

令和 3 年 6 月 14 日現在

機関番号：34315

研究種目：若手研究

研究期間：2018～2020

課題番号：18K18337

研究課題名（和文）拓本の時空間データベースの作成及び潜在的知識の発見

研究課題名（英文）Creating Spatiotemporal Rubbing-database for Discovering Potential knowledge

研究代表者

孟 林 (meng, lin)

立命館大学・理工学部・准教授

研究者番号：60615938

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究は漢字を中心とした日本と中国の拓本を時空間データベースにし、そして、キーワードの可視化により潜在的知識を抽出し、歴史の整理、経済の推移、気候の変動などの研究に貢献する。主な研究成果を以下に示す。(1)深層学習の不足を補い、深層学習とビッグデータ解析技術を用いて、高精度の拓本文字の認識を実現した。(2)時空間拓本データベースを作成することにより、拓本という古代文献の整理を達成した。(3)キーワードを用いた時空間拓本データベースで検索と可視化を実現したことにより、古代文献の整理と潜在的な知識の抽出に貢献できたと考える。(4)上記の機能をサーバーに取り込んで、APIとしての応用も実現した。

研究成果の学術的意義や社会的意義

研究成果の学術的意義としては、(1)深層学習の不足を補う技術を提供し、高精度な拓本文字の認識を実現した。(2)拓本と古代文献解析において、時空間データベースと可視化の機能を提供している。それにより、歴史に関する政治、災害、文化などの研究に対して、研究時間の短縮と効率化などに貢献できる。社会的な意義は、時空間拓本データベースで検索と可視化を実現したことにより、古代文献の整理と潜在的な知識の抽出に貢献できたと考える。

研究成果の概要（英文）：This project aims to create a spatiotemporal rubbing database and apply keyword search for realizing the keyword visualization. The project may help for discovering potential knowledge such as historical politics, economics, and culture. The contributions are shown as follows.

(1) A deep learning and big-data analysis combined approach has been proposed for realizing the high accuracy rubbing character recognition. The proposed approach also supplementing the shortcomings of deep learning.(2) A spatiotemporal rubbing database has been created, which may help to re-organize the ancient documents.(3) The visualization of keywords is realized, the results may be displayed the changes of keywords in spatial and temporal. (4) All of the functions are equipped on the server. And the API is realized, can be used freely for all the users.

研究分野：画像認識

キーワード：深層学習 古代文献整理 時空間データベース 可視化 潜在的知識の発見

1. 研究開始当初の背景

拓本は漢字を中心とした石、骨、金属などに刻まれた文献の重要な保存資料で、歴史が長い、書体が多い、潜在知識が豊富などの特徴がある。これらの認識、解読、及び分析は歴史の整理、経済の推移、気候の変動などの研究に重要である。しかし、拓本に対する情報の整理技術がまだ確立されていなく、時空間情報を含めた拓本データベースも存在しない。尚且つ、拓本は長い年月を経て、劣化が激しくなっているうえ、文字欠損も多く存在していることにより、文字認識が重要な課題となる。本研究は、文字認識から膨大な拓本を整理し、拓本の時空間データベースを作成し、そして、キーワードを検索することによりキーワードの時空間上の可視化を実現する。それらの研究により、潜在的知識を抽出し、歴史の整理、経済の推移、気候の変動などの研究に貢献することを目指す。

2. 研究の目的

本研究は漢字を中心とした日本と中国の拓本を時空間データベースにし、可視化を行うことにより潜在的知識を抽出し、歴史の整理、経済の推移、気候の変動などの研究に貢献する。これらを実現するために、下記の目的を達成することを目指す。

