古代文字フォント字形の特徴抽出に基づく蔵書印の検索支援

李　康穎（立命館大学 情報理工学研究科）
Batjargal Biligsaikhan（立命館大学 衣笠総合研究機構）
前田　亮（立命館大学 情報理工学部）

蔵書印とは，図書の所蔵者が図書の所有権と自分の個性や趣味などの情報を表すための印である．古典籍に特有の要素である蔵書印には，多数の重要な情報が含まれている．篆書で彫られた蔵書印の印文情報を専門家ではない現代人に伝えるため，本研究では，複数特徴量の抽出に基づく篆書体による蔵書印の検索ベース文字認識手法を提案する．

Retrieval of ownership stamp characters based on feature extraction of ancient character typeface

Kangying Li (Graduate School of Information Science and Engineering, Ritsumeikan University)
Biligsaikhan Batjargal (Kinugasa Research Organization, Ritsumeikan University)
Akira Maeda (College of Information Science and Engineering, Ritsumeikan University)

In order to prove the ownership of a book, book collectors and organizations usually seal their books with an ownership stamp. As a unique element in Asian ancient books, the seal carries a lot of information such as versions of the books, the owner of the books, etc., and it is an important part of book collection culture. To convey those information to people who are not professional scholars, in this research, we propose a character notation recognition system of each element in ownership stamps, using a clustering algorithm and multi-feature extraction of ancient character typeface images.

１．まえがき

古典籍に特有の情報である蔵書印には，多数の重要な情報が含まれている．現在では，大量の古典籍の画像データが Web サイトで公開されており，多くの著作には蔵書印が捺されている．しかしながら，篆書で彫られた蔵書印が伝える情報は，専門家でない現代人にとって，その内容を理解するのは難しい．また，既存の日本語を対象にした文字認識システムは，楷書の漢字を認識対象とするものが多く，現在の漢字である楷書と字形に大きな違いがある篆書に対応したシステムはほとんど存在しない．

蔵書印の内容を自動的に認識できれば，所有者などの書籍の関連情報を多数の Web サイトのメタデータから取得できる．これらの情報に基づいて，書承関係，書物の所蔵履歴など様々な関連情報の取得もできるようになり，人文系の研究に大きく貢献できると考えられる．

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」[1]には篆文 2,590 字のフォントデータが含まれており，このデータを活用した蔵書印の自動解析システムの構築が期待されている．本研究では，大量の学習データを用いず，「白川フォント」の字形情報の複数特徴量の抽出に基づく篆書体による蔵書印の検索ベース文字認識手法を提案する．

２．関連研究

蔵書印を自動的に解析するシステムとして，富士通による蔵書印検索技術のプロジェクト[2]がある．中国の蔵書印画像を対象とし，構築されたデータベースを用い，蔵書印の印文全体をマッチングなどの技術で認識する実験が進められている．近年，オフライン文字認識の手法として，畳み込みニューラルネットワーク（CNN: Convolutional Neural Network）の応用が注目されている．人の神経回路網を数式的なモデルで表現し，「畳み込み層」により画像からのエッジ等の特徴を抽出し，画像の抽象化を行い，分類を行う．大量の学習データに基づき，手書き文字を高い精度で認識することができる．CNN を用いた手書き漢字の認識システムとして，スタンフォード大学の学生により「ETL 文字データベース」[3]の漢字データを 99.64%の精度で認識できたとの報告がある[4]．また，古代文字に対する解析では，甲骨文字を 114 種 820 枚から 184 種 2000 枚に文字種を増やして実験を行い，94.4%の認識率を達成した研究もある[5]．他の認識手法として，古代文字データベースを構築し，候補テンプレートを検索することにより文字認識を行う研究がある[6]．

蔵書印の印文内容には，人物名や所蔵機構名が多く存在する．すなわち，蔵書印の印文全体を認識対象とする場合，多種類の印文データを同等の量で集めるのは難しいと考えられる．また，同じ印文文字に対応する篆書体のバリエーションも

多種類が存在するため，全てのバリエーションを同じ種類にして学習データを作成するのは問題がある．そのため，本研究では，深層学習の技術を用い，同じ文字のバリエーションの違いを考慮した手法を提案する．

