

GAN を活用した個別データによる手書き認識精度の向上

Improving Handwriting Recognition Using GAN-Enhanced

Techniques with User-Specific Data

広島県立広島叡智学園高等学校 2年A組 中野 竜希 2年B組マーシュ 隆民

Abstract: This research investigates how GANs can enhance OCR accuracy for individual handwriting. Using rotated and GAN-augmented datasets, the models showed improved accuracy compared to default setups. The findings highlight the potential of applying GANs to reduce communication barriers and optimize practical use cases like medical documentation and personalized handwriting synthesis.

Keywords: Digit recognition, GAN, augmentation, handwriting.

1. 研究背景

医療・福祉現場には発声・聴覚障害を持つ患者が存在する。聴覚障がい者のコミュニケーション手段は、厚労省のデータでは、補聴器が約 70%、筆談・要約筆記が約 30%、手話が約 19%となっている(加古川市役所, n.d.)。とあり、筆談が占める割合が高い。しかし、筆談に慣れていない場合や、手の震えなどで文字が不明瞭になる場合、筆談自体が新たな障壁となることがある。

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できる。(中略) 応用研究や理論的研究も急速に進んでおり、今後の発展が大いに期待されている(アイマガジン, 2018)。とあるため、この技術を OCR (光学文字認識) に応用することで、筆談の障壁を打破することが可能だと考え、本テーマを選択した。

2. 研究目的・意義

GAN について研究することで、「個人の筆跡を学習する OCR ツール」の開発につながる。このツールによって、筆談による意思疎通が容易になり、患者の精神的負担の軽減や、医療従事者の業務効率化に寄与できるため、研究を実施する。

3. 研究方法

GAN を用いた OCR の精度を研究するため、既存の OCR 技術との比較を実施する。回転処理を施さないものの、回転処理は施すが、GAN 強化学習を実施せず、個人に適合していないものの 2 つを対照として設定する。

表 1 仮説と変数

| | |
|------------|------------------|
| H_0 帰無仮説 | 学習方法は識字率に影響を与えない |
| H_1 対立仮説 | 学習方法は識字率に影響を与える |
| 独立変数 | 学習方法 (No Unit) |
| 従属変数 | 手書き数字の識字率 (%) |

1. 機械学習の準備:

- ・大規模手書き数字モデル MNIST (Deng, 2012) から、学習データを用意
- ・5 人のユーザーが各数字(0~9)を 50 回書き、ラベル付きのデータとして、ユーザーごとに保存
→評価用データセットとして使用

2. 機械学習の実施:

- ・対照①[default]: MNIST のデータセットを用いて学習
- ・対照②[rotatel]: MNIST のデータセットをランダムで $\pm 30^\circ$ 回転させて学習
- ・[rotate+gan]: 上記の手法に加え、GAN を MNIST のデータセットで学習させ各数字 1000 ずつ画像を生成し学習用データセットに追加して学習
- ・[rotate+gan+{n}samples]: 上記の手法に加え、評価用データセットの各数字からランダムに{n}個データを取得し、上記の GAN を強化学習した後に画像を再度生成し学習

