

深層学習による木簡欠損部位の画像復元

塩野 健治（東京電機大学） 大山 航（東京電機大学）

概要: 木簡は、過去の生活状況を伝える重要な資料であるため、出土した削り屑木簡の補完、解読が課題とされている。本稿では、木簡文字読解の支援を目的とした、深層ニューラルネットワークによる木簡欠損部位のデジタル画像上での復元手法を提案する。また、復元画像の画質の定量評価および復元された木簡文字画像を用いた画像検索実験により提案手法の有効性を評価した。

キーワード: 木簡, 画像復元, 機械学習, 深層学習, 敵対的生成ネットワーク

Image Completion for Missing Region of Mokkan Fragments by Deep Learning

Kenji Shiono (Tokyo Denki University) Wataru Ohyama (Tokyo Denki University)

Abstract: Mokkan are important materials that convey information about daily life in the past, so it is a challenge to supplement and decipher excavated shaved mokkan. In this paper, we propose an image completion method using a deep learning network for the missing parts of mokkan images to support the decipherment of mokkan images. We also prove the reliability of the completed images through qualitative and quantitative evaluation experiments.

Keywords: Mokkan, Image Completion, Machine Learning, Deep Learning, Adversarial Generative Network,

1. はじめに

木簡とは、文字や記号を記した木片古文書の総称である。木簡は主に日本、中国や韓国などの東アジア地域で使用されており、国内では藤原京や奈良平城京跡などの遺跡から2011年末までに38万点以上が出土している[1]。これらの資料の解読により、当時の社会構造、政治、経済、文化、宗教だけでなく、言語や文字の変遷、社会の発展、知識伝達の方法などについての情報を得ることができる。

本来、木簡は図1(a)のように短冊形であるが、表面を削りとり再利用されていたため、削りとられた木簡の断片（削り屑木簡と呼ばれる）が図1(b)のように記載文字が欠損した状態で大量に出土している。歴史学、考古学上、木簡は過去の生活状況を伝える重要な資料であるため、削り屑木簡を含め、記載文字や内容の解読が必要不可欠である。出土した削り屑木簡の多くは木簡形状や墨痕、筆跡などの情報が欠損しているため、記載されている文字の読み取りは困難である。資料点数が膨大であるため手作業での網羅的な解読は困難であり、迅速かつ正確性の高い作業を実現するために、情報技術、人工知能技術に基づく作業支援が必要とされている。

本稿では木簡文字読解の支援を目的に、深層ニューラルネットワークを活用して木簡欠損部位



図1 木簡(a)と削り屑木簡(b)の例

をデジタル画像上で復元する手法を提案する。また、復元画像の画質の定量評価および復元された木簡文字画像を用いた画像検索実験により提案手法の有効性を評価した。

2. 関連研究

画像処理技術を活用した木簡文字の可読性向上に関する先行研究例が存在する。ここでは関連する研究の代表例を紹介しつつ、本研究との違いを説明する。

大山らは深層学習による画像変換で木簡の実測図を自動生成する手法を提案した[2]。この手法は木簡の木目を除去し文字だけを残した画像を

生成するが、木簡文字が欠損していた場合には、その欠損部分を補うことはできない。

Truyen らは色に注目した特徴を用いて削屑木簡の画像をグループ分けした上で、局所的な輪郭形状を用いた削屑木簡の再構成手法を提案した[3]。この手法は削屑木簡どうしを繋ぎ合わせることで欠損部分を補おうとするアプローチであり、繋ぎ合わせる相手が存在しない削屑木簡の欠損部位の復元は不可能である。

近年、深層学習技術の活用によりデジタル画像処理を自動化する研究開発が進んでいる。画像に写り込んだ不要な物体を削除したり、欠損した画像領域を補完したりするイメージインペイティング技術[4]はその典型例である。本研究ではこのインペイティング処理を活用して、木簡文字画像の欠損部位の画像復元を試みる。