(1) 高精度な拓本文字認識の実現

画像処理、深層学習、ビックデータ解析技術を併用し、劣化された文字を含めた未開拓な漢字高精度多書体の認識を実現する

(2) 時空間拓本データベースの実現

解読された大量の拓本とその文章に対して、時空間情報を用いて、時空間データベースに登録し、時空間拓本データベースを構築する。

(3) 時空間拓本データベース検索によりキーワードの可視化の実現

キーワードを用いて拓本の時空間データベースを検索し、キーワードに関する時間・空間などの情報を地図上に整理し、潜在的知識を発見する。

3. 研究の方法

図1にはシステムの全体図を示す。主に、拓本文字認識、拓本の時空間データベースの作成、及びキーワードの可視化により構成される。

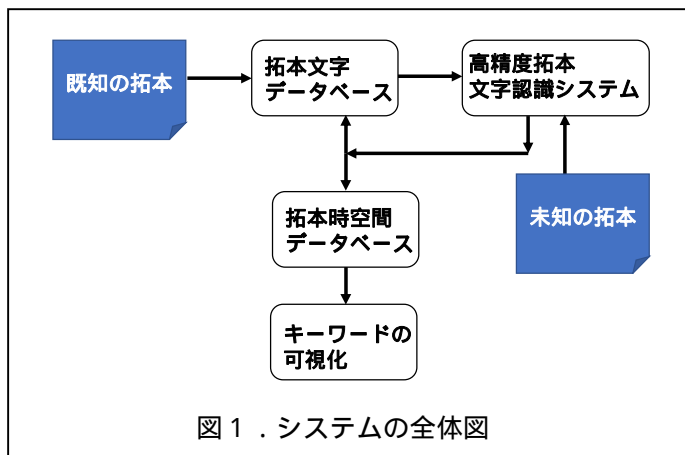


図1. システムの全体図

(1) 高精度な拓本文字認識の実現

図2は高精度な拓本文字認識を示す。まず、既知の拓本文字を用いて文字認識用の拓本文字データベースを作成する。そして、拓本文字データベースを用いて、文字認識用の深層学習のモデルを作成し、学習させ、拓本文字認識用の深層学習モデルを実現する。

しかし、劣化などにより欠損した文字が生じるため、深層学習の性能

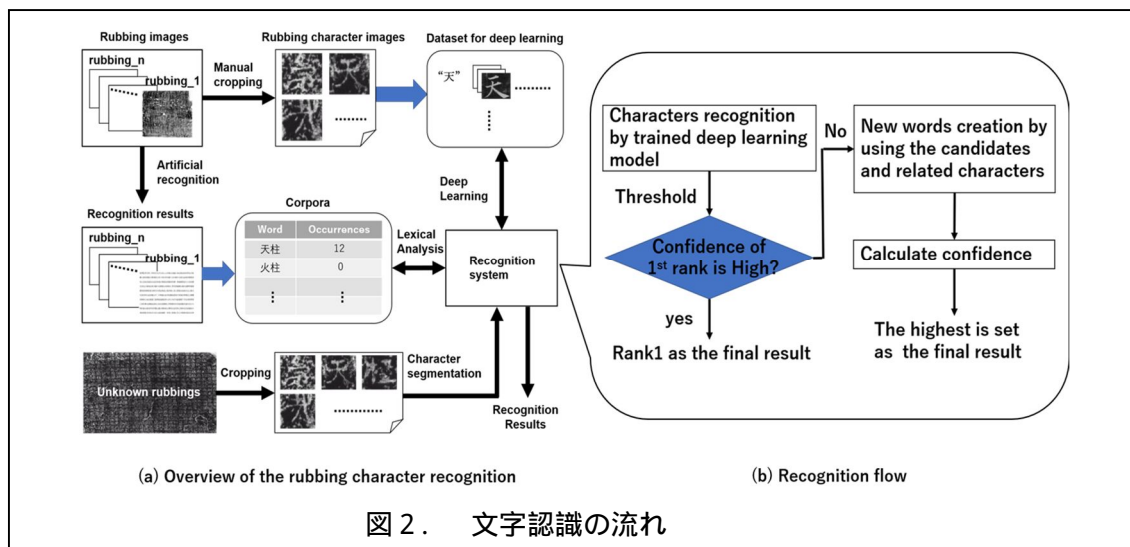


図2. 文字認識の流れ

を十分に発揮できないケースが存在している。それらの文字認識に対応するために、深層学習とビッグデータ解析技術を用いて、高精度の拓本文字認識の実現を目指す。具体的には、既知の文字の単語を記録し、文字の出現頻度と組み合わせ、深層学習の認識結果を再評価し、高精度な拓本文字の認識を実現する。

(2) 時空間データベースの構築とキーワードの可視化

我々は時空間拓本データベースを構築し、検索することにより、時間空間上でのキーワードの可視化を実現した。主に、各拓本の記載時間を時間情報にし、拓本の発見場所を空間情報にし、拓本の内容をデータベースに登録した。また、便利な GUI 機能を追加し、キーワードを検索することにより、時間空間上でのキーワードの可視化を実現した。具体的には、以下の機能を実現した。

- ・ 時空間拓本データベースにキーワードを検索すると、空間上(地図上)にキーワードの出現頻度数を表示する。また、わかりやすくするために、色を尺度にし、キーワードの出現頻度に合わせて、場所に色でマークされる。
- ・ 時間を選択することで、細かい時間帯においたキーワードの出現頻度を地図上に表示もできる。
- ・ 便利な機能により、情報の表示をわかりやすくする。具体的には地図上での表示のみならず、場所をクリックすると、該当場所に、検索したキーワードが各時期の出現頻度をテーブルに表すこともできる。それにより、キーワードの変化が一目瞭然となる。その上に、横軸は場所、縦軸は出現頻度を表す棒グラフで表示する。統計学の視点でキーワードの可視化を実現した。

