エッジデバイスを用いた簡易顔認識示システムの試作

宮崎英一[1]・坂井　聡[2]・谷口公彦[3]
佐野将大[4]・近藤　創[5]

要旨

本研究では、エッジデバイスとして安価なシングルボードコンピュータ（Raspberry Pi）を用いた顔認識システムの試作を行った。ここでは、推論アクセラレータを組み合わせる事で、非力なエッジデバイスでも速度及び精度の面でも実用に耐えられる顔認識システムが構成できた。また通信プログラムを自作する事により、ネットワークを介してエッジデバイスとコンピュータ間でデータ共有が可能になった。これにより、エッジデバイスからクラウドコンピューティングの可能性も考えられる。

1．はじめに

現在、我々の生活に密着し、文部科学省の科学白書[i]においてもSociety 5.0（人工知能（AI），ビッグデータ，Internet of Things（IoT），ロボティクス等）の１つとして記載されるようになってきたAIであるが、AI等の新たなICTを、人間が「できること」を代替して雇用を奪うものと捉えるのではなく、人間の様々な能力を「拡張」（手足の機能、視覚・聴覚、理解・習得能力等を向上）することで、人間の「できること」を強化するものと捉える視点[ii]が重要である。

香川大学教育学部　宮崎英一・坂井聡（を中心とする）研究室では、視覚障害者支援として「ディープラーニングシステムを応用して動画からリアルタイム物体認識行うシステム」[iii]や、運動機能障害支援として「ディープラーニングシステムを用いてモーションヒストリーの解析を行うシステム」[iv]の開発を行ってきた。これらの研究は主として一般的なパーソナル・コンピュータを用いて行って来たが、家庭や学校あるいは病院といった一般的な環境下では、コンピュータのようなデバイスを机上あるいはベッド横に置くことはスペースや管理上好ましい事では無い。更にコンピュータに関しては、アップデートやソフトウェア・メンテナンスの問題があり、障害者やその家庭等で利用するのは、負担が大きいと考えられ、これが普及を妨げる原因の１つでもある。そこで、本研究ではパーソナル・コンピュータの代わりにエッジデバイスを用いる事で、小型、メンテナンスが不要、電源を接続するだけで動作する簡易顔認識シス

1　技術教育
2　特別支援教育
3　善通寺養護
4　高松養護養護
5　香川県教育委員会

テムを試作した。

本研究では、エッジデバイスとしてRaspberry Piを用いた。これは、ARMプロセッサを搭載したシングルボードコンピュータであり、元来は教育用コンピュータとして開発されたり、安価なシステムである。性能としては、価格と比例してあまり高速な処理は行えない。そこでここでは、Neural Compute Stickを組み合わせる事で、計算能力の向上を図った。またエッジデバイスは、物体認識において、モデルを利用して予測のみを行う事で、高いリソースを必要としないシステムとした。ここでは、このエッジデバイスで画像認識だけを行う事で、センサのように取り扱える画像認識システムとした。

2．エッジデバイス

エッジデバイス型AIの対義語としては、クラウドコンピューティング型AIが想定出来る。クラウドコンピューティング型AIはネットワーク環境で接続された高速で巨大なコンピュータリソースを有し、高い計算機スペックを提供するものである。一方、エッジデバイス型AIとしては、インターネットには接続されているが、安価で小型、計算機スペックは概して低い。

しかし、IoTにおいてエッジデバイスは大きな意味を持ち、分散コンピューティングとして様々な分野からデータを取得する事が可能になる。本来は、エッジデバイスから得られたデータ群をビックデータとしてクラウドコンピュータで解析するものであり、お互いが共同的にふるまうものである。

コンピュータを用いた画像認識における大まかな処理の流れとしては、

1）大量のデータを利用して予測モデルを作成（学習フェーズ）
2）作成した予測モデルを利用して実際に予測を行う（推論フェーズ）

となる。1）においては極めて多量のデータを利用して学習を行うので、コンピュータに高い計算性能が要求される。このため、クラウドコンピュータのような高性能なコンピュータ環境が要求される場合が多い。一方、2）において「モデルを利用して予測を行う」は、リアルタイム性が要求されるもの、演算には高い計算性能は要求されない。そのため、自動運転等ではエッジデバイスを用いて画像認識を行う事が行われている。本研究では、このエッジデバイスを用いる事で、家庭内や病院といった一般的な環境におけるシステムの実用化を目指すものである。

