

## 深層学習の分散実行に関する一検討

福島 悠太\*<sup>1</sup> 三浦 太樹\*<sup>1</sup> 濱谷 尚志\*<sup>1</sup> 山口 弘純\*<sup>1</sup> 東野 輝夫\*<sup>1</sup>\*<sup>1</sup> 大阪大学 大学院情報科学研究科

{ y-fukushima, d-miura, h-takasi, h-yamagu, higashino }@ist.osaka-u.ac.jp

## 1 はじめに

世界のデジタルデータは飛躍的に増大しており、2020年にはインターネットに接続する機器が全世界で500億台を超え、特にセンサーなどのIoT機器が普及すると予想されており、センサーが収集するデータ量の増加とそれに伴う電力消費が懸念される。そこでスマートシティをはじめとする持続可能な環境の実現には、増大するIoT機器からのデータをいかに効率よく処理するかが重要な課題となっている。

IT機器による消費電力増大は全世界的な課題であり、その消費電力量は2025年には、2015年の約2.5倍になるとの報告もなされている。近年省エネルギー化の努力がなされているものの、依然、2020年には世界全体でおよそ1,400テラワット時(全世界の消費エネルギーの1.1~1.5%)の電力がIT機器によって消費されるといわれている。従来のIoTデータ処理は、多数の機器やセンサーからのデータをビッグデータとしてクラウドストレージに蓄積し、計算時間と計算資源を大量に投入して分析を行い、知識やパターンを抽出する集中型クラウド処理を前提とした「クラウドヘビー型」であった。しかし、クラウドヘビー型は大量の電力を消費するため、IoTデータ・デバイスの爆発的増大の時代を迎えた今日において、電力課題を抜本的に解決するIoTデータアーキテクチャを開発していくことは極めて重要である。

そこで本研究では、これまでクラウドが実施していた機械学習などのデータ解析処理を分散IoTセンサーネットワーク内のエッジノードに移行し、網内でローカル無線センサーネットワークを介してIoTセンサーノードが集約計算などの高次の学習処理、およびデータ転送を実施し、学習とデータ解析を自律分散的に実行可能な「エッジヘビー型」となるアーキテクチャの実現を目指す。エッジコンピューティングにより、センサーネットワーク自身が知的、かつ省電力でデータ処理、転送ならびに学習を行うことで、クラウドベースの処理と比較し、学習性能を維持しつつ、消費電力およびトラフィック量を削減することを目指す。

また近年、物体認識、画像検索など、データ解析において、深層学習が主流となりつつあり、畳み込みニューラルネットワークを用いた研究が数多くなされている[1]。そこで本稿では、豊富な計算資源を前提としている深層学習について、その処理をエッジノード、およびIoTセンサーノードに分散して実装し、深層学習をセンサーネットワーク内で分散実行可能なプロトコルを提案する。

## 2 関連研究

深層学習を分散処理する手法として、GPUを用いて処理を並列化する手法が提案されている[2]。この手法は、入力データを集約し学習する際に、大量に必要な計

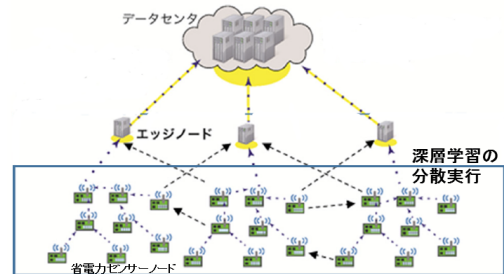


図1: エッジヘビーコンセプト

算をGPUを用いることにより、処理を高速化する手法である。

他に、SINGAと呼ばれる学習システムを用いた手法が提案されている[3]。SINGAはニューラルネットワークの分割スキームを提供し、分割したニューラルネットワークで学習を行うことにより、大規模なモデルの学習を可能にしている。

これらの手法に対し、本稿で提案する手法は、センサーネットワークを構成するノードが深層学習におけるユニットの役割をすることで、センサーネットワーク内で深層学習の分散実行を行い、省電力化を図る手法である。

