

FCN に基づく 3D 冊子体 CT データの可視化

韓 忠江[○] (京都大学), 土井 章男 (岩手県立大学), 小山田耕二 (京都大学)

Visualization of 3D booklet CT data based on FCN

Zhongjiang HAN, Akio DOI, Koji KOYAMADA

ABSTRACT

CT technology has been widely used in medical, geographic, architectural, and other industries. In recent years, significant contributions have also been made to archaeological and ancient cultural relics research. We have found that in the face of fragile paper ancient literature, how to digitize its content on the basis of protecting it from being destroyed, and provide valuable information for scholars becomes a problem. Therefore, we have digitally reconstructed the paper text in which metal-containing inks are applied for the printing process in CT machines, and used FCN-based deep learning to solve the problem of page bending. We experimentally demonstrated the feasibility of obtaining the internal information of the booklet through CT scanning, using deep learning to clear the printed text, and demonstrating the effectiveness of deep learning technology in restoring CT data of affine transformation books.

Keywords: CT, Booklet, Page extraction, Deep learning, 3DFCN

1. 序 論

CT 技術、特に X 線ベースの c t スキャンイメージング技術の誕生以来、様々な分野で広く使用されてきた。X 線は、非常に短い波長および大きなエネルギーを有する電磁波である。X 線の波長は、可視光の波長 (0. 0 0 1 ~ 1 0 n m 程度、葉に照射される X 線の波長は 0. 0 0 1 ~ 0. 1 n m 程度) よりも短い。その光子エネルギーは、可視光の光子エネルギーよりも数万から数十万倍大きい。それは 1895 年にドイツの物理学者 WK レントゲンによって発見されたので、レントゲン光線とも呼ばれる。X 線は高い浸透能力を持ち、インク紙、木、そして紙のような可視光線に対して不透明な多くの材料を通り抜けることができる。私たちはこの X 線の特性のために視覚研究のための CT 装置の使用に触発され、私達は CT 装置が古代文学に基づいて内部のテキストを視覚化するという研究上の問題に対処できると信じている。実際には、CT 機を使用した古

代のシープスキンロールの仮想的な拡大に関する研究があり、それは使用が保証されている限り、そのインクが X 線によって認識されることができる画像を与えることができる金属元素を含みます紙やシープスキンとは異なる。

事実、古代中国と古代日本に代表される古代文学は西洋とは異なり、ロールではなく小冊子に似ている。素材は動物の皮とは異なり、非常に薄いライスペーパーなので、基礎からの研究が必要である。

我々は実験的に CT スキャンを通して小冊子の内部情報を得ること、印刷されたテキストをクリアするためにディープラーニングを使用すること、およびアフィン変換ブック CT データを復元することにおけるディープラーニングの有効性を実証した。

2. FCN における深層学習

これまでの研究で、特殊インクを使用して冊子体をデジタル化するためにCTマシンを使用することが実行可能であることを実証した。また、Active grid 法はページ抽出とページ曲げ問題をテストするために使用され、実験結果によって、この方法が実行可能であることを示したが、改良すべきところも多数であった。これまでの研究では、様々な曲線ページに直面して、様々なパラメータを一つ一つに対応することによって実験時間を長くし過ぎることが大きな問題であることがわかった。そこで、この問題に対処するための共通の方法を見つけようとし、深層学習を選択した。

このセクションでは、3DFCN の原理とディープラーニングを使用する理由について詳しく説明し、実験の設計と結果の議論をする。

2.1 関連研究

中国では、歪んだページの回復に関する多くの研究がある。従来の方法は、多数のページグレースケールデータと曲線モデルを組み合わせ、ページのデータベースを作成することである。それからデータベースのモデルに対してページをスキャンし、データベースに対応することによって復元される。いくつかの研究は実験で高い精度を示しているが、実際にデータベースはすべてのページの曲げを含んでいるわけではない。

私たちはすべてのページの折り曲げ問題に対処するための一般的な方法を見つけようとしてきました。これが、次の実験で3Dモデル抽出を実行するためにディープラーニングを使用することにした理由です。次節では3DFCNを用いた画像処理実験について紹介する。

2.2 FCN フレームワーク

3DFCN は2DFCN に基づいており、コンピュータは3Dモデルを直接操作できる。つまり、コンボリユーションカーネルは3次元行列になるため、CPUに膨大な量の計算を行わせることになる。GPUを主な操作として使用しても、メモリがオーバーフローしないようにするためにまだ小さいデータを入力する必要がある。

本実験で使用した3DFCNディープラーニングフレームワークは、128×128×128サイズのボリュームデータを入力とし、コンボリユーションとデコンボリユーションによって最終的に出力する。