それらの便利な機能により、時間と空間におけるキーワードの出現頻度と変化を確認できる。また、いろいろなその他の研究に貢献できると見込まれる。例えば、“地震”、“洪水”などのキーワードの出現頻度により、災害の出現頻度を表示でき、災害の研究に役に立つと考える。また、“米”をキーワードにする場合は、最初の出現時期と場所を表すことができる。そして、時間と共に、“米”出現頻度の変化により、“米”がどんどん広がることもわかる。つまり、“米”が主な食糧として、最初の産地や広がり方などの研究に役に立つと考える。

(3) 時空間データベースの構築とキーワードの可視化の API の実現

本システムは、サーバー上にインストールし、API を実現した。ユーザーは、Web API を通じて、サーバーにアクセスし、キーワードの検索とそのキーワードの可視化を実現できた。図 3 は Web API の構造となる。

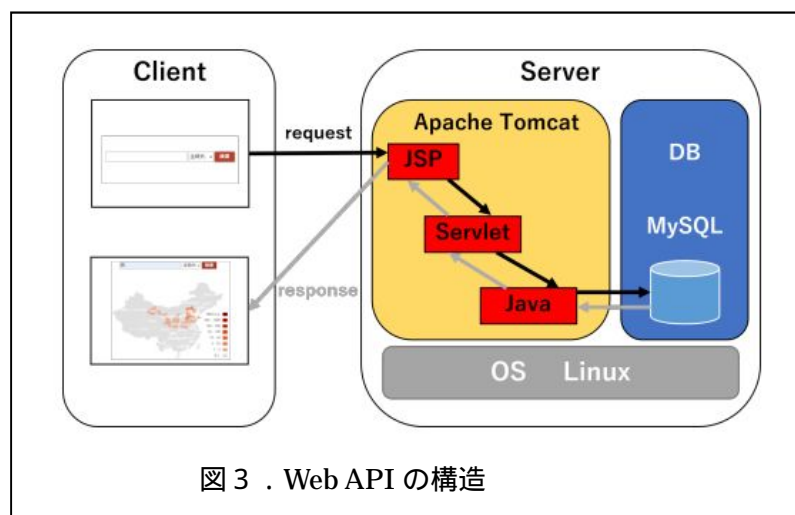


図 3 . Web API の構造

4 . 研究成果

本研究は、システム下記の論文発表を実現した。

孟 林, 孟 澤林, 王 志辰, 富山 宏之, "横断分野での深層学習について," 電気学会・制御研究会, 2020.3.

Zhiyu Zhang, Hiroyuki Tomiyama, Lin Meng, "Combining Deep Learning and Lexical Analysis method for Rubbing Character Recognition , "人文科学とコンピュータシンポジウム, 2019.11.

孟林, "Culture Heritage Protection based on Artificial Intelligence,"第九回全国(中国)文字と計算機学術研討会, 華中師範大学, 武漢, 2019.10.(招待講演)

Lin Meng, Bing Lyu, Zhiyu Zhang, C.V. Aravinda, Naoto Kamitoku, Katsuhiko Yamazaki, "Ocrable Bone Inscription Detector Based on SSD," International Workshop on Pattern

Recognition for Cultural Heritage(PatReCH 2019), in conjunction with ICIAP2019, 2019.9.

Zhiyu Zhang, Yuxi Chen, Hiroyuki Tomiyama and Lin Meng, "Rubbing Characters Recognition based on Deep Learning and Lexical Analysis," Ninth Annual Conference of Japanese Association for Digital Humanities (JADH2019) , 2019.8.

Zhiyu Zhang, Zhichen Wang, Hiroyuki Tomiyama, and Lin Meng, "Deep Learning and Lexical Analysis Combined Rubbing Character Recognition," The 2019 International Conference on Advanced Mechatronic Systems (ICAMechS 2019), 2019.8.

Bing Lyu, Hiroyuki Tomiyama and Lin Meng, "Deep Learning based Japanese Early Books Understanding,"Ninth Annual Conference of Japanese Association for Digital Humanities (JADH2019), 2019.8.

Lin Meng, "Deep Learning based Ancient Literature Recognition and Preservation," the SICE Annual Conference International Conference (SICE 2019), 2019.8.