## 3 提案手法

## 3.1 プロトコルのアーキテクチャ

提案する手法は、今までの「クラウドヘビーコンセプト」とは異なり、図1のように分散する大量のIoT機器がセンサーネットワークを構成し、センサーネットワーク内で各ノードが深層学習におけるユニットの役割をし、深層学習の分散実行を行う手法である。近年、計算機能を搭載した消費電力の少ないIoT機器が普及しており、それらをセンサーノードとして用いることで、エナジーハーベスティングに近い電力で動作させることができる。深層学習において、畳み込み処理の場合、近隣にユニットに対してフィルタをかけるため、各ユニットは近隣のユニット間のデータ通信のみで、畳み込み処理を行うことができる。したがって、センサーネットワーク内でできる処理は、センサーネットワーク自身で行い、その処理電力を環境エネルギーで賄い、最低限の電力でデータを処理し、更には処理済みのデータをクラウドに送信することで、各センサーノードがクラウドへデータを送信する場合と比較して、通信のトラフィック量・クラウドサーバに蓄積されるデータ量の両方において効率化を図る。

本稿では、分散実行する深層学習のネットワーク構成として、畳み込み2層、プーリング1層、全結合2層の計5層のネットワークを用いる。深層学習の分散実行において、各ノードで計測したデータに対し、畳み込みプーリング処理をセンサーネットワーク内のノード間で通信することによって実行する。各ノードでの畳み込み、プーリング結果はエッジノードに集約され、エッジで全結合

A Study on Distribution and Execution of Deep Learning  
Yuta Fukushima \*<sup>1</sup> Daiki Miura \*<sup>1</sup> Takashi Hamatani \*<sup>1</sup>  
Hirozumi Yamaguchi \*<sup>1</sup> Teruo Higashino \*<sup>1</sup>  
\*<sup>1</sup> Graduate School of Information Science and Technology, Osaka Univ., Osaka Japan

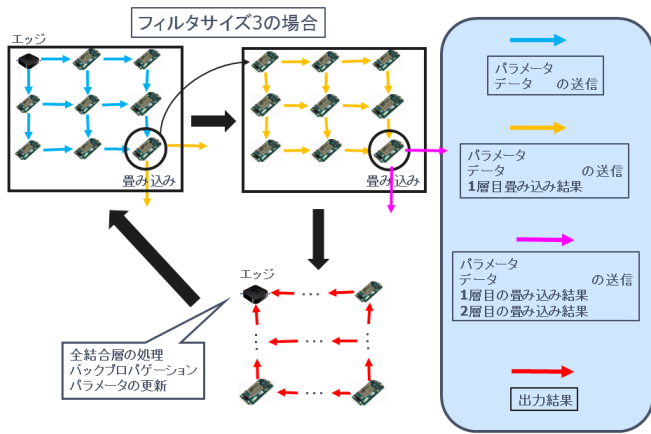


図 2: 深層学習の分散実行の流れ

層の処理・バックプロパゲーションを行う。これを繰り返すことにより、センサーネットワーク内での深層学習の分散実行を可能にする。しかし、センサーネットワーク内で学習をする際に、ノード間のデータの受け渡しのための通信コストの問題が発生する。そこで、センサーネットワーク内で深層学習を分散実行する際のノード間の通信コストを削減する手法について提案する。

### 3.2 ノード間の通信コスト削減方法

深層学習を行う上での仮定として、まず各センサーノードがメッシュ状に配置されており、それぞれが1つの入力データを持っているとし、ノード間はワンホップで通信を行うものとする。通信コスト削減のために、エッジから各ノードへのフィルタの更新パラメータの送信処理と各畳み込み・プーリング層でのデータの周辺ノードへの送信処理を並列して行う。まず図2のように、エッジからパラメータを隣のノードに送信する。受信したノードは、隣の次のノードにパラメータとデータを送信する。このとき、同じデータを重複して送信しないように、送信データを制御する。これを繰り返し、周辺のノードのデータを取得したノードがその時点で畳み込み処理を行う。次に、畳み込み処理を終えたノードはパラメータ・データと同時に畳み込み結果も次の隣のノードに送信する。これも繰り返していくことで、周辺のノードの畳み込み結果を取得することができ、この時点でこのノードが次の畳み込み処理を行うことが可能となる。以上の処理をエッジから端のノードまで繰り返すことで、1回のデータの流れて畳み込みからプーリングまでの処理を行うことができる。その後は、逆のデータの流れてエッジにすべての出力結果を集約する。そしてエッジで全結合層の処理・バックプロパゲーションを行い、パラメータの更新を行う。これで1回の学習が完了する。2回目以降も同様にエッジから更新したパラメータをノードに送信しながら、データを送信し、畳み込み・プーリングを行う。