3. 提案手法

本研究では画像生成を行う深層学習技術の一種である敵対的生成ネットワーク (GAN) [5]を用いて、木簡文字画像の欠損部分を画像復元する手法を提案する。

提案する手法の基本アプローチを図 2 に示す。提案手法の目的は、文字が欠損した削り屑木簡画像を入力とし、欠損部分を復元した木簡画像を出力とする画像生成器 (図中の G) を機械学習により構築することである。訓練用の木簡文字画像に対して、マスキング処理により擬似的な欠損画像を生成し、GAN を用いて画像生成器を学習する。

GAN は、生成器 (generator) と識別器 (discriminator) の二つのネットワークから構成される。このモデルは、生成器と識別器が対立することから "敵対的" と呼ばれる。生成器は、識別器が生成画像を本物の画像であると誤認識させるような画像を生成するように最適化される。識別器は生成画像を本物と誤認識することを防ぐように最適化される。このように、生成器と識別器が互いに敵対しつつ改善し合いながら機械学習が行われる。

具体的な手法として、本研究では Nazeri らの提案した 2 段階の GAN で画像のインペイティング処理を行う手法である edge-connect[5]を用いて木簡画像の復元を試みる。図 3 に提案手法の流れを示す。この手法は輪郭生成器と画像補完ネットワークの 2 段階で構成される。第 1 段階の輪郭生成器では画像マスクで指示された木簡画像の欠損部分の輪郭を推定した輪郭マップを生成し、第 2 段階の画像補完ネットワークでは生成された輪郭マップと原画像を入力として、ネットワークが欠損領域の画素値を補完する。

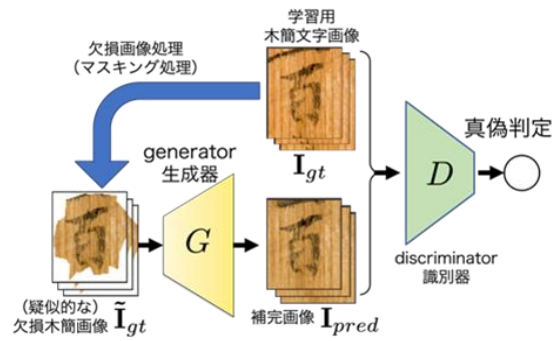


図 2 提案手法の基本構成

3.1 輪郭生成器

輪郭生成器 (G_1) は、グレイスケール化された欠損した木簡画像、残存部位の輪郭画像、欠損部位を示すマスク画像の 3 点を入力し、画像の欠損部位に含まれる輪郭を推定した輪郭マップを出力する。木簡画像において輪郭マップは主に木目や文字輪郭の情報を含み、欠損部位の画像復元に有用な情報となる。

輪郭生成器は前述の通り GAN を構成し学習する。輪郭生成に関する識別器 (D_1) は、学習用画像データに対して生成器が生成した輪郭マップが本物の輪郭線画像と同一か否かを判定する。同時に、識別器は識別器内で生成した輪郭マップと本物の輪郭線画像との間で特徴量のマッチングを行い、マッチング誤差を算出する。生成機は識別器の判定エラーを最大化しつつマッチング誤差を最小化するような輪郭マップを生成するように学習される。

3.2 画像補完ネットワーク

画像補完ネットワーク (G_2) は、欠損部位を含む木簡のカラー画像と輪郭生成器で生成された輪郭マップを入力し、欠損部位を復元した木簡のカラー画像を出力する。

画像補完ネットワークも GAN を構成し学習する。画像保管に関する識別器 (D_2) は、復元された木簡画像が本物の木簡画像と見なせるか否かを判定する。画像補完ネットワークは識別器の判定エラーを最大化しつつ、復元された画像と学習用画像との違いを最小化するように学習される。

4. 実験用データセット

画像復元を行う深層ニューラルネットワークの学習には、欠損を含まない画像と欠損部位を示す欠損マスクが必要である。学習段階で欠損のない画像を欠損マスクでマスキング処理して仮想的に欠損部位を含む画像を作成し、欠損部位の画像情報を復元する学習を行う。

