

CNNを用いた段階的な一枚超解像 Stepwise Single-Image Super-Resolution Using Convolutional Networks

朴君[†] 白石 壮馬[†] 井下 哲夫[†] 岩元 浩太[†]
Jun Piao Soma Shiraishi Tetsuo Inoshita Kota Iwamoto

1. はじめに

近年、カメラ映像などの高精細化や画像認識精度の向上を目的に超解像技術が注目されている。超解像技術は低解像度で入力された画像を高解像度化する技術であり、従来から様々な手法が提案されている[1-3]。例えば、事例ベース学習型超解像は、パッチ単位で高解像度画像と低解像度画像間の対応関係を辞書とし、辞書から入力画像に適合するパッチを探索して高解像度画像を合成する。一方近年、画像認識等の様々な領域で注目されている深層学習をベースにした超解像化手法 Super Resolution Convolutional Networks(SRCNN)[4]が提案されている。SRCNNは低解像度画像と高解像度画像間のマッピングをCNNにより end-to-end で直接学習することで、従来の事例ベース学習型超解像化手法に比べ、画質を飛躍的に向上した。

SRCNNでは、主に2~4倍の超解像化において評価を行っており、これらの倍率の超解像化では、構築された画像の画質が高いことが報告されている。しかしながら、例えば広域監視カメラを対象にした場合、小さく映る人・物等の被写体の細部の再現には、より高い倍率の超解像化が望まれる。このような高倍率超解像化に end-to-end で低解像度画像と高解像度画像間のマッピングを直接学習するSRCNNを適用すると、構築すべき高周波成分が多くなり、構築された画像の画質が極端に低くなる問題がある(図1(a))。

そこで本稿では、高倍率な超解像化においても高品質な高解像度画像の構築を実現するために、段階的に複数のCNNを用いる超解像化手法を提案する。以降、第2章で従来手法 SRCNN について述べ、第3章で提案手法について説明する。第4章で実験を行い、第5章でまとめとする。

2. 従来手法

本章では、従来手法であるCNNを用いた超解像化手法SRCNN[4]の概要について述べる。SRCNNは、単一のCNNを用いて、画像を分割することで得られるパッチ単位で、低解像度画像を直接超解像化することで、高解像度画像を構築する。CNNのネットワークの構造は、3層の畳み込み層を持ち、フィルタサイズはそれぞれ 9×9 、 5×5 、 5×5 であり、チャンネル数は64、32、1である。

学習では、前章で述べたように低解像度画像と高解像度画像間のマッピングを end-to-end のCNNで学習することで、ネットワークのパラメータ $\theta = \{W_1, W_2, W_3, B_1, B_2, B_3\}$ を推定する。ここで、 W_1, W_2, W_3 は、各畳み込み層で畳み込みを行うフィルタ群(入力チャンネル数 \times フィルタサイズ \times 出力チャンネル数)であり、 B_1, B_2, B_3 は、各畳み込み層でのバイアスを表す。高解像度画像 \mathbf{X} と、 \mathbf{X} をダウンサンプリン

