

手書き文字を用いた認識ツール

文京学院大学女子高等学校 2年 桐山夢叶 守玲奈



要旨 皆さんは、提出物を名無しで提出してしまっただけではありませんか？一度はあるかもしれません。そんなとき誰が書いたか先生方は困っていると思います。解決策として、**人の文字を見ただけですぐに誰が書いた文字かを判断できる**ように、機械学習を用いて人の文字を判断できるか実験しました。最終的には、個人を判別することが81%できました。



目的 同じ人が書いた文字を NNC (ニューラルネットワーク) で判断し、どのくらい識別することができるかを実験した。データ数を10倍、20倍、30倍、40倍、50倍と増やすことで手書き文字が識別できる確率が上昇していくのではないかと考えた。このことから、誰が書いた文字か判断できるアプリ開発を作ることではないだろうか。

研究方法

① データ収集

「**Birthday**」の文字を高校生11人に1人50回ずつ手書きで紙に書いてもらい(図1)合計550文字のデータを集めた。

Birthday	Birthday
Birthday	Birthday
Birthday	Birthday
Birthday	Birthday

図1.書き文字

② スキャニング

書いてもらった紙を一枚ずつコピー機でスキャンしPCでPDF化した。

増やす倍率	学習データ数	学習データ数	検証データ数
10倍	110	88	22
20倍	220	176	44
30倍	330	264	66
40倍	440	352	88
50倍	550	440	110

図2.データ数

③ 切り取り

PCでMicrosoft Paintを使用して一文字ずつ200mm×200mmの大きさに一枚サイズにし、BMPファイルとして保存した。

④ データセットの作成

BMPのデータを10倍から50倍の五つの数のデータに分けた後、NNCで画像分析するために学習データと検証データに分類分けをし、NNCのデータセットアップロードツールを使ってディープラーニングを行う準備をした。(図2)

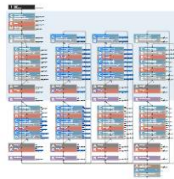


図3.ニューラルネットワーク

⑤ 機械学習

NNCを用いて、ニューラルネットワークを構築し、識別させた。(図3)

結果

NNCを使って、110, 220, 330, 440, 550の文字数を順番に識別の実験をした結果、10倍<20倍<30倍<40倍<50倍と識別率が上り、高い判別精度を得ることができた。

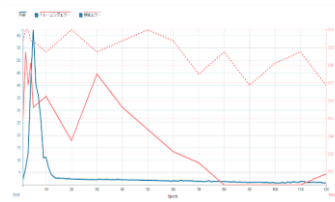


図4.10倍33%のグラフ

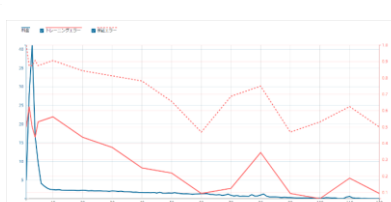


図5.20倍59%のグラフ



図6.30倍71%のグラフ

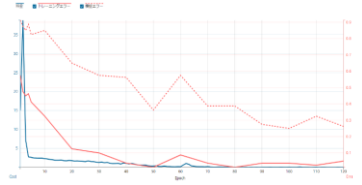


図7.40倍76%のグラフ



図8.50倍81%のグラフ

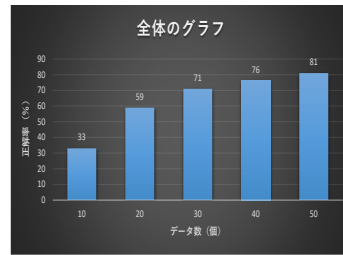


図9.33%<59%<71%<76%<81%

初めは、データ量を少なくし10倍から実験した。110個のデータを使ったときは33%正解することができたが、正解率が低かったためデータ数を倍にして実験を行なった。20倍の220個のデータで実験した際には、59%に上昇した。その後、30倍、40倍、50倍とデータ量を増やすことにより正解率が上がり81%まで上昇することができた。データ量が少なければ比較できるデータも少ないので正確と判断がしづらいものの、データ数を増やすことによってより正確に見分けることができるようになった。

結論

AIを活用して手書きで書いた文字を一人一人認識させることによって、中高生が書いた文字を**81%**識別することが可能になった。より、正確に人の文字を判断するためにはより多くデータベースを集めることが重要となってくる。そのため、高い精度を得るにはデータ量を増やせば増やすだけ期待精度が向上し正確に文字の識別が可能になると考えられる。このことから、データ量が少なく認識できたとしても精度が低くなるのが分かる。(図10)

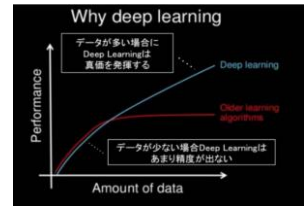


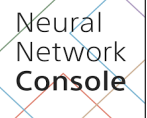
図10.データ量 <https://support.dl.sony.com/>

今後の展望

この実験では、11人の生徒が対象だったためもっと人数を増やして実験したいと思う。この実験を応用して、中高生が書いた文字を先生が見たときに、どの生徒が書いた文字かをすぐに判断できるアプリが作りたい。

使用したソフトウェア

Neural Network Console (Sony) <https://dl.sony.com/ja/>



参照

- {1}<https://cir.nii.ac.jp/crid/1110845735109350784> 統計的性質に基づく文字の高精度認識に関する研究
- {2}<https://core.ac.uk/download/pdf/268257619.pdf> 手書き文字の感性印象と筆跡から推測されたおよび 実際の書き手のパーソナリティ特性との関連について
- {3}<file:///C:/Users/yumek/Downloads/ot3329.pdf> 手書き文字の機械認識における特徴抽出と識別関数の研究
- {4}file:///C:/Users/yumek/Downloads/japanesecultureandai_uchida_2019_1111.pdf 文字認識研究の 過去・現在・未来