

画像超解像と劣化除去の同時反復学習

吉田 智樹^{1,a)} 秋田 和俊¹ 浮田 宗伯^{1,b)}

概要: 深層学習を利用する超解像 (SR) 手法の多くは, 高解像度画像 (HR 画像) が双三次補間法などの特定の過程によってのみ縮小された低解像度画像 (LR 画像) に対する SR モデルを学習している. このような手法では, その特定の劣化以外の LR 画像に対しては, 性能が落ちてしまう. 実応用のためには, 多様な任意の劣化に対応可能な手法が望まれる. 任意の劣化に対応するための SR として, 入力 LR 画像が得られた劣化の種類を推定し, その劣化に対応した SR を実行する手法が多く提案されている. しかし, 既存手法の多くは, 劣化推定の学習には劣化カーネルと真値との損失のみで学習している. これに加え, 得られる SR 画像と劣化推定の結果を使って擬似 LR 画像を生成し, それと入力 LR 画像との誤差 (差分) も用いることで, 劣化推定の性能を上げられると考える. そこで, 本研究では, この差分を損失に加え, さらに誤差をモデル内でも使用し, 画像特徴量に付与することで, 多様な劣化に対応した SR 画像生成の精度も上げることを目指す. また, SR と劣化除去 (ボケ除去) を反復的に行うことで, より性能の高い SR を目指す.

キーワード: 超解像, ブラインド超解像, 劣化除去

1. はじめに

画像超解像 (super-resolution, 以下 SR) とは, 低解像度画像 (LR 画像) から高解像度画像 (HR 画像) を生成する画像処理技術の 1 つである. LR 画像から HR 画像への変換は不良設定問題として長年研究されているが, 近年は深層学習の進歩により, それを応用した画像超解像も飛躍的な進歩を遂げている [1, 2, 4-7]. しかし, ほとんどの SR 手法は, 双三次補間法などの特定の画像縮小過程によって生成された LR 画像のみを学習している. このため, 学習されていない未知の劣化を含む LR 画像に対しては, 性能が落ちてしまう. しかし, 実応用のためには多様な劣化に対応する必要がある.

多様な劣化に対応する SR に取り組む時, 対象の LR 画像 I_{LR} は次式 (1) のように表され, 真値 HR 画像 I_{HR} に劣化カーネル k が畳み込まれた後, 任意の縮小過程 \downarrow_s によって得られたものとされる. $*$ は畳み込み演算, s は倍率, n はノイズである.

$$I_{LR} = (I_{HR} * k) \downarrow_s + n \quad (1)$$

(1) 式に基づき, 多様な劣化を考慮した [8, 9] のような SR

手法では, 既知劣化カーネル k を用いている. さらに, 一般に劣化カーネル k は未知であるため, 現在では [10, 12] のように, 劣化カーネルを推定した上で SR を行う手法, ブラインド SR も盛んに研究されている. 以上のような, 多様な劣化に対応する SR 手法における劣化カーネルは, モデル内で LR 画像とともに入力したり, [12] のように劣化カーネルと中間特徴量との演算結果を特徴量とするといった使用方法が主流である. このようにして劣化を考慮した SR は, 多様な劣化に対して性能を上げることに成功しており, このことから, 多様な劣化を考慮する SR では, 劣化カーネルの推定および利用が重要なことが分かる. しかし, 劣化カーネルを使用する SR において, 劣化カーネルの推定精度の評価は, 真値との誤差のみであること多く, またそれを損失関数でしか使っていない. このような評価の他にも, 推定劣化カーネルを式 (1) に当てはめて, LR 画像を生成し, それを入力 LR 画像と比較する方法もある. この評価法も用いれば, さらに精度の高い劣化推定が可能になるだろう.

