

Pix2Pixを用いた古典籍くずし字画像の裏抜け除去

植田ちひろ[†] 藤岡寛之[†] 日高章理^{††}[†]福岡工業大学情報工学部 ^{††}東京電機大学理工学部

1 はじめに

本研究では、機械学習により文字の芸術的品質を探ることを念頭におき、その準備として古典籍のくずし字画像データ [1] から墨の裏抜けを除去する問題を考える。この種の裏抜け除去問題は、従来から両面印刷された出版物などの OCR の読み取り問題などで扱われてきた。例えば、表裏面の画素値ペアの輝度値の分布に対してクラスタリングを行い、裏抜けを除去する方法などが開発されている [2]。しかし、これらの従来法では古典籍のくずし字画像のように、表面は劣化等による色斑がある和紙上に筆文字独特のかすれやにじみをもつ文字が書字されており、さらに裏面からの和紙への墨汁の裏抜けが強い場合があり、うまく裏抜き除去することが難しい。

そこで、本研究では古典籍くずし字画像の裏抜け除去問題を解決するために、Generative Adversarial Network (GAN) の一種であり、画像データのペア集合間の変換方法を学習可能な Pix2Pix[3] を導入する。その際の学習用の画像データとして、毛筆文字フォント画像とそれに人工的に生成した裏抜けモデルを付加した文字画像のペア集合データを用いる。本手法の有用性を古典籍のくずし字画像を用いた実験により検討する。

2 Pix2Pix による古典籍くずし字画像の裏抜け除去法

ここでは、図1の例に示すような、墨の裏抜けを含むかもしれない文字画像 (図1(a)) が与えられており、これから裏抜き除去を行う、と同時に後に文字の芸術的品質を探ることを念頭に紙材部分も除去した文字* (図1(b)) を得る問題を考える。以下では、2.1 節において裏抜け除去を行うために導入する Pix2Pix について述べ、その実装方法について 2.2 節で述べる。

(a) 元画像 x (b) 裏抜き除去画像 y

図1: 古典籍くずし字画像の裏抜き除去例

2.1 Pix2Pix

今、墨の裏抜きを含むかもしれない文字の元画像を x (例: 図1(a))、裏抜き除去文字画像を y (例: 図1(b)) とし、それらのペア画像集合が与えられているとする。このとき、Pix2Pix[3] により、画像 x から集合 y への写像の学習を行い、その写像を裏抜き除去法として利用する。

x から y への生成写像を $G: x \rightarrow y$ とする。加えて、裏抜き除去文字画像 y が生成器 G により生成された偽のデータか否かを判別する鑑別器を D と定義する。このとき、文字画像データ集合 $\{(x, y)\}$ を利用して、生成器 G^* を次のように得る。

$$G^* = \arg \min_G \max_D L_{cGAN}(G, D) + \lambda L_{L_1}(G) \quad (1)$$

ただし、 $L_{cGAN}(G, D)$ および $L_{L_1}(G)$ は以下のように与えられる。

$$L_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))] \quad (2)$$

$$L_{L_1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_1] \quad (3)$$

(1) 式では、(2) 式の Conditional GAN (cGAN) の評価関数 $L_{cGAN}(G, D)$ により本物らしい画像を生成するため (3) 式の L_1 正則化項を加えている。このときの L_1 正則化項のみでは画像の高周波成分をうまく生成できず、そのため鑑別器を D には与えられた文字画像を小さいサイズに分割して与える PatchGAN を採用する。なお、生成器 G には U-net[4] を採用し、そのエンコーダ側の複数層にドロップアウトという形でノイズ z を加える。

2.2 実装

2.1 節の Pix2Pix の生成器 G および鑑別器 D のネットワーク構造に関しては、[3] と同一のものを使用す

Removing Ink Bleed-Through of Cursive Character Images Using Pix2Pix

[†] Chihiro UEDA (s15m1013@bene.fit.ac.jp)

[†] Hiroyuki FUJIOKA (fujioka@fit.ac.jp)

^{††} Akinori Hidaka (hidaka.akinori@mail.dendai.ac.jp)

Faculty of Information Engineering, Fukuoka Institute of Technology ([†])

Faculty of Science and Engineering, Tokyo Denki University (^{††})

* 本研究では、裏抜きと紙材部分を除去した文字画像を“裏抜き除去文字画像”と呼ぶ。

る。Pix2Pixの学習用画像ペア (x, y) は図2のような方法で生成する。教師画像となる表文字 y には、衡山毛筆フォント [5] からランダムに選んだ文字画像を採用する。入力画像となる裏抜き文字画像 x は以下の流れで合成する。まず、[1]の作品「好色一代男」の背景部分から無作為に切り取った画像 b に対し、次式によって表文字 y を重ねた合成画像 y_b を作る。

$$y_b(r, c) = \begin{cases} b(r, c) & (y(r, c) = 1) \\ y(r, c) & (y(r, c) \neq 1) \end{cases} \quad (4)$$