本研究では，敵対的ニューラルネットワークにより拡張された字形データを畳み込みニューラルネットワークの学習データとして訓練させる．学習済みモデルでは，転移学習の原理による特徴抽出器を作成する．作成された特徴抽出器では，画像の抽象化により，画像ノイズ，字形の変形に対して柔軟性が高いという特徴があり，その特徴の検索枝刈り演算への応用を検討する．さらに，画像処理による複数特徴の抽出に基づく，検索ベース認識手法を提案する．

3．提案手法

本研究で提案する手法は，(1) 蔵書印画像データからの単一文字領域の判定，(2) フォント字形からの複数特徴情報の抽出，(3) ランキング計算結果による文字の推定の三つのステップによって構成される．

3.1 データの取得および前処理

表 1 に「白川フォント」における篆文の例を示す．

表 1　白川フォントの例
Table 1　Examples of Shirakawa font.

現代文字	人	文	科	学
白川フォント				

より良い字形の特徴量を抽出するため，「白川フォント」から抽出されたデータの前処理を行う．前処理は，画像のサイズの規格化，画像の二値化，字形の細線化を含む．画像のサイズは全て 224×224ピクセルに統一し，画素の色特徴と空間分布特徴を利用し，k-means による画像の二値化を行う．k-means クラスタリングは画像中で使用される色を削減する処理の手法としてよく使われる．入力画像には RGB の三つの色特徴があり，ここで，入力された画像の画素を RGB の特徴で表現し，画素数×3 の配列に変換する．古典籍の画像中の蔵書印がある領域では，背景色，印文色，ノイズ色の三つの特徴色を持つと仮定し，クラスタリングの際の RGB 特徴の重心の数 k を 3 に設定する．k-means クラスタリングの計算を通じて，画像の全画素を，計算された三つの重心の RGB 値により色の量子化を行う．

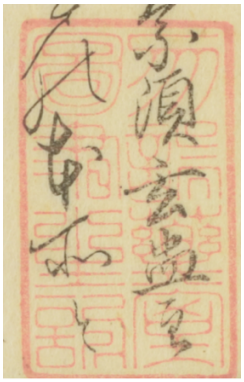


図 1　蔵書印と手書き漢字が重なる例
Figure 1　Examples of handwritten words embedded in a ownership stamp.

図 1 に示すように，蔵書印のスキャンデータには蔵書印領域と手書き漢字が重なる画像が多く存在する．上述の色の量子化手法により，蔵書印画像領域と手書き文字領域を分離できる．

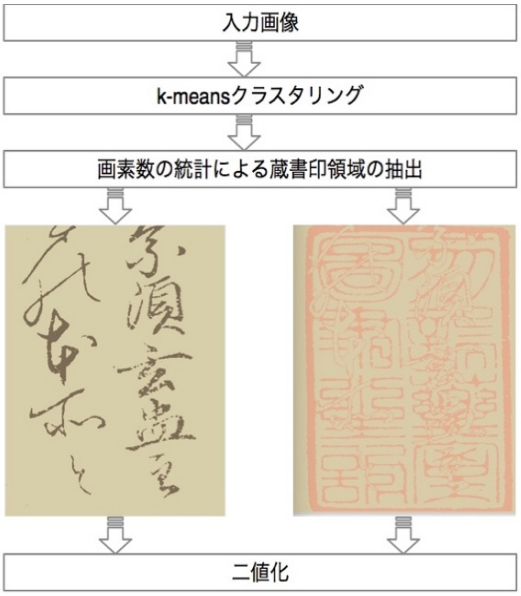


図 2　前処理の流れ
Figure 2　The flow of preprocessing.

二値化処理までの前処理の流れを図 2 に示す．クラスタ数を 3 に設定することで，結果画像の全画素は三つの集合に分けられる．画素数が最も多いのは背景画素と考えられる．残り二つの画素の集合のうち，クラスタ重心の RGB 値が RGB 空間モデル中でもっとも赤色空間に近いものを蔵書印領域として抽出する．抽出された領域を二値画像に変換し，次の計算に用いる．

3.2 蔵書印画像データからの単一文字領域の判定

古典史料の画像の輝度ヒストグラムを解析し，

輝度の頻度分布間の谷の部分を閾値とし，文字領域を抽出する手法がよく用いられる．また，近年注目されるようになった深層学習を用い，文字境界識別器を訓練させ，古代文字のくずし字を抽出する研究もある[7].

