

客観品質を考慮した敵対的生成ネットワークによる画像超解像

山本 宏哉[†] 北原 大地[†] 平林 晃[†]

[†] 立命館大学情報理工学部 〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: [†] is0335vx@ed.ritsumei.ac.jp, d-kita@fc.ritsumei.ac.jp, akirahr@media.ritsumei.ac.jp

あらまし 敵対的生成ネットワーク (GAN) を用いた画像超解像において、ネットワークが生成する画像に直交射影を適用し、更に損失関数を改善した手法を提案する。従来手法では、正解画像と一致しない不自然な高周波成分 (アーチファクト) が表れるという問題があった。提案手法では、ダウンサンプリング結果が入力画像と一致する線形多様体に超解像画像を直交射影することにより、アーチファクトの生成を抑制する。更に直交射影による移動量を損失関数に加えてネットワークの学習を行う。数値実験により、提案手法が従来手法より高い客観品質を保ちつつ、高周波成分をより適切に生成できることを示す。

キーワード 画像超解像, 敵対的生成ネットワーク, 直交射影

Image Super-Resolution via Generative Adversarial Network Considering Objective Quality

Hiroya YAMAMOTO[†], Daichi KITAHARA[†], and Akira HIRABAYASHI[†]

[†] College of Information Science and Engineering, Ritsumeikan University
1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577 Japan

E-mail: [†] is0335vx@ed.ritsumei.ac.jp, d-kita@fc.ritsumei.ac.jp, akirahr@media.ritsumei.ac.jp

Abstract We propose a super-resolution method based on a conventional technique using the generative adversarial network (GAN). The conventional method suffers from pixel-wise black or white artifacts. This is because the number of training data was not sufficient for a network with huge parameters. To solve this problem, we exploit the observation error by adding it to the conventional cost function. This guarantees the consistency for training examples, but not for a novel input data after training. Hence, we further introduce the orthogonal projection onto a linear manifold, in which the observation error is completely zero. These two ideas enable us to produce high quality super-resolution results without the artifact. Simulation results show the effectiveness of the proposed method.

Key words Image Super-Resolution, Generative Adversarial Network, Orthogonal Projection

1. はじめに

画像超解像とは、画素数の少ない低解像度画像から、高周波成分が復元された画素数の多い高解像度画像を生成する技術である [1]。犯罪捜査における防犯カメラ影像の拡大 [2] や、電子顕微鏡画像の生成 [3] など多様な分野で必要とされている技術である。近年は、既存の画像や映像を 4K・8K といった膨大な画素数へ拡大することも増えている [4]。いずれの分野においても画像超解像では、単に画素数を増やすだけでなく、高周波成分が復元された鮮明な画像を生成することが求められる。

単一画像のみからの超解像は、解が一意に定まらない劣決定逆問題である。画素数を増やす単純な手法として、最近傍補間、双線形補間、双三次補間などの代数的補間法を用いることが

考えられる [5]。これらの手法は非常に高速ではあるが、最近傍補間と双線形補間では高周波成分が全く復元されず、双三次補間でも高周波成分が十分に復元されないため、超解像結果 (超解像画像と呼ぶ) が鮮明でないという問題がある。

低解像度画像から失われた高周波成分を高精度に復元するために、大量の低解像度画像と高解像度画像の組からなる学習データを利用する手法が提案されている [6-14]。Yang らは、学習データを用いて低解像度画像と高解像度画像に対して同一のスパース表現を与える辞書の組を作成し、この辞書に基づき超解像を行った [6]。更に、この手法における辞書学習アルゴリズムの改良も提案されている [7]。これらの手法では代数的補間法と比較して鮮明な超解像画像が得られるが、辞書学習にも画像生成にも時間がかかるという欠点がある。