そこで本研究では, モデル内処理の途中段階で, 推定した劣化カーネルと生成した SR 画像を, 式 (1) に従って畳み込んで擬似 LR 画像を生成する. そして, 擬似 LR 画像と入力 LR 画像の誤差を画像特徴量に付与して, 変化させることで, モデル内で推定誤差を考慮したような特徴量抽出を行わせる. これにより, 精度の高い劣化カーネル推定

¹ 豊田工業大学

Toyota Technological Institute, Hisakata, Tempaku-ku, Nagoya, Aichi, 468-8511, Japan

a) sd19455@toyota-ti.ac.jp

b) ukita@toyota-ti.ac.jp

とそれを用いた劣化除去（ボケ除去）を行う SR を提案する。また、擬似 LR 画像と入力 LR 画像の差分は損失関数としても用いることで、より高精度な劣化推定を行う SR を目指す。さらに、以上のような SR と劣化除去を交互に複数回行うことで、一度の SR と劣化除去では除ききれなかった劣化を除き、より性能のよい SR を実現することを目指す。

2. 関連研究

2.1 深層学習を用いた画像超解像

SR における大きな発展は、SR に初めて畳み込みニューラルネットワークを用いた、Dong らの SRCNN [1] に始まる。そこから、深層学習モデルの研究に伴い残差学習を導入した [2] や Goodfellow らが提案した敵対的生成ネットワーク [3] を用いた Ledig らの SRGAN [4] などが提案され、飛躍的に SR の性能を上げてきた。最近ではモデル構造以外に、損失関数にも深層学習モデルを使用した [5,6] も提案されている。このような SR モデルの内、我々の手法に近いモデルに、Haris らが提案した DBPN がある [7]。この DBPN は、画像特徴量の拡大と縮小を複数回繰り返す、その都度得られる SR 特徴量群を全て用いて、最終的な SR 画像を生成する SR モデルである。DBPN の特徴は、複数回の縮小により縮小過程に伴う様々な劣化を擬似的に抽出し、それらに対する拡大を学習することである。このようにして、様々な縮小過程による劣化に頑健な SR モデルを学習している。こうした、拡大と縮小を複数回行うというアルゴリズムが、本研究の、SR と劣化除去を交互に複数回行うアルゴリズムと近いものになっている。ただし、DBPN を含めたこれらの SR モデルは、双三次補間法のような特定の縮小のみを学習しており、それ以外の画像劣化を考慮するようなモデル構造や学習をしていないため、実応用には難しい手法となっている。

画像劣化を考慮するため、Riegler らの CAB [8] や Zhang らの SRMD [9] では、式 (1) に基づいて、LR 画像とともに、LR 画像を生成した劣化カーネル k も SR モデルに入力することで、画像劣化に対応しようとした。このようにすることで、[8,9] 以前と違い、特徴量に劣化カーネルに関する数値が付与されるようになるため、縮小過程に付随する劣化以外の画像劣化にも対応した SR モデルを学習させることができた。しかし、CAB や SRMD では劣化カーネルを既知とし、真値劣化カーネルを用いている。一般に劣化カーネル k は未知であることから、CAB や SRMD でも実応用に適したものではないと言える。

2.2 ブラインド超解像

2.1 節で述べたとおり、実応用に向けた SR に取り組む場合、一般に劣化カーネルは未知である。よって、画像劣化は、入力 LR 画像から推定しなければならない。このよう

な問題設定での SR が、ブラインド SR である。[10] のように、多くのブラインド SR 手法は、入力された LR 画像から劣化カーネル k を推定し、その劣化カーネルを CAB や SRMD のように、LR 画像とともにモデルに入力することで多様な劣化に対応しようとしている。Gu らの [12] も近い方法で画像劣化を考慮しているが、彼らは、入力 LR 画像とともに劣化カーネルを入力するのではなく、Wang らの SFT 層 [11] を用い、推定劣化カーネルを付加情報として、中間の画像特徴量を変化させるという改良をして、SR の性能を上げている。さらに IKC では、推定した劣化カーネルを反復的に補正し、補正後のカーネルを SR モデルで利用する、ということも複数回行っている。Gu らが推定劣化カーネルを複数回補正を行うアルゴリズムは、文献 [13,14] に基づいており、不適切な劣化カーネルを用いて SR をしたために、SR 画像には劣化が残ってしまったたり、余計なアーティファクトが生じたりすることを防ぐためのものである。