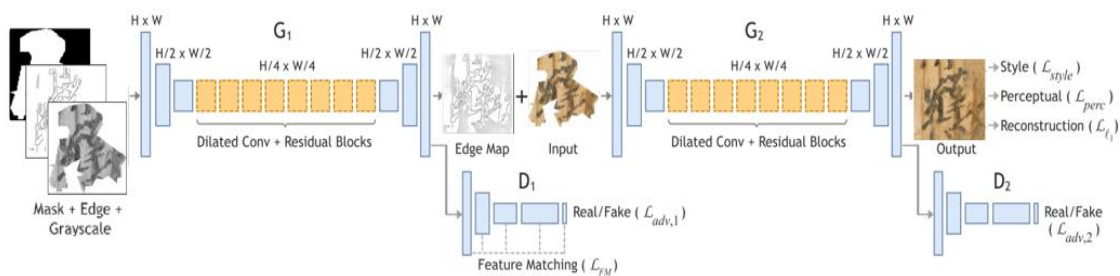


図3 提案手法の流れ

本研究では、欠損のない画像として奈良文化財研究所から提供された木簡文字画像を用いた。これらの文字画像は、図4に示すように木簡から1文字ずつ切り出された19,741文字の画像である。これらの画像をランダムに訓練用(10,000点)、検証用(5,000点)、評価用(4,741点)の3グループに分割し、ニューラルネットワークの学習と復元画像の精度評価実験に用いた。

欠損マスクは、現実起こりうる欠損を反映することが望ましい。本研究では奈良文化財研究所から提供された削屑木簡の画像を用いた。これらの画像は削屑木簡1点ずつを撮影した画像である。削り屑木簡画像に対して、木簡領域、背景領域をそれぞれ白色(残存部位)、黒色(欠損部位)とする2値化処理を適用し欠損マスクを作成した。欠損マスクの例を図5に示す。使用した削屑木簡画像は36,136点である。これらを訓練用(18,568点)、検証用(9,784点)、評価用(8,784点)の3グループに分割した。

画像復元の精度は欠損部位の大きさに影響を受けると考えられる。欠損部位の大きさが画像復元の性能に与える影響を検証できるように、欠損マスクを画像中で欠損領域が占める面積の割合に基づいてグループ化した。欠損マスク全体における各グループの内訳は、欠損領域の大きさが大きい順に欠損領域70%以上(180点)、欠損領域40%以上(10,993点)、欠損領域10%以上(25,948点)である。欠損領域が10%未満となる欠損マスクは評価には使用しない。

ネットワークの学習に使用するデータセットは、木簡文字画像と欠損マスク画像をランダムに組み合わせて作成した疑似的な欠損木簡画像で構成される。図6に生成された疑似欠損木簡画像の例を示す。

5. 評価実験及び結果

前節で述べたデータセットを用いて、欠損木簡の画像復元実験を行った。深層ニューラルネットワークの学習には学習用および検証用グループの木簡画像と欠損マスクを用いた。一方、評価用グループを対象に、学習が完了した深層ニューラ



図4 木簡文字原画像の例



図5 削屑木簡画像から作成した欠損マスクの例

ルネットワークによる欠損部位復元を適用した。学習用および検証用グループと評価用グループとのそれぞれに含まれる木簡画像と欠損マスクは重複していない。この実験条件において深層ニューラルネットワークは未見の欠損木簡画像を復元することになり、現実の利用場面を想定した実験条件と言える。実験では、復元画像の画質評価と、復元画像の木簡研究における有用性評価、および木簡研究者による定性評価を行った。

5.1 復元画像の画質評価

まず、復元画像の目視評価と画質の定量評価を行う。図7に提案手法による欠損木簡の画像復元処理結果の例を示す。図中(a)から(d)はそれぞれ欠損領域の大きさが画像サイズの約30%、40%、60%、80%の例である。それぞれの結果において、左側、中央、右側の画像はそれぞれ、原画像（木簡文字画像）、マスクにより欠損させた画像（入力画像）、復元結果画像である。これらの画像例から、木簡の欠損部位が大きい場合でも、文字の概形が推定できる程度には画像を復元できていることがわかる。一方で、欠損領域の大きさが結果画像の精度に大きな影響を与えることも確認できる。文字の欠損領域が大きいほど、木簡文字に対する復元精度が低い。