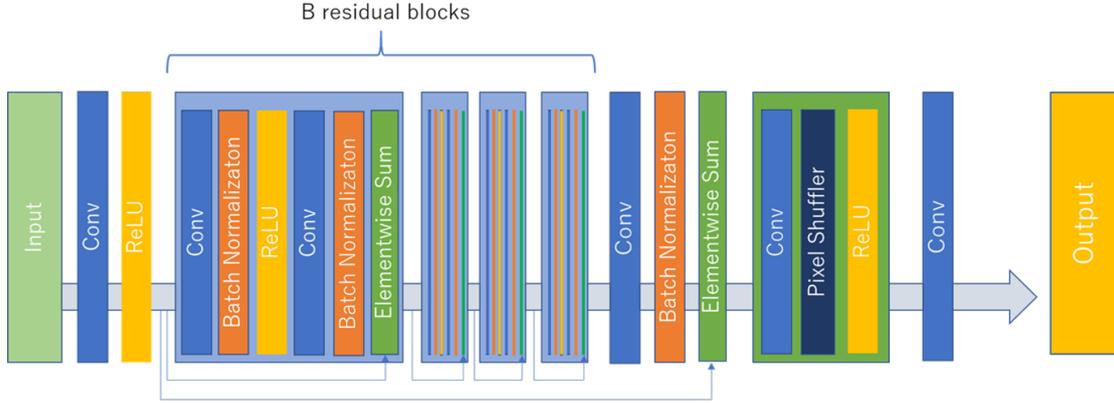


図1 SRGANの生成ネットワーク (SRResNet)

Dong らは、畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) を用いた超解像手法 SRCNN を提案した [8]. この手法では、低解像度画像を入力、高解像度画像を出力とする CNN を学習する. 画像生成時には、学習した CNN に低解像度画像を入力すれば、超解像画像が瞬時に出力される. SRCNN を改良した手法として、入力層における双三次補間を削除して高速化したもの [9] や、残差学習を導入しながら畳み込み層の数を増やした VDSR [10], 任意の拡大率で超解像画像を生成する ESPCN [11] 等がある. これら CNN による超解像手法では、超解像画像を瞬時に生成できるだけでなく、超解像画像の客観品質の値も文献 [6, 7] の手法より優れている. 一方、高解像度画像と超解像画像の二乗誤差を損失関数としてネットワークを学習するため、超解像画像が平滑化され、結果として高周波成分を復元しきれないという問題があった.

そこで、Ledig らは敵対的生成ネットワーク (Generative Adversarial Network: GAN) を用いた超解像手法 SRGAN を提案した [12]. GAN とは、生成ネットワーク (Generator) と識別ネットワーク (Discriminator) の2つを敵対的に学習させることで、複雑な信号の生成を可能にするネットワークモデルである [15]. 図1に示す Generator は従来と同様に、低解像度画像を入力し、高解像度画像が出力されるように学習を行う. 図2に示す Discriminator は、高解像度画像または Generator が生成した超解像画像を入力し、高解像度画像である確率を出力する. Discriminator の出力を Generator の損失関数の一部に用いることで、単一の CNN では困難である高周波成分の生成が可能となる. しかし、高周波成分の正確な復元は保証されないため、超解像画像に歪みが生じてしまい客観品質は低下する. 一般に、歪みの抑制と高周波成分の生成はトレードオフの関係にある.

Cheon らはこのトレードオフを緩和するために、EUSR-PCL を提案した [13]. SRGAN の生成ネットワークを SRResNet から EUSR [14] と呼ばれる CNN に変更し、更に損失関数も変更することで、高周波成分を十分に生成しつつ、歪みの発生を SRGAN より抑えることに成功した. しかしながら、EUSR は SRResNet と比較してパラメータ数が膨大であり、学習効率の観点では優れているとは言えない. また、歪みの発生は減少しているものの、客観品質である PSNR や SSIM の値は、単一の EUSR に比べて、依然として優位に低くなってしまっている [13].