2.1　エッジデバイス

本研究では使用したエッジデバイスを図1に示す。これが上述したRaspberry Piシリーズの「Raspberry Pi 3 Model B+」であり、5000円程度で入手可能である。元来、教育用として開発されたが、安価な事、取り扱いが簡単な事、RaspbianというLinux系OSで運用出来る事、ハードウェア（GPIOポート）が制御できる事等から、ホビーユースだけに留まらず、プロトタイプの試作等では、企業でも利用される事が多い。このため多くの資料が存在し、いわゆるエコシステムとして構成されるようになった。

本研究では、これをエッジデバイスとして用いる事で、安価なシステムが実現出来た。ただ、リアルタイムで画像認識を行うにはどうしても基本的な計算能力では不足する。そこでこのデバイスでも高速処理を行うために、次に述べるアクセラレータを採用した。

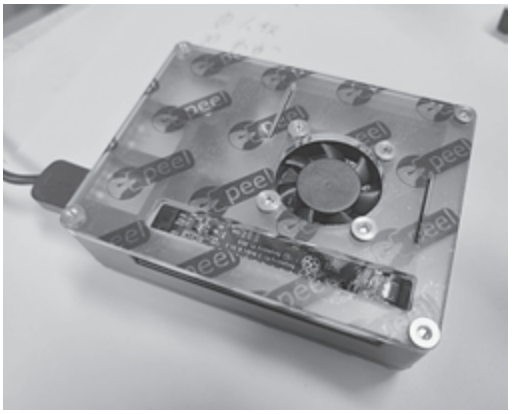

図1　Raspberry Pi 3 Model B+

2.2　Neural Compute Stick

本研究で用いたエッジデバイスのみで（Raspberry Pi）リアルタイムな画像認識を行う事は可能であるが、精度や応答性を向上させるためにNeural Compute Stickと呼ばれるアクセラレータを用いた。これは1万円程度で購入可能なUSBスティック型の推論アクセラレータで，高速な推論処理を実現するものであり、「AIプログラミングの外部演算装置として扱うことができるため、クラウドに頼らず、ローカル環境で画像処理やディープラーニング開発を行うことができるようになります。」と説明[v]されている。これを図2に示す。

本デバイスは現在生産中止であり、新しくNeural Compute Stick 2が発売されている。これは初期型に比較して8倍の性能を持つと説明されており、今後はこちらを導入予定である。しかもこれが、ソフトウェア上は旧アクセラレータとソフトウェア的にコンパチブルになっており、本研究で試作した初期型のプログラムがそのまま動作するので、単にデバイスの差し替えだけで計算の高速化が実現可能である。

3．物体認識システム

本研究では、物体認識システムにおいて独自に学習フェーズを行わず、OpenVINO[vi]ツールキットを用いた「ディープラーニング推論」を行った。

ディープラーニングによる画像認識においては

1）自分のオリジナルなモデルを学習

2）パブリックな学習済みモデルを使用

3）インテルの学習済みモデルを使用

の3つのアプローチがある。この中で1）はTensorflow等のフレームワークを使用して自前で計算するので手間が掛かる。2）と3）は事前に学習する必要は無いものの、2）はモデルをNeural Compute Stick用に変換する必要がある。そのため本研究ではインテル学習済みモデル[vii]を利用した。

3.1　顔認識システム

本研究では物体認識において上記の学習済みモデルを用いた顔認識システムを試作した。システムの詳細は参考文献[viii]に記載されており、概ねこの手順に沿ってシステムを構成させてもらった。この中に感情分析モデルがあり、WEBカメラで撮影した動画からエッジデバイスだけでリアルタイムで感情を分析する事が可能である。

その感情判別例を図3に示す。ここでは、撮影された顔動画からanger（怒っている）と判定されているが、本人的には怒っていないので、誤判断と推定されるか、もしくは他人から見ると怒っている顔として見えている可能性もある。