ここで、 $z(r, c)$ は画像 z の画素 (r, c) の輝度値とする。 y は二値画像であり、背景部分の輝度値が1 (白色) であることに注意する。次に、 y とは別にランダムに選んだ衡山毛筆フォント文字 x_b を左右反転した画像を \bar{x}_b とする。これを次式で y_b に重ねた画像 x を作る。

$$x(r, c) = \begin{cases} y_b(r, c) & (\bar{x}_b(r, c) = 1) \\ \alpha \bar{x}_b(r, c) + (1 - \alpha)y_b(r, c) & (\bar{x}_b(r, c) \neq 1) \end{cases} \quad (5)$$

ここで $\alpha = 0.25$ とした。これは \bar{x}_b の有色部分のみ y_b との α ブレンドを行ったことに相当する。以上の方法により、画像サイズが 256×256 [pix] の学習用画像ペア (x, y) を 1000 組生成した。

Pix2Pixの学習・評価は NVIDIA GeForce GTX960 搭載 PC (CPU: Intel Core-i7-4790, RAM: 32GB) 上

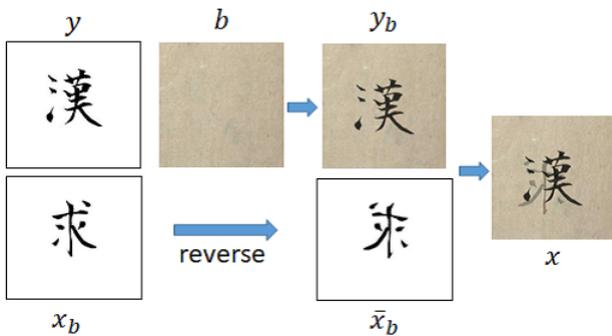


図 2: 学習用ペア画像データ (x, y) の作成法

で Python および Tensorflow[6] によって行う。なお、学習におけるバッチ数は 1, エポック数は 1000 とする。

3 実験

2節の実装および学習により得られた生成器 G を用いて裏抜け除去を行い、本手法の有用性を検討する。テスト画像として、古典籍くずし字画像 [1] の 63,956 文字分の画像を生成器 G への入力として用いる。図3に実験結果の一例を示す。図3の上段は入力、下段は本手法による結果である。これらの結果から、本手法は、墨汁による裏抜けがかなり強い場合であってもそれらを十分にうまく除去でき、文字のみをうまく抽出できていることがわかる。

参考文献

- [1] 人文学オープンデータ共同利用センター, 日本古典籍データセット, <http://codh.rois.ac.jp/charshape>.
- [2] ハリムサンディ, 和田俊和, 文書表裏面スキャン画像の輝度値分布変換による裏写り除去法, 電子情報通信学会技術研究報告 PRMU, パターン認識・メディア理解, Vol.109, No.470, pp.495-500, 2010.
- [3] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Image-to-Image Translation with Conditional Adversarial Networks, *arXiv.org*, Computer Vision and Pattern Recognition, arXiv:1611.07004, Nov, 2016.
- [4] O. Ronneberger, P. Fischer and T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, *arXiv.org*, Computer Vision and Pattern Recognition, arXiv:1505.04597, May, 2015.
- [5] 青柳 衡山, 衡山毛筆フォント, <https://opentype.jp/kouzanmouhitufont.htm>.
- [6] T. Hope, Y. S. Resheff and I. Lieder, *Learning TensorFlow: A Guide to Building Deep Learning Systems*, O'Reilly Media, 2017.

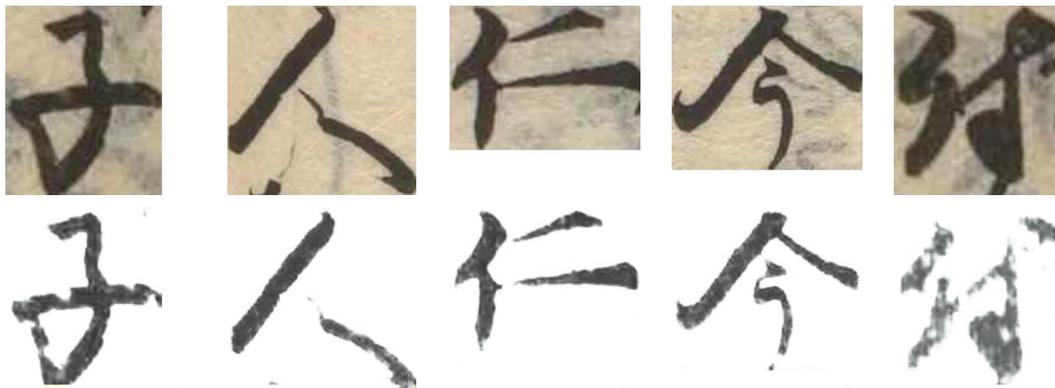


図 3: 古典籍くずし字画像の裏抜き除去例