本研究では、客観品質を維持したまま高周波成分を生成するために、GAN の生成ネットワークの出力層 (最終層) に新たに直交射影層を追加することを提案する. 提案手法では、ダウンサンプリング結果が入力低解像度画像と一致する超解像画像全体の集合を考える. この集合は線形多様体であり、直交射影の計算も容易である. 提案手法では、出力層で直交射影を適用することで、入力低解像度画像との整合性が保証された超解像画像を生成できる. これにより、白飛びや黒飛びを含む歪みの発生が大幅に低減する. 更には、生成ネットワークの学習時の損失関数として、EUSR-PCL で用いていた「高解像度画像と超解像画像の絶対誤差」、「差分画像の絶対誤差」、「識別ネットワークの出力」の他に、「射影前の線形多様体との距離」を新たに用いることを提案する. 数値実験により、提案手法は、単一 CNN による超解像結果と同等以上の客観品質を保ちつつ、真の高解像度画像と酷似した高周波成分を生成することを示す.

2. GAN を用いた超解像手法

2.1 ダウンサンプリングの定式化

超解像を行う低解像度画像を Y で表す. 低解像度画像 Y の縦方向の画素数を I , 横方向の画素数を J とし、画素値を $Y_{i,j}$ ($i = 1, 2, \dots, I; j = 1, 2, \dots, J$) とする. 本研究では、低解像度画像 Y は、縦方向の画素数が IK , 横方向の画素数が JL の高解像度画像 X をダウンサンプリングすることで得られるものとする. また、ダウンサンプリングは重み付き平均でモデル化されるものとする. 具体的には、高解像度画像 X の画素値を $X_{i,j}$ ($i = 1, 2, \dots, IK; j = 1, 2, \dots, JL$) とし、重みを $w_{k,l} > 0$ ($k = 1, 2, \dots, K; l = 1, 2, \dots, L$) とすれば、低解像度画像は

$$Y_{i,j} = \sum_{k=1}^K \sum_{l=1}^L w_{k,l} X_{(i-1)K+k, (j-1)L+l} \quad (1)$$

のように与えられる. ここで、重み $w_{k,l}$ は $\sum_{k=1}^K \sum_{l=1}^L w_{k,l} = 1$ を満たしている. 低解像度画像 Y と高解像度画像 X をベクトル化したものを、それぞれ $\mathbf{y} \in \mathbb{R}^{IJ}$, $\mathbf{x} \in \mathbb{R}^{IJKL}$ とすれば、適当な行列 $A \in \mathbb{R}^{IJ \times IJKL}$ を用いて、式 (1) は

$$\mathbf{y} = A\mathbf{x} \quad (2)$$

と表される.

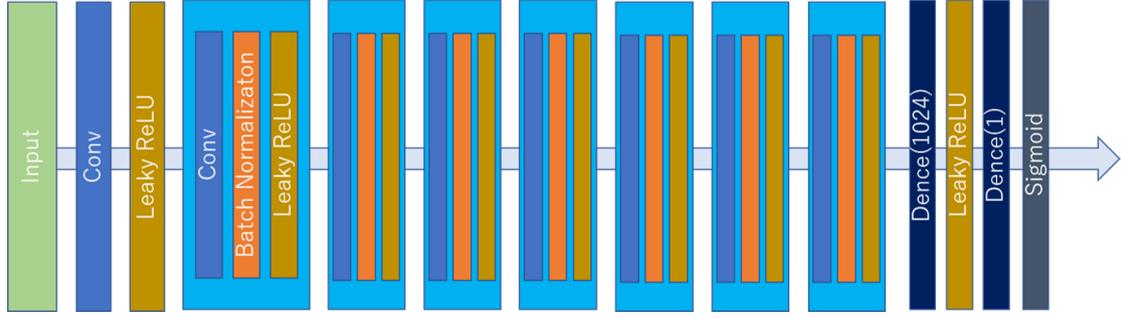


図2 SRGANの識別ネットワーク

2.2 SRGAN

学習用の低解像度画像と高解像度画像の組をベクトル化したものを $\{(\mathbf{y}_n, \mathbf{x}_n)\}_{n=1}^N$ とする。低解像度画像 \mathbf{y}_n を入力としたときの生成ネットワークの出力超解像画像を $\hat{\mathbf{x}}_n$ で表す。LedigらはGANを用いた単一画像からの超解像手法を提案した [12]。この手法はSRGANと呼ばれ、生成ネットワークでは低解像度画像 \mathbf{y}_n から高解像度画像 \mathbf{x}_n への写像を学習するため、新規の低解像度画像を生成ネットワークに入力すれば、畳み込み演算により超解像画像が瞬時に生成される。識別ネットワークでは高解像度画像または超解像画像を入力し、入力が高解像度画像であれば1に近い値を、超解像画像であれば0に近い値を出力するように学習を行う。SRGANでは図1・図2の生成・識別ネットワークを用いており、特に図1は、優れた画像認識精度を誇るResNet [16]を改変したものでSRResNetと呼ばれる。