続いて、復元結果画像の定量評価実験を行った。定量評価尺度として、復元画像のピーク信号雑音比(PSNR)、元画像と復元画像との間の構造類似性尺度(SSIM)と平均絶対誤差(l_1)、および原画像グループと復元画像グループとの間の Fréchet Inception Distance(FID)の4指標を用いた。これらの評価値は画像のインペインティング処理手法の評価に広く用いられている指標であり、PSNRとSSIMはその値が大きいほど、 l_1 とFIDは値が小さいほど復元画像の画質が優れていることや復元画像と原画像との違いが小さいことを示す。図7(a)から(d)の復元結果の例には、PSNR、SSIM、 l_1 の値を合わせて示した。表1に評価用画像グループに対して算出した各評価尺度の平均値を示す。参考に一般画像を対象とした先行研究[5]における評価尺度値を併記した。木簡文字画像に対して、一般画像より優れる評価結果となり、木簡文字画像の復元に深層ネットワークが適用できることが示唆される。

表 1 定量評価実験結果

評価指標	木簡文字画像	先行研究
PSNR(↑)	24.8	21.8
SSIM(↑)	0.897	0.823
l_1 (%) (↓)	3.02	3.86
FID(↓)	6.20	8.16

↑, ↓はそれぞれ、値が大きい、小さいほど復元画像の画質が優れていることを示す。

5.2 輪郭生成器の必要性

前述の通り、提案手法は2段階の処理で欠損木簡を画像復元する。木簡画像においては、木目や文字輪郭が重要な情報となることを想定し、第1段階の輪郭生成器を使用して欠損部位の画像輪郭を推定するが、この第1段階の推定処理の有無が、画像復元結果に与える影響を調査した。



図 6 作成した欠損木簡画像



(a) 欠損面積 28.0%
PSNR:25.6; SSIM:0.972; l_1 (%):2.12



(b) 欠損面積 37.2%
PSNR:20.6; SSIM:0.862; l_1 (%):4.37



(c) 欠損面積 58.0%
PSNR:20.4; SSIM:0.711; l_1 (%):5.41



(d) 欠損面積 80.0%
PSNR:22.5; SSIM:0.442; l_1 (%):9.39

図7 提案手法による欠損木簡の画像復元結果例。各図左から右へ元画像、欠損画像、復元画像である。

図8に輪郭生成あり(a), なし(b)の場合の復元結果画像の例を示す. 輪郭推定なしの例では欠損部位において画像のぼけが発生するが, (b)の輪郭推定ありの例ではぼけが軽減され, 文字が判読しやすくなっていることが確認できる.

輪郭推定の有無の影響を先述の定量評価尺度を用いて評価した. 表2にその結果を示す. 輪郭推定ありの場合が全ての評価尺度において良好な結果となり, 輪郭生成器が復元結果の画質向上に効果的に働くことが示された.

表2 輪郭生成あり/なしの場合の復元画像画質の定量比較

評価指標	輪郭推定あり	輪郭推定なし
PSNR(↑)	24.8	19.0
SSIM(↑)	0.897	0.823
l_1 (%) (↓)	3.02	8.29
FID(↓)	6.20	33.9

5.3 補完画像の有用性評価

欠損部位の画像復元が削り屑木簡に記された文字の翻刻に応用可能かを検証するために, 定量評価として画像復元を併用した木簡画像検索の実験を行った. 使用した画像データは評価用にした4,741点である. 欠損がある入力画像(クエリ画像)に対して, 提案手法で画像復元を行った後, 畳み込みニューラルネットワークであるVGG16モデル[7]で画像特徴を抽出する. 同様に画像特徴を抽出した登録画像から, クエリ画像の画像特徴との類似度が高い順に画像を検索する.

検索結果の1例を図9に示す. 図中の上段, 中段, 下段はそれぞれ, クエリ画像に対する復元を提案手法(2段階処理)により行った場合, 復元を輪郭推定なし(1段階処理)で行った場合, 復元を行わない場合の検索結果である. 各結果の左端画像(青色の枠)がクエリ画像である. 検索結果は左から順にクエリ画像との類似性が高い. 検索結果の青色枠は検索結果画像とクエリ画像の文字コードが一致していることを示す. この例において, 欠損した画像では正しく文字画像を検索できないが, 提案手法による画像補完により正しく検索できるようになった.