実際、小冊子の単語数が少ないほど、CTマシンの精度を向上させる必要がある。さもなければ、ノイズの影響で、結果として生じる画像は区別するのが困難になるでしょう。同時に、高精度CTマシンの使用は画像ピクセルを増加させ、それは3D再構成の時間を浪費し、再構成された3Dモデルデータサイズは数GB以上になる。モデル全体を一度に入力するのは明らかに非現実的だ。

用意された実験冊子は100枚のA4用紙で構成されている。ディープラーニングを正常に進めることができることを確認するために、128×128×128の小さなブロックにダイシングすることを実験した。幸いなことに、そのような条件下ではそれ自体小さい文字の印刷されても、はっきりして識別可能になるでしょう。

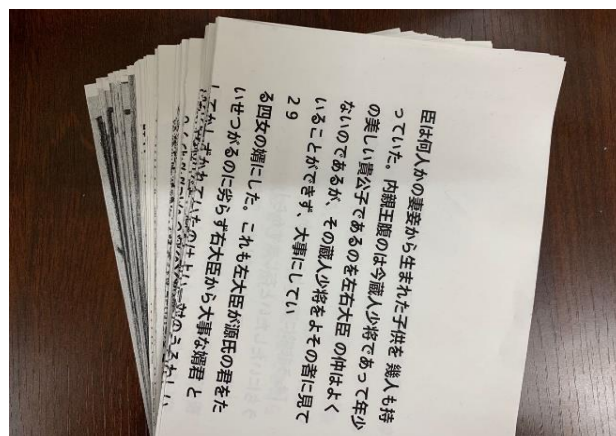


Fig. 1. Experimental booklet (A4 100 pages double-sided printing)

3. 実験

本章では、FCNに基づくディープラーニングに用い、アフィン変換モデルの復元という実験を設計した、実験結果はアフィン変換小冊子モデルを復元するため、深層学習の有効性を実証できる。

CTスキャン処理中に、ページの重なり、ノイズ、ジッターなどが原因でCT画像が歪んだり、コンテンツの一部が失われたり、文字の明るさが不均一になったりする可能性がある。基礎として3DFCNを使用して、我々は再構成された小冊子3Dモデルを訓練し、きれいなスライスを得るために正しいデータとして小冊子のきれいな画像を使用した。ページ数が多いため、スキャン時にページが平らになり、結果として得られる3Dモデルスライスの文字を直接認識できる。深層学習によって簡単な本のページ抽出

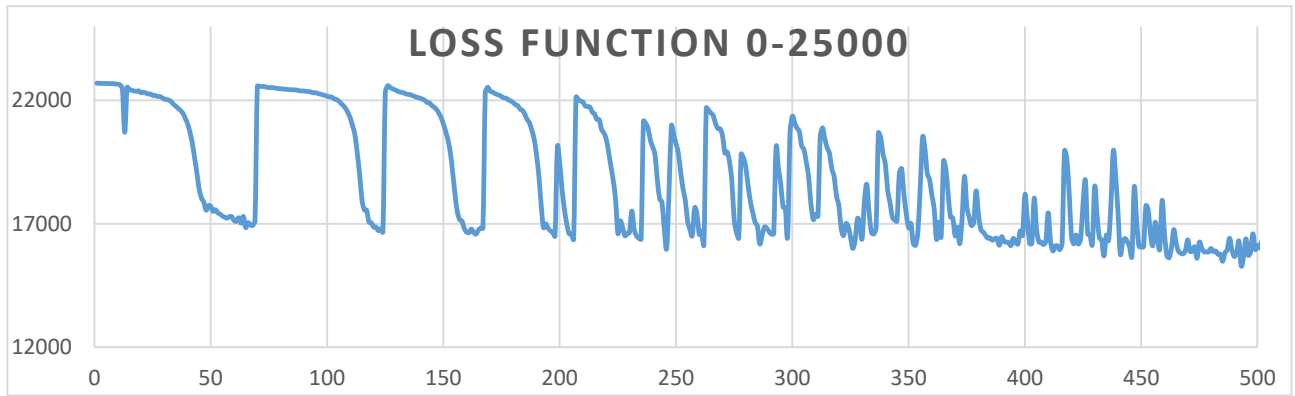


Fig.2. Loss function in the deep learning process of FCN framework

の問題を解決できることが証明できる。

単純なモデルを処理するために学習を使用することは研究目的ではない。我々は、さまざまな問題に対処するため、用途の広い方法を開発することを目指す。小冊子の視覚的研究では、ページの不規則な曲がるは最も一般的な問題です。前の研究では、アクティブなグリッドを使用してページの曲率を処理し、曲がったページを抽出して平面上に投影し、完全に識別可能なページを取得した。また、アクティブグリッドを使用する際の制限についても議論し、このジレンマを解決するためにディープラーニングを使用することにした。