図2　Neural Compute Stick

図3　顔認識結果

図４　エッジデバイスから他のコンピュータとの連携

４．エッジからクラウドへ

本研究では、WEBカメラで撮影した動画からエッジデバイスだけでリアルタイムな顔認識が実行できる事が示され、研究の骨子は実現できた。ここでは、エッジデバイスで判定された解析結果を他のコンピュータに渡す事で、クラストコンピューティングの可能性を探すものである。

このシステムを図４に示す。ここでは、エッジデバイスとコンピュータ間をWi-Fiを介して接続し、ソケット通信を用いてデータのやり取りを行っている。送信するデータは顔認識結果よる判定された「anger」等の短いテキストだけなので、通信帯域をひっ迫するものでは無い。このデータをコンピュータ側で受信し、データ（顔認識結果）に応じた画像を表示するプログラムをprocssing言語で作成した。

その結果例を図５に示す。同図は画面構成上、同一画面に写っているが、実際にはエッジデバイス出力画面（WEBカメラで撮影された本人の顔）の上にコンピュータ上で表示されている図示された解析結果（笑い顔）を重ねている。このようにエッジデバイスからコンピュータへのデータ通信が行えたので、今後は、このデータをクラウドコンピューティングで利用するようなシステムに応用できる事が分かった。

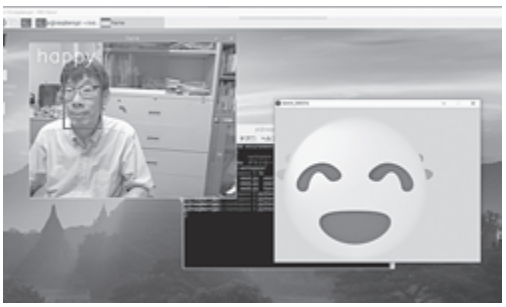

図５　PCとの顔認識結果連携

５．今後に向け

今回はエッジデバイスを用いて顔認識による、感情判別を行い、この結果がネットワークを介して他のコンピュータで処理できることが分かった。

しかし今回は、学習モデルは既に計算済みの学習モデルを用いたため、今後はオリジナルな学習モデルの構築を行う必要がある。

６．おわりに

本研究では，エッジデバイスを用いてリアルタイムで顔認識を行い，感情表現を判断する事が出来た。

今回は試作という事で、研究の中心となる部分は参考文献[vii]通りの手法を使用させて頂いたため、ここに新規性は無い。しかし、エッジデバイスを用いた顔認識システムをツールとして使用する事は出来た。

今後はこのツールを活用し、ディープラーニングの持つ特性を生かした障害をもった方用の入力インタフェースにつなげていけるようなシステムの構築を目指すものである。

７．謝辞

本研究は，令和元年度「重度重複障害児のコミュニケーション獲得性に向けた支援者支援ディープラーニングシステムの試作（課題番号：19K11417）」の一部として行われたことを記して謝意を示す。

８．参考文献

ⅰ　平成30年度文部科学白書＞第11章　ICTの活用の推　進，https://www.mext.go.jp/b_menu/hakusho/html/hpab201901/detail/1422160.htm，2020

ⅱ　令和元年版　情報通信白　人間とICTの新たな　関　係，https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/html/nb000000.html，2020

ⅲ　視覚障がい者に向けたディープラーニングを用いた 物体提示システムの試作，宮崎英一，坂井聡，谷口公彦，佐野将大，近藤創，香川大学教育学部研究報告（２）pp，125−129，2020

ⅳ　ディープラーニングを用いたモーションヒストリー解析の応用，宮崎英一，坂井聡，谷口公彦，佐野将大，近藤創，香川大学教育学部研究報告（２）pp.119−123，2020

ⅴ　Movidius Neural Compute Stick，https://www.switch-science.com/catalog/3501/, 2020

ⅵ　OpenVINO　Deploy High-Performance Deep Learning Inference，https://software.intel.com/content/www/us/en/develop/tools/openvino-toolkit.html，2020

ⅶ　Pretrained Models，https://software.intel.com/content/www/us/en/develop/tools/openvino-toolkit/pretrained-models.html，2020

ⅷ　OpenVINO™ でゼロから学ぶディープラーニング推論，https://jellyware.jp/openvino/#04, 2020