3. テストの実施

- ・各モデルの精度を各ユーザーの評価用データセットを用いて評価(試行 1~5)。

4. 結果・考察

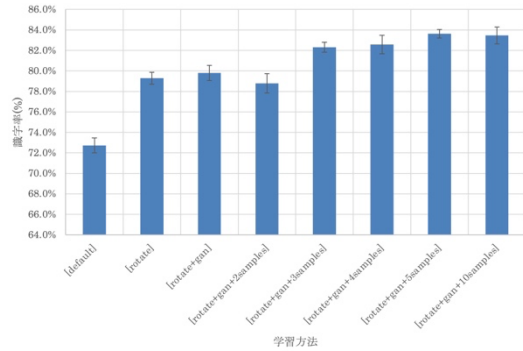


図 1 学習方法ごとの識字率

表 2 実験の総括的データ

| 学習方法 | 試行 1 | 試行 2 | 試行 3 | 試行 4 | 試行 5 | 平均 | 標準偏差 |
|------------------------|--------|--------|--------|--------|--------|--------|----------|
| [default] | 0.7207 | 0.7302 | 0.7365 | 0.7194 | 0.7299 | 0.7273 | 0.007205 |
| [rotate] | 0.7958 | 0.7924 | 0.7843 | 0.7918 | 0.8004 | 0.7929 | 0.005904 |
| [rotate+gan] | 0.7860 | 0.7968 | 0.8055 | 0.8011 | 0.8008 | 0.7980 | 0.007404 |
| [rotate+gan+2samples] | 0.7875 | 0.7944 | 0.7785 | 0.7794 | 0.7999 | 0.7879 | 0.009319 |
| [rotate+gan+3samples] | 0.8185 | 0.8299 | 0.8187 | 0.8265 | 0.8225 | 0.8232 | 0.004953 |
| [rotate+gan+4samples] | 0.8271 | 0.8183 | 0.8155 | 0.8381 | 0.8299 | 0.8258 | 0.009102 |
| [rotate+gan+5samples] | 0.8378 | 0.8409 | 0.8387 | 0.8312 | 0.8329 | 0.8363 | 0.004086 |
| [rotate+gan+10samples] | 0.8335 | 0.8484 | 0.8331 | 0.8321 | 0.8265 | 0.8347 | 0.008157 |

結果より、対照①と比較して、回転処理を施した群の識字率が高くなっていることが読み取れる。また、対照②と比較すると、GAN 強化学習を行った群の識字率は概ね同程度高くなっていることが読み取れる。一元配置分散分析表より、対照①、②について、P-値が有意水準 0.05 よりも非常に小さく、F 値は F 境界値よりも大きいことがわかる。したがって、帰無仮説が棄却され、学習方法は識字率に影響を与えるため、考察の正当性を示す根拠となる。

表 3 一元配置分散分析表

| 対照 | 変動要因 | 変動 | 自由度 | 分散 | F 値 | P-値 | F 境界値 |
|-----|-------|------------|-----|------------|------------|------------|------------|
| 対照① | グループ間 | 0.04560906 | 7 | 0.00651558 | 124.389622 | 2.2768E-21 | 2.31274119 |
| | グループ内 | 0.00167617 | 32 | 5.238E-05 | | | |
| 対照② | グループ間 | 0.01264438 | 6 | 0.0021074 | 40.1817707 | 1.7065E-12 | 2.4452594 |
| | グループ内 | 0.0014685 | 28 | 5.2447E-05 | | | |

5. 結論及び今後の展望

本研究を通じて、個別データを用いた GAN の強化学習と、回転処理を施したデータを学習に用いることで OCR 技術の精度の向上に寄与できると結論づけることが可能だ。しかし、GAN 強化学習におけるサンプル数が識字率に与える影響についてはデータポイントが少ないため明確には分からなかった。この点が実験の限界であり、今後より多くのデータを収集し、影響について分析する必要がある。

ローカルで実行可能かつ、個々の特徴的な筆跡を反映した OCR 技術の開発は筆談による意思疎通が容易になり、患者の精神的負担の軽減や、医療従事者の業務効率化に寄与することができる。他にも、GAN を採用した OCR 技術は、それを用いてユーザーの文字から特徴的な筆跡を反映した合成画像を作ることにも可能であるため、デジタル上の文字をユーザーの筆跡に似た手書きの文字として表示することができる。これにより、手の震えなどで文字が不明瞭な人などが、過去に書いた文字を読み込ませることで、自らの文字で手紙や年賀状の作成といったことが行える。このように OCR 技術に GAN を採用することにより、精度を向上させると同時に、モデルの用途を大幅に広げることができる。

このような展望に向け、モデルの精度向上のために GAN の代替として VAE、拡散モデルといったモデルの使用、モデルのサイズ・構成の変更、学習時間の調節などがモデルの精度にもたらす影響についても分析する必要がある。

引用文献

- Bose, A. (2019, February 22). *A Must Read Intro To Neural Networks Using PyTorch — Handwritten Digit Recognition*. Medium. <https://towardsdatascience.com/handwritten-digit-mnist-pytorch-977b5338e627>
- Deng, L. (2012). The mnist database of handwritten digit images for machine learning research. *IEEE Signal Processing Magazine*, 29(6), 141–142.
- Elanwar, R., & Betke, M. (2024). Generative adversarial networks for handwriting image generation: a review. *The Visual Computer*. <https://doi.org/10.1007/s00371-024-03534-9>
- Inkawhich, N. (2018, July 31). *DCGAN Tutorial — PyTorch Tutorials 1.6.0 documentation*. Pytorch.org. https://pytorch.org/tutorials/beginner/dcgan_faces_tutorial.html
- アイマガジン. (2018, September 4). GAN：敵対的生成ネットワークとは何か ～「教師なし学習」による画像生成. I Magajine. <https://arc.net//quote/hwdjuqml>
- ・加古川市役所. (n.d.). コミュニケーションを図ることについての 現状と課題について. 加古川市 KAKOGAWA CITY. Retrieved December 22, 2025, from https://www.city.kakogawa.lg.jp/material/files/group/45/shiryoku4_danntaitouhiarinngumatome.pdf