IKC は、反復的な SR と劣化推定によって、多様な劣化に対応した SR を実現した。ただし、IKC の劣化カーネル推定の学習は、主に推定劣化カーネルと真値劣化カーネルとの損失関数のみを基にしている。推定劣化カーネルの整合性は、真値の劣化カーネルとの差以外にも、生成 SR 画像と推定劣化カーネルから得られる擬似 LR 画像と入力 LR 画像との差でも、推定誤差を測ることは可能である。この差分は式 (1) に基づいたものでもあるため、物理的にも正しく、かつ強い制約となる。よってこの差分も利用すれば、より良い劣化カーネル推定を行うことができると考える。さらに、この誤差は SR 画像生成精度も含んでいるため、SR の性能を上げることが見込めるような非常に有用な測定値であると考えられる。

3. 提案手法

本手法では、SR モデルと、劣化カーネル推定によって画像劣化を除くモデルの 2 つを使って、多様な劣化に対応する SR を行う。それらを一つのモジュール（KB モジュール）に内包し、それを連結させることで、SR と劣化カーネル推定を反復的に行う。本手法の大きな特徴は 3 つある。1 つ目は式 (1) に従って、推定したカーネルを SR 画像に適用して得られる擬似 LR 画像と、入力 LR 画像との差分を学習に用いること。2 つ目は、SR と画像劣化除去を反復的に行うこと。3 つ目は SR と劣化除去を同時に学習することである。特に 1 つ目の特徴は他手法に見られないもののため、本章で詳しく述べる。以下、3.1 節では図 1 中の KB モジュールについて、3.2 節では劣化カーネル推定器について述べる。