従来のCNNを用いた手法 [8–11] では、生成ネットワークの学習時の損失関数として高解像度画像と超解像画像の二乗誤差

$$L(\Theta) = \frac{1}{N} \sum_{n=1}^N \|\hat{\mathbf{x}}_n - \mathbf{x}_n\|_2^2 \quad (3)$$

を採用し、この値が小さくなるようにネットワークパラメータ Θ を更新する。二乗誤差を用いると、歪みがなく、客観品質であるPSNRの値が高い超解像画像を生成できるが、一方で画像全体が平滑化されてしまうため高周波成分の復元に限界がある。そこで、SRGANでは損失関数として、式(3)の代わりに

$$L(\Theta) = \frac{1}{N} \sum_{n=1}^N \|\phi(\hat{\mathbf{x}}_n) - \phi(\mathbf{x}_n)\|_2^2 - \frac{\alpha}{N} \sum_{n=1}^N \log D(\hat{\mathbf{x}}_n) \quad (4)$$

を用いる。ここで、 $\phi(\mathbf{x})$ は画像認識ネットワークVGG19 [17]によって得られる特徴ベクトル、 $D: \mathbb{R}^{IJKL} \rightarrow (0, 1)$ は識別ネットワーク、 $\alpha > 0$ は第2項の重みを表している。式(4)では、VGG19と識別ネットワークが学習した真の高解像度画像らしさを表す特徴を、生成ネットワークの学習に利用できる。これにより、従来困難であった高周波成分の生成が可能となる。

2.3 EUSR-PCL

CheonらはSRGANが生成した超解像画像には多くの歪みが生じていることを指摘し、これを緩和するEUSR-PCLを提案した [13]。EUSR-PCLでは、生成ネットワークにEUSR [14]を使用し、その損失関数を新たに定義した。識別ネットワークに関しては損失関数も含めてSRGANと同じものを使用した。提案された生成ネットワークの損失関数は以下の通りである。

Content loss l_c

高解像度画像と超解像画像の画素値の差を、 ℓ_2 ノルムによる二乗誤差ではなく、 ℓ_1 ノルムによる絶対誤差

$$l_c(\Theta) = \frac{1}{N} \sum_{n=1}^N \|\hat{\mathbf{x}}_n - \mathbf{x}_n\|_1 \quad (5)$$

で評価する。

Differential content loss l_d

高解像度画像と超解像画像それぞれの水平・垂直方向の差分画像を、Content lossと同様に絶対誤差

$$l_d(\Theta) = \frac{1}{N} \sum_{n=1}^N \left(\|\Delta_h \hat{\mathbf{x}}_n - \Delta_h \mathbf{x}_n\|_1 + \|\Delta_v \hat{\mathbf{x}}_n - \Delta_v \mathbf{x}_n\|_1 \right) \quad (6)$$

で評価する。ここで、 Δ_h と Δ_v は水平と垂直方向の差分行列を表す。画像の構造を差分を用いて詳細に比較することで、過度の平滑化を防ぐとともに、高周波成分の復元性能を向上させる。

