

AI を用いたデジタルアーカイブ手法の考察

Consideration of digital archiving methods using AI

○清水 能理

○Yoshimasa Shimizu

八戸工業大学

Hachinohe Institute of Technology

キーワード: デジタルアーカイブ(digital archive), 画像処理(image processing), GAN (Generative Adversarial Network), 超解像(super-resolution), 最適化(optimization)

連絡先: 〒031-8501 八戸市大字妙字大開 88-1 八戸工業大学工学部工学科システム情報工学コース

清水能理 Tel.: (0178)25-8135, Fax.: (0178)25-1691, E-mail: shimizu@hi-tech.ac.jp

1. はじめに

デジタルアーカイブは知的資源を対象とし、その所蔵および利用機関は文書館・公文書館・資料館・博物館・美術館・図書館等である。デジタル化知的資源を公開することにより、ネットワークを通じた利用が容易となる。一方、超解像 (Super-Resolution) は、画像や動画の解像度を擬似的に上げる技術のことで、解像度が低い画像の高周波成分を予測・復元して、解像度を高める。この技術は、テレビなどに関わるデジタルの画像処理技術の一つで、入力信号の解像度を高めて出力信号を作る技術を指す。具体的な製品では、入力された動画や静止画の信号を高解像度化して出力し、高解像度の画像を表示する。この技術を、従来の画像処理に代わる新しい技術として、絵画などのデジタルアーカイブに応用することを検討する 1)- 3)。

2. 問題の記述

絵画のデジタル化にはカメラで撮影する方法と、スキャナーで保存するという2つの方法がある。AI の超解像手法を用いると、さらに高画質に保存が可能になるのではと考える。本研究では、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可能な Generative Adversarial Network (GAN, 敵対的生成ネットワーク) と呼ばれる AI

の一種を用いる。GAN の中でも超解像に焦点を置き、浮世絵画像を対象として絵画の高解像を試みる。また、複数の手法を比較しながら超解像の精度向上を行い、超解像とノイズ除去に着目し、絵画画像の保存手法として有効か検証し、考察していく 4)- 9)。

3. 画像処理

画像処理は、コンピュータを使用してデジタル画像を処理する技術で、その目的は、画像の品質を向上、特定の特徴を抽出、または画像を他の形式に変換するなど、様々な目的に応じて異なり多岐にわたる 7)- 10)。

画像処理の一般的なタスクは、以下になる。

- 3-1. ノイズ除去: 画像から不要なノイズを取り除き、画像の品質を向上させる。
- 3-2. 解像度改善: 画像の解像度を向上させ、画像の詳細や鮮明さを向上させる。
- 3-3. 特徴抽出: 画像から特定の特徴を抽出し、物体認識やパターン認識などのタスクを実行する。
- 3-4. セグメンテーション: 画像を複数の領域に分割することで、異なるオブジェクトや背景を識別する。
- 3-5. 物体検出: 画像内の物体の位置や境界を検出し、識別する。

3-6. モザイク処理: 画像の一部をモザイク状にぼかすことで、個人情報の保護や不正利用の防止などで使用される。

4. デジタルアーカイブ

デジタルアーカイブは、デジタル技術を用いて文書、画像、音声、動画などの情報資源を保存、整理、公開するためのシステムまたはプロセスである。従来のアーカイブは主に物理的な形態で保存されていたが、デジタルアーカイブは情報を電子的な形式で保存し、オンラインでアクセス可能にすることができる。主な目的は、情報資源を長期間保存し、容易に検索・閲覧できるようにすることである。歴史的な文書や芸術作品、学術研究データなど、重要な情報を保存し、後世に伝えることが可能になる。次のような要素から構成される 11)- 13)。

4-1. デジタル化: アナログ形式の情報をデジタル形式に変換する。スキャンや録音などの手法を使用し、紙の文書や写真、テープなどをデジタルデータに変換する。

4-2. デジタル保存: デジタルデータを適切な形式で保存し、データベースやクラウドストレージなどのシステムを使用し、データの安全性と可用性を確保する。

4-3. メタデータ管理: デジタルアーカイブ内の情報資源を特定するためのメタデータ (情報に関する情報) を作成・管理する。作成者、作成日、キーワード、関連するテーマなどの情報が含まれ、効果的な検索と分類が可能になる。