検索結果において, クエリ画像と同じ文字種の画像を挙げた場合に検索成功とし, 異なる欠損面積に対してそれぞれの検索順位における検索成功率を求めた結果を図10に示す. 併載されている欠損木簡画像は該当する欠損サイズの例である. 図中折れ線グラフの縦軸と横軸はそれぞれ, 検索成功率と検索順位である.

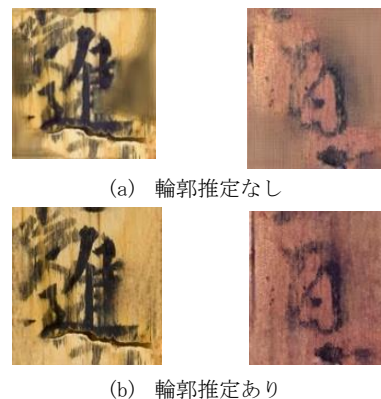


図8 輪郭推定の有無が復元結果に与える影響

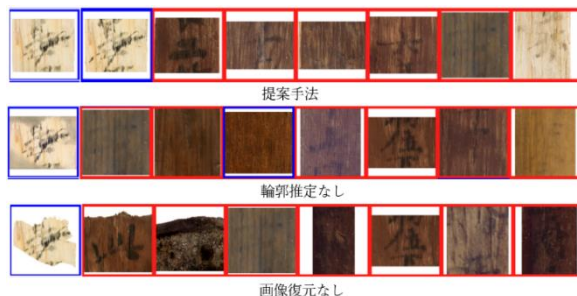


図9 欠損木簡を用いた木簡画像検索結果の例

欠損面積の拡大により, 補完結果の検索正解率は低くなるが, いずれの場合でも提案手法(2段階手法)による画像復元を用いた場合が全ての検索順位において最も大きい検索成功率となった. 以上のことから, 提案手法による画像復元の木簡文字翻刻への応用可能性が示唆された.

5.4 復元画像の定性的評価

復元画像の定性的評価として, 提案手法による木簡復元画像と上述した実験結果を複数の木簡研究者に提示し, 復元画像の正確性や木簡研究への有用性に関するコメントを依頼した.

- (1) 木簡の視認効果向上の点では, 背景の復元(補完)により, 完形を想像しやすくなる.
- (2) 墨の濃淡を強調させるような処理があると, 文字判別の補助になる可能性がある.
- (3) 断片どうしの綴合(接続)の推察する補助となる可能性がある

等のコメントが得られ, 専門家の観点からも木簡研究への貢献に期待が寄せられている.

6. まとめ

本研究では, 敵対的生成ネットワークを利用して, 欠損木簡に対して画像復元を行った. 提案手法により, 背景や筆跡, 墨の濃淡など木簡研究者にとって役立つ情報を復元できる機能が確認さ

