

# ディープラーニングによるくずし字認識組み込みシステムの開発

## Development of Embedded System for Recognizing *Kuzushiji* by Deep Learning

竹内正広<sup>\*1</sup> 早坂太一<sup>\*1</sup> 大野亙<sup>\*1</sup> 加藤弓枝<sup>\*2</sup> 山本和明<sup>\*3</sup> 石間衛<sup>\*4</sup> 石川徹也<sup>\*4</sup>  
Masahiro Takeuchi, Taichi Hayasaka, Wataru Ohone, Yumie Kato, Kazuaki Yamamoto, Mamoru Ishima, and Tetsuya Ishikawa

<sup>\*1</sup> 豊田工業高等専門学校 <sup>\*2</sup> 鶴見大学 <sup>\*3</sup> 国文学研究資料館 <sup>\*4</sup> TRC-ADEAC 株式会社  
National Institute of Technology, Toyota College, Tsurumi University, National Institute of Japanese Literature, and TRC-ADEAC, Inc.

**Abstract:** Effective utilization of “Pre-modern Japanese book database” constructed by the project supervised by Center for Collaborative Research on Pre-Modern Texts, National Institute of Japanese Literature, will push forward the development of the inter-field study. It may become an obstacle for the researchers with a little knowledge of classical literature, however, because historical Japanese texts have been written by *Kuzushiji* (*Hentaigana* and cursive kanji). In this article, we report an attempt of recognizing *Kuzushiji* by deep learning. Using Raspberry Pi which is the low-cost small single board computer, we developed the embedded system with enough recognizing rate and speed. In addition, manifold images of *Kuzushiji* written in a text image of the spread can be extracted. Our system will be effective for school children or elderly people not possessing mobile devices who want to come in touch with *Kuzushiji*.

### 1. はじめに

国文学研究資料館により平成 26 年度より開始された「日本語の歴史的典籍の国際共同研究ネットワーク構築計画」[1]では、研究基盤整備として約 30 万点の歴史的典籍を画像データ化し、既存の書誌情報データと統合させたデータベースの構築を行っている。これを有効活用できれば、人文科学のみならず、自然科学系分野を融合させた研究の展開も期待される。しかしながら多くの研究者にとっては、それらに書かれている文字が「くずし字」であることが障壁となる。

本研究では、ディープラーニングを用いたくずし字の自動翻刻の試みを行っている。本稿では、既に WWW アプリケーションとして実装されている学習モデル[2]を応用し、小型で比較的安価なシングルボードコンピュータ Raspberry Pi を用いて開発されたくずし字認識組み込みシステムについて報告する。

### 2. システム概要

図 1 に本研究で開発したシステムの概要を示す。Raspberry Pi 用カメラモジュール Picamera から対象の古典籍画像を撮影する。本研究で用いるディープラーニングにより学習されたモデルは、くずし字1文字に対して認識を行うため、Python3 および

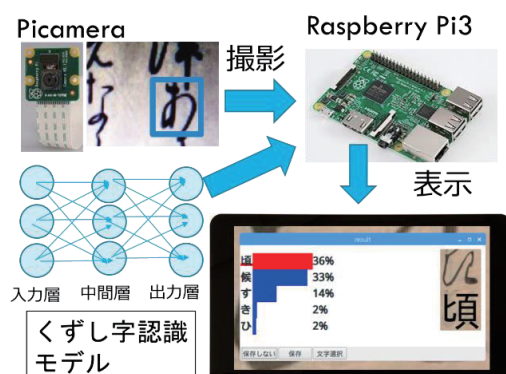


図 1 システムの概要図

連絡先: 早坂太一, 豊田工業高等専門学校 情報工学科,  
〒471-8525 愛知県豊田市栄生町 2-1, TEL(0565)36-  
5861, FAX(0565)36-5926, hayasaka@toyota-ct.ac.jp

OpenCV3 を用いて認識する文字の選択, 画像処理などを行う。認識は機械学習用ライブラリ TensorFlow Lite を用いて行われ、結果は7インチのタッチディスプレイに表示される。