4-4. アクセス制御: デジタルアーカイブには、一般の人々や特定のユーザーに対してアクセスを制限するためのセキュリティ対策が必要で、機密性の高い情報や著作権が保護された情報については、アクセス制限が設けられる場合もある。

4-5. オンライン公開: デジタルアーカイブは、インターネットを通じてオンラインで公開されることが一般的で、一般の人々がアーカイブのコンテンツを閲覧・検索できるようになる。国立公文書館デジタルアーカイブ、立命館大学 ARC などがある 1)- 2)。

5. 敵対的ネットワーク (Adversarial Network)

敵対的ネットワークは、主に敵対的生成ネットワーク (Adversarial Generative Network) と敵対的識別ネット

ワーク (Adversarial Discriminative Network) の 2 つがある 4)。

5-1. 敵対的生成ネットワーク (Adversarial Generative Network) : 敵対的生成モデル (Adversarial Generative Model) のことで、GAN (Generative Adversarial Network) などがあり、生成器と識別器の 2 つのネットワークが互いに競い合いながら学習する。

5-2. 敵対的識別ネットワーク (Adversarial Discriminative Network) : 敵対的識別モデル (Adversarial Discriminative Model) を指し、入力データを分類する識別モデルを意味し、学習において敵対的な要素が組み込まれ、生成されたデータと本物のデータを区別するため GAN の識別器が使用される。

生成モデルと識別モデルを敵対的に学習させることで、より高度な表現や生成能力を実現するために使用され、生成モデルがより本物に近いデータを生成できるようになり、識別モデルがより洗練された識別能力を持つように期待される。

6. GAN (Generative Adversarial Network)

GAN は生成モデルの一種で、データを生成するモデル (生成器) と、生成されたデータと本物のデータを識別するモデル (識別器) の 2 つのネットワークで構成される。GAN の学習は、生成器と識別器が互いに競い合うような形で進む。生成器はランダムノイズからデータを生成し、識別器は生成されたデータと本物のデータを区別しようとする。識別器は生成されたデータと本物のデータを正しく識別することで報酬を得る一方、生成器は識別器を騙すようなデータを生成することで報酬を得る。この競争が進む中で、生成器は本物のデータに似たデータを生成するように学習し、識別器は生成されたデータと本物のデータをより正確に区別できるようになる。

画像生成や音声生成などの例として、Deepfake (ディープフェイク) 技術があり、GAN で生成されたデータは、本物と区別がつかないほど高い品質になる。一方、GAN は深層学習の一部として研究され、データの補完・修復、データ拡張など応用範囲は広がっている 4), 7), 14)- 15)。

7. 超解像 GAN (Super Resolution GAN)

超解像とは、低解像度の画像を高解像度の画像に変換することである。単純に画像を大きくするのではなく、ひずみを抑えて人間が見ても自然な画像として認識できるような形で高解像度の画像を生成するというのが大きな特徴である。従来の超解像の手法としては、SRCNN、FSRCNN、SRGAN がある 7)。

超解像 GAN (Super Resolution GAN) は、超解像技術を実現するために GAN を利用したアプローチで、通常の GAN と同様に、ジェネレータとディスクリミネータの 2 つの主要な要素で構成される。低解像度画像から高解像度画像を生成する手法で、GAN の能力を活用することで詳細な特徴やテクスチャの回復が可能となる。この技術は、画像の拡大、ビデオの品質向上、映像処理、医療画像解析などの領域で広く応用される 16)- 18)。

7-1 ジェネレータ：低解像度の画像を高解像度に変換する役割を果たし、畳み込みニューラルネットワーク (Convolutional Neural Network) や深層学習モデルで構築される。入力として与えられた低解像度画像を取り、高解像度の画像を生成する。

7-2 ディスクリミネータ：ジェネレータが生成した画像と本物の高解像度画像を区別する役割を果たす。畳み込みニューラルネットワークや識別モデルとして構築され、与えられた画像が本物の高解像度画像か、それともジェネレータが生成した偽の画像であるかを判断する。