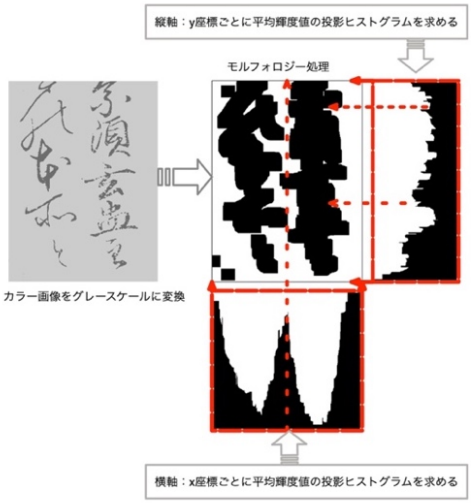


図 3　投影ヒストグラムによる古典資料からの単一文字の抽出

Figure 3　Extraction of a single character from an ancient document by histogram projection.

前節で抽出された手書き文字の画像を例として説明する．画像の平均輝度値の投影ヒストグラムを求め，それらの結果により単一文字を抽出する．その結果を図 3 に示す．

図 4　投影ヒストグラムによる蔵書印からの単一文字の抽出

Figure 4　Extraction of a single character from an ownership stamp by histogram projection.

比較的整った手書き文字からなる古典史料のスキャンデータとは異なり，図 4 に示すように，蔵書印画像の文字分布は，特定の領域に偏ることがなく，文字の分布範囲が広くなるため，中心領域および文字の分布辺縁の境界の判定ができない場合が多くあり，既存手法を用いた個々の古代文字の抽出は困難である．本研究では，画像の画素分布密度特徴による単一文字領域の抽出を行う．その前に，まず 3.1 節で述べた手法を用い，画像の二値化画像を取得する．伝統的な方法と比較して，k-means に基づく画像の二値化処理方法は，元の画像のより多くの情報を保持することができる．

単一文字領域の抽出について，本研究では漢字の形態と構成には重心の構造的安定性と単語間の隙間があることを踏まえ，空間分布特性の密度を考慮したカーネル密度推定を用いる．画像の横軸におけるカーネル密度推定量を数式 1 に示す．

$$\hat{f}_{bandwidth}(x) = \frac{1}{n*bandwidth}\sum_{i=1}^{n} K\left(\frac{x-x_i}{bandwidth}\right) \quad (1)$$

ここで，$x_1,x_2,x_3,\ldots x_n$は，二値化された画像の印文を示す黒画素の x 座標の集合を表す．$y_1,y_2,y_3,\ldots y_n$は，それらの x 座標に対応する y 座標の集合を表す．bandwidth はバンド幅と呼ばれる平滑化のためのパラメータであり，K は数式 2 に示した Gaussian Kernel を用いる．画像の縦軸におけるカーネル密度推定量も数式 1,2 を通じて計算する．

$$K(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right) \quad (2)$$

カーネル密度推定は，画素の分布特性を分析するために使用できる．

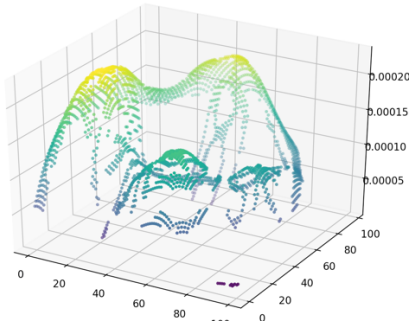


図 5　カーネル密度推定

Figure 5　Kernel density estimation.

図 5 に，図 4 と同じ対象画像に対するカーネル密度推定の例を示す．図の中央に示した 3 次元画像の表現において，各軸は，画像の水平および垂直座標およびその密度特徴の予測値をそれぞれ表す．3 次元座標図の縦軸は密度特徴の予測値であり，蔵書印画像の個々の文字は空間分布において明確な分布特性を有していることがわかる．そこで，濃度に基づく Mean Shift Clustering 法[8]を

用い，図 6 に示すように画素クラスタの局所最大濃度点をクラスタの重心として求める．クラスタリングの結果により単一文字領域を予測する．

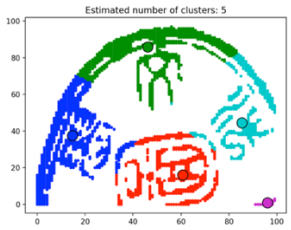


図 6　Mean Shift Clustering による蔵書印からの単一文字の抽出
Figure 6　Extraction of a single character by Mean Shift Clustering.

