

DeepLearningによる次元圧縮を用いた時系列行動認識

Sequential activity recognition by dimension reduction with deep learning

横山 晃 岡留 剛 角所 考
Akira Yokoyama Takeshi Okadome Koh Kakusho

関西学院大学理工学部人間システム工学科
Kwansei Gakuin University, Department of Human System Interaction

Recurrent Neural Networks(RNN) are an extremely precise class for time-series classification, but times it needs to learn from input sequences exponentially increases as in higher dimensions. To overcome this shortcoming, we use deep autoencoders to reduce input dimension, and classify sequential activity acceleration datasets for evaluation. We firstly obtain the hand-crafted features for acceleration data, and give it to our own window integrating process. Finally, the classification accuracy marked almost the same score as that with no reduction, although the computation time exponentially decreases.

1. はじめに

リカレントニューラルネットワーク (RNN) は、ループ構造を含むニューラルネットワークである。データの時間情報を分析することが可能で、音声認識や手書き文字認識に有効であることが知られている。しかし、RNN は学習の際計算時間が $O(N^2T)$ と非常に長く必要となる (N :ノード数, T :時系列長)。

2. 研究目的

本研究では DeepAutoencoder を用いて次元圧縮した特徴量を RNN の入力とすることで行動分類を行なう。次元削減によってノード数 $N \cdot$ 時系列長 T を減少させることで、RNN の学習計算が容易になる。DeepAutoencoder と RNN を組み合わせた複合型ニューラルネットワークによってこれを実装する [図 1]。動画像に対する同様のアプローチを用いた実験が先行研究 [Baccouche 11] で行なわれており、以下のような利点が示されている。

1. 深層学習での特徴抽出により、汎用性がありコストの低い分類が可能
2. RNN による時間情報の処理を通して、高精度な予測が可能

本研究では 3 軸加速度データからの行動分類を同様のアプローチで行ない、その有効性を確かめる。

さらに、本研究では時間軸上で重なり合ったウィンドウを用いて二重のスタッガー処理を行ない、ネットワークへの入力とする。これによりデータの時間情報が保持され分類の予測精度が上昇することを期待する。

3. 実験

3.1 実験データ

実験には単一のセンサーから取得した 3 軸加速度データを用いた。1 つ目のデータセットが HASC が提供する“HASCC2011corpus”で、『静止(直立)』『歩行』『ジョギング』『スキップ』『階段を上る』『階段を下りる』の 6 行動を含み

連絡先: 横山 晃, 関西学院大学理工学部人間システム工学科,
Akira.Yokoyama@kwansei.ac.jp

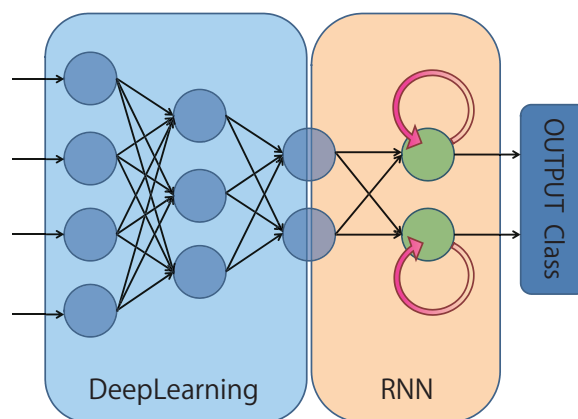


図 1: 複合型ニューラルネットワーク

86 名の被験者から周期 100[Hz] でサンプリングされた 3581 個の加速度時系列データから成る。2 つ目は Machine Learning Repository で公開されている日常行動 (Activities of Daily Living) データセットで、『歩く』『椅子に座る』『椅子から立ち上がる』『グラスに水を注ぐ』『グラスから水を飲む』『ベッドに寝る』『階段を上る』の 7 つの行動を含む。サンプリング周期は 32[Hz] で、16 名の被験者から取得した 716 個のデータを用いて分類を行なった。

3.2 実験条件

加速度データを識別器に入力する前に、前処理として [Matsushige 14] に倣い 3 軸加速度データから 11 次元の特徴量を抽出し新たな時系列データを生成する。重力ベクトル \mathbf{v}_g を算出し、時刻 t のセンサーデータから重力ベクトルを減算して得られた正規化運動加速度ベクトル $\mathbf{v}_n(t)$ を得る。この正規化加速度ベクトルから、ベクトル長 $F_1(t) \cdot$ 重力ベクトルとの内積 $F_2(t) \cdot$ 重力ベクトルとの外積のノルム $F_3(t)$ を算出する。

- $F_1(t) = \|\mathbf{v}_n(t)\|$
- $F_2(t) = \mathbf{v}_n(t) \cdot \mathbf{v}_g$
- $F_3(t) = \mathbf{v}_n(t) \times \mathbf{v}_g$

上記3種の特徴量について、外れ値の影響を受けにくい25%値・中央値・75%値の3種類の統計量を求め、計9種類の特徴量を得る。

さらに、周波数領域における特徴量としてエネルギーと周波数領域エントロピー (FDE) の2種類を使用する。エネルギー E_n は、各データにおける正規化加速度ベクトルのノルムに関してFFTを行ない、得られた周波数成分 F_i の絶対値の合計で表される。

$$E_n = \frac{1}{N} \sum_{i=2}^n |F_i|^2, \quad (1)$$

$$FDE = - \sum_{i=2}^n p(i) \log p(i), \quad p(i) = \frac{|F_i|^2}{\sum_{i=2}^n |F_i|^2}. \quad (2)$$

