生成モデルによる篆書体の文字認識手法の提案

李 康穎† Batjargal Biligsaikhan‡ 前田 亮†‡
†立命館大学情報理工学研究科 〒525-8577 滋賀県草津市野路東 1-1-1
‡立命館大学総合科学技術研究機構 〒525-8577 滋賀県草津市野路東 1-1-1
†‡立命館大学情報理工学部 〒525-8577 滋賀県草津市野路東 1-1-1
E-mail: †gr0319ss@ed.ritsumei.ac.jp ‡biligsaikhan@gmail.com †‡amaeda@is.ritsumei.ac.jp

あらまし 篆文は今でも幅広く使われている古代文字である．しかし，篆文の字形は現代の漢字とは大きく異なっており，専門的な知識を持たない人にとって篆文字形を読むのは非常に難しい. 国文学研究資料館が公開している蔵書印データベースには，篆文で彫られた蔵書印のデータが多く含まれており，活用が期待されている．本研究は，篆文のフォントデータを使用し，蔵書印データベースから取得した画像の分割と認識を行う手法を提案し，今後の蔵書印ディジタル図書館の検索支援の実現に向けた基礎を構築することを目指す.

キーワード 古代文字認識，情報検索，生成モデル，ディジタルアーカイブ

1. まえがき

篆文は現在でも印鑑などに幅広く使われている古代の書体であるが，現代人にとって篆文を読むのは難しい．蔵書印とは，図書所蔵者が図書の所有権と自分の個性趣味を表すための印である．昔から今まで，蔵書を好き好んだ有名人であれば，自分の蔵書に押印しない人はいないと考えられている．蔵書印の表面には名前と雅名だけでなく，ほかの文字もつけられている場合がある．蔵書印には様々な形態があり，漢篆書を主体とした漢字，図案，あるいは，西洋文字も含まれている場合がある．その中で，篆書で印を彫った蔵書印が篆字で伝えた情報は，専門家でない現代人にとって，理解するのは難しい．現在の文字認識システムは楷書を対象とするものが多く，篆書や甲骨文などの古代文字を対象としたシステムはあまり存在しない．一方，国文学研究資料館が公開している蔵書印データベース[1]には，篆文で彫られた蔵書印のデータが多く含まれており，これらの活用が期待されている.

立命館大学白川静記念東洋文字文化研究所が公開している「白川フォント」及びその検索システムには篆文 2,590 字のフォントデータが含まれる．本論文では，「白川フォント」[2]の篆文フォントデータと「說文解字 True Type 字型」[3]を用いて蔵書印中の篆文の文字認識システムの構築を試みる.本研究は，篆文のフォントデータを使用し，蔵書印データベースから取得した画像の分割と認識を行う手法を提案し，今後の蔵書印画像の検索支援の実現に向けた基礎を構築することを目指す.

2. 関連研究

2.1 古代文字の文字認識

篆文と同じ形式の象形文字には，甲骨文や金文などの書体が含まれる．古代文字認識の研究では，碑刻，木簡，史料などに書かれた古代文字の領域を計算し，画像に含まれる文字をテキストとして抽出することを目指している．例えば，Lin らによる研究[4]では，甲骨文を 2 ステップで認識する手法を提案した．ライン特徴の分析とテンプレート類似度の計算を用いて，甲骨文の認識を行った．鈴木ら[5]は線分抽出パラメタの自動最適化による甲骨文の認識手法を提案した.

2.2 文字領域の抽出

画像からの文字の領域検出は，コンピュータビジョンアプリケーションで重要な役割を果たしている．既存のテキスト領域の検出方法は主に英語の文字に焦点を当てており，漢字のテキスト領域の検出に関する研究は少ない．近年，自然環境とノイズが含まれる画像からの文字領域の抽出に関する研究が注目されている. 渡邉ら[6]は，ガボールフィルタを用いて甲骨拓本から文字の領域を抽出した．Ren による自然画像からの漢字の抽出手法[7]では Multi-input-layer Deep Belief Network（DBN）を用いて自然環境の漢字テキストを検出した.

2.3 GAN による訓練データ量の拡張

Zhu ら[8]は Generative Adversarial Networks（GAN）を使用して，表情画像のデータの増強方法を提案した. GAN の性能を検証するために，3 つのベンチマークデータセットに対するいくつかの実験をしており，結果として，分類結果の精度で 5%～10%の向上を得ることができることを示した.

