

Mobile-aware Convolutional Neural Network for Sensor-based Human Activity Recognition

小林 慧[†]
福井大学大学院[†]

長谷川 達人[‡]
福井大学大学院[‡]

1 はじめに

センサベースの行動認識では畳み込み層を3層ほど重ねた Convolutional Neural Network (CNN) や VGG などの 2014 年頃に提案された構造の適用に留まっている。そのため、最新のモデル構造がセンサベースの行動認識に有効であるか定かではない。また、深層学習は多くの計算リソースを必要とするが、スマートフォンはリソースが限られている。本研究では、Neural Architecture Search (NAS) を用いて、センサベースの行動認識向けに最新のモデル構造を取り入れた軽量なモデル MarNASNet を提案する。

2 提案手法

本研究では Keras Tuner を用いてベイズ最適化によりモデルアーキテクチャを探索する。

2.1 探索範囲

本研究では CNN モデルをブロックに分割しブロックごとに畳み込み層の種類などを別々に探索する探索範囲を採用した。例として、 i 番目のブロックは次に示すような探索空間で探索される。

- **ConvOps**: regular conv (Conv), separable conv (SepConv), and mobile inverted bottleneck conv (MBConv) [1]
- **KernelSize**: 2 - 5
- **SkipOps**: pooling, identity residual, no skip
- **層の数 N_i** : 2 - 5
- **出力フィルタ数 F_i** : 32 - 192 (4 steps)

全体で共通のパラメータとして Squeeze-and-

excitation ratio (*SERatio*: 0.0, 0.25) を探索する。畳み込みブロックの数は 4 に固定した。今回は、出力フィルタ数の探索範囲を変えた 3 種類の探索範囲 (A/B/C) を設定しモデルを探索した。A は $F_i = 16 \cdot 2^{i-1}$ ($i = 1, 2, 3, 4$) に固定した探索範囲で、B は F_i を 32 - 128 に設定した。C は上記に示す探索範囲である。

2.2 探索設定

探索に使用したデータセットは、HASC データセット [2] である (表 1)。学習データとして 100 名 (D_{train})、検証データとして 50 名 (D_{valid})、テストデータとして 20 名 (D_{test}) の計 170 名のデータを使用する。探索時には D_{train} と D_{valid} を用いる。 D_{valid} に対する Accuracy を最大化することをモデル探索の目的に設定し、探索回数は 1000 とした。探索時のエポック数は 20 とした。学習時のバッチサイズは 1024、最適化手法は Adam、学習率は 0.001、損失関数は多クラス交差エントロピーとした。また、 D_{train} に対しデータ拡張として flipping と channel shuffling を適用している。探索は Intel Core i9-9900X, RAM 64GB, NVIDIA TITAN RTX を搭載した計算機で行い、各探索範囲で約 5 日ずつかかった。

2.3 MarNASNet アーキテクチャ

図 1 に探索によって得られた 1 つのモデルである MarNASNet-C のアーキテクチャ概要図を示す。探索されたモデルは、*ConvOp* として MBConv がよく使われており *SkipOp* として identity residual がよく使われている。また、*SERatio* は

表 1 使用した HASC データセットの詳細

行動	6 種類
サンプリング周波数	100Hz
デバイス	iPhone, iPod touch
被験者数	170 名
ウィンドウサイズ	256
ストライド	256

Mobile-aware Convolutional Neural Network for Sensor-based Human Activity Recognition

[†] Satoshi Kobayashi, Graduate School of Engineering, University of Fukui

[‡] Tatsuhito Hasegawa, Graduate School of Engineering, University of Fukui

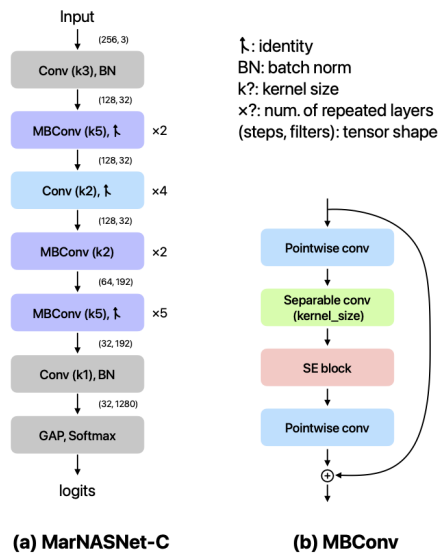


図1 MarNASNet-C アーキテクチャ
(a): MarNASNet-C; (b): MBCConv (mobile inverted bottleneck conv)

