

エッジ増強と深層学習を用いた低階調画像の高階調化 High Toning with Edge Augmentation and Deep Neural Networks

笹平 啓介¹⁾ 原 健二¹⁾ 井上 光平¹⁾
Keisuke Sasahira Keiji Hara Kohei Inoue

1 はじめに

通信量やデータ容量に制限があるとき、画像の階調数(色数)を減らすことは有効な手段であり(図 1)、これまで数多くの低階調化(減色)手法が提案されてきた。これとは逆に、近年は高速かつ大容量のデータ通信が可能なることから、低階調画像を逆に元の高階調画像に復元する高階調化手法の確立が求められる。本稿では、深層学習(Deep Learning)の適用可能性を検証し、擬似エッジを増強することで深層学習による画像の高階調化を改善する手法を提案する。



図 1 高階調画像(左)と低階調画像(右)

2 Pix2Pix[1]

画像変換のモデルとして Pix2Pix[1] と呼ばれる深層学習手法が知られている。Pix2Pix は近年研究が盛んに行われている敵対的生成ネットワーク(Generative Adversarial Network, GAN, [2]) の一つである。GAN は画像を生成する Generator と画像の真贋を弁別する Discriminator の二つのネットワークから構成される。Generator は Discriminator を欺こうと学習し、Discriminator はより正確に弁別しようと学習することでより精度の高い画像生成が可能になる(図 2)。Pix2Pix は、GAN の中でも特に入力に条件を課す条件付き敵対的生成ネットワーク(Conditional GAN, [3]) であり、画像ペアを入力として与えて、カラリゼーションや絵画風画像生成など所望の画像を生成する。しかし、後述するように画像の高階調化の問題に対しては Pix2Pix をそのまま用いても良好な結果は得られないことが実験で確認された。

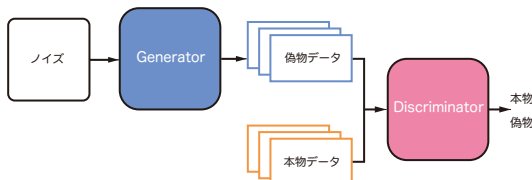


図 2 GAN のアーキテクチャ

3 提案手法

同一シーンの低階調画像と高階調画像のペアからなるデータセットを使用して Pix2Pix を画像の高階調化に用いると、色相が大きく変化したり偽のエッジ(擬似エッジ)が除去されずに残ったりすることが多い。この問題を回避するため、低階調/高階調画像ペアを用いる代わりに、低階調画像の擬似エッジが強調された画像と高階調画像のペアからなるデータセットを用いることを考える。ここで、低階調画像を高階調化する際、元の高階調画像も所有していれば低階調画像と高階調画像のそれぞれのエッジ画像の差分をとることで擬似エッジが正確に得られるが(以下、参照型手法)、高階調画像を所有していることは実際にはほとんどないため、この差分を取る手法は利用できない。そこで、ニューラルネットワークのテスト時は低階調画像のみ所有していることを仮定し、以下のようにして擬似エッジを検出する(以下、非参照型手法)。

低階調画像のエッジ付近に二つの特徴が現れる頻度が高いことより、その二つの特徴を用いて低階調画像のみから擬似エッジ検出する。まず、擬似エッジは偽輪郭とも呼ばれるように輪郭状をなすことが多いことから、各エッジ画素ごとにその近傍ウィンドウ内に属するエッジ画素集合に楕円を当てはめたときに第 1 主成分の寄与率が高くなる。また、擬似エッジ画素の近傍カラー画素集合の画素値の分散は小さくなる傾向にある。これらの性質に基づき、次式を満たすエッジ画素を擬似エッジ画素と判定する。

$$\frac{r_i}{\sigma_i} \geq T \quad (1)$$

ここで、 r_i は i 番目のエッジ画素の近傍に属するエッジ画素からなる 2 次元プロットに対する第一主成分の寄与率であり、 σ_i は i 番目のエッジ画素の近傍カラー画素集合の画素値の分散である。本実験では $T=0.0008$ に設定している。図 3 の左右に、参照型手法と非参照型手法を図 1 の画像に適用した結果をそれぞれ示す

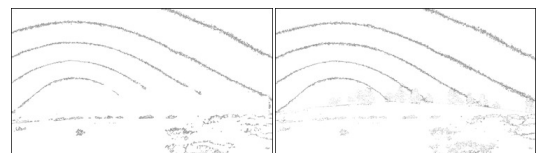


図 3 参照型手法(左)と非参照型手法(右)

4 実験結果

本実験では $\alpha=1.2$ と設定し、ニューラルネットワークのトレーニングでは参照型手法を用いて入力された低階調画像の擬似エッジを検出する。本実験では、エッジ検出のために Canny 法を使用している。本手法により推定された高階調画像に対し、定性評価と定量評価を行う。Pix2pix のトレーニングでは画像サイズが 256×256

