

K-007

## 機械学習を用いた手書き文字の正誤判定システムの開発

## Development of Grading System of the Handwritten Character Using the Machine Learning

宮下 知也<sup>†</sup>  
Tomoya Miyashita縄手 雅彦<sup>†</sup>  
Masahiko Nawate伊藤 史人<sup>†</sup>  
Fumihito Ito多々納 俊治<sup>‡</sup>  
Toshiharu Tadano

## 1. 背景

文字の読み書きに困難のある児童の発見のために簡易的なスクリーニング検査が求められており、書字の検査においては「小学生のための読み書きスクリーニング検査 (STRAW)」[1]などがある。しかし、検査で書いた文字の正誤判定は検査者が自分で行わなければならないため、不便さが残っている。そこで、検査者の負担を減らすために手書き文字の正誤判定を自動で行うシステムの開発についての検討を行う。

一般的な文字認識には、時系列の筆点座標列から手書き文字認識をするオンライン文字認識と、画像から文字認識をするオフライン文字認識の 2 つがある。それぞれの手法には長所と短所があり、オンライン文字認識は、筆順の続けや崩しに頑健であるが、筆順誤りや重ね書きに不安定であり、オフライン文字認識は、筆順誤りや重ね書きに影響されないが、続けや崩しに不安定といった点がある[2]。

また、今までの文字認識は文字の一部が間違った文字を入力しても、機械の中で補完を行って正しい結果を返すというものであったため、文字の正誤判定が正しく行えなかった。しかし、最近になり手書き文字の自動正否判定ができるツールとして「手書き電子ドリル」[3]がリリースされた。このことにより、書字検査を手軽にできるようになったが、文字の書き順で正否判定を行っているため、書いた文字の書き順が不正確で字形が正しいといった場合に正しく正否判定ができないという問題がある。

## 2. 目的

本研究では、機械学習を用いて筆順に影響されない手書き文字の自動正誤判定システムを構築する手法についての検討を行うことを目的とする。

## 3. システムの構成

今回用いたシステムの構成の概略図を図 1 に示す。今回用いたシステムは、文字認識モデルと詳細分析モデル、正誤判定モデルの 3 つのモデルで構成されている。文字認識モデルは、入力された文字画像がどのような漢字であるかを大まかに分類する役割、詳細分析モデルは、入力された

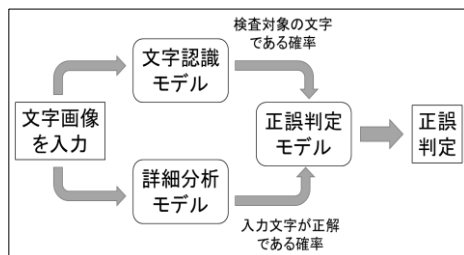


図 1 システムの概略図

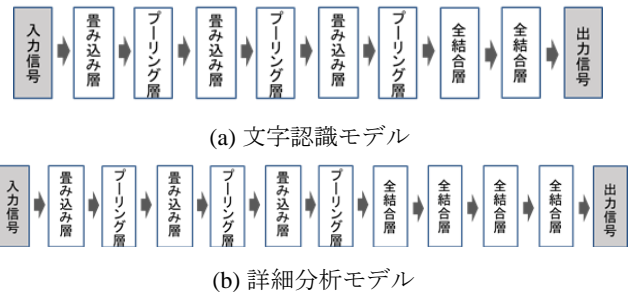


図 2 ネットワークの構成図

表 1 間違い漢字画像の例

漢字の違い内容	漢字画像
1・2画抜け	会 全 公
1・2画追加	会 会 会
鏡文字・その他文字	会 会 字

文字画像が正しい文字かどうかを細部まで判定する役割、正誤判定モデルは、最終的な正誤判定をする役割がある。

文字の正誤判定の流れとしては、まず始めに、文字画像を文字認識モデルと詳細分析モデルに入力し、文字認識モデルから得られた検査対象の文字である確率と詳細分析モデルから得られた入力文字画像が正解である確率を正誤判定モデルに入力し、最終的な正誤判定を行うといったものである。

文字画像の処理には、画像処理ライブラリである「OpenCV」[4]を用いた。文字認識モデルと詳細分析モデルは畳み込みニューラルネットワーク (CNN) で構成を行い、モデルの構築にはディープラーニングライブラリである「Chainer」[5]を用いた。また、正誤判定モデルはサポートベクターマシン (SVM) で構成を行い、モデルの構築には機械学習ライブラリである「scikit-learn」[6]を用いた。

## 3.1 文字認識モデル

文字認識モデルは、2012 年の大規模画像認識コンテスト ImageNet Large Scale Visual Recognition Challenge (ILSVRC) で優勝したアルゴリズムの AlexNet[7]を参考に構築を行った。文字認識モデルのネットワーク構成は、畳み込み層とプーリング層をそれぞれ 3 層、全結合層を 2 層、入力画像は 64 × 64 ピクセルのサイズにリサイズをし、2 値化処理を行った。ネットワークの構成図を図 2(a) に示す。また、モデルの学習用の手書き文字画像には、産業技術総合研究所の

<sup>†</sup> 島根大学 大学院総合理工学研究科, Graduate School of Science and Engineering, Shimane University