3.3 フォント字形からの複数特徴情報の抽出

蔵書印には所蔵機関名，所有者の名前などの情報が刻印されており，文字ごとの出現頻度に大きな違いがある．すなわち，多種類の古代文字に対応する学習データを同じ数で集めるのは困難である．中心線特徴，コーナー特徴は字形データを用いたテンプレートマッチングの研究によく使われる特徴量であり，本研究では中心線特徴，コーナー特徴を含む複数の特徴量を抽出し，特徴量のデータベースを構築することで，字形データの拡張性を考慮した検索ベース古代文字認識手法を提案する．図 7 の結果例に示すように，細線化については，Zhang ら[9]が提案した細線化手法で中心線特徴の抽出を行う．

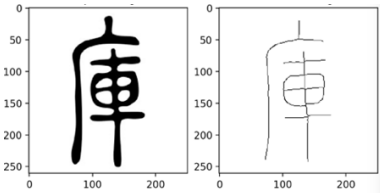


図 7　細線化手法による中心線特徴の抽出
Figure 7　Extraction of edge features by thinning method.

古代文字の字形には多くのバリエーションがあり，同じ文字でも字形が変わることで，輪郭，中心線特徴，コーナー特徴による特徴の抽出結果にも大きな影響がある．畳み込みニューラルネットワークの構造により，画像を低次元に抽象化することができる．Kavukcuoglu [10]らの研究によると，入力層に近い中間層の出力は抽象的な特徴量であることがわかる．入力層に近い中間層の情報を用いることで，同一の種類に属する画像の共通の特徴を抽出することができると思われる．字形の変形の影響を抑えるため，本研究では，転移学習の原理に基づき，生成モデルにより拡張された字形データを VGG19[11]の訓練データとし，その学習済みモデルを特徴抽出器として用いる．

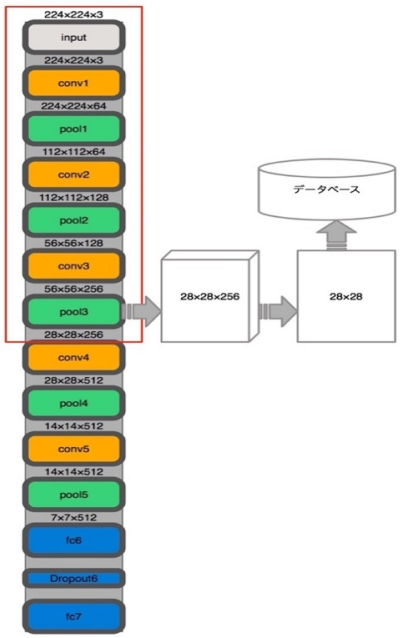


図 8　VGG19 の学習済みモデルによる特徴量の抽出
Figure 8　Extraction of a feature by pre-learned model of VGG 19.

図 8 に示すように，抽出された左側は VGG19 モデルの一部を示す．本研究では，学習済みモデルのプーリング 3 層の出力を特徴抽出器の出力とする．学習データの取得の流れを図 9 に表す．「白川フォント」の字形画像だけではなく，フォントの字形データから文字の本質的な構造を持つ中心線特徴を抽出し，学習済みモデルに入力し，三番目のプーリング層（Pooling layer pool3）特徴を取得する．

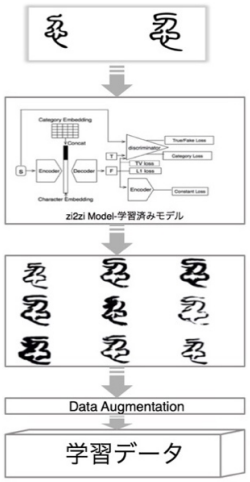
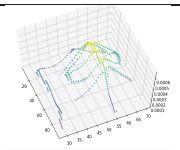
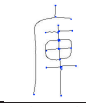


図 9　zi2zi モデルによる字形画像の拡張
Figure 9　Augmentation of font typeface images by zi2zi model.

