

深層学習を用いた手書き文字のフォーマルとカジュアルの分類

Classification of Formal and Casual Handwritten Characters by means of Deep Learning

突廻恵太^{1*} 上野未貴¹

Keita Tsukimawashi¹ Miki Ueno¹

¹ 大阪工業大学

¹ Osaka Institute of Technology

Abstract: 近年のデジタル化の進行により、手書きで文字を書く機会が減少している。履歴書を手書きで書くようなフォーマルな場面に適した字と、メモのような見た目を意識せずにカジュアルに書く字の書き分けは、日頃から字の形などを意識して書いていないと難しい。本研究では、CNN を用いて人が書いた字をフォーマルな字かカジュアルな字かの識別を行う。識別率を提示することで、場面に応じた字であるかを判断し、人が書き直すことで求める字体に近づけることを目的とする。

1 はじめに

近年のデジタル化の進行により、手書きで字を書く機会が減少している。手書きする機会が減少する一方で、手書きに対する特別感は一層増える。手紙や履歴書を手書きする際には、丁寧に正しく整った字（フォーマルな字）を意識して書く。しかし、日頃はメモ書きのような粗雑で崩した字（カジュアルな字）で書くことが多い。手書き離れが進む現代では、状況に応じた字を書き分けるのは難しくなっている。本研究は、深層学習を用いたフォーマルな字かカジュアルな字かの識別を行い、識別結果を提示するシステムを作成する。識別率を提示することで、場面に応じた字であるかを判断し、人が書き直すことで求める字体に近づけることを本研究の目的とする。

2 関連研究

関連研究として、CNN (Convolutional Neural Network) を用いた文字の識別等に関する研究は多数存在する。二ノ宮ら [1] は書道作品における難読文字の識字に CNN を用いており、村上ら [2] は CNN を用いて手書き文字から性格や感情を推測している。また、字の上達を支援するシステムに関する研究も多数存在する。久保田ら [3] は手書き文字とお手本文字を融合させ、自動美化された自身の字を意識することで字の上達を狙う練習支援システムを提案している。本研究は手書き文

字の場面に応じた字形の書き分けに着目し、手書き文字のフォーマルとカジュアルの分類をする。また、作成した分類器を用いて手書き文字のフォーマルさとカジュアルさを提示するシステムを作成する。

3 データセット構築

3.1 フォーマルな字とカジュアルな字の定義

本実験における、フォーマルな字とカジュアルな字の定義を以下に示す。フォーマルな字は教科書体を基準として、丁寧に書かれ字の形が整った文字とする。カジュアルな字はまるもじ体を基準として、荒々しく雑に書かれ字の形が崩れた文字とする。

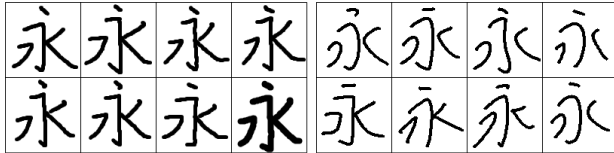
3.2 データ収集方法

本実験では、漢字の「永」をフォーマル/カジュアル分類の題材とする。表 1, 表 2 に本実験のデータセットの内容および設定を示す。本実験の CNN による分類には「永」の手書きデータ 100 枚を使用し、フォーマルとカジュアルでそれぞれ 50 枚ずつ用意した。100 枚の手書き画像データは、筆者がタッチ対応ディスプレイ搭載の PC にスタイラスペンを用いて筆記し、表 2 に示す設定で用意した。図 1 にフォーマルとカジュアルのデータ例をそれぞれ 8 枚ずつ示す。

*連絡先：大阪工業大学工学部電子情報通信工学科
〒535-8585 大阪府大阪市旭区大宮 5 丁目 16 番 1 号
E-mail: e1617061@st.oit.ac.jp

フォーマル	50 枚
カジュアル	50 枚

サイズ	100px×100px
線の幅	3 ~ 8 px
線の色	黒
背景	白



[1] フォーマル

[2] カジュアル

図 1: データセット例

4 提案手法

本研究では、手書きデータのフォーマルとカジュアルの2クラス分類にCNNを用いる。CNNによって作成した分類モデルを使用して、手書き入力データをリアルタイムで分類を行うWebアプリケーションの実装を提案する。

4.1 CNNによる学習モデル作成

1. 文字画像を 100×100 px から 28×28 px に圧縮。
2. 画像データを Numpy 配列に変換し、正規化する。
3. CNN のモデルを作成し、モデルをコンパイルする。
4. モデルにデータを入力し、モデルの訓練をする。
5. 訓練したモデルの評価をし、パラメータ調整を繰り返すことでモデルの改善をする。
6. 作成したモデル構造と学習済みの重みを保存する。

