

Multi-label CNN による部首検出

周楽陶* Brian Iwana* Zhu Anna* 内田 誠一* 田中 久美子**
(*九州大学, **東京大学)

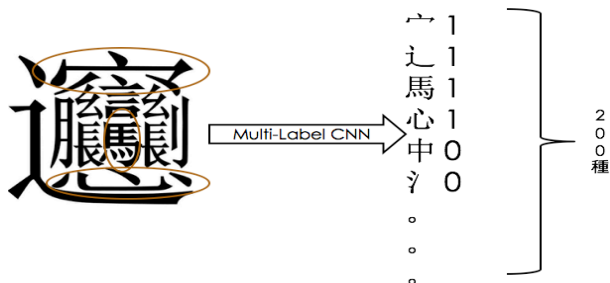


図 1 Multi-label CNN による部首検出

1 はじめに

近年, CNN (Convolutional Neural Network) が画像認識において精度がよいとされ, 注目されている. 2014 年に ILSVRC (ImageNet Large Scale Visual Recognition Challenge) のコンペティション, 1000 種類のクラス分類タスクにおいて[1]が, これまでの手法に大差をつけて 1 位となり, 画像認識分野において広く知られるようになった. 特に, 近年 CNN は Multi-label Image Classification にも高い認識率が得られている[2]. しかし, CNN による分類手法の原理と分類能力の限界は引き続き研究されている.

CNN の分類能力をテストするため, 本研究で着目しているのは, 漢字のように, 局所的に見ても単なる線構造でしかない文字を, 正しく部首単位に分離できる能力を CNN が持つかどうかを見極めることである. 本研究に関連し, 情景内文字検出・認識の研究が数多く試みられている. しかしながら, 同技術では, 一文字全体の形状を考慮して検出・認識することを前提としているため, 一文字がどのような部首から構成されているかを求めることはできない.

そこで, 発表者は CNN を利用して, 漢字を構成する部首検出の問題に取り込んでいる. 通常的情景内文字検出と部首検出では, 問題が異なるために, 従来とは異なる CNN の利用形態を考える必要がある. 具体的には, 通常形態では, CNN の出力層を構成するノードのうち, 「1」となるノードは一つ (例えば文字検出なら文字か非文字かのどちらかが 1 となり, 文字認識なら, 対象クラスの中の一つだけが 1 となる) であった. これに対し, 部首検出の場合, 「1」文字を構成する部首は複数であるため, 複数のノードが「1」(Multi-Label) を出力する必要がある. (図 1)

2 CNN による部首検出の意味

本研究は, 表面的には「漢字からの部首検出」という問題であるが, その本質は CNN の構造分離能力を漢字というわかりやすいパターンで検証することにある. 例えば, 「まだれ (广部)」や「かこい」のようにその内部に別のストロークがある場合に正しい検出ができるか. また, 「林」のように同じ部首が二つ以上存在する場合や, 同じ「木」でも, 「村」や「禁」のように形状が異なるとどうなるか, 「火へん」や「さんずい」「れっか」のような複数の連結成分からなる部首 (図 2) が検出できるかについても検証する必要がある.



図 2 部首のパターン例

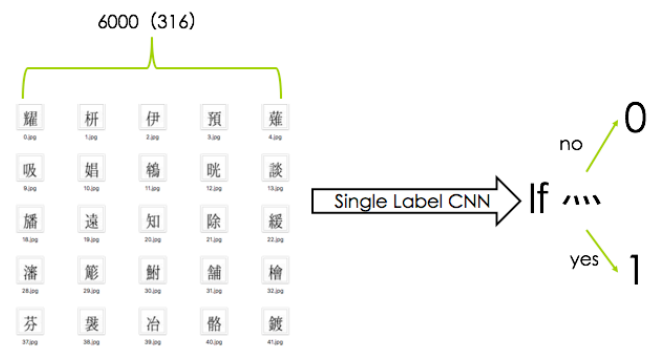


図 3 CNN による単数の部首の検出

表 1

部首	Train_data	Valuate_data	認識率
宀	4000	2000	98.86%
广	4000	2000	99%

3 実験

3.1 Single-label CNN による実験

発表者は CNN の部首検出能力をテストするため, まず Single-label CNN を用いて印刷体の漢字 (約 6000) で検証した. CNN の学習にはオープンソースの Deep Learning 用フレームワークの Caffe を利用した. 本研究では LeNet をベースにしたネットワークを利用した. 入力画像のサイズを 32x32 で日本語のビットマップフォントから学習を行った. 学習の流れを図 3 で表す.

3.2 実験結果

実験の結果 (表 1) として, CNN は「れっか」のような複数の連結成分からなる部首を検出可能で, また「まだれ」のようにその内部に別のストロークがある場合にも正しい検出が可能であることが確認できる.

4 まとめと今後の研究

今回の実験を通じて, CNN は単数の部首に対し, 良い分解能力を持っていることが確認された. 今後はより複雑の部首とパターンを Single-Label CNN で試みた後 Multi-Label CNN による部首検出を実験したいと考えている.

参考文献

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems, 2012.
- [2] Jiang Wang, Yi Yang. CNN-RNN: A Unified Framework for Multi-label Image. CVPR 2016