

家庭用サービスロボット向け組み込み一般画像認識の検討

神谷 亮範[†] 田中 宙夫[†] 森江 隆[†] 田向 権[†]
[†] 九州工業大学大学院生命体工学研究科

1. はじめに

近年、掃除ロボットや介護ロボットなど人々の生活に役に立つ家庭用サービスロボットが注目を浴びている。このサービスロボットの要素技術発展のために、ロボカップ@ホームリーグが開かれている。@ホームリーグでは、ロボットに頑健な画像認識機能が求められる。たとえば課題の中には、指定されたモノをロボットが探し出して掴むというタスクが存在し、その実現にはロボットが物体を正確にかつリアルタイムに認識する必要がある。

画像認識の分野において、Deep Learning は、認識精度を従来技術と比べ飛躍的に伸ばすことに成功したことで注目を浴びている [1]。Deep Learning を@ホームリーグのロボットに取り入れることで、@ホームロボットの物体認識精度の向上につながるかと期待される。しかしDeep Learning は計算量が多く、膨大な処理時間が必要である。そこで本研究では、Deep Learning の演算量やメモリ使用量を削減し、サービスロボットへの組み込み化が可能かどうか検討する。

2. Deep Learning の組み込み化に向けて

サービスロボットへの一般画像認識の組み込みでは、少ない演算量で実用的な認識率が得られることが重要である。Deep Learning は入力層と中間層、出力層で形成されたニューラルネットワークを多層に組み合わせて構成されており、それぞれの層に多数のノードが存在する。このノード数と認識率には密接な関係があり、またノード数が増加すると処理時間は多くなる。適切なノード数で実用的な認識率を得るには、これらの関係を十分に把握する必要がある。また、組み込み化においては、固定小数点演算の導入や学習アルゴリズムの簡略化といった検討が必要である。

そこで本稿では、Deep Learning の組み込み化の最初のステップとして、ノードの数と認識率及び学習時間の関係を調査する。Deep Learning での検証の前準備として、本稿では3層のMulti-Layer Perceptron (MLP)を用いる。

3. 実験結果

3層のMLPで、手書き文字の画像のデータセットであるMNISTの学習を行った。入力は、28×28ピクセルのグレ

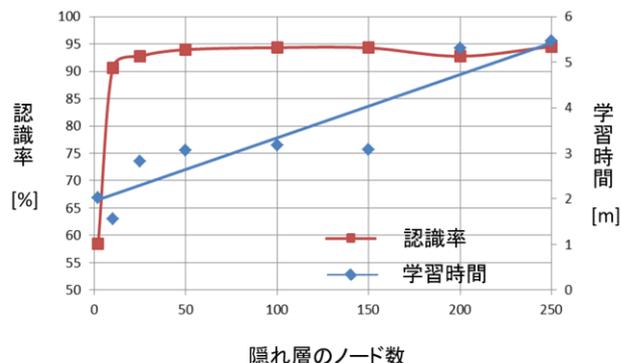


図1:ノード数に対する学習時間と認識率の関係

ースケール画像を60,000枚用いた。教師データは0~9までの10クラスの情報をを用いた。すなわち、入力層のノード数は784、出力層のノード数は10となる。

実験環境として、CPU Intel Core i7-4820k 3.7 GHz 4Cores, GPU NVIDIA GeForce GTX780を用いた。GPUの動作周波数は863[MHz]である。Theanoライブラリ [2]を利用してMLPを構築し、NVIDIA社のCUDAによるGPGPUコンピューティングでその動作を確認した。

図1に隠れ層のノード数に対する認識率とその学習時間を示す。認識率はノード数50以上で一定の値に収束する。一方で学習時間は隠れ層のノードの数に比例する傾向があることが分かる。よって、隠れ層のノード数を適切に調整することで無駄な学習時間を省くことができる。

4. むすび

本稿ではDeep Learningの組み込み化の前準備として、MLPにおけるノード数と学習時間、認識率の関係を調べた。今後は、サービスロボット向け一般画像認識を対象として、固定小数点化やアルゴリズムの簡略化を行う。さらに、Deep Learningの研究では、問題の規模に対する適切な構造獲得も話題になっており[1]、組み込み化の視点からネットワーク構造の最適化にも取り組んでいく。

参考文献

- [1] Jürgen Schmidhuber, "Deep Learning in Neural Networks: An Overview," Neural Networks, vol. 61, pp. 85-117, 2015.
 [2] Theano, <http://deeplearning.net/software/theano/>, (2015/02/05 アクセス).