3. 提案手法

本研究では，データ量拡張による古代文字認識の手法を提案する．認識手法として、多層ニューラルネットワーク CNN（Convolutional Neural Network）を使う. CNN では，画像などの情報をフィルタ内の情報が畳み込まれて抽象化できるモデルである．画像認識に高い

性能があるため，特に文字認識によく使われる．画像を自動エンコーダで予め訓練し，漢字画像の情報量を大幅に圧縮して抽象化することができる．

本研究で提案する手法は，(1)　蔵書印データの前処理，(2)　生成モデルによる訓練データの生成，(3)篆文文字認識の試みの三つのステップによって構成する．

3.1　蔵書印データの前処理

図 1 に示すように、蔵書印に様々な形態があり，篆文を主体とした漢字，図案，あるいは，西洋文字も含まれている場合がある．

図　1　様々な蔵書印

本研究では，篆文で彫られた蔵書印を対象とした文字認識の手法を提案する．データベースから抽出した画像データから，篆文の蔵書印を自動的に検出する必要がある．そのため，主成分分析（PCA）とユークリッド距離の計算を用いたデータ前処理手法を提案する．抽出した画像を 255×255 ピクセルのグレイスケール画像に変換し，手動で選んだ篆文画像を主成分分析する．これにより多次元の画像を圧縮できる．次元削減の目標次元数は以下の式 1 で計算する:

$$\frac{\frac{1}{m}\sum_{i=1}^{m}\left\|x^{(i)}-x_{approx}^{(i)}\right\|^2}{\frac{1}{m}\sum_{i=1}^{m}\left\|x^{(i)}\right\|^2}\leq 0.01 \qquad (1)$$

射影先の篆文蔵書印データをx_{approx}と置き，99%以上の特徴情報を保存する．次に，圧縮した篆文蔵書印画像の特徴ベクトルの平均値ベクトルを計算する．さらに図案も含む全ての画像データを決められた同じ次元数に投影する．投影したベクトルと蔵書印の投影ベクトルのユークリッド距離を計算する．

図 2 に示すように，各サンプルの投影ベクトルと篆文蔵書印の投影ベクトルのユークリッド距離が表される．横軸は表示されるサンプル数，縦軸は各サンプルを削減された次元に投影した低次元ベクトルと同じ次元に圧縮した篆文蔵書印ベクトルの平均ベクトルのユークリッド距離類似度である．ユークリッド距離類似度の閾値を 0.8 に設定することで，篆文蔵書印と他の蔵書印を区別することができる．

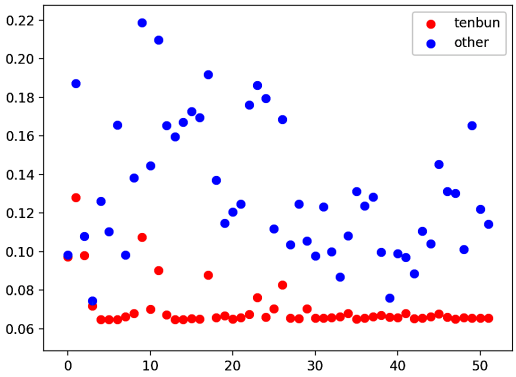

図　2　ユークリッド距離による非篆文蔵書印画像の除去

3.2　生成モデルによる訓練データの生成

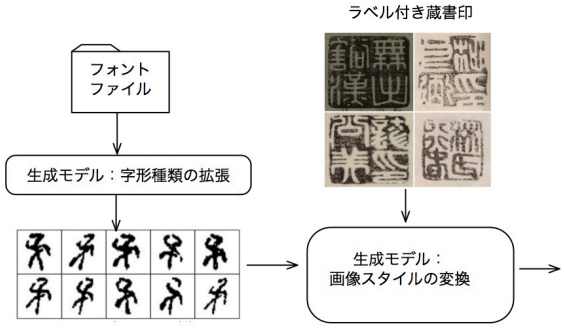

図　3　訓練データ生成の流れ図

これまでの我々の研究[9]では，生成モデル zi2zi を用いてフォントから異なる形態を持つ画像を抽出した．図 3 に示すように，本研究ではこれらの画像とラベル付き蔵書印画像を Cycle GAN[10]（Cycle-Consistent Adversarial Networks）への入力として使い，その出力を認識実験の訓練データとして利用する．