4.2 Webアプリケーションの実装

図 2 に Web アプリケーションの実装画面例を示す。GUI ライブラリの Gradio[4] を用いて、Web アプリケーションの実装をする。

1. 学習済みモデルを Gradio で読み込み、手書き入力フォームと分類結果の表示領域を作成する。
2. アプリケーションを開きスタイラスペンなどを用いて、手書き入力を与え識別する。
3. 識別率が表示され、書いた字の最適な場面を知る。
4. 求める識別率が出力されるように書き直し繰り返し練習する。



図 2: Gradio 実行画面例

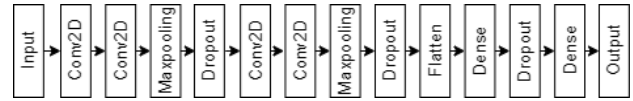


図 3: CNN の構造

表 3: パラメータと識別率

epoch	batch	識別率	損失率
20	10	0.95	0.18

5 実験

5.1 数値実験: CNN による分類

データセットの8割を訓練用データとしてCNNで学習してモデルを作成し、2割をテストデータとして評価した。また、CNNで学習したモデルは10分割交差検証によって最適なモデルを選択した。図3に本実験に用いたCNNの構造の内容を示す。本実験のCNNは入力層と出力層に挟まれた中間層が12層からなっており、Convolution層を4層、Pooling層を2層含む構造にした。表3にパラメータとテスト時の識別率を、図4、図5に学習時と検証時の識別率の推移と損失率の推移を示す。epoch数やbatchサイズなどのパラメータは、学習不足または過学習にならない値を選択した。識別率はepoch数が15で訓練用データと検証用データともに100%に達し、損失率はepoch数が20付近で0%に限りなく近く推移した。

5.2 ユーザ実験: Webアプリによる練習

研究室の2人に依頼し、本実験で作成したWebアプリケーションを用いて、フォーマルな字とカジュアルな字の練習過程を取得した。図6、図7に協力者2人の練習過程を示す。初めに書いた字がフォーマルな字と判定されれば、カジュアルな字になるように練習し、初めがカジュアルな字であればフォーマルな字を目指して練習するよう指示した。提示される識別率の割合を確認しながら練習をすることにより、フォーマルな字とカジュアルな字の練習をすることができた。

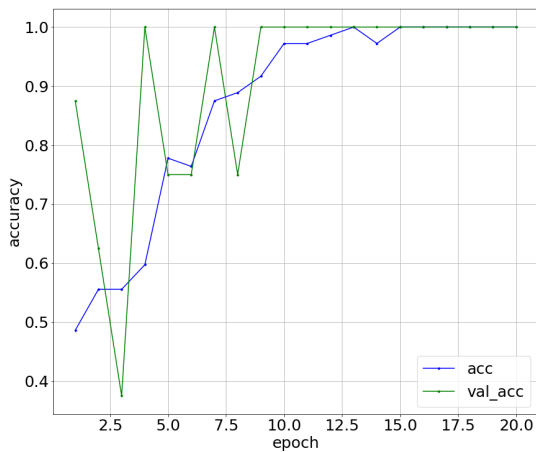


図 4: 識別率の推移

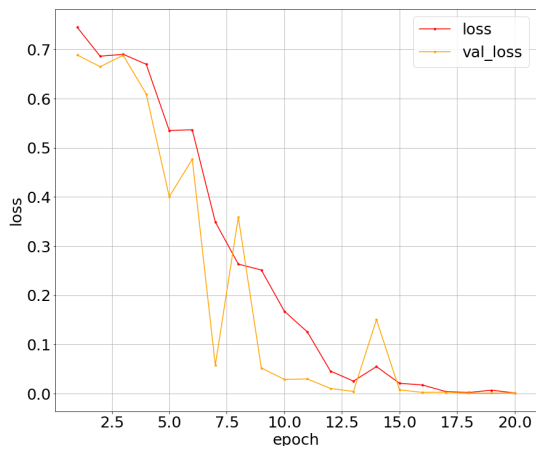


図 5: 損失率の推移

6 考察

6.1 CNN による分類

本実験より、筆者の字に対するフォーマルとカジュアルの識別率は高いと分かった。筆者以外の字に対する識別を調べるため、研究室内の 5 人の字で識別した結果、筆者の評価基準では全て正しく識別された。しかし、人によって評価のずれが生じるため、多数の人の評価の平均を求める必要があると考える。また、「永」以外の異なる字における識別結果を調べた。図 8 に「永」以外の異なる字による識別に用いた文字画像を、表 4 にその異なる字での識別結果を示す。「水、水、大、阪」の 4 種の漢字で、それぞれフォーマルとカジュアルの筆者による手書き文字画像を用意して識別をした結果、いずれも正しく識別された。「永」に形が似た「水」や「水」だけでなく、「阪」のような構造の異なる字においても高い識別率で識別が可能であった。これらより、本実験で用意したデータセットのみでも、筆者の字に対するフォーマ

