

学習済みネットワークを用いる重なりのある文字の認識性能の評価に関する研究

日大生産工(院) ○四井大悟 日大生産工 原一之

1. はじめに

ディープネットワークとは人間の神経構造を模倣して、アルゴリズムを発展させたものである。従来の機械学習のアルゴリズムでは人の手でデータから特徴量の抽出を行わなければいけないのに対し、ディープネットワークでは自動でデータにどのような特徴があるのか習得し、学習を行うことができる。そこで今回は、普段我々が目にする数字を学習させたネットワークに、重ね合わせた数字の認識性能に関する研究を行なった。

ディープネットワークに学習させた手書き数字とは違う、重ね合わせた手書き数字が認識できるかが本研究の目的となる。

2. 方法

まず初めに、使用する手書き数字の判別を行なった。判別結果の確信度が高く第一候補に判定される数字と、判別結果の確信度が低く第一候補に判定される数字を 10 枚ずつ用意した。重ね合わせた手書き文字の作成であるが、画像透過ツールを用いた。

透過した文字の重ね方は Fig.1 に示す。Fig.1 の左の画像は、判別結果の確信度が高く第一候補に判定された手書き数字を重ね合わせた図であり、判別結果の確信度が低く第一候補に判定された手書き数字を重ね合わせた図である。またこれ以降、判別結果の確信度が高く第一候補に判定された手書き数字を重ね合わせたものを手書き文字 1、判別結果の確信度が低く第一候補に判定された手書き数字を重ね合わせたものを手書き文字 2 とする。



Fig.1 重ね合わせた手書き数字

手書き文字 1、手書き文字 2 は計 90 枚ずつ用意し、重ね合わせた文字がすべて別の文字となるように作成した。また Fig.2 のように、重ね合わせ方は前後関係が分かるように作成を行った。



Fig.2 前後関係が分かる重ね合わせ方

60000 枚で学習を行い、10000 枚の手書き数字でテストを行った学習済みネットワークで判別を行った。判別は確信

度の高いもの 3 つが表示される。画像はすべてグレースケールモードで、 28×28 のピクセル数に統一して行った。

3. 結果

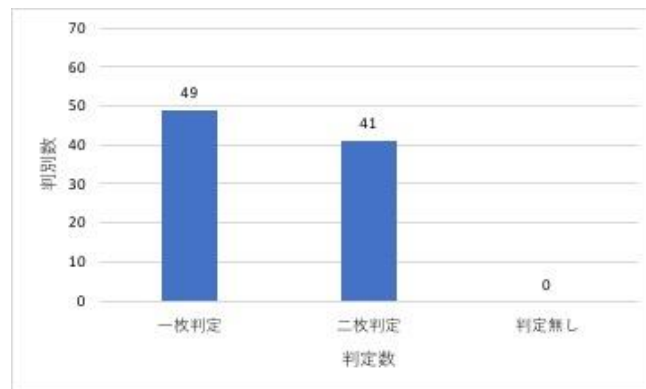


Fig.3 手書き数字 1 の判定数に対する判別数

Fig.3 は手書き文字 1 の判定数に対する判別数である。手書き文字 1 に含まれる文字を 1 枚判定、2 枚判定、判定無しの総数を表している。Fig.3 より重ね合わせた画像のうち 1 枚判定した数は 49 枚、2 枚判定した数は 41 枚、判定無しの数は 0 枚であった。

二枚判定できた画像の殆どが前後に関係なく第三候補までに判定されたが、4 に 8 を貼った画像は第一候補に 8 (確信度 99.5%)、第二候補に 4 (確信度 0.2%) と、僅かながらではあるが 4 が含まれていると判別したが、8 に 4 を貼った画像では 4 の判別はされず、第一候補に 8 が判別され確信度が 99.9% となった。

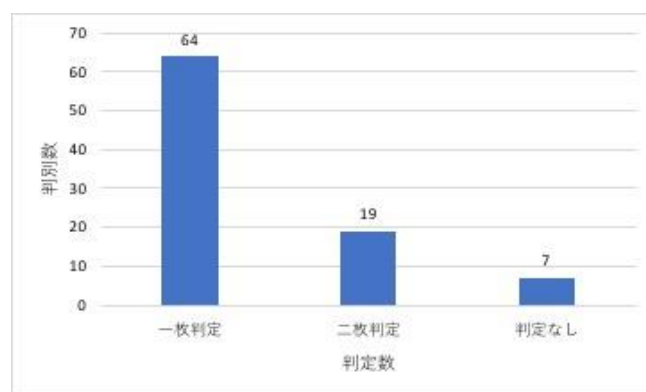


Fig.4 手書き数字 2 の判定数に対する判別数

Fig.4 は手書き文字 2 の判定数に対する判別数である。手書き文字 1 に含まれる文字を 1 枚判定、2 枚判定、判定無しの総数を表した。Fig.4 より重ね合わせた画像のうち 1 枚判定した数は 64 枚、2 枚判定した数は 19 枚、判定無しの数は 7 枚であった。

二枚判定できた画像 19 枚のうち 14 枚は前後に関係なく第三候補までに判定された。5 に 9 を貼った画像は第一候

補に 5(確信度 48.9%)と 1 文字のみ判別をしたが、9 に 5 を貼った画像では、第二候補に 5(27.5%)、第三候補に 9(22.3%)と判別することが出来た。