0.25 が選ばれている。

3 評価実験

探索されたモデルを評価するために、推定精度の評価実験とオンデバイス性能の評価実験を行う。ベースラインとして関連研究で用いられているシンプルな CNN [3] と画像分類分野で提案された 10 種のモデルと比較する。

探索されたモデルの推定精度の評価は、 $D_{train} + D_{valid}$ で訓練し D_{test} で評価したときの Accuracy によって行う。評価実験では 5 回試行の平均 Accuracy を最終的な評価値とする。5 回の試行において、各試行ごとに被験者はシャッフルされている。エポック数は 1000 とし、バッチサイズ等はモデル探索時の訓練時設定と同じである。

デバイス上での性能を評価するために、モデルをスマートフォン上に実装し実環境下における推論時間を測定する。今回は iOS 15.0.1 を搭載した iPhone 12 mini を使用し、モデルの実装には Core ML を用いた。作成した iOS App では、加速度データを 0.01 秒間隔 (100Hz) で収集し 3 軸 256 サンプルが収集されたら推論するという動作を、1 分間行い平均推論時間を計測する。

4 実験結果

図 2 に推定精度、平均推論時間、モデルサイズをバブルチャートで示す。DenseNet 121 や Effi-

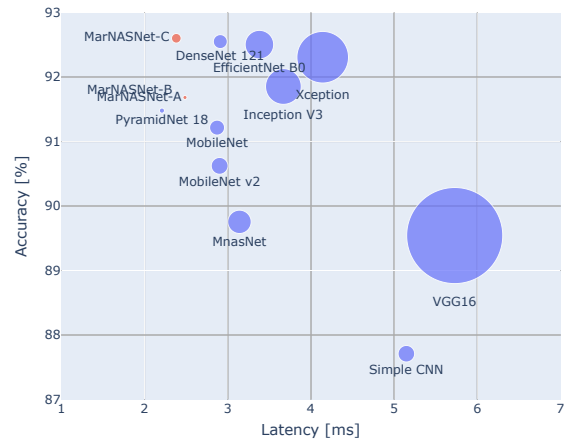


図2 Accuracy vs. Latency vs. Model size

cientNet B0 と比較して、MarNASNet-C によって同等の精度をより小さいモデル (=より少ないパラメータ数) かつより小さい推論時間で達成していることが分かる。また、PyramidNet 18 と同等のモデルサイズと推論時間でより高精度を達成していることが分かる。

5 まとめ

本研究では、センサベースの行動認識のため CNN モデルをベイズ最適化による NAS によって探索し、得られたモデルの有効性を評価した。既存のモデルアーキテクチャと比較して、探索されたモデルは同等の推定精度をより小さいモデルサイズかつ小さい推論時間で達成した。今後は、他のプラットフォームでの検証や探索プロセスに推論時間の情報を取り入れることを行いたい。

参考文献

- [1] M. Tan, et al. MnasNet: Platform-Aware Neural Architecture Search for Mobile. In *Proc. of the CVPR*, pp. 2820–2828, 2019.
- [2] N. Kawaguchi, et al. HASC Challenge: gathering large scale human activity corpus for the real-world activity understandings. In *Proc. of the AH*, pp. 1–5. ACM, 2011.
- [3] F. Li, et al. Comparison of Feature Learning Methods for Human Activity Recognition Using Wearable Sensors. *Sensors*, Vol. 18, No. 679, pp. 1–22, 2019.