#### 2.1 くずし字領域の検出

古典籍画像に対する前処理を OpenCV3 によって行う。1文字の選択は、タッチスクリーンに指を触れてから離すまでの x-y 座標を取得し、文字領域を 62×62 ピクセルにリサイズし、2値化を行った後に、学習モデルに入力する。しかしながら、1文字ずつ選択して認識する方法では、時間と手間がかかるため、OpenCV3 のモルフォロジ処理によって、古典籍画像から自動で1文字ずつの領域を検出することを試みた。

まず入力画像に対して、図 2(a)に示すように縦方向に対して膨張処理を行い、上下の字をつなげることで、1行ごとの集まりを作成し、領域検出を行う。次に、検出した領域範囲を用いて、図 2(b)に示すように元の入力画像を横方向に膨張処理を行い、1行の中から1文字の領域を検出する。図 2(c)に検出例を示す。

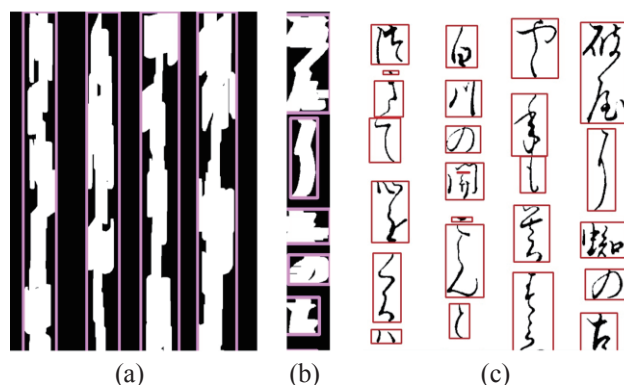


図 2 文字領域の検出過程と結果例

#### 2.2 くずし字認識モデル[2]

本研究では、ディープラーニングで学習したくずし字認識モデルを用いる。このモデルは、4層の畳み込み層と3層の全結合層から構成される畳み込みニューラルネットワークである。このモデルに対して、ディープラーニング用ライブラリ Caffe を用

いて学習を行った。学習には 3,985 種類の 418,831 文字(平均 105 文字/種類)のくずし字画像データ[3]を用いた。そのうち、変体仮名(48 文字+濁点・半濁点)は 75 種類の 262,547 文字(平均 3,501 文字/種類)、その他の漢字・カタカナ・記号などは 3,910 種類の 156,284 文字(平均 40 文字/種類)である。このモデルは、テストデータの源氏物語絵巻「桐壺」[4]11,286 文字(変体仮名 10,026 文字、その他 1,260 文字)に対して、正解率 92.4%(変体仮名 95.2%, その他 70.1%)を得ている。

## 2.3 モデルの量子化

本研究では当初、Raspberry Pi で学習モデルを実行させるために、代表的な機械学習ライブラリである TensorFlow を使用した[5]。しかしながら、学習モデルのサイズは 298.2MB、認識時間は約 3.0 秒と、モデルの読み込みおよび認識にある程度の時間がかかった。そこで、モバイル端末や組み込み機器向けの TensorFlow Lite を使用した。TensorFlow Lite では、学習済みモデルの重みなどを浮動小数点形式から固定小数点形式に量子化し、モデルサイズを小さくすることで、能力がやや劣る機器での演算の高速化、および読み込みにかかる時間を短縮することができる。本研究では、入力のみ 32bit の浮動小数点形式とし、他のパラメータは 8bit の固定小数点形式に量子化した。その結果、先行研究と同精度のまま、モデルサイズは 78.4MB に軽量化され、1 文字あたりの認識にかかる時間は約 0.4 秒と、TensorFlow による認識時間から約 1/8 に短縮することができた。

## 3. 実行例

開発したくずし字認識システムの表示画面を図 3 および図 4 に示す。開発したシステムでは、GUI アプリケーションから、カメラ撮影、ファイル読み込み、文字検出、切り取り、自動認識のそれぞれのボタンを操作することで処理を行うことができる。