3.1 KB モジュール

本手法で使う、SR モデルと劣化カーネル推定器を内包す

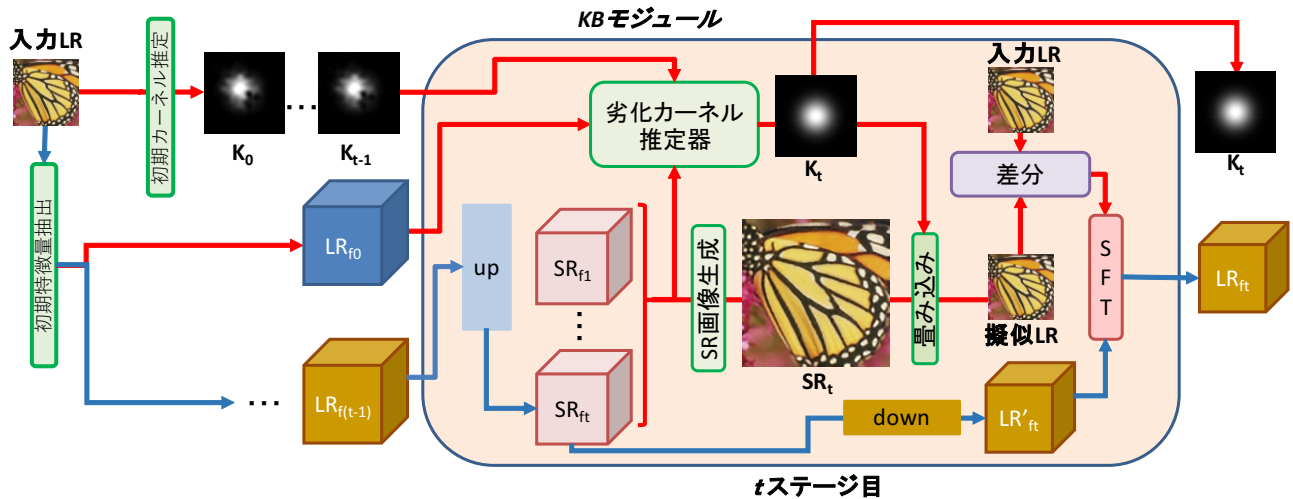


図1 提案手法概要図. 主な処理はKBモジュール(肌色枠内)で行い, このKBモジュールを連結することでSRと劣化推定を反復して行う. $t-1$ 個目のKBモジュールで得られたLRサイズの特徴量を, DBPN [7]のUp-projectionによって拡大して, SR特徴量 SR_{ft} を得る. 劣化カーネル推定は赤矢印に沿って行われ, 既に抽出されたSR特徴量群, 初期特徴量抽出で得られたLR特徴量 LR_{f_0} , $t-1$ ステージ目の推定カーネル K_{t-1} を入力とした, 劣化カーネル推定器で行う. そして, 推定劣化カーネル K_t と, 先程と同じSR特徴量群から生成した暫定SR画像 SR_t を用いて擬似LR画像を生成し, 入力LR画像との差分を取る. 最後にSFT層を使い, この差分を青矢印に沿って得られたLR特徴量 LR'_{ft} に付与してLR特徴量 LR_{ft} を得る.

るモジュールを kernelized back-projection モジュール (以下, KBモジュール)と呼ぶことにし, 図1中の肌色枠内を指す. このKBモジュールでは, DBPN [7]のup-projectionモジュールとdown-projectionモジュールを使用する(それぞれ図1中のupとdown. 以下upモジュール, downモジュールと呼ぶ). upモジュールとdownモジュールを使う理由は, DBPNの拡大縮小を繰り返すアルゴリズムと, 本手法のSRと劣化除去を繰り返すアルゴリズムが似ているからであることと, 反復的なアルゴリズムで縮小過程に伴う劣化除去に大きな成功をしているからである.

KBモジュールでは, 主に3つの処理を行う. 1つ目はSR画像生成に必要な特徴量抽出, 2つ目は推定SR画像と推定劣化カーネルから得られる擬似LR画像と入力LRとの差分(二乗誤差)を取ることで, 3つ目はその差分を画像特徴量に付与することである. 特に, 擬似LR画像と入力LR画像との差分は, SR画像生成精度と劣化カーネル推定精度の二つを評価するために重要である. なぜなら, この差分が小さければ, 得られたSR画像や推定劣化カーネルは真値と類似している可能性が高くなるためである. 逆に, 差分が大きい場合は, その差分を特徴量に付与することで, SR画像推定精度が悪い領域を強調することができ, SR画像生成にとって良い特徴量を抽出することが期待できる. また, このKBモジュールを反復的に応用する本手法では, KBモジュールを経由する度に有用な特徴量が得られると考える.

以下, KBモジュール内での処理を述べていく. 最初に $t-1$ 個目のKBモジュール($t-1$ ステージ目)で得られたLR特徴量 $LR_{f_{(t-1)}}$ をupモジュールに入力し, SR特徴量 SR_{ft} を得る. 次に, 二つに分岐して画像特徴量抽出をしていき, 一方では図1中青矢印に沿って, SR_{ft} をdownモジュールに入力してLR特徴量 LR'_{ft} を得る. 他方では, SR特徴量群 SR_{f_1} から SR_{ft} を使って, 劣化推定を行う. まずは, SR特徴量群からSR画像 SR_t を生成する. 次に, 同じSR特徴量群, 初期特徴量抽出で得られたLR特徴量, $t-1$ ステージ目の推定劣化カーネルを劣化カーネル推定器に入力して, 推定劣化カーネル K_t を得る. そして, 得られたSR画像 SR_t と推定劣化カーネル K_t を, 式(1)に従って畳み込み, 擬似LR画像を生成する. そして, 擬似LR画像と入力LR画像との差分を取る.

この差分と, 図1中青矢印に沿って得られたLR特徴量 LR'_{ft} を, [11]のSFT層に入力して LR_{ft} を得る. SFT層による LR'_{ft} から LR_{ft} の変換により, 擬似LR画像と入力LR画像との差分が大きい箇所の特徴量の値が大きくなり, 次のステージで, SR画像生成精度を上げるような特徴量を抽出するようにさせる. 使用するSFT層は, 図3と4.1で詳細を述べる.