7-3 超解像 GAN の学習プロセス：ジェネレータとディスクリミネータが相互に競い合う。ジェネレータはディスクリミネータをだまそうとして高品質な高解像度画像を生成し、ディスクリミネータは本物画像とジェネレータ生成画像を区別しようとする。この競争により、ジェネレータはより高品質な高解像度画像を生成するよう進化し、ディスクリミネータはより正確に識別するように学習する。

超解像 GAN を、絵画のデジタルアーカイブで利用する場合は、古い作品や低解像度のスキャンなど質の低い元画像に使用することで、高品質な高解像度画像に復元することができる。生成器は低解像度の画像を高解像度に変換する役割を果たし、識別器は生成器が生成した画像が本物の高解像度画像かどうかを判別する役割を果たす。2 つの要

素は競合的に学習され、互いに向上し続ける。

8. 研究方法

8-1. デジタルアーカイブの絵画を超解像 GAN で処理するには、まず学習データセットが必要で、高解像度な絵画画像とそれに対応する低解像度な画像のペアを使って GAN を学習させ、生成器が低解像度画像を高解像度に変換する能力を獲得させる。

8-2. 学習が完了した GAN を使用し、デジタルアーカイブの絵画を処理する。まず低解像度の画像を入力として生成器に与える。生成器はそれを高解像度に変換し、より鮮明な画像を生成する。生成された高解像度画像は、元の低解像度画像よりも詳細な情報を含んでいる。

8-3. 超解像 GAN は、画像の詳細な復元や補完に利用されるが、完璧を保証するものではない。絵画の場合、GAN によって生成された高解像度画像は、元の絵画とは異なる可能性がある。また、GAN のパフォーマンスは学習データセットやパラメータ設定に依存するため、調整が必要である。デジタルアーカイブの絵画を、注意深く評価する。

9. ESRGAN

ESRGAN とは、SRGAN での課題であるアーティファクトというノイズが現れることを改善した手法である。ESRGAN の生成器のネットワークは、畳み込み層、23 個の Residual-in-Residual (RRDB)、畳み込み層、アップサンプリングレイヤー (畳み込みと Leaky ReLU と PixelShuffle)、そして 2 つの畳み込み層から構成されている 16)- 18)。

ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks 5

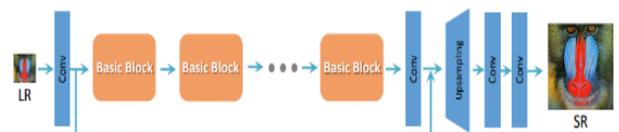


図 1 ESRGAN の生成器のネットワーク構造

10. 実験方法

ESRGAN、Real-ESRGAN、さらに改良のため SCUNet を組み合わせ、学習を行った。

10.1 ESRGAN

- (1) 画像データ、アノテーションデータを読み込む。
- (2) ラベルを作成する。
- (3) 128×128 でランダムに画像を切り取る。
- (4) ESRGAN で 15epoch 学習させる。

10.2 Real-ESRGAN

超解像を行う画像をフォルダに入れ、Real-ESRGAN で、サイズ倍率 4 倍で学習させる。

10.3 SCUNet

ノイズ除去を行う画像をフォルダに入れ、SCUNet で学習させる。

同様に、以下の方法を用いる。

10.4 ESRGAN 学習後 Real-ESRGAN の学習

10.5 Real-ESRGAN 学習後 ESRGAN の学習

10.6 Real-ESRGAN 学習後 SCUNet の学習

10.7 SCUNet 学習後 ESRGAN の学習

10.8 SCUNet 学習後 Real-ESRGAN の学習

11. 実験結果

学習結果を口、目、髪で比較し、保存手法として最も有効な手法は、Real-ESRGAN 学習後 SCUNet を学習した手法であった。



図 2 Real-ESRGAN 学習後 SCUNet 学習画像

GAN を利用し超解像を行い、絵画の保存に有効か考察を行った。超解像手法の ESRGAN は、従来の SRGAN の課題であったアーティファクトを解決し、自然な画像生成を可能にした。Real-ESRGAN は、ESRGAN の低解像度のデータ生成が改良され、元の情報を損なわずに綺麗に高解像にできた。ESRGAN の学習後 Real-ESRGAN を学習させると、ESRGAN で生成できなかった部分があるまま