Bing Lyu, Ryo Akama, Hiroyuki Tomiyama, and Lin Meng, "The Early Japanese Books Text Line Segmentation base on Image Processing and Deep Learning," The 2019 International Conference on Advanced Mechatronic Systems (ICAMechS 2019),2019.8.

Yuxi Chen, Zhiyu Zhang, 富山宏之, 孟林, " 深層学習と語彙分析を用いた高精度な拓本文字認識の実現 , "第 18 回情報科学技術フォーラム(FIT2019),2019.9.

渡邊 清威, 孟 林, 泉 知論, "甲骨拓本からの自動文字抽出手法と支援環境,"画像電子学会 第 288 回研究会,2019.3.

Lin Meng, Masahiro Kishi, Kana Nogami, Michiko Nabeya and Katsuhiro Yamazaki, "Unlocking Potential Knowledge Hidden in Rubbing:Multi-style Character Recognition using Deep Learning and Spatiotemporal Rubbing Database Creation,"EuroMed2018 7th International Conference on DIGITAL HERITAGE, LNCS11196, pp.741-751,2018.10.

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 孟 林, 辻 翼, 泉 知論, 落合 淳思, 山崎 勝弘	4. 巻 47
2. 論文標題 ハフ空間での依存マトリクスを用いた主線分抽出による甲骨文字認識	5. 発行年 2018年
3. 雑誌名 画像電子学会誌	6. 最初と最後の頁 457-466
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Bing Lyu, Zhichen Wang, Hengyi Li, Ami Tanaka, Katsuyuki Funumoto, and Lin Meng	4. 巻 187
2. 論文標題 Deep Learning based Medicine Packaging Information Recognition for Medication Use in the Elderly	5. 発行年 2021年
3. 雑誌名 Procedia Computer Science Volume	6. 最初と最後の頁 194-199
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.procs.2021.04.108	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Lehan Chen, Bing Lyu, Hiroyuki Tomiyama and Lin Meng	4. 巻 174
2. 論文標題 A Method of Japanese Ancient Text Recognition by Deep Learning	5. 発行年 2020年
3. 雑誌名 Procedia Computer Science	6. 最初と最後の頁 276-279
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.procs.2020.06.084	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計23件（うち招待講演 4件／うち国際学会 15件）

1. 発表者名 Lin Meng
2. 発表標題 Culture Heritage Protection based on Artificial Intelligence
3. 学会等名 第九回全国(中国)文字と計算機学術研討会（招待講演）
4. 発表年 2019年

1 . 発表者名 Bing Lyu, Lin Meng and Hiroyuki Tomiyama
2 . 発表標題 Frame Detection and Text Line Segmentation for Early Japanese Books Understanding
3 . 学会等名 In Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods (国際学会)
4 . 発表年 2020年

1 . 発表者名 Bing Lyu, Zhichen Wang, Lin Meng, Katsuyuki Furumoto and Ami Tanaka
2 . 発表標題 Deep Learning based Medicine Packaging Information Recognition for Medication Use in the Elderly
3 . 学会等名 International Conference on Identification, Information and Knowledge in the Internet of Things (国際学会)
4 . 発表年 2020年

1 . 発表者名 Masahiko Atsumi, Syunsuke Kawano, Tomoki Morioka, Lin Meng
2 . 発表標題 Deep Learning Based Ancient Asian Character Recognition
3 . 学会等名 The 2020 International Conference on Advanced Mechatronic Systems (国際学会)
4 . 発表年 2020年

1 . 発表者名 Lin Meng
2 . 発表標題 Culture Heritage Protection based on Artificial Intelligence
3 . 学会等名 第九回全国(中国)文字と計算機学術研討会 (招待講演)
4 . 発表年 2019年

1. 発表者名 Zhiyu Zhang, Zhichen Wang, Hiroyuki Tomiyama, and Lin Meng
2. 発表標題 Deep Learning and Lexical Analysis Combined Rubbing Character Recognition
3. 学会等名 2019 International Conference on Advanced Mechatronic Systems (ICAMechS 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Lehan Chen, Bing Lyu, Hiroyuki Tomiyama and Lin Meng
2. 発表標題 A Method of Japanese Ancient Text Recognition by Deep Learning
3. 学会等名 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI) (国際学会)
4. 発表年 2019年

1. 発表者名 Zhiyu Zhang, Yuxi Chen, Hiroyuki Tomiyama and Lin Meng
2. 発表標題 Rubbing Characters Recognition based on Deep Learning and Lexical Analysis
3. 学会等名 Ninth Annual Conference of Japanese Association for Digital Humanities (国際学会)
4. 発表年 2019年