Adversarial loss l_a

SRGANにおける式(4)の損失関数の第2項と同様に、

$$l_a(\Theta) = -\frac{1}{N} \sum_{n=1}^N \log D(\hat{\mathbf{x}}_n) \quad (7)$$

を用いる。識別ネットワークの出力を考慮することで、真の高解像度画像と類似した高周波成分を復元できるようになる。

以上の3つの指標を用いて、最終的な損失関数を

$$L(\Theta) = l_c(\Theta) + \alpha l_d(\Theta) + \beta l_a(\Theta) \quad (8)$$

と定義した。ここで、 $\alpha, \beta > 0$ は第2項、第3項の重みである。生成ネットワークEUSRを式(8)の損失関数で学習させることで、歪みの発生を抑えながら高周波成分を生成できる。しかしながら、客観品質であるPSNRやSSIMの値は、式(3)の二乗誤差を用いた場合と比較して、優位に低くなる。また、式(8)の損失関数をSRGANにそのまま適用させてみると、重みを適切に設定しても、図3のように白飛びや黒飛びが発生してしまう。

3. 直交射影層を導入したGANによる超解像

従来手法 [6–14] では、生成ネットワークの学習時に、高解像度画像 \mathbf{x}_n と超解像画像 $\hat{\mathbf{x}}_n$ の差を主に評価している。しかし、超解像画像 $\hat{\mathbf{x}}_n$ を再度ダウンサンプリングして得られる画像 $A\hat{\mathbf{x}}_n$ が元の入力低解像度画像 \mathbf{y}_n に近いかどうかは考慮されていない。一方で、真の高解像度画像 \mathbf{x}_n に関しては、式(2)より

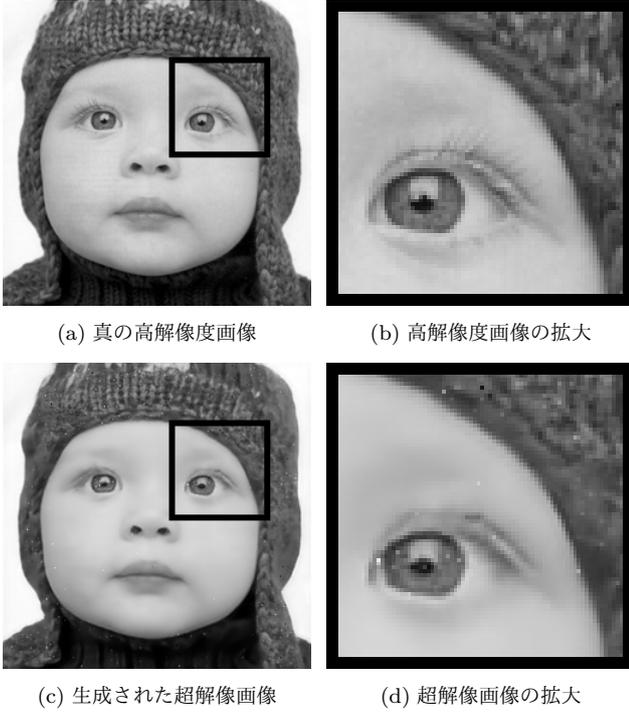


図3 SRGAN に式 (8) の損失関数を適用させた結果

$Ax_n = y_n$ が常に成立することから、学習データの情報を十分に活用するには、超解像画像のダウンサンプリング結果も考慮する必要がある。そこで本研究では、客観品質を維持したまま高周波成分を生成するために、 $A\check{x}_n = y_n$ を満たす超解像画像 \check{x}_n が生成ネットワークから常に出力される手法を提案する。

ダウンサンプリング結果が入力低解像度画像 y_n と一致する超解像画像の集合

$$\mathcal{A}_n := \{x \in \mathbb{R}^{IJKL} \mid Ax = y_n\} = \{x_n + z \in \mathbb{R}^{IJKL} \mid Az = 0\}$$

を定義すると、 \mathcal{A}_n は線形多様体であり、 $x_n \in \mathcal{A}_n$ が常に成り立つ。本研究では、文献 [18, 19] と同様に、 \mathcal{A}_n への直交射影 $P_{\mathcal{A}_n} : \mathbb{R}^{IJKL} \rightarrow \mathcal{A}_n$ を従来の生成ネットワークの出力結果 \hat{x}_n に適用することで、低解像度画像 y_n との整合性が保証された超解像画像 $\check{x}_n := P_{\mathcal{A}_n}(\hat{x}_n)$ を生成する。 \check{x}_n は具体的に

$$\begin{aligned} \check{x}_n &= \operatorname{argmin}_{x \in \mathcal{A}_n} \|\hat{x}_n - x\|_2 = \hat{x}_n - A^T(AA^T)^{-1}(A\hat{x}_n - y_n) \\ &= \hat{x}_n - \frac{A^T(A\hat{x}_n - y_n)}{\sum_{k=1}^K \sum_{l=1}^L w_{k,l}^2} \end{aligned}$$