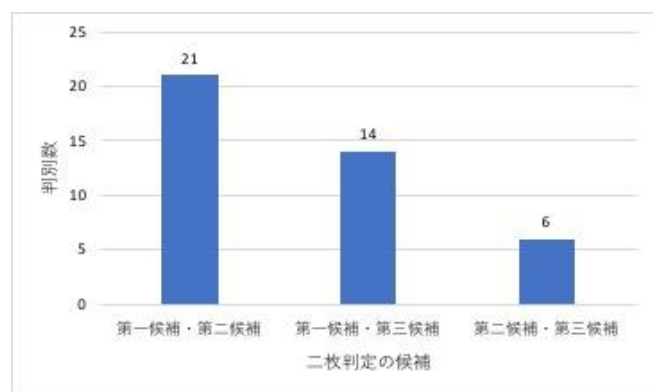


Fig. 5 手書き文字 1 に関する二枚とも判定が行えた画像の候補と判別数

Fig.5 は手書き文字 1 のうち、判定が二枚とも行えた画像の各候補の数を表したグラフである。手書き文字 1 に含まれた数字が、第一候補・第二候補と判別された枚数は 21 枚であり、そのうち前後関係に関係無く判別されたものには 8 が含まれていることが多いことがわかった。また、先行研究では 1 を重ね合わせた画像から 1 の判別が行えなかったが、今回のデータでは 1 に 4 を貼った画像の第一候補に 4(確信度 99.1%)、第二候補に 1(0.5%)と僅かながらであるが、1 を判別している事が結果から読み取れた。また 4 に 1 を貼った画像では、第一候補に 4(確信度 99.8%)、第三候補に 1(0.1%)と、1 の確信度が下がったが判別出来た。

Fig.5 の第一候補・第三候補の結果から、前後関係とは関係なく、判別される数字は変化が少ないことがわかった。また第二候補には 8 の判別が多いことがわかった。第二候補・第三候補の結果から、第一候補に 8 と判別されることが多いことがわかった。

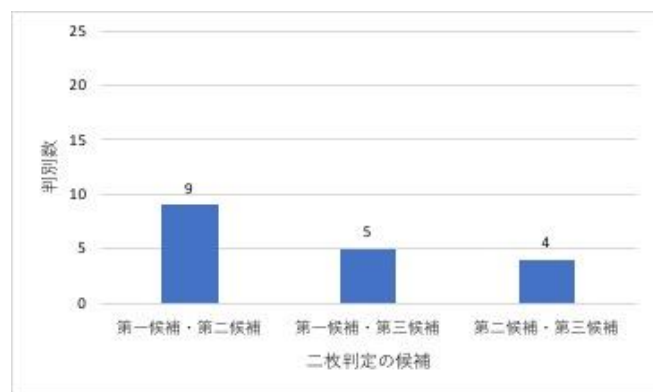


Fig.6 手書き文字 2 に関する二枚とも判定が行えた画像の候補と判別数

Fig.6 は手書き文字 2 のうち、判定が二枚とも行えた画像の各候補の数を表したグラフである。手書き文字 2 に含まれた数字が、第一候補・第二候補と判別された枚数は 9 枚であり、そのうち前後関係に関係無く判別されたものには 8 が含まれていることが多いことがわかった。

Fig.6 の第一候補・第三候補の結果から、前後関係に関係し、下地にした数字の確信度が残る傾向が見られた。第二候補・第三候補の結果から、4 に 6 を貼った画像は第二候補に 4(確信度 34.0%)、第三候補に 6(確信度 10.5%)と判定することができ、6 に 4 を貼った画像は第二候補に 6(確信度 47.6%)、第三候補に 4(確信度 2.5%)とどちらも下地になった数字の確信度が高くなった。

4. 考察

今回使用した判別結果の確信度が高く第一候補に判定された手書き数字は、確信度 90%を超える数字を使用しており、確信度が低く第一候補に判定された手書き数字は、確信度が 40%から 60%程度の数字を使用した。しかし、8 はどの数字も 80%を超えてしまうため、比較的低い確信度の数字を用いて行った。

Fig.3 では、一枚判定は 54%で二枚判定は 46%と約半数二枚判定できており、判定なしは出てこなかった。また Fig.4 では、一枚判定が 70%以上であり、二枚判定は 21%、判定なしは 8%と Fig.3 と比べて精度はよくなかった。これは、元となる文字の確信度が高いものを重ね合わせた際は二枚判別することができ、確信度が低いものを重ね合わせた際は二枚判別するのが難しくなると考えられた。その為、手書き文字 2 では 8 が含まれる画像に対し二枚判定することが多かったと考えられた。

Fig.5 では重ね合わせた二つの文字が第一候補・第二候補として判別されたのは 21 枚、第一候補・第三候補として判別されたのは 14 枚、第二候補・第三候補として判別されたのは 6 枚となった。二枚判別した結果で特に 0 と 8 が多かった理由としては、一文字の確信度が両方とも 100%であったことが考えられる。また、1 を判別できた理由としては、今回手書き文字 1 で使用した 1 は手書き文字 2 に比べてサイズの大きいものを使用したため、1 としての情報が多く残ったのではないかと考えられた。また Fig.6 も同様、第一候補・第二候補としては判別されたのは 9 枚、第一候補・第三候補として判別されたのは 5 枚、第二候補・第三候補として判別されたのは 4 枚と Fig.5 に比べて判別数が少ない結果が得られた。これは、Fig.4 同様、画像一枚一枚の確信度が小さいためだと考えられた。

6. 参考文献

- Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", *Nature*, vol. 521, pp. 436--444 (2015)
- MNIST:
<http://yann.lecun.com/exdb/mnist/>
- caffe
<http://caffe.berkeleyvision.org/>
- AI/ディープラーニング/機械学習
<https://www.souya.biz/>