ここで用いられる学習済みモデルの学習データでは，「白川フォント」の検索システムで公開されている篆文フォントデータと「說文解字 True Type 字型」[12]をソース画像と目標画像として用い，Zhang[13]らが提案した zi2zi モデルでフォントの字形データから多種類の字形データを取得する．生成された字形データを VGG19 の学習データとして訓練させる．このほか，本研究では，字形が類似する文字を区別するため，データの字形構造のコーナー特徴と密度特徴を組み合わせて利用し，フォントの二値化画像から表 2 に示した特徴を取得する．

表 2　使用する特徴量
Table2　Multi-feature extraction.

特徴名	例	データベースにおける記述
カーネル密度特徴		3×n ベクトル(n は二値化画像における黒画素の数)
中心線特徴の Harris コーナー特徴		2×n ベクトル(n は検出されたコーナー数)
検出された Harris コーナー数	(同上)	数値

3.3 ランキング計算結果による文字の推定

3.2 節で述べた手法により特徴量を抽出し，データベースを構築する．ユーザが入力した認識対象になる画像も，上述の手法により特徴量を抽出し，表 3 に示した距離計算手法により，データベースに保存されている特徴量との距離を計算してランキングを行う．カーネル密度などの特徴量は二次元を超えるため，多次元にも応用できるハウスドルフ距離の計算によりランキングを行う．他の特徴量に対してはユークリッド距離の計算を行う．

表 3　距離計算の手法
Table3　Distance calculation methods.

特徴	距離計算手法
Harris コーナー数	Hausdorff
Harris コーナー特徴	Hausdorff
カーネル密度特徴	Hausdorff
「Pooling layer pool3」特徴	Euclidean
「Pooling layer pool3」中心線特徴	Euclidean

ユーザが入力した認識対象の画像に対しても上述の手法により特徴量を抽出し，表 3 に示した距離計算手法により，データベースに保存されている特徴量との距離を計算してランキングを行う．

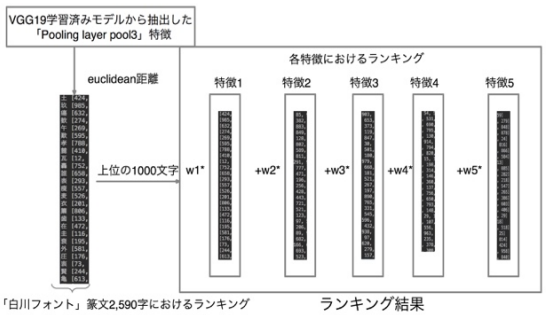


図 10　ランキング計算
Figure 10　Ranking calculation

取得された「Pooling layer pool3」特徴では，字形に対する抽象化ができ，同じ文字の字形における小さな違いによる影響を抑えることができる．この特徴を用いて，ランキング結果に対する枝刈り演算を行う．図 10 に示すように，「Pooling layer pool3」特徴の距離計算によるランキング結果を用いた枝刈り演算を行うことで，上位 1,000 字の結果画像が残され，次の計算に用いられる．[w1,w2,w3,w4,w5]は，検索結果を改善するため設定する重みであり，初期設定は[1,1,1,1,1]である．

4．実験

国文学研究資料館の「蔵書印データベース」[14]から収集したデータを用い，以下の実験を行った．

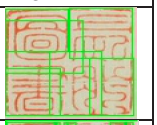
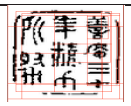
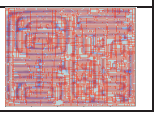
図 11　実験データの例
Figure 11　An example of experimental data.

抽出されたデータは，図 11 に示すように，蔵書印には様々な形態があり，篆文を主体とした漢字が使われていることがわかる．

4.1 単一文字領域の抽出実験

よく使われるセグメンテーション手法としてモルフォロジー演算に基づいた手法と Selective search[15]と呼ばれる手法がある．表 4 に示すように，本研究の提案手法でスタイル 1，2 のような蔵書印画像の処理はより良い結果が得られることがわかる．スタイル 3 に示された文字間の隙間が狭い蔵書印画像に対する単一文字の検出を検討する必要があると考えられる．

表 4　単一文字領域の抽出の結果例
Table 4　Examples of extraction results of a single character.