のように与えられる。

式 (3) の「高解像度画像 x_n と従来の出力結果 \hat{x}_n の二乗誤差」は、図 4 で示したように、「 \mathcal{A}_n に垂直な成分の二乗誤差」と「 \mathcal{A}_n に平行な成分の二乗誤差」に分解できるため、

$$\begin{aligned} \|\hat{x}_n - x_n\|_2^2 &= \|\hat{x}_n - P_{\mathcal{A}_n}(\hat{x}_n)\|_2^2 + \|P_{\mathcal{A}_n}(\hat{x}_n) - x_n\|_2^2 \\ &\geq \|P_{\mathcal{A}_n}(\hat{x}_n) - x_n\|_2^2 = \|\check{x}_n - x_n\|_2^2 \quad (9) \end{aligned}$$

が成立する。式 (9) から確認できるように、 $A\hat{x}_n \neq y_n$ であれば、直交射影を適用することで高解像度画像 x_n との二乗誤差が小さくなり、客観品質である PSNR が必ず向上する。また、

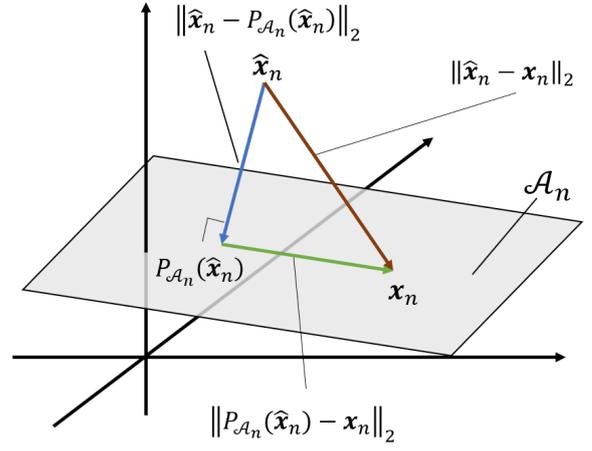


図4 提案手法で用いる線形多様体 \mathcal{A}_n とその空間内で算出する誤差

出力が $P_{\mathcal{A}_n}(\hat{x}_n)$ になるため、Content loss を式 (5) から

$$l_c(\Theta) = \frac{1}{N} \sum_{n=1}^N \|P_{\mathcal{A}_n}(\hat{x}_n) - x_n\|_1 \quad (10)$$

に変更する。更に、射影前の段階でなるべく線形多様体 \mathcal{A}_n に近い画像を生成するために、「 \mathcal{A}_n に垂直な成分の二乗誤差」を **Projection loss** l_p

$$l_p(\Theta) = \frac{1}{N} \sum_{n=1}^N \|\hat{x}_n - P_{\mathcal{A}_n}(\hat{x}_n)\|_2^2 \quad (11)$$

として評価する。式 (6), (7), (10), (11) を用いて、損失関数は

$$L(\Theta) = l_c(\Theta) + \alpha l_d(\Theta) + \beta l_p(\Theta) + \gamma l_a(\Theta) \quad (12)$$

となる。ここで、 $\alpha, \beta, \gamma > 0$ は各項の重みで、次節の実験では $\alpha = 0.001$, $\beta = 0.5$, $\gamma = 0.001$ とした。最適化アルゴリズムには、adaptive moment estimation (ADAM) [20] を用いた。