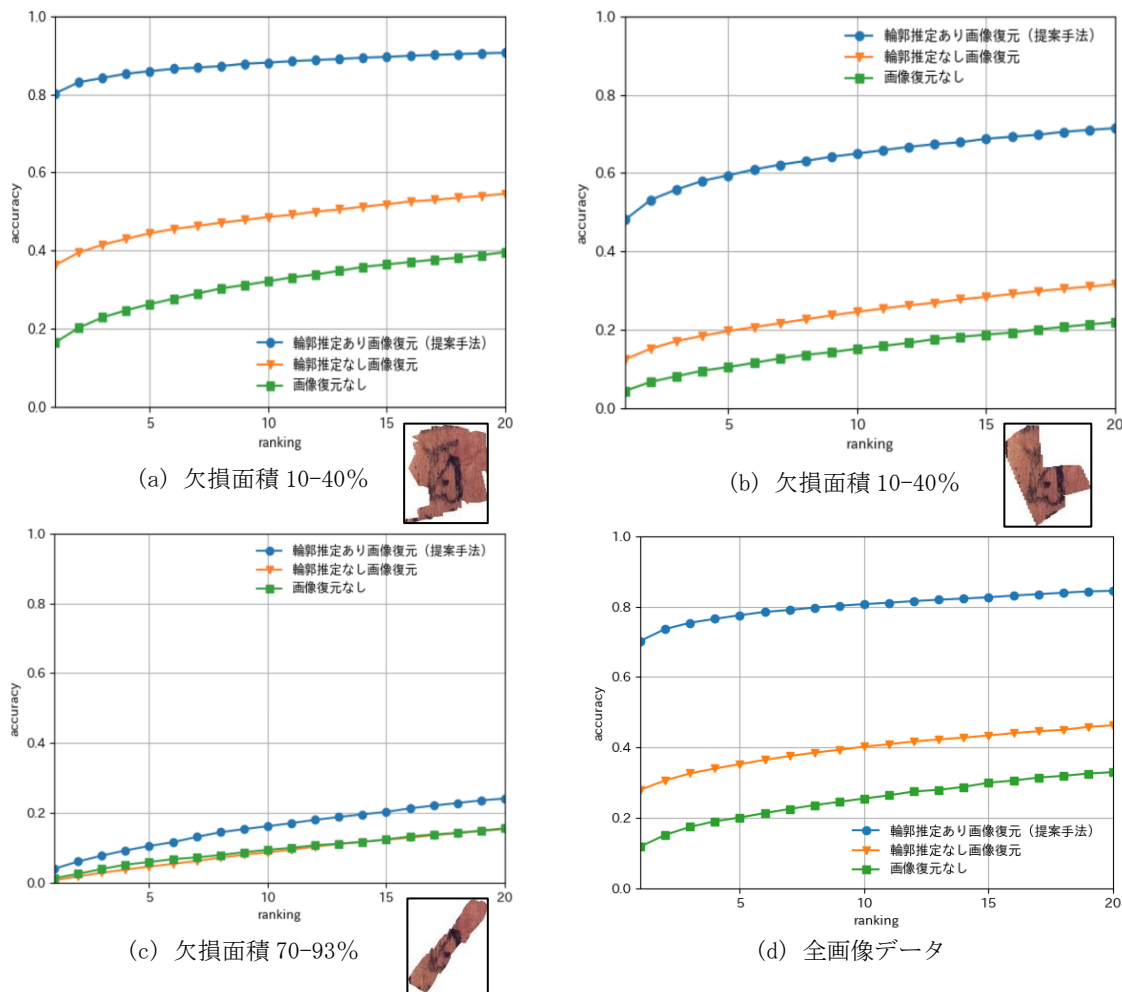


図 10 画像復元を併用した木簡画像検索実験の結果

れた。また、異なる欠損程度による生成結果の精度変化についても分析を行った。

今後の研究課題は、より多くの欠損木簡画像に対して欠損部分を利用者が手入力で指示して補完処理を行う機能の実現や、その他のインペイティング手法の補完性能の比較、複数の補完結果候補を提示する機能の実現、一文字ずつの判読の結果を組み合わせて連続する文字として文脈情報を付加した分析などが挙げられる。

謝辞

本研究の一部は JSPS 科研費 18H05221 の助成を受けたものです。また、奈良文化財研究所より木簡画像データの提供と定性評価への協力をいただきました。謹んで感謝の意を表します。

参考文献

[1] 市大樹. 飛鳥の木簡 古代史の新たな解明. 中央公論新社〈中央新書〉, 2012, ISBN 9784121021687

[2] 大山 航, 畑野 吉則, 馬場 基. 深層学習による木簡実測図の自動作成. 人文科学とコンピュータシンポジウム(じんもんこん)2020 論文集, 2020, pp.235-240.

[3] Truyen Van Phan, Hajime Baba, Akihiro Watanabe, Masaki Nakagawa. A re-assembling scheme of fragmented Mokkan images. Proceedings of the 2nd International Workshop on Historical Document Imaging and Processing (HIP '13), 2013, pp.22-28.

[4] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester. Image Inpainting. Proc. SIGGRAPH2000, 2000, pp.417-424.

[5] Kamyar Nazari, Eric Ng, Tony Joseph, Faisal Z. Qureshi, Mehran Ebrahimi. Edge Connect: Generative Image Inpainting with Adversarial Edge Learning. arXiv:1901.00212, 2019.

[6] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative Adversarial Networks. Advances in neural information processing systems 27, 2014.

[7] Karen Simonyan, Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv:1409.1556, 2014.