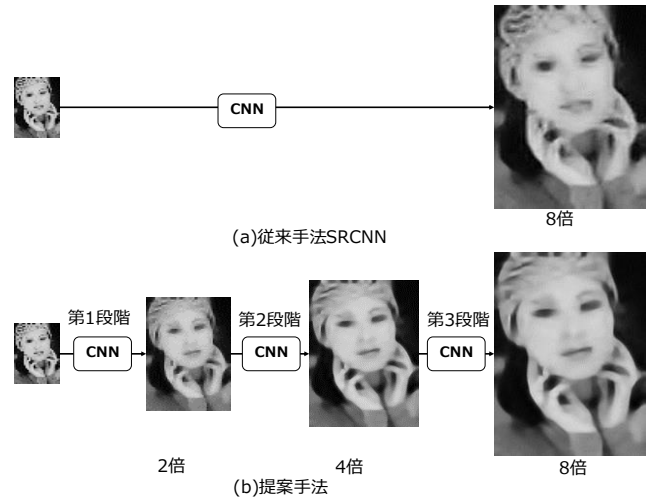


図1. 提案手法と従来手法の比較。(8倍超解像化)。

グすることで得られた低解像度画像 \mathbf{Y} を用い、ネットワークパラメータ θ を学習する。具体的には、低解像度画像から構築された画像 $F(\mathbf{Y}; \theta)$ と高解像度画像 \mathbf{X} 間の損失関数 $L(\theta)$ を下式で定義する。

$$L(\theta) = \frac{1}{N} \sum \|F(\mathbf{Y}; \theta) - \mathbf{X}\|^2 \quad (1)$$

ここで、 N は画像の総数を表す。定義された損失関数 $L(\theta)$ が最小となるように、パラメータ θ を更新する。

3. 提案手法

提案手法では、単一のCNNを用いた高倍率の超解像化問題を、複数の低倍率超解像化の問題に置き換え、複数のCNNを用いて、低解像度画像から段階的に超解像化を行うことで、最終的な高倍率の超解像化を実現する(図1(b))。各段階での超解像化には、従来手法 SRCNN のネットワークを用いる。このような段階的な超解像化を行うことで、各段階での超解像化の倍率を低く抑えることができ、各段階で高品質な画像が構築され、最終的に得られる高解像度画像の画質の向上が実現できる。

学習では、高解像度画像から複数の低解像度画像を生成し、それらを用いて、各段階におけるネットワークのパラメータを推定する。ここで、段階的な超解像化での段階の数を M とし、各段階 i により構築される画像の倍率を S_i ($i = 1, 2, \dots, M$)とすると、学習時の低解像度画像を生成するためのダウンサンプリングスケール k_i は、 $k_i = S_{i-1}/S_M$ となる。各段階で行う超解像化の倍率は S_{i-1}/S_i となる。ただし、 $S_0 = 1$ である。高解像度画像を \mathbf{X} とし、 \mathbf{X} をスケール k_i で、ダウンサンプリングして得られる複数の低解像度画像を \mathbf{Y}_{k_i} とする。ここで、 \mathbf{Y}_{k_1} ($k_1 = 1/S_M$)は解像度が最も低い低解像度画像である。各段階 i で超解像化を行うためのパラメ