4. 数値実験

ImageNet^(註1) から 3,286 枚の自然画像を取得し、それぞれの画像を 90 度、180 度回転させて得られる画像を更に追加した、合計 9,858 枚の画像を用いてネットワークの学習を行う。学習データの高解像度画像として、データセット内の画像を双三次補間でサイズ 96×96 に縮小した後に、グレースケールに変換したものを用いる。式 (1) において、 $K = L = 2$, $w_{k,l} = 1/4$ ($k, l = 1, 2$) として高解像度画像をダウンサンプリングすることで、サイズ 48×48 の学習データの低解像度画像を用意する。テスト画像には、超解像性能の比較によく用いられる Set5 [21] (baby, bird, butterfly, head, woman) を利用する。テスト画像は、高解像度画像のサイズを 256×256 、低解像度画像のサイズを 128×128 として、SRGAN の生成ネットワーク SRResNet を二乗誤差で学習したもの (従来手法 1)、SRGAN に式 (8) を適用させたもの (従来手法 2)、SRGAN に直交射影層を追加して式 (12) を適用させたもの (提案手法) による超解像結果を比較する。SRGAN の生成ネットワークを学習する際に、従来手法 1 で学習したパラメータを初期値として設定する。

(註1) : URL: <http://www.image-net.org/>

表 1 PSNR / SSIM による画質評価

| | 従来手法 1 | 従来手法 2 | 提案手法 |
|-----------|-----------------------|----------------|------------------------------|
| baby | 33.05 / 0.9212 | 31.13 / 0.8897 | 32.91 / 0.9327 |
| bird | 36.58 / 0.9662 | 34.65 / 0.9457 | 36.27 / 0.9673 |
| butterfly | 29.48 / 0.9382 | 28.67 / 0.9188 | 29.59 / 0.9448 |
| head | 34.59 / 0.8679 | 33.07 / 0.8399 | 35.19 / 0.8963 |
| woman | 33.08 / 0.9522 | 31.97 / 0.9348 | 32.88 / 0.9576 |
| 平均 | 33.35 / 0.9273 | 32.29 / 0.9057 | 33.36 / 0.9397 |

図 5 はそれぞれの手法によって生成された超解像画像であり、表 1 は客観的な評価指標である PSNR 及び SSIM の値を示している。PSNR は二乗誤差が小さいほど良好な値をとる。このため、二乗誤差を損失関数としている従来手法 1 が良好な値を示しやすいが、butterfly と head では提案手法が最も良好な値を示しており、平均では従来手法 1 を 0.01 dB 上回った。人間の知覚に近い指標である SSIM においては、提案手法が全ての画像に対して最も良好な値を示した。表 2 は非参照画質スコアである NIQE [22] の値を示している。NIQE は単一画像からの画質評価であり、値が小さいほど自然画像らしいことを意味する。提案手法がいずれの画像においても最小値を示しており、特に baby, butterfly, head は真の高解像度画像とほぼ同じ値である。また、超解像画像を見ると、従来手法 1 は歪みの少ない画像を生成できているが、画像全体が平滑化されて高周波成分が失われている。従来手法 2 は従来手法 1 に比べて高周波成分を生成できているが、白飛びや黒飛びが発生してしまっている。一方、提案手法では真の高解像度画像と類似した高周波成分を復元することができ、白飛びや黒飛びも除去できている。

5. おわりに

本研究では、客観品質を維持しつつ高周波成分が復元された超解像画像を生成するために、敵対的生成ネットワークに直交射影層を導入した超解像手法を提案した。従来の CNN による画像超解像では画像全体が平滑化されて高周波成分が失われてしまい、GAN を用いた画像超解像では高周波成分が生成されるが、画像に歪みも生じていた。提案手法では、GAN による超解像画像を、ダウンサンプリング結果が入力画像と一致する線形多様体に射影することで、歪みの発生を大幅に低減させた。数値実験によって、提案手法の超解像結果は、高い客観品質を示すだけでなく、視覚的にもアーチファクトが除去されている、真の高解像度画像に酷似した鮮明な画像となることを示した。