手法	蔵書印スタイル 1	蔵書印スタイル 2	蔵書印スタイル 3
提案手法			
モルフォロジー演算			
Selective search			

4.2 画像検索実験

本実験で用いたテスト用の単一文字の種類は 10 種類であり，データベースから抽出したラベル付き画像の文字の出現頻度の計算により決められる．テストデータの一部を図 12 に示す．

図 12　実験用の単一文字データの一部
Figure 12　Experimental data for a recognition experiment.

テスト画像は文字の種類ごとに 20 枚を用いる．評価は平均逆順位 (Mean Reciprocal Rank, 数式 3) を用いて行う．Q は検索された画像の総数，i は画像番号，$rank$ はランキングの順位である．

$$MRR = \frac{1}{|Q|}\sum_{i=1}^{|Q|}\frac{1}{rank_i} \quad (3)$$

実験結果の MRR のスコアを表 4 に示す．

表 4　検索結果の MMR スコア
Table4　MRR scores of retrieval results.

文字	MRR スコア	文字	MRR スコア
印	0.0566	書	0.1073
文	0.1340	庫	0.0371
記	0.0322	之	0.1086
木	0.1013	氏	0.1880
図	0.0427	山	0.2150

今回の学習済みモデルではわずか 100 種類の文字で訓練されているため，種類の拡張および重みの調整による検索精度の向上が期待される．

5．あとがき

本研究では，筆者らによる先行研究[16]の提案手法を特徴抽出手法の一部として使い，蔵書印画像における単一文字の抽出および複数特徴を用いた蔵書印の字形検索を試みた．検索精度を高めるため，今後は，重みまたはパラメータの自動調整を考える必要がある．

参考文献

[1] “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”. http://www.ritsumei.ac.jp/acd/re/k-rsc/sio/shirakawa/index.html, (参照 2018-6-1)
[2] “Seal Retrieval Technique for Chinese Ancient Document Images, Fujitsu Research & Development Center Co. Ltd”. http://www.fujitsu.com/cn/en/about/resources/news/press-releases/2016/frdc-0330.html, (参照 2018-2-10)
[3] “ETL 文字データベース”. http://etlcdb.db.aist.go.jp/, (参照 2018-2-10)
[4] Tsai, Charlie. Recognizing handwritten Japanese characters using deep convolutional neural networks. Report of Stanford University, 2016, p.1-7
[5] 紙徳直生，伊藤大喜，多田晃己，孟林，山崎勝弘．深層学習を用いた甲骨文字認識．第 80 回全国大会講演論文集, 2018,p.513-514
[6] 柴田睦月，辻翼，紙徳直生，山形卓也，孟林，山崎勝弘．甲骨文字データベースの構築と候補テンプレートの検索．第 79 回全国大会講演論文集, 2017,p.453-454
[7] 長井歩．畳み込みネットワークによる仮名 3 文字のくずし字の文字認識．人文科学とコンピュータシンポジウム 2017 論文集, 2017,p213-218
[8] Comaniciu, D., Meer, P. Mean shift analysis and applications. The Proceedings of the Seventh IEEE International Conference on Computer Vision,1999, Vol. 2, p.1197-1203
[9] Zhang, T Y, Suen, C Y. A fast-parallel algorithm for thinning digital patterns. Comm ACM, 1984, Vol.27, No.3, p.236-239
[10] Kavukcuoglu, Koray, et al. Learning convolutional feature hierarchies for visual recognition. Advances in neural information processing systems, 2010, p.1090-1098
[11] Simonyan, K., Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint, 2014, arXiv:1409.1556
[12] “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”. http://www.cns11643.gov.tw/MAIDB/welcome.do（参照 2017-5-1）
[13] “Tian, Y.C.: zi2zi Master Chinese Calligraphy with Conditional Adversarial Network”. https://github.com/kaonashi-tyc/zi2zi（参照 2017-7-1）
[14] “国文学研究資料館:「蔵書印データベース」”. http://base1.nijl.ac.jp/~collectors_seal/（参照 2018-9-25）
[15] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders. Selective search for object recognition. International Journal of Computer Vision, 2013, p.154-171
[16] 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，人文科学とコンピュータシンポジウム論文集，2017, p.125-130