[†] NEC データサイエンス研究所

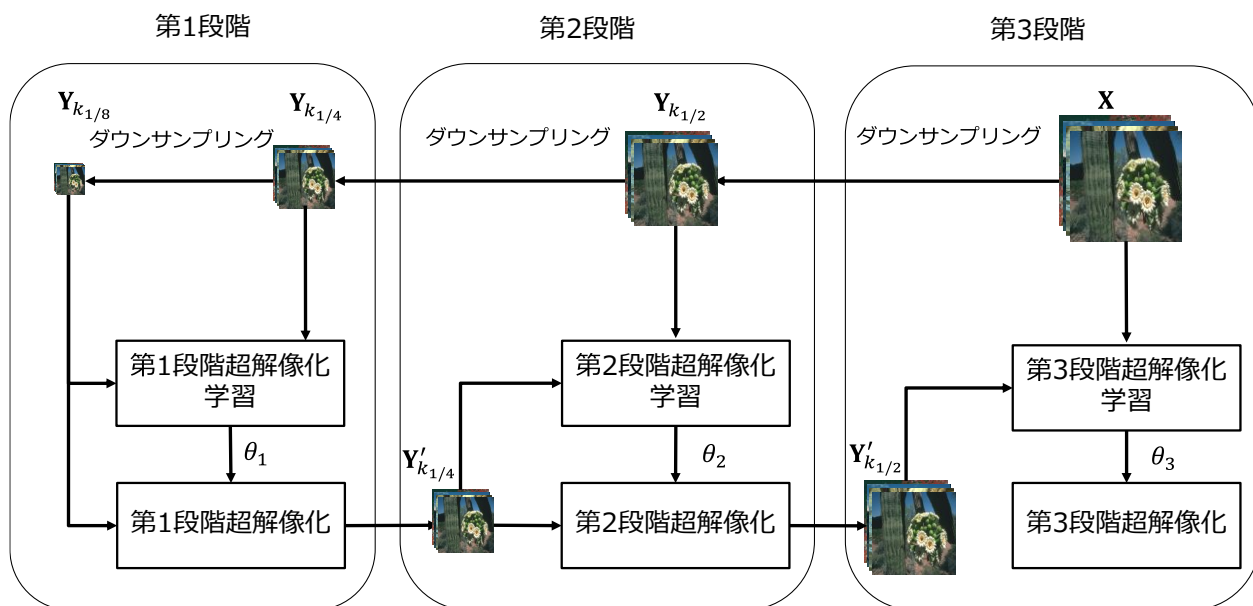


図 2. 段階的な超解像化の学習プロセス.

一タ θ_i は、画像 Y'_{k_i} と $Y_{k_{i+1}}$ 間のマッピングを式(2)の損失関数 $L(i; \theta_i)$ の最小化により推定される。

$$L(i; \theta_i) = \frac{1}{N} \sum \|F(Y'_{k_i}; \theta_i) - Y_{k_{i+1}}\|^2 \quad (2)$$

ここで、 Y'_{k_i} は前の段階で構築された画像であり、次式で表される。

$$Y'_{k_i} = F(Y'_{k_{i-1}}; \theta_{i-1}) \quad (3)$$

ただし、 $Y'_{k_1} = Y_{k_1}$ であり、第 M 段階の学習において、 $Y_{k_{M+1}} = X$ である。以上より、各段階の CNN のネットワークパラメータ θ_i が学習される。

図 2 に、例として、2 倍、4 倍、8 倍の 3 段階超解像化における学習プロセスの概要を示す。ここで、 $M = 3$ 、 $S_1 = 2$ 、 $S_2 = 4$ 、 $S_3 = 8$ 、 $k_1 = 1/8$ 、 $k_2 = 1/4$ 、 $k_3 = 1/2$ となる。

4. 実験

本実験では、提案手法の有効性を確認するため、従来手法 SRCNN と Bicubic との比較を行う。

4.1. 実験条件

本実験では、4 倍及び 8 倍の超解像化で評価する。従来手法 SRCNN[4] と Bicubic は低解像度画像から直接 4 倍及び 8 倍の超解像化を行う。提案手法では、4 倍超解像化においては、2 倍→4 倍 ($M = 2, S_1 = 2, S_2 = 4$) の順に 2 段階超解像化を行う。8 倍超解像化では、図 2 に示す通り、2 倍→4 倍→8 倍 ($M = 3, S_1 = 2, S_2 = 4, S_3 = 8$) の順に 3 段階超解像化を行う。

学習データと評価データは、超解像化分野においてベンチマークとして広く使われているデータセットを用いる。学習データとしては、4 倍超解像化については、91 枚の画像[4]を用い、8 倍超解像化については、解像度が比較的高い 100 枚の画像 Urban100 を用いる。評価データは Set5[2]、Set14[5] を用いる(図 3, 4)。実験では、Set5 と Set14 の原画像をそれぞれ 1/4、1/8 にダウンサンプリングした画像を低解像度として、それぞれの手法で超解像化し、構築された画像の画質を Peak Signal to Noise Ratio(PSNR)で評価する。