この11次元のデータの連続する10点を纏めて110次元の入力データとし、クラス分類を行なう。

さらに、深層学習ネットワークの入力には以下のような(特徴抽出に用いる時と別の)スタagger処理を施したものをを用いる。

1. 窓関数としてデータの連続した一定時間分を取り出す
2. 窓をずらして次の区間を取り出す
3. 1, 2を繰り返して得られたデータ系列を新たな時系列データとする。

ここで、深層学習に際しては1つの窓から取り出されたデータを一つのデータとして扱う。次元削減を行ないRNNに入力する際に、これらのデータを繋げて時系列データとする。このような処理を行なう目的は、RNNの入力となる時系列データが出来るだけオリジナルデータの時間情報を強く保持するようにするためである。

3.3 実験結果

HASC2011corpus に対する分類実験の結果を図2に表す。同データについて、先行研究[Matsushige 14]での3近傍NN分類器による分類精度が89.8%である。

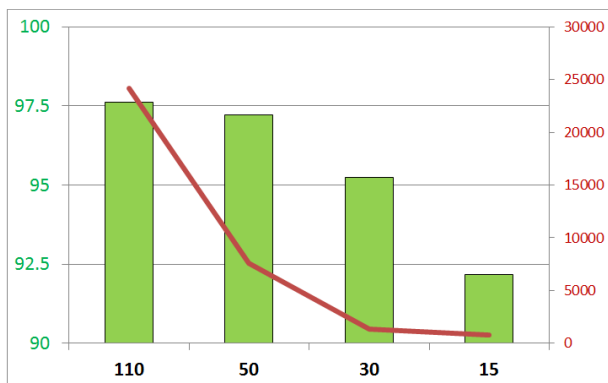


図2: HASC2011corpus に対する分類実験の結果。棒グラフが正答率 [%], 折れ線が学習時間 [sec.] を表す。

続いて Activities of Daily Living データセットに対する分類実験の結果を図3に示す。このデータセットにおける、

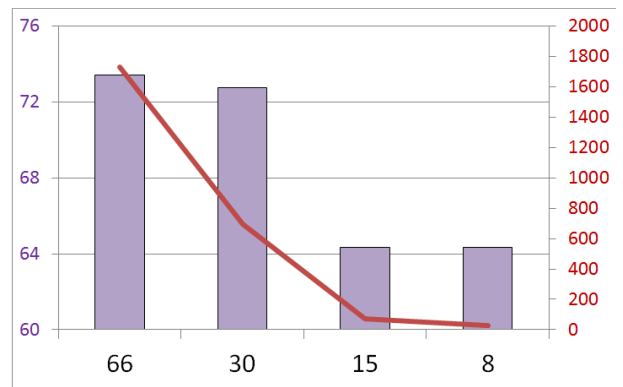


図3: ADL データセットに対する分類実験の結果。棒グラフが正答率 [%], 折れ線が学習時間 [sec.] を表す。

Dynamic Time Warping を用いて学習データとの距離による分類実験を行なった際の予測精度は54.3%である。

いずれの場合も、学習データを保持する従来手法と比較して高い水準の精度であり、加速度データに対するRNNによる時系列情報予測能力が高いことが窺える。入力次元数 D を110次元から50次元に削減した場合、正答率は97.21%と次元削減を行なわない場合と比較して0.4%程度の低下で、学習速度は3倍以上高速化された。但しこれ以上の次元削減を行なうと、予測精度が大きく低下し始めた。

また認識精度に対する二重のスタagger処理の寄与を評価するため、これを行なわずに HASC2011corpus のデータをネットワークに入力して分類を行なった。ここで、窓関数で取得した110次元データを深層学習で50次元に圧縮したものと、11次元データを5時刻毎にまとめた55次元データをそれぞれRNNに入力し比較を行った。結果、スタagger処理を行なわない場合は96.24%の分類精度であった。上記の結果と比較すると1.0%程度精度が低下する結果となった。

4. 議論

入力データのうち予測に有用な特徴量を残し不要な特徴量を削減すれば、精度を大きく落とすことなく学習が高速化できると考えられる。入力に冗長な情報量が多い場合は、提案手法がより有効になるとと思われる。例えば、複数のセンサデータを用いて分類を行なう場合などがこれに当たると考えられる。

参考文献

[Baccouche 11] Baccouche M. et al. (2011). Sequential deep learning for human action recognition. *Human Behavior Understanding*, 29-39.

[Matsushige 14] 松重龍之介・角所考・岡留剛. (2014) 半教師あり擬RVMによる加速度データからの行動推定. *人工知能学会全国大会論文集*, 28, pp.1-4

[Iketani 08] 池谷直紀・菊池匡晃・長健太・服部正典. (2008). 3軸加速度センサを用いた移動状況推定方式. *電子情報通信学会研究報告*, ユビキタス・センサネットワーク, 108, 138, 75-80.