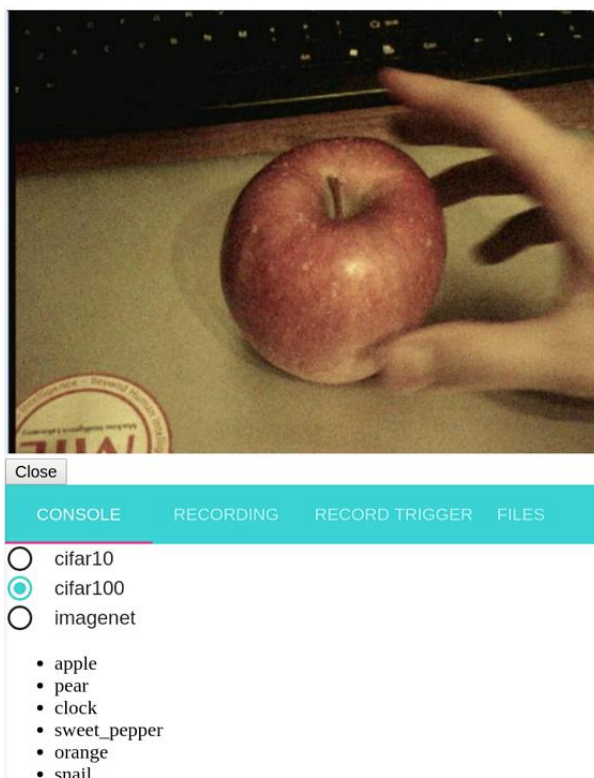


図 3 “apple”を認識

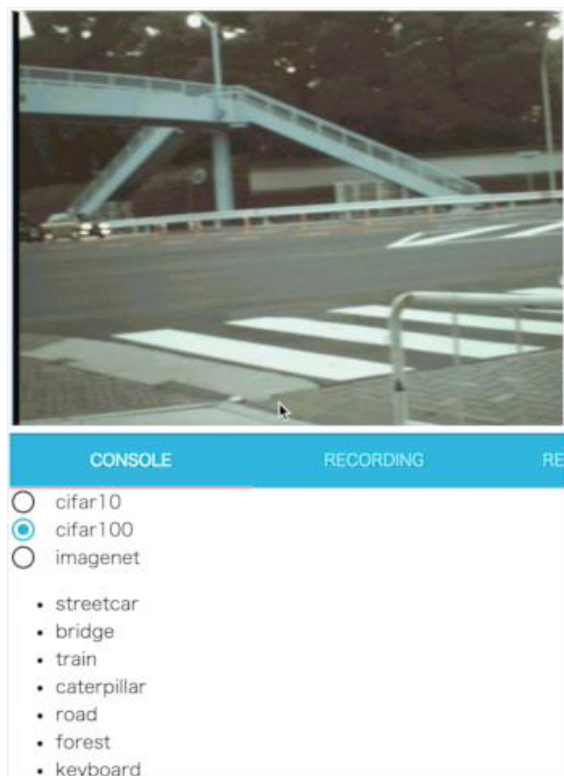


図 4 “streetcar”, “bridge”を認識

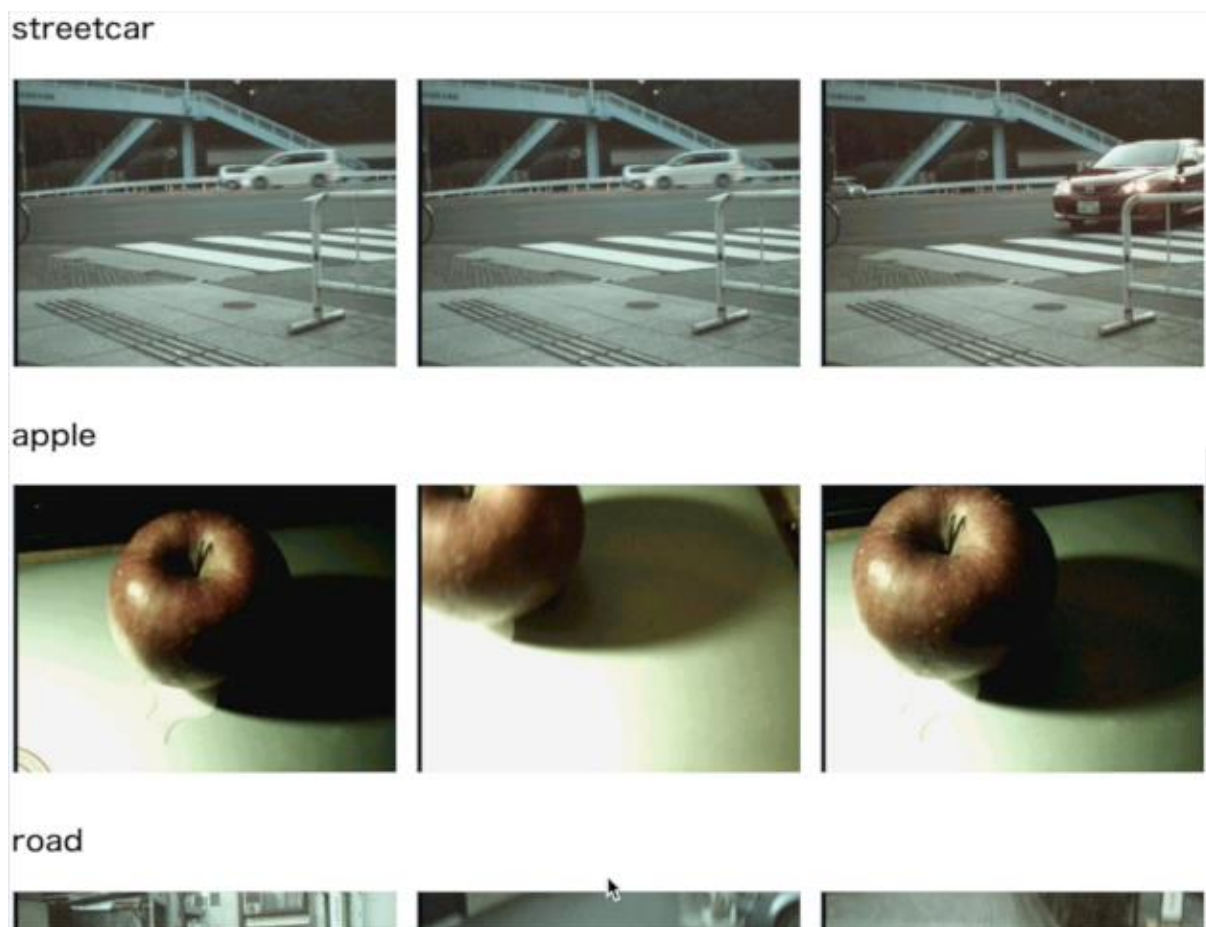


図 5 カテゴリ名による検索

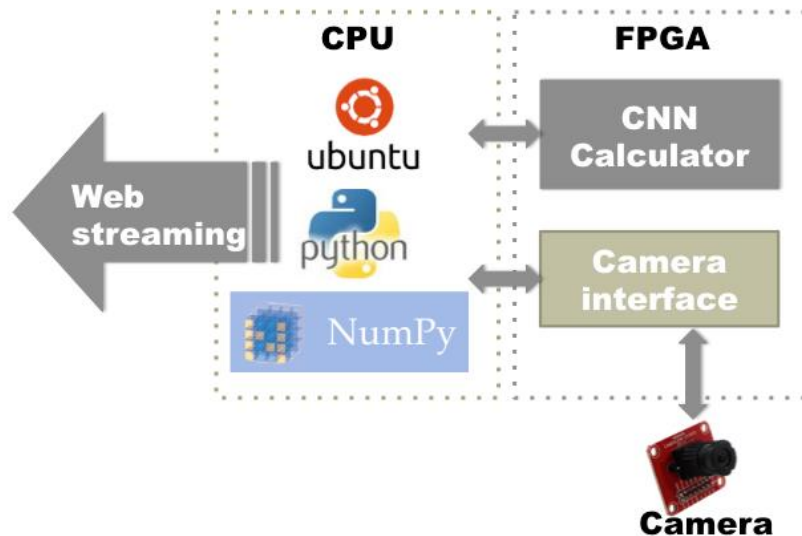


図 6 ソフトウェアとハードウェアの連携

4. 従来の技術(または機能)との相違

FPGA を利用して CNN の計算を高速化したプロジェクトは過去にいくつかあるが, Nano Deep は既存プロダクトに比べてコスト対性能比で有利である. これは, 既存プロダクトの多くがサーバサイドでの CNN 計算を高速化することを企図したものであるのに対し, Nano Deep は小型デバイスへの搭載を念頭に置き, 経済性や消費電力の低さなどの制約条件を満たしつつ最大限の性能を発揮させることを目指したからである.

また, ソフトウェア面での充実も重要な側面である. Linux や Python といった, デベロッパが一般的に利用する環境を Nano Deep 上に用意することで, アプリケーション開発のハードルが極めて低くなっている.

5. 期待される効果

小型デバイスで高性能な画像認識システムが動作することで, IoT デバイスやロボット・ドローンへの組み込みなど, Nano Deep には様々な応用先が期待される. Nano Deep を利用すれば, 一般的なソフトウェア開発者であれば, 深いハードウェアの知識なくアプリケーションを開発できるので, 幅広いデベロッパへの普及も期待される.

6. 普及(または活用)の見通し

デベロッパ向けプラットフォームとして本システムをよりブラッシュアップした上で, Kickstarter などのクラウドファウンディングサイトへの投稿を行い, 資金調達を兼ねた本システムのニーズの調査を行っていく.

7. クリエータ名(所属)

土屋祐一郎(東京大学 大学院)