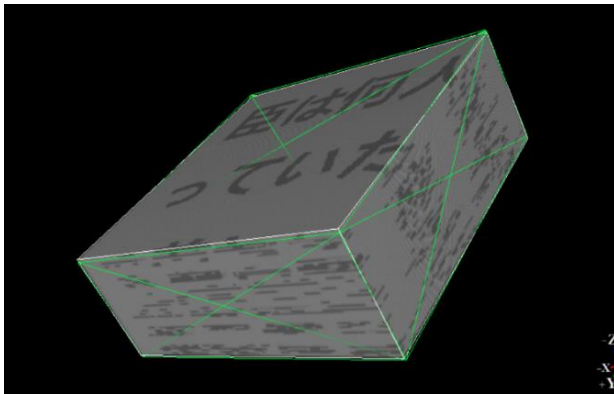


Fig.3 The dicing of the experimental model
(size 128×128×128)

半教師つき学習法として、FCN はかなりの数のトレーニングデータセットと正しいデータセットを提供する必要があります。曲がったページの問題を解決する際に、我々は深層学習実験のために以下の実験計画を設定しました：

- 1) 正しいデータとして小さなダイシングブロックを択する

- 2) アフィン変換を使用して、x 方向に 5°、y 方向に 10°のようにデータを調整する

- 3) 学習プロセスが収束するように、回転していない同じ正しいデータを正解データとし、複数の回転後モデルデータをトレーニングする

- 4) トレーニングモデルの後に他のデータの入力をテストする

4. 結果

6 セットのデータについて 10 万回以上トレーニングした。損失関数は直接ピークを受け取らないが、収束の全体的な傾向は大きく、そして小さい範囲で大きな変動がある。これは、6つのデータセットが同時に訓練されるのではなく順番に訓練されるためだ。つまり、6つのデータセットを順番に得ることができます(これらの結果は同じデータになる傾向があります)。学習は終了します。結果によると、データの基本的な回復の滑らかさを決定し、元のデータのテキスト部分を抽出することができ、これはアフィン変換後に書籍データを復元するためにディープラーニングを使用することが有効であることを示した。



Fig.4 x, y, z direction rotation restoration experimental result

5. 結論

我々の方法は、データ処理のためにCTマシンを通して小冊子をスキャンすることが実行可能であることを実証した。それはまた3DFCNベースのディープラーニングフレームワークが単純なページ抽出問題と曲がったページ問題を解決できることを示した。実際の古代文学に関する視覚的な研究はないが、我々は深い学習実験を用いて将来の研究の基礎を提供するための一般的な解決策を提案しようとする。

今後の課題として、各ページのデータをより完全に可視化するなど、現段階で存在するいくつかの問題を解決し、新しいモデルを再構築し、関連する文献作者が研究を行えるようにする。より複雑な小冊子データを処理するために既存のモデルを改良するために、新しいデータセットを作成する。可視化された結果ページにテキスト認識を追加し、視覚化の結果を評価するためのより科学的なアプローチを使用しようとする。より複雑な問題を解決するための3DFCN以外の深層学習フレームワークを構築し、さらに古代中国と日本の古代文学を研究し、対応する解決策を提案すること。

参考文献

- 1) Xu Y, Xu C, Kuang X, et al. 3D-SIFT-Flow for atlas-based CT liver image segmentation. Med Phys. 2016; 43:2229–2241.
- 2) E. Bulska, B. Wagner, B. Stahl, M. Heck, H.M. Ortner R. Van Grieken, K. Janssens, L. Van't dack, G. Meersman (Eds.), Seventh International Conference on Non-destructive Testing and Microanalysis for the Diagnostics and Conservation of the Cultural and Environmental Heritage Proceedings, University of Antwerp, Belgium (2002)
- 3) R. Bade, J. Haase, B. Preim, Comparison of fundamental mesh smoothing algorithms for medical surface models, in: Proceedings of the Simulation and Visualization, 2006, pp. 289–304.
- 4) Shafarenko, L, Petrou, M, and Kittler, J., Histogram-based segmentation in a perceptually uniform color space, IEEE Trans. on Image Processing, IP-7(9), pp. 1345–1358, Sep. 1998.
- 5) P. Campadelli, E. Casiraghi, G. Lombardi, Automatic liver segmentation from abdominal CT scans, in: Proceedings of the 14th International Conference on Image Analysis and Processing, ICIAP '07, 2007, pp. 731–736.
- 6) X. Tang, J. Hsieh, R. Nilsen, S. Dutta, D. Samsonov,

and H. Hagiwara, “A three-dimensional weighted cone beam filtered backprojection (CB-FBP) algorithm for image reconstruction in volumetric CT - helical scanning,” Phys. Med. Biol. 10.1088/0031 - 9155/51/4/007 51, 855–874 (2006).

- 7) Tompson, J., Jain, A., LeCun, Y. & Bregler, C. Joint training of a convolutional network and a graphical model for human pose estimation. In Proc. Advances in Neural Information Processing Systems 27 1799–1807 (2014).