1) 九州大学大学院芸術工学研究院

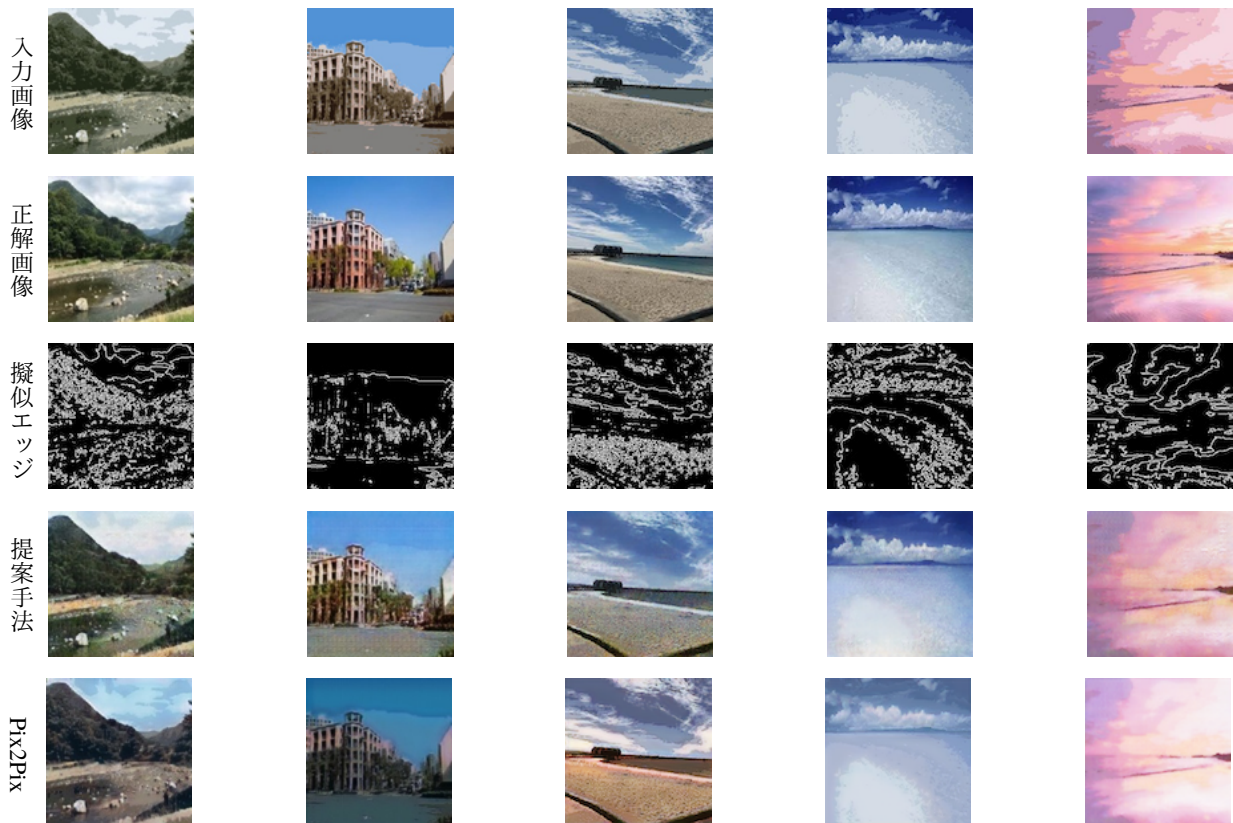


図 4 実験結果

の 24bit カラーの高階調画像を k-means 法で 8bit カラーに低調化した画像に対して擬似エッジを参照型手法で検出・増強した画像、及び元の高階調画像からなる画像ペアを 487 個入力し、NVIDIA GPU を使用した。テストではトレーニングに用いなかった低階調画像を入力とし、擬似エッジ検出には非参照型手法を用いた。エッジ増強倍率はいずれも $\alpha=1.2$ 、netG は unet-256、optimizer には Adam optimizer を適用した。

4.1 定性評価

図 4 の 1 行目～3 行目に、テスト時の入力に用いられた低階調画像、正解の高階調画像、擬似エッジ画像をそれぞれ示す。図 4 の 4 行目と 5 行目に、提案手法と従来の Pix2Pix の結果画像をそれぞれ示す。両者を比較すると、Pix2Pix では色相が大きく変化したり擬似エッジが残ったりしているのに対して、提案手法では正解画像に近い色が復元されて低階調画像の擬似エッジも除去されており、良好に高階調化されていることが分かる。

4.2 定量評価

表 1 は、14 種類の低階調画像に Pix2Pix と提案手法をそれぞれ用いて得られた結果について、PSNR、SSIM、MS-SSIM の各画質評価指標の平均値と標準偏差をまとめたものである。表 1 より、どの画質評価指標においても提案手法が上回っていることが分かる。また、標準偏差を比較すると提案手法の方がいずれも小さくなっており、推定値のばらつきが抑えられていることが分かる。

表 1 定量評価

評価指標	Pix2Pix	提案手法
PSNR	17.55 ± 2.99	22.20 ± 1.75
SSIM	0.61 ± 0.07	0.63 ± 0.06
MS-SSIM	0.81 ± 0.05	0.82 ± 0.04

5 結論

Pix2Pix を低階調画像の高階調化に応用する際に生じる問題を解決するため、エッジ増強画像を Pix2Pix のトレーニングとテストに用いる改善手法を提案した。色相の変化の抑制と擬似エッジの除去が確認された。しかし、擬似エッジの増強に関してまだ不明瞭な部分が多く、またパラメータの値を実験的に求めなければならない。今後の課題として、増強するエッジの抽出方法、色、太さなど、より優れた手法を確立する必要がある。また、特に人工物が多い画像では、エッジがぼやけた画像が生成されてしまうこともあるため、本手法を超解像手法と組み合わせることを検討している。

参考文献

- [1] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, “Image-to-image translation with conditional adversarial networks”, CVPR (2017).
- [2] M. Mirza and S. Osindero, “Conditional generative adversarial nets”, arXiv preprint arXiv:1411.1784 (2014).
- [3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets”, In NIPS, pp. 2672-2680 (2014).