文 献

- [1] 奥富 正敏, 田中 正行, 竹島 秀則, 松本 信幸, “画像超解像処理技術の最新動向,” 電子情報通信学会誌, vol. 93, no. 8, pp. 693–698, 2010.
- [2] 大倉 直, 高橋 友和, 井手 一郎, 村瀬 洋, “超解像処理を用いた低解像度文字の認識,” 情報科学技術フォーラム講演論文集, vol. 7, no. 3, pp. 65–66, 2008.
- [3] B. Hauang, W. Wang, M. Bates, and X. Zhuang, “Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy,” *Science*, vol. 319, no. 5864, pp. 810–813, 2008.
- [4] 松尾 康孝, “4K・8K テレビと超解像技術,” 映像情報メディア学会誌, vol. 69, no. 7, pp. 548–552, 2015.

表 2 NIQE による画質評価

| | 従来手法 1 | 従来手法 2 | 提案手法 | 真の高解像度画像 |
|-----------|--------|--------|--------------|----------|
| baby | 6.825 | 6.884 | 4.569 | 4.569 |
| bird | 6.194 | 5.472 | 5.268 | 4.746 |
| butterfly | 6.138 | 6.607 | 5.453 | 5.458 |
| head | 5.712 | 6.824 | 4.705 | 4.635 |
| woman | 7.685 | 8.053 | 5.231 | 4.377 |
| 平均 | 6.510 | 6.768 | 5.045 | 4.760 |

- [5] 高木 幹雄, 下田 陽久, 画像解析ハンドブック. 東京大学出版会, 2004.
- [6] J. Yang, J. Wright, T. S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [7] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY: Springer, 2010.
- [8] C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” *IEEE TPAMI*, vol. 38, no. 2, pp. 295–307, 2016.
- [9] C. Dong, C. C. Loy, and X. Tang, “Accelerating the super-resolution convolutional neural network,” in *Proc. ECCV*, 2016, pp. 391–407.
- [10] J. Kim, J. K. Lee, and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks,” in *Proc. CVPR*, 2016, pp. 1646–1654.
- [11] W. Shi, J. Caballero, F. Huszár, J. Totz, A. Aitken, R. Bishop, D. Ruecker, and Z. Wang, “Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network,” in *Proc. CVPR*, 2016, pp. 1874–1883.
- [12] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, “Photo-realistic single image super-resolution using a generative adversarial network,” in *Proc. CVPR*, 2017, pp. 105–114.
- [13] M. Cheon, J. H. Kim, J. H. Choi, and J. S. Lee, “Generative adversarial network-based image super-resolution using perceptual content losses,” in *Proc. ECCV*, 2018, pp. 51–62.
- [14] J. H. Kim and J. S. Lee, “Deep residual network with enhanced upscaling module for super-resolution,” in *Proc. CVPR Workshops*, 2018, pp. 800–808.
- [15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Proc. NIPS*, 2014, pp. 2672–2680.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. CVPR*, 2016, pp. 770–778.
- [17] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *Proc. ICLR*, 2015, 14 pages.
- [18] L. Condat and A. Montanvert, “A framework for image magnification: induction revisited,” in *Proc. ICASSP*, 2005, pp. 845–848.
- [19] 馬場 敦之, 片岡 秀公, 北原 大地, 平林 晃, “直交射影層を用いた畳み込みニューラルネットワークによる画像超解像,” vol. 117, no. 516, pp. 347–352, 2018.
- [20] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” in *Proc. ICLR*, 2015, 15 pages.
- [21] M. Bevilacqua, A. Roumy, C. Guillemot, and M. L. Alberi-Morel, “Low-complexity single-image super-resolution based on nonnegative neighbor embedding,” in *Proc. BMVC*, 2012, 10 pages.
- [22] A. Mittal, R. Soundararajan, and A. C. Bovik, “Making a “completely blind” image quality analyzer,” *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, 2013.



(a) 従来手法 1

(b) 従来手法 2

(c) 提案手法

(d) 真の高解像度画像

図 5 従来手法 1 (SRResNet と式 (4)), 従来手法 2 (SRGAN と式 (8)), 提案手法 (SRGAN と直交射影と式 (12)) による超解像結果の比較