3.3　生成画像を用いた篆文文字の認識

図 4 に示すように，前の段階で収集したデータの抽象化情報を抽出する．既存のモデル VGG-16 を用い，転移学習を行い、コサイン類似度の計算により，文字領域を推定し，領域情報を画像に付与する．単位文字の領域情報付き画像データ R-CNN(Regions with Convolutional Neural Networks)モデルで分類して認識結果を出力する．

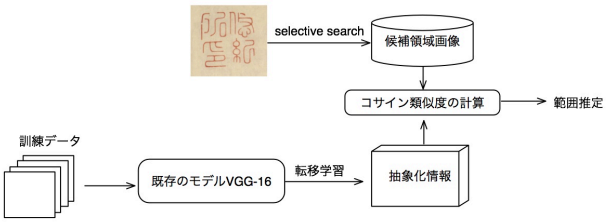

図 4　篆文文字認識の流れ図

4. 実験

現在，提案手法について以下の実験を行っている．

4.1　篆文蔵書印の抽出

データベースから抽出したデータの非篆文蔵書印の部分を主成分分析（PCA）とユークリッド距離で抽出した．抽出したデータの総数は 5086 枚であり、除去した非篆文蔵書印は 367 枚、残った篆文蔵書印は 4719 枚であった．本研究では、ピアソン類似度とコサイン類似度による特徴ベクトル関係も計算した．

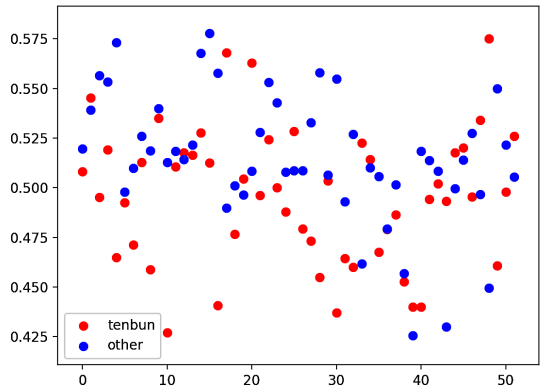

図 5　ピアソン類似度による非篆文蔵書印画像と篆文蔵書印画像の特徴ベクトル関係

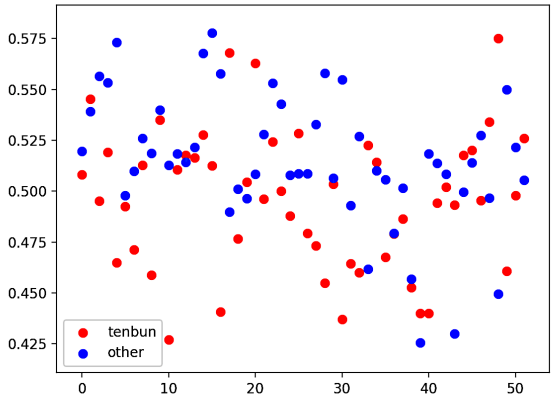

図 6　コサイン類似度による非篆文蔵書印画像と篆文蔵書印画像の特徴ベクトル関係

ピアソン類似度とコサイン類似度による特徴ベクトル関係を図 5 と図 6 に示す．図 2 と同じように，横軸は表示されるサンプル数，縦軸は各サンプルを削減された次元に投影した低次元ベクトルと同じ次元に圧縮した篆文蔵書印ベクトルの平均ベクトルのユークリッド距離類似度である．図 5 と図 6 から見ると，篆文蔵書印と他の蔵書印の計算値が混在しており、ピアソン類似度とコサイン類似度の計算では、非篆文蔵書印画像の除去に適当ではないことがわかった．

4.2　生成モデルによる訓練データの生成

生成モデルを用いて，訓練データの生成実験を行っている．

4.3　蔵書印画像の文字認識

zi2zi モデルを用いて生成した画像に基づく手書き篆文画像の認識の実験結果を表 1 に示す．

手法	正解率（%）
zi2zi モデルからの生成画像	60.00
形態変換による生成画像	77.00

表 1 生成画像データを用いた認識モデルによる手書き画像に対する実験結果

訓練データは 300 種類であり、テストデータは図 7 に示すような手書き画像 300 枚である．

図 7　手書きテストデータ

生成したデータを訓練データの一部として用い，篆文で彫った蔵書印の認識実験を試みている．

スタイル変換により変換したデータを図 8 に示す。

図 8 スタイル変換による結果例

実験用訓練データは図 9 に示すように，五種類のスタイルがあり，毎種類はランダム化アフィン変換

(affine transformation)，ランダム射影変換 (projective transformation)，膨張化処理(dilation)，収縮化処理(erosion)によりデータ量を拡張する。