## 4 深層学習の結果

本研究では、深層学習の分散実行を適用する環境として、およそ 1,400m<sup>2</sup> のラウンジを想定しており、この空間を 17 × 25 のメッシュ状に分割したときの 2次元の温度データを入力データとし、そのデータが異常であるか否かを判定する。温度データはタイムスロットごとの 17 × 25 の 2次元の温度データであり、1タイムスロットを 30

分とした。また、温度データの上下左右隣り合うデータの温度差が 3.0℃以上であるとき、異常とラベル付けた。以上の設定のもと、まず深層学習自体が適切に行うことができるか否かを調査するため、単一の計算機において深層学習を実行した。訓練データ数 2,400、テストデータ数 561、バッチサイズ 100、中間層のユニット数 200、学習回数 20、フィルタサイズ 3、フィルタ数 5 で深層学習を行った場合、訓練データの数が少ないため、精度の収束が遅くばらつきがあったが、およそトレーニング精度 94.4%、テスト精度 92.8% となった。また、フィルタサイズを変えて深層学習を行ってみると、4 の場合、トレーニング精度 95.0%、テスト精度 93.7%、5 の場合、トレーニング精度 94.5%、テスト精度 93.4%、6 の場合、トレーニング精度 93.0%、テスト精度 90.0% となった。フィルタサイズ 4、5 が精度が良く、それ以外では精度が下がることが確認された。

また、入力データを  $m$  行  $n$  列の 2次元のデータ、フィルタサイズ  $k$ 、フィルタ数  $h$  としたときの、1回の畳み込みにおける通信コストは、およそ  $(k^2 - 1)mnh$  となる。フィルタサイズが増えるにつれて、通信コストはフィルタサイズの二乗に比例して増加するが、精度の大きな向上は見られないため、今回の環境では、フィルタサイズ 3 で十分といえる。

## 5 おわりに

本稿では、深層学習を IoT センサーネットワーク網内で自律的に実行するプロトコルを提案した。

今後、深層学習を分散実行可能な環境を仮想的に構築し、仮想環境内で深層学習を分散実行し、分散処理による深層学習と分散処理によらない深層学習を比較することによって、分散実行の動作確認を行う。その後、実環境で起こりうるパケットロス・センサーノードの故障などを想定した場合における学習精度の評価を行う予定である。

仮想空間上での深層学習において、分散実行による通信コストの削減が認められれば、現在ラウンジに展開している IoT センサーによるデータ収集環境に深層学習を分散して実装し、異常な温度を検知する実証実験を行う予定である。

## 謝辞

本研究成果の一部は、国立研究開発法人情報通信研究機構 (NICT) の委託研究「未来を創る新たなネットワーク基盤技術に関する研究開発」ならびに JSPS 科研費 15K12019 の助成を受けたものです。

## 参考文献

- [1] Wang, P., Li, W., Gao, Z., Zhang, J., Tang, C. and Ogunbona, P. O.: Action recognition from depth maps using deep convolutional neural networks (2015).
- [2] Xiaqing Li, Guangyan Zhang, K. L. and Zheng, W.: Deep Learning and Its Parallelization: Concepts and Instances.
- [3] Ooi, B. C., Tan, K.-L., Wang, S., Wang, W., Cai, Q., Chen, G., Gao, J., Luo, Z., Tung, A. K., Wang, Y. et al.: SINGA: A distributed deep learning platform, *Proceedings of the 23rd ACM international conference on Multimedia*, ACM, pp. 685–688 (2015).