図 3. Set5

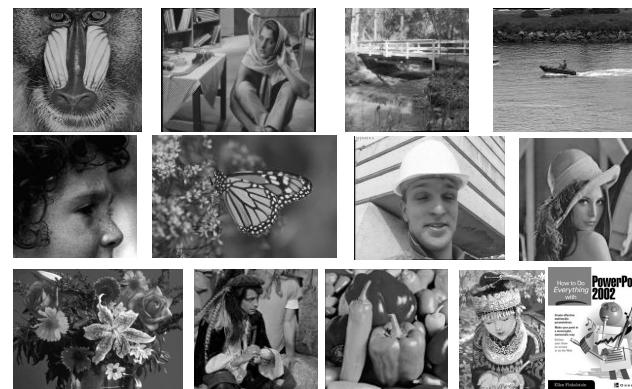


図 4. Set14

4.2. 実験結果

4 倍及び 8 倍の超解像化の PSNR の結果を表 1, 2 に示す。表 1, 2 から、評価した全ての画像について、提案手法の PSNR の値が従来手法 SRCNN から向上したことが確認できる。4 倍超解像化については、Set5 で PSNR の値が平均 1.35dB 向上し、Set14 では平均 1.27dB 向上した。8 倍の超解像化については、Set5 で PSNR が平均 1.33dB 向上し、Set14 で平均 1.29dB 向上した。

図 5 に、8 倍超解像化における各手法で構築された画像の例を示す。baboon(図 5(a))の結果を見ると、目の一部の構造が破綻している SRCNN に比べ、提案手法では、目の構造を保ちながら再構築されていることが確認できる。ppt3(図 5(b))の結果においても同じく、SRCNN に比べ、提案手法は数字の構造が精細に構築されていることが確認できる。他の画像結果からも、同じような傾向が見える。以上から、高倍率の超解像化において、段階的に超解像化する提案手法の有効性が確認できた。

表 1. 4 倍超解像化の PSNR の結果

Dataset	Image	Bicubic	SRCNN	Proposed method
Set5	head	29.33	29.88	30.67
	baby	29.80	30.02	30.98
	butterfly	19.91	21.07	23.06
	woman	24.41	25.11	26.61
	bird	28.12	28.60	30.12
	平均値	26.31	26.93	28.28
Set14	lenna	27.71	28.55	29.55
	monarch	25.37	26.29	27.92
	flowers	23.26	23.85	25.17
	bridge	21.69	22.02	23.03
	pepper	26.54	27.33	29.03
	zebra	22.25	22.89	24.19
	ppt3	19.28	19.78	21.98
	baboon	19.40	19.93	20.79
	coastguard	22.02	22.59	23.99
	face	29.32	29.83	30.64
	man	23.20	23.94	25.07
	barbara	22.54	22.75	23.59
	foreman	25.05	25.64	27.09
	comic	19.05	19.66	20.81
	平均値	23.33	23.93	25.20

表 2. 8 倍超解像化の PSNR の結果.

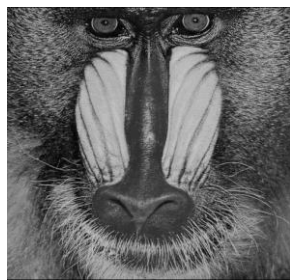
Dataset	Image	Bicubic	SRCNN	Proposed method
Set5	head	26.39	26.88	27.98
	baby	24.56	25.31	26.54
	butterfly	15.25	15.65	17.11
	woman	19.75	20.20	21.75
	bird	22.71	22.78	24.07
	平均値	21.73	22.16	23.49
Set14	lenna	23.55	24.27	25.67
	monarch	20.94	20.85	22.20
	flowers	19.52	19.92	21.09
	bridge	19.08	19.70	20.70
	pepper	23.43	23.77	25.08
	zebra	16.94	17.72	18.85
	ppt3	15.90	16.47	17.99
	baboon	18.02	18.88	19.69
	coastguard	19.59	20.34	21.97
	face	26.26	26.55	27.76
	man	20.11	20.83	21.98
	barbara	20.17	20.93	22.23
	foreman	20.78	21.05	22.98
	comic	16.26	16.84	18.01
	平均値	20.03	20.58	21.87