1. 発表者名 Bing Lyu, Hiroyuki Tomiyama and Lin Meng
2. 発表標題 Deep Learning based Japanese Early Books Understanding
3. 学会等名 Ninth Annual Conference of Japanese Association for Digital Humanities (国際学会)
4. 発表年 2019年

1 . 発表者名 Lin Meng
2 . 発表標題 Deep Learning based Ancient Literature Recognition and Preservation
3 . 学会等名 the SICE Annual Conference International Conference (国際学会)
4 . 発表年 2019年

1 . 発表者名 Lin Meng, Bing Lyu, Zhiyu Zhang, C.V. Aravinda, Naoto Kamitoku, Katsuhiro Yamazaki
2 . 発表標題 Oracle Bone Inscription Detector Based on SSD
3 . 学会等名 International Workshop on Pattern Recognition for Cultural Heritage(PatReCH 2019), in conjunction with ICIAP2019 (国際学会)
4 . 発表年 2019年

1 . 発表者名 Bing Lyu, Ryo Akama, Hiroyuki Tomiyama, and Lin Meng
2 . 発表標題 The Early Japanese Books Text Line Segmentation base on Image Processing and Deep Learning
3 . 学会等名 The 2019 International Conference on Advanced Mechatronic Systems (国際学会)
4 . 発表年 2019年

1 . 発表者名 Zhiyu Zhang, Hiroyuki Tomiyama, Lin Meng
2 . 発表標題 Combining Deep Learning and Lexical Analysis method for Rubbing Character Recognition
3 . 学会等名 人文科学とコンピュータシンポジウム (国際学会)
4 . 発表年 2019年

1. 発表者名 呂冰, 富山宏之, 孟林
2. 発表標題 ARU-NetとLeNetを用いた日本古典籍の自動解読
3. 学会等名 人文科学とコンピュータシンポジウム (国際学会)
4. 発表年 2019年

1. 発表者名 Lin Meng, Masahiro Kishi, Kana Nogami, Michiko Nabeya and Katsuhiro Yamazaki
2. 発表標題 Unlocking Potential Knowledge Hidden in Rubbing: Multi-style Character Recognition using Deep Learning and Spatiotemporal Rubbing Database Creation
3. 学会等名 EuroMed2018 7th International Conference on DIGITAL HERITAGE (国際学会)
4. 発表年 2018年

1. 発表者名 Lin Meng, C.V Aravinda, K R Uday Kumar Reddy, Tomonori Izumi and Katsuhiro Yamazaki
2. 発表標題 Ancient Asian Character Recognition for Literature Preservation and Understanding
3. 学会等名 EuroMed2018 7th International Conference on DIGITAL HERITAGE (国際学会)
4. 発表年 2018年

1. 発表者名 Lin Meng, Naoki Kamitoku and Katsuhiro Yamazaki
2. 発表標題 Recognition of Oracle Bone Inscriptions Using Deep Learning based on Data Augmentation
3. 学会等名 2018 IEEE International Conference on Metrology for Archaeology and Cultural Heritage (IEEE MetroArchaeo 2018) (国際学会)
4. 発表年 2018年

1. 発表者名 孟 林
2. 発表標題 画像処理と深層学習を用いた古代文献の解読と整理
3. 学会等名 電気学会、制御研究会
4. 発表年 2019年

1. 発表者名 渡邊 清威, 孟 林, 泉 知論
2. 発表標題 甲骨拓本からの自動文字抽出手法と支援環境
3. 学会等名 画像電子学会 第288回研究会
4. 発表年 2019年

1. 発表者名 Lin Meng
2. 発表標題 Artificial Intelligence for Computer Science and Image Recognition
3. 学会等名 ICETE-2018 (International Conference on Emerging Trends in Engineering) (招待講演)
4. 発表年 2018年

1. 発表者名 Lin Meng
2. 発表標題 Oracle Bone Inscriptions Recognition by Image Processing and Deep Learning
3. 学会等名 the 1st Int. Symp. On Oracle Bone Inscriptions Information Processing (招待講演)
4. 発表年 2018年

1. 発表者名 孟 林, 孟澤林, 王志辰, 富山宏之
2. 発表標題 横断分野での深層学習について
3. 学会等名 電気学会・制御研究会
4. 発表年 2020年

1. 発表者名 森岡 智樹, 河野 竣介, 渥美 柁彦, 孟 林
2. 発表標題 深層学習モデルを用いた甲骨字認識システムの構築
3. 学会等名 電気関係学会関西連合大会
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

甲骨文字（拓本）・くずし字認識システム http://www.atait.se.ritsume.ac.jp/CharacterRecognition/
--

6. 研究組織		
氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------