図 9　スタイル変換により生成した訓練データ

表 2 に示すように，今回位置情報推定と文字認識に用いられる文字種類は 10 種類であり,データベースから抽出した画像のラベルの単一文字の出現頻度の計算により決められる。

文字	出現頻度
印	1145
文	968
記	842
改	738
表	735
書	450
蔵	278
之	226
氏	203
山	183

表 2　単一文字の出現頻度の計算

今回 R-CNN モデルの訓練データとして全て真実データを選択するため，出現頻度が 100 枚を超えた単一文字に注目し，提案手法によりそれぞれの位置情報を付けて R-CNN モデルの訓練データとして使う。

R-CNN 分類学習ごとの訓練データ数はおよそ画像 160～250 枚であり，テスト画像は種類ごとに 20 枚である。

評価の手法を式(2)に示す:

正確率（%）＝（領域推定＋カテゴリ判定）正解数/テストデータ数×100　　（2）

実験結果は表 3 の通りとなった.

文字	正解率（%）	正解数(枚)
印	90	18/20
文	90	18/20
記	85	17/20
改	85	17/20
表	85	17/20
書	80	16/20
蔵	75	15/20
之	85	17/20
氏	90	18/20
山	95	19/20

表 3 実験結果

5. まとめ

本研究では，フォントから生成された文字画像を用い，公開されている蔵書印データベースの篆文文字に対する認識手法を提案した．今後の課題として、訓練ネットワークのよりよい調整と，文字領域を抽出する手法の検討が必要であり、認識率を高められるように，文字認識のモデルに対するより質的な分析と評価が必要である．

参　考　文　献

[1] 国文学研究資料館：「蔵書印データベース」，入手先：＜http://base1.nijl.ac.jp/~collectors_seal/＞（参照 2017-12-25）

[2] 立命館大学白川静記念東洋文字文化研究所：白川フォント，入手先〈http://www.ritsumei.ac.jp/acd/re/k-rsc/sio/shirakawa/index.html〉（参照 2017-10-15）

[3] 台灣地區行政院主計處電子處理資料中心: 說文解字 True Type 字型，入手先〈http://www.cns11643.gov.tw/MAIDB/welcome.do〉（参照 2017-5-1）

[4] Meng, L.:"Two-Stage Recognition for Oracle Bone Inscriptions." International Conference on Image Analysis and Processing. Springer, Cham, pp. 672-682.（2017）

[5] 鈴木達也，孟林，泉知論. "線分抽出パラメタの自動最適化による甲骨文字認識率の向上（画像工学）." 電子情報通信学会技術研究報告= IEICE technical report: 信学技報 116.464 pp.315-320（2017）

[6] 渡邉清威，孟林，泉知論. "ガボールフィルタを用いた甲骨拓本からの文字領域の抽出（パターン認識・メディア理解: 様々なメディア）." 電子情報通信学会技術研究報告= IEICE technical report: 信学技報 117.105pp.45-50.（2017）:

[7] Ren, X., Zhou, Y., He, J., Chen, K., Yang, X., & Sun, J.: "A Convolutional Neural Network-Based Chinese Text Detection Algorithm via Text Structure Modeling." IEEE Transactions on Multimedia 19.3 pp.506-518.（2017）

[8] Zhu, X., Liu, Y., Qin, Z., et al. Data Augmentation in Emotion Classification Using Generative Adversarial Networks[J]. arXiv:1711.00648v5 [cs.CV]（2017）

[9] 李康穎，Batjargal Biligsaikhan，前田亮: “古代文字フォントの画像データに基づく手書き篆文文字の検索支援” じんもんこん 2017 論文集，pp.125-130（2017）

[10] Zhu, J. Y., Park, T., Isola, P., Efros, A. A.:"Unpaired image-to-image translation using cycle-consistent adversarial networks." arXiv preprint arXiv:1703.10593（2017）.