高解像になり、和紙に描いたような画像が生成された。Real-ESRGAN 学習後 ESRGAN の学習を行うと、主に髪の部分が赤みがかかった画像が生成される結果となった。結果画像の口、目、髪の部分のみで比較を行うと、Real-ESRGAN の学習が情報を損なわずに高解像にできた。

12. おわりに

デジタルアーカイブに用いる GAN による超解像の精度向上を目的として、SCUNet を学習後 Real-ESRGAN を学習させる方法と、Real-ESRGAN 学習後にノイズ除去ネットワークアーキテクチャ SCUNet を学習させる方法の 2 種類の手法を考察した。Real-ESRGAN の学習前後どちらも安定した学習結果を得ることができた。Real-ESRGAN の学習後に SCUNet を学習させると、拡大させた際元の部分を維持したまま GAN のみの場合より、鮮明ではっきりとした画像結果となった。

本研究で試みた実験からは、Real-ESRGAN の学習後に SCUNet の学習を行った提案手法が情報を損なわず細部まで正確に高解像に変換でき、絵画の保存に有効な手法ではないかと考察する。

参考文献

- 1) 国立公文書館 デジタルアーカイブ:
<https://www.digital.archives.go.jp>
- 2) 立命館大学 ARC. 浮世絵データベース. 国立情報学研究所 DSC (2020):
<https://doi.org/10.32130/rdata.2.1>
- 3) 毛利ほか:GAN ディープラーニング実装ハンドブック, 秀和システム (2021)
- 4) Jakub Langr/Vladimir Bok, 実践 GAN 敵対的生成ネットワークによる深層学習, マイナビ出版 (2020)
- 5) 臼井孝俊, 清水能理: 敵対的サンプルに対する異常検知の研究, 情報処理学会東北支部研究会 研究報告, Vol. 2022 No. 1-1 (2022)
- 6) 相馬徹平, 清水能理: 画像データに対する異常検知手法の特徴および性能の比較, 情報処理学会東北支部研究会 研究報告, Vol. 2022 No. 1-4 (2022)
- 7) チーム・カルポ:物体検出と GAN, オートエンコーダー,

画像処理入門, 秀和システム(2021)

- 8) 袴田晨比, 清水能理: Learning-to-See-in-the-Dark を用いた暗所における物体検出の研究, 情報処理学会東北支部研究会 研究報告, Vol. 2022 No. 1-5 (2022)
- 9) 山田果林, 清水能理: ランドマーク検出法を用いた顔輪郭による個人認識の研究, 情報処理学会東北支部研究会 研究報告, Vol. 2022 No. 2-3 (2022)
- 10) 大村龍生, 清水能理: Python を用いた AI アプリ開発および最適化の研究, 情報処理学会東北支部研究会 研究報告, Vol. 2022 No. 1-2 (2022)
- 11) ARC 浮世絵顔データセット. Yingtao Tian, ROIS-DS CODH:
<https://doi.org/10.20676/00000394>
- 12) 人文学オープンセンター共同利用センター. 浮世絵顔データセット:
<http://codh.rois.ac.jp/ukiyo-e/face-dataset/>
- 13) Yingtao Tian, Tarin Clanuwat, Chikahiko Suzuki, Asanobu Kitamoto: Ukiyo-e Analysis and Creativity with Attribute and Geometry Annotation, arXiv. 2106.02267, 2021.
- 14) 坂本祐介, 清水能理: 敵対的サンプルに対して堅牢な AI モデルの構築, 情報処理学会東北支部研究会 研究報告, Vol. 2022 No. 1-3 (2022)
- 15) 村上綾香, 清水能理: ESRGAN を用いた絵画画像のデジタルデータ保存の研究, 情報処理学会東北支部研究会 研究報告, Vol. 2022 No. 2-2 (2022)
- 16) ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks (2018):
<https://arxiv.org/abs/1809.00219>
- 17) Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data (2021):
<https://arxiv.org/abs/2107.10833>
- 18) Practical Blind Denoising via Swin-Conv-UNet and Data Synthesis (2022):
<https://arxiv.org/abs/2203.13278>