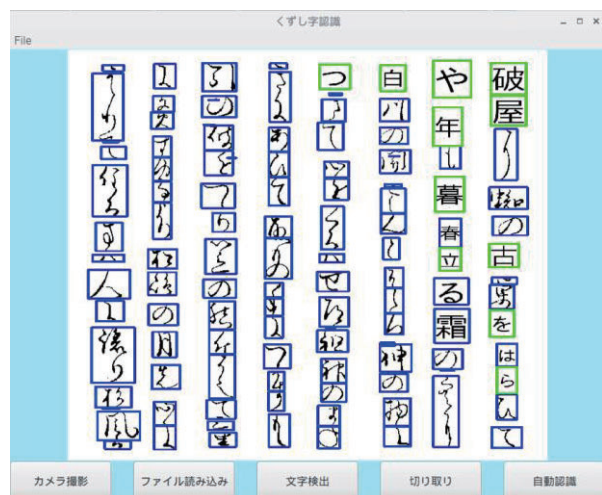


図 3 開発したシステムの実行例

ディスプレイの背面に設置した Raspberry Pi 用カメラモジュールは、焦点距離を約 30cm とし、3 秒間のプレビュー後に古典籍画像を撮影するよう設定した。既に保存済みの古典籍画像データ、もしくは保存を行った過去の画像データのどちらかを選択して、ファイル読み込みを行うこともできる。認識後は図 4 に示すように結果を表示し、認識結果の保存と元画像のくずし字を認識結果に置き換えることができる。その後、認識した文字を再びタッチすることで結果を何度でも参照することが可能であり、利用者が候

補となる 5 文字の中から認識結果を選択することが可能である。認識結果はテキストデータとして保存され、画像と同時に読み込むことで、いつでも認識結果を参照することができる。

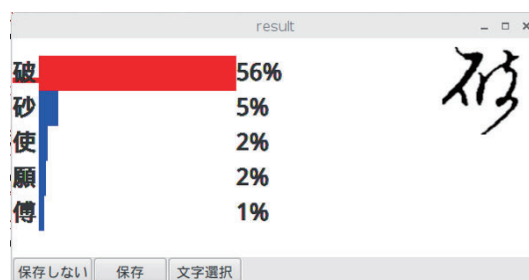


図 4 認識結果の表示例

また、ディスプレイを 2 本指でタッチすることで、図 3 に示すように、認識を行った文字は緑色枠で、文字検出を行った結果の領域は青色枠で囲まれて表示される。検出した領域をタッチすることで認識が行われ、結果が表示される。任意の領域の切り取りを行いたい場合は、ディスプレイからタッチおよびドラッグにより選択し、その後に表示される認識ボタンを押すと認識を行うことができる。ボタンにより再度選択を行うことも可能である。

## 4. おわりに

本研究では、Raspberry Pi を用い、カメラで撮影した画像などからくずし字の認識を可能とするシステムを開発した。量子化ニューラルネットワークを用いたところ、1 文字を認識するのにかかる時間や精度は先行研究とほぼ変わらず、高性能なコンピュータを用いなくても、Raspberry Pi のようなシングルボードコンピュータで問題なく動作することを示すことができた。

スマートフォンを持ち込むことができない小・中学校や、普段モバイル機器を持ち歩かない高齢者の方々でも、くずし字に触れたいという場面に遭遇することは少なくないと考えられる。ネットワーク環境のない場所も含めて、そうした現場で支援ツールとして活躍することができるよう、本システムの操作性や認識精度をより高めることが今後の課題である。

## 謝辞

本研究は JSPS 科研費 JP16K02433 および TRC-ADEAC 株式会社からの助成を受けたものです。

## 参考文献

- [1] 国文学研究資料館：歴史的典籍に関する大型プロジェクト, <https://www.nijl.ac.jp/pages/cijproject/>, 2015 年 10 月 14 日参照。
- [2] 早坂太一 他：2017 年度人工知能学会全国大会(第 31 回)論文集, 3Q1-2in1, 2017.
- [3] 人文学オープンデータ共同利用センター：日本古典籍字形データセット(国文研所蔵・CODH 加工), <http://codh.rois.ac.jp/char-shape/>, 2016 年 12 月 8 日参照。
- [4] 人文学オープンデータ共同利用センター：日本古典籍データセット(国文研所蔵), 源氏物語, <http://codh.rois.ac.jp/pmjt/book/200003803/>, 2017 年 2 月 15 日参照。
- [5] 竹内正広 他：第 17 回情報科学技術フォーラム講演論文集, vol.4, pp.293-294, 2018.