

GAN に基づいた手書きフォントの生成に関する研究

甘 偉永

指導教員：何 立風

1 研究背景と目的

現代社会では、パソコンの普及に伴って文字を手で書く機会が明らかに減っていることが分かる。オンラインで少しでも相手の誠意を感じるように個人スタイルの手書き文字生成について研究したいと考えている。スタイルを学習し生成するために GAN を使用するのが一般的である。GAN は生成モデルの一種であり、データから特徴を学習することで、実在しないデータを生成したり、存在するデータの特徴に沿って変換することができる。本研究は zi2zi ネットワーク [1] 利用してフォントから手書き文字のフォントに変換することを目的とする。

2 提案手法

2.1 使用するネットワーク

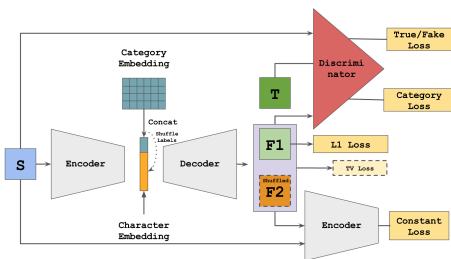


図 1: ネットワーク

本研究に使用するネットワークは zi2zi, zi2zi は pix2pix を元に文字の生成に改善したネットワーク。zi2zi は pix2pix と同様に、生成器は skip-connection を使用した U-net 構造, 判別器は patchGAN に基づく。しかし, zi2zi は文字生成を特化するために category Embedding, category loss と constant loss を導入した [1]。本研究は生成器の性能を上昇させるため, SE-block を追加した。

2.2 データセット作成

データセットの作成において, 手書き文字の写真を作成し, Tesseract-OCR を使用して文字レイアウト box を作成する。手書き文字の写真と box ファイルを jTessBoxEditor 内で枠修正をする。最後は手書き文字の写真と box ファイルを用いてデータセットを生成する。データセットの作成において簡単な漢字を使用している。漢字の一部である部首も複数使用している。また, 人が書く文字には不安定性があるため, データセットに同一文字を各五回書いている。拡張する前のデータ数は 400, 拡張する方法は拡大, 縮小と回転, 拡張後のデータ数は 1600 になる。

2.3 本研究の流れ

1. データセットを作成する。
2. 訓練データと検証データは 9:1 の比率で分ける。
3. 訓練用データを用いて zi2zi ネットワークと zi2zi+SE-block ネットワークをそれぞれ訓練する。
4. 検証データを用いて提案手法を評価する。

3 実験結果

3.1 データ訓練

zi2zi ネットワーク と zi2zi+SE-block ネットワーク をそれぞれ 600 回訓練を行う。訓練終了後, 生成効果の良い 11000 の checkpoint を選択して文字生成に用いる。

‘干’, ‘旱’, ‘回’, ‘品’, ‘凡’ と ‘早’ の手書き文字, zi2zi ネットワークと zi2zi+SE-block ネットワーク の生成結果はそれぞれ図 2a, 図 2b と図 2c に示す。

3.2 実験評価



(a) 手書き文字



(b) zi2zi による生成結果



(c) zi2zi+SE による生成結果

図 2: テストデータの比較

表 1: テストデータの評価結果

評価項目	zi2zi	zi2zi+SE-block
評価 (dice)	0.9194	0.9229

図 2 から, 漢字 ‘干’ の生成効果が一番いい, 他の漢字の生成効果は漢字 ‘干’ より悪いが, どの漢字であるかは分かる。漢字 ‘旱’ は zi2zi より zi2zi+SE-block 追加のほうが効果がいい。一方, ‘早’ の生成結果では, zi2zi+SE-block のほうがぼやけるため, zi2zi の生成効果が良い。

また, 2 つの集合がどれだけ似ているかを測定するための統計指標の一つである dice 係数を用いて評価を行った。手書き文字と zi2zi ネットワークと zi2zi+SE-block ネットワークがそれぞれ生成した文字間の dice 係数は表 1 に示す。表 1 から, 全体を見ると, SE-block 追加のほうが少し性能がいいことが分かる。

4 考察

図 2 から, 漢字 ‘干’ は ‘一’ と ‘十’ の組み合わせからできているが, データセットには漢字 ‘士’ もあり, それが原因で他の字より効果が良いと考えられる。SE-block 以外に transformer-block にも試していたが, 元の結果より悪くなる。漢字の一つの点でも違ったら全く異なる字になることもあるため, より良い効果の字を生成するにはよりいい性能のネットワークに改善する方法があると考えられる。

5 おわりに

本研究は手書きフォントの生成を目標として, 一定な効果を達成しているが, 文字がぼやけることが起きる。平均 dice 係数は 0.9 以上もある, このことから, 漢字の誤差許容率が低いことが考えられる。

参考文献

- [1] Yuchen Tian, “Master chinese calligraphy with conditional adversarial networks,” <https://github.com/kaonashi-tyc/zi2zi>, 2017.