<sup>‡</sup> 出雲医療看護専門学校, Izumo Medical Care Nursing School

ETL8B の手書き文字データベース[8]の漢字 881 字のデータを用いた。

### 3.2 詳細分析モデル

詳細分析モデルも AlexNet を参考に構築を行った。詳細分析モデルのネットワーク構成は、畳み込み層とプーリング層をそれぞれ 3 層、全結合層を 4 層、入力画像は文字認識モデルと同様に、 $64 \times 64$  ピクセルのサイズにリサイズし、2 値化処理を行った。ネットワークの構成図を図 2(b)に示す。モデルの学習には、正解の手書き漢字画像と不正解の手書き漢字画像を作成し、それらを用いた。

また、今回は詳細分析モデルの構成の違いによる認識精度の違いについて検証するため、入力漢字画像の状態を 2 つのクラス (正解漢字, 不正解漢字) に分類するモデルと、入力漢字画像の状態を 4 つのクラス (正解漢字, 1・2 画抜け漢字, 1・2 画追加漢字, 鏡文字・その他文字) に分類するモデルの 2 つのモデルを作成した。学習に用いた間違い漢字画像の一例を表 1 に示す。

### 3.3 正誤判定モデル

文字認識モデルと詳細分析モデルを作成した後、正誤判定モデルの作成を行った。正誤判定モデルの作成の流れは、まず始めに、正解漢字画像と不正解漢字画像を文字認識モデルと詳細分析モデルにそれぞれ入力し、次に、文字認識モデルから得られた検査対象の文字である確率と、詳細分析モデルから得られた入力文字画像が正解である確率を正誤判定モデルに学習させるといったものである。

### 4. 作成したモデルの精度検証

今回の研究では、「会」という漢字の正誤判定を行うシステムを作成し、実際に漢字を認識させ、精度の検証を行った。まずは、詳細分析モデルを 2 つのクラスに分類するモデルにしてシステムを構築した場合について検証を行った。表 2 に正解漢字の判定結果、表 3 に不正解漢字の判定結果を示す。表 2 と表 3 より、詳細分析モデルを 2 つのクラスに分類するモデルとした場合、正解漢字に関しては、全て「正解」の正しい判定を行えているが、不正解漢字に関しては、5 枚中 2 枚「正解」の誤判定を出している。

次に、詳細分析モデルを 4 つのクラスに分類するモデルにしてシステムを構築した場合について検証を行った。表 4 に正解漢字の判定結果、表 5 に不正解漢字の判定結果を示す。なお、表 5 は精度の検証をより細かく行うために 10 枚の漢字画像を認識させた。表 4 と表 5 より、詳細分析モデルを 4 つのクラスに分類するモデルとした場合、正解漢字に関しては、全て「正解」の正しく判定を出しているが、不正解漢字に関しては、10 枚中 1 枚「正解」の誤判定を出している。

検証結果より、正解漢字画像の認識では、詳細分析モデルが 2 つのクラスに分類するモデルの場合と 4 つのクラスに分類するモデルの場合で認識精度の差は見られなかった。不正解漢字画像の認識では、詳細分析モデルが 2 つのクラスに分類するモデルの場合に正しく判定できなかった漢字画像であっても、詳細分析モデルを 4 つのクラスに分類するモデルに変更を行うと正しく判定できていた。さらに、4 つのクラスに分類するモデルで追加検証を行った結果、「正解」の誤判定を出したものは 10 枚中 1 枚であった。こ

表 2 正解画像の判定結果(2 クラス分類)

画像	会	会	会	会	会
判定結果	正解	正解	正解	正解	正解

表 3 不正解画像の判定結果(2 クラス分類)

画像	云	会	会	会	習
判定結果	不正解	正解	不正解	正解	不正解

表 4 正解画像の判定結果(4 クラス分類)

画像	会	会	会	会	会
判定結果	正解	正解	正解	正解	正解

表 5 不正解画像の判定結果(4 クラス分類)

画像	云	会	会	会	習
判定結果	不正解	不正解	不正解	不正解	不正解
画像	会	会	会	会	会
判定結果	不正解	不正解	不正解	不正解	正解

のことより、間違い漢字の間違いパターンをある程度分類することで、判定精度が向上するのではないかと考えられる。

### 5. 今後の課題

今回は、「会」という漢字の正誤判定を行うことを目標として正誤判定システムを作成したが、今後は他の漢字に対しても正誤判定ができるようにシステムの拡張を行い、正誤判定の精度を向上させていく必要がある。正誤判定の精度を向上させていくにあたって、機械学習させる正解文字と不正解文字のデータ数を増やしていく必要がある。また、今回は精度検証を限られた枚数の漢字画像で行ったが、今後は検証する漢字画像の枚数を増やして精度の検証を行う必要がある。

#### 参考文献

- [1] 宇野 彰, 春原 則子, 金子 真人, Taeko N Wydell, “小学生の読み書きスクリーニング検査-発達性読み書き障害(発達性 dyslexia)検出のために”, インテルナ出版, (2006).
- [2] 朱 碧欄, 中川 正樹, “オンライン手書き文字認識の最新動向”, 電子情報通信学会誌, Vol.95, No.4 (2012).
- [3] 富士通株式会社, 手書き電子ドリル, <http://www.fujitsu.com/jp/solutions/industry/education/school/learning/handsstudy/>
- [4] Itseez, OpenCV, <http://opencv.org/>
- [5] Preferred Networks, Chainer, <https://chainer.org/>
- [6] David Coumapeau, scikit-learn, <http://scikit-learn.org/>
- [7] A. Krizhevsky, I.Sutskever, G.E.Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”, NIPS, (2012)
- [8] 産業技術総合研究所, Japanese Technical Committee for Optical Character Recognition, ETL 文字データベース, (1973-1984).