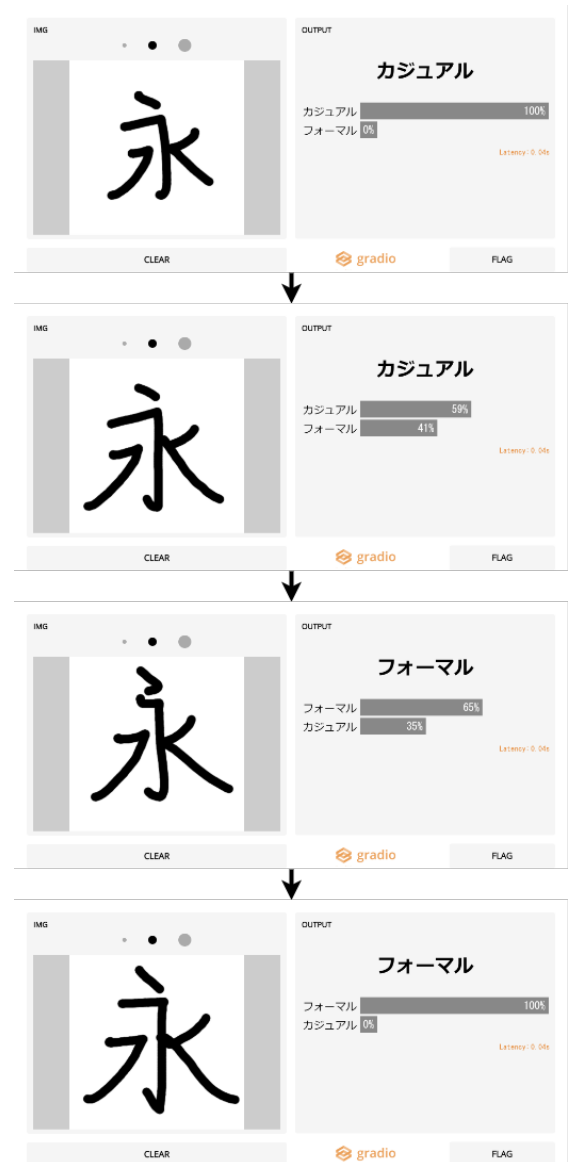


図 6: Gradio での練習過程 [1]

ルとカジュアルの識別は高水準で可能であると考えられる。

表 4 に示す識別率は整数型で求め、小数点以下は切り捨てにしている。そのため識別率の合計が 100 % になっていない。

6.2 練習過程における字形バランスの崩れ

本研究では、字形分類の識別結果の提示に従って求める字形の練習をした。しかし、図 7 に示す練習過程では、カジュアルに寄せようと過度に書き崩して、文字のバランスが失われる問題が見られた。これは提示される識別率の値に注視するあまり、文字のバランスを無視して書いてしまうという事例であると考えられる。したがって、文字の筆画の特徴を損なわず、文字のバランスを保つように支援する機能を追加する必要がある。



図 7: Gradio での練習過程 [2]

表 4: 異なる字による識別結果

	フォーマル率	カジュアル率
「氷」フォーマル①	99 %	0 %
「氷」カジュアル②	6 %	93 %
「水」フォーマル③	99 %	0 %
「水」カジュアル④	0 %	99 %
「大」フォーマル⑤	82 %	17 %
「大」カジュアル⑥	0 %	99 %
「阪」フォーマル⑦	98 %	1 %
「阪」カジュアル⑧	3 %	96 %



図 8: 異なる字による識別

7 まとめ

本研究では, CNN による手書き文字のフォーマルとカジュアルの分類をした. また, CNN で学習したモデルを用いて手書き文字のフォーマルとカジュアルの識別率を提示するシステムを作成した. 作成したシステムで, 求める字体を練習することによりフォーマルな字とカジュアルな字の書き分けを支援することを本研究の目的とした. 今後の課題点としては, 実験に用いなかった漢字に対しても分類可能にする点や, 練習過程における字形バランスの維持を支援する機能を追加する点が挙げられる. また, フォーマルとカジュアルの書き分け支援として, 書き手の字の特徴を残しつつ求める字体に近づけた字を生成するシステムの作成を検討する.

参考文献

- [1] 二ノ宮梢平, 中山雅紀, 宮澤篤, 藤代一成. 深層学習を用いた書道作品における書体翻訳. 第 82 回全国大会講演論文集, pp. 135–136. 情報処理学会, 2020.
- [2] 村上拓麻, 戴瑩. CNN を用いた手書き文字画像から人の性格を推測するシステム. 第 82 回全国大会講演論文集, pp. 145–146. 情報処理学会, 2020.
- [3] 久保田夏美, 新納真次郎, 中村聡史, 鈴木正明. Mojivator: 手書き文字の自動融合により書きたくなる練習支援システム. 第 24 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2016). 1-A07. 日本ソフトウェア科学会, 2016.
- [4] Abubakar Abid, Ali Abdalla, Ali Abid, Dawood Khan, Abdulrahman Alfozan, and James Zou. Gradio: Hassle-free sharing and testing of ml models in the wild. In *ICML Workshop on Human in the Loop Learning (HILL 2019)*, 2019.