5. まとめ

本研究では、CNN を用いた段階的な超解像法を提案した。具体的には、単一の CNN を用いた高倍率の超解像化問題を複数の CNN を用いた低倍率超解像化問題に置き換え、低解像画像から段階的に超解像化を行うことで、最終的な高倍率の超解像化を実現する。実験により、4 倍超解像化については、Set5 と Set14 で PSNR の値が、それぞれ 1.35dB と 1.27dB 向上し、8 倍の超解像化については、Set5 と Set14 で PSNR の値が、それぞれ 1.33dB と 1.29dB 向上したことを確認した。さらに、8 倍超解像化により構築された画像の比較からも提案手法の方がより高品質であることが確認できた。

参考文献

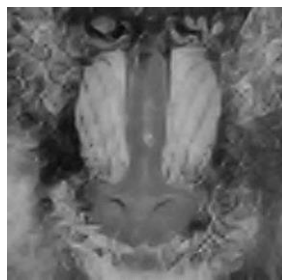
- [1] R. Timofte, V. D. Smet, and L. V. Gool, "Anchored neighborhood regression for fast example-based super resolution," IEEE International Conference on Computer Vision, 2004.
- [2] M. Bevilacqua, A. Roumy, C. Guillemot, and M. L. A. Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," British Machine Vision Conference, 2012.
- [3] R. Timofte, V. D. Smet, and L. V. Gool, "A+: Adjusted anchored neighborhood regression for fast super resolution," Asian Conference on Computer Vision, 2014.
- [4] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," European Conference on Computer Vision, 2014.
- [5] Zeyde, R., Elad, M., Protter, M., "On single image scale-up using sparse representation," Curves and Surfaces, 2012.



Original/PSNR



Bicubic/18.02

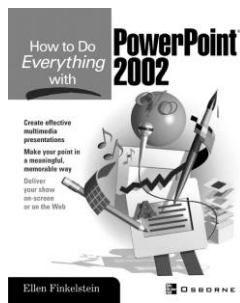


SRCNN/18.88



Proposed method/19.69

(a) baboon



Original/PSNR



Bicubic/15.90

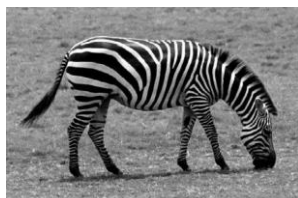


SRCNN/16.47



Proposed method/17.99

(b) ppt3



Original/PSNR



Bicubic/16.94



SRCNN/17.72

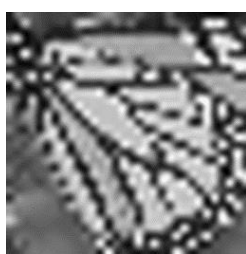


Proposed method/18.85

(c) zebra



Original/PSNR



Bicubic/15.25



SRCNN/15.65

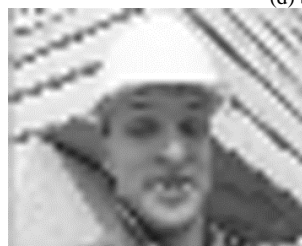


Proposed method/17.11

(d) butterfly



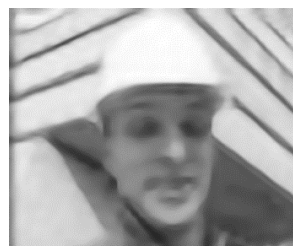
Original/PSNR



Bicubic/20.78



SRCNN/21.05



Proposed method/22.98

(e) foreman



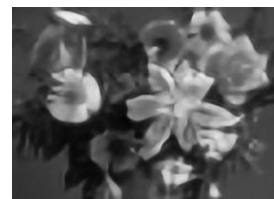
Original/PSNR



Bicubic/19.52



SRCNN/19.92



Proposed method/21.09

(f) flowers

図 5.8 倍超解像化の結果.