このようにすることで, 複数のKBモジュールによって得られる特徴量は, 擬似LR画像と入力LR画像の差を小さくするような特徴量となっていく. それらから得られたSR特徴量によって, 求めるべきHR画像に近いSR画

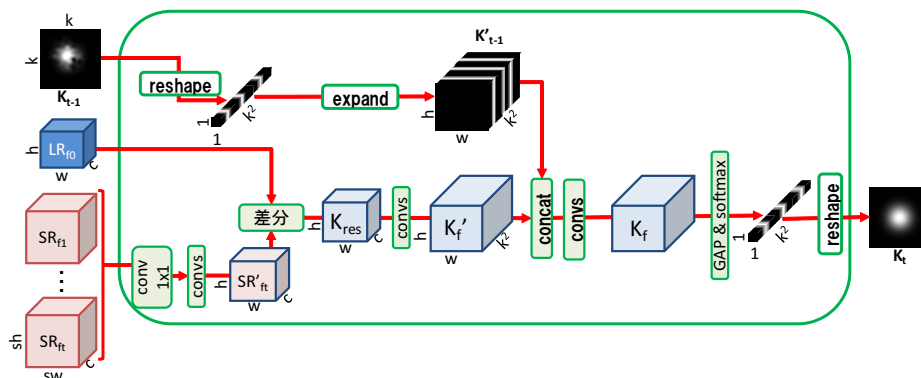


図 2 劣化カーネル推定器概要図。まず、 $1 \sim t$ ステージまでで得られた SR 特徴量群を畳み込み、初期特徴量抽出から得られた LR 特徴量 LR_{f0} と同じ大きさの特徴量 SR'_{ft} を生成し、それらの差分を取る。これを畳み込み層に入力し、 K'_f を得る。次に、 K'_f と前ステージで得られた劣化カーネルを変形させたものをチャンネル方向に連結し、畳み込むことで劣化カーネル特徴量 K_f を得る。 K_f に global average pooling と softmax を取ったベクトルを整形したものが、 t ステージ目の推定劣化カーネル K_t 。

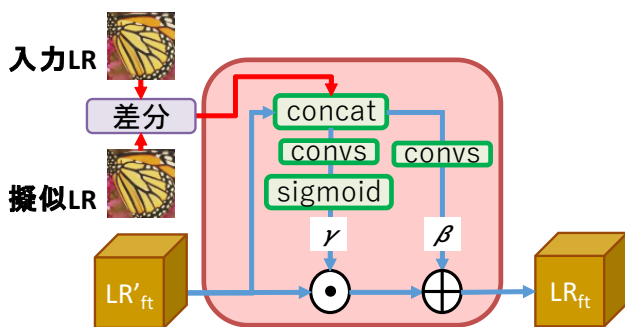


図 3 本手法で使う SFT 層。 γ 、 β は共に、フィルタサイズ 3 の畳み込み層 2 つから出力される。 \odot は要素積、 \oplus は要素和。

像が得られるようになると思われる。また、この差分を損失関数としても用いることで劣化カーネル推定精度の向上も望める。最終的な SR 画像の出力は、全ステージで得られた SR 特徴量群全てをチャンネル方向に連結し、一つの畳み込み層に入力することで得る。

また、この KB モジュールの拡張として、IKC [12] のように SFT 層を使って、劣化カーネル自体を LR 特徴量 LR_{ft} に付与することも考えられる。この拡張は劣化カーネルも考慮しつつも、差分の大きさに応じた特徴量抽出が可能になると考えられる。実際にこの拡張を行って、5.2 節で実験をしており、各 KB モジュールの直後に劣化カーネルを付与する SFT 層を入れた。

3.2 劣化カーネル推定器

本手法で用いる劣化カーネル推定器を図 2 に示した。本手法では、主に初期特徴量抽出で得られた LR 特徴量 LR_{f0} と t ステージまでで得られた SR 特徴量から得られる特徴量 SR'_{ft} から劣化カーネルを求める。特に LR_{f0} を用いた理由は、得られた SR 特徴量群が求める HR 画像を表現す

るのに十分であれば、解像度以外の SR 特徴量と初期特徴量 LR_{f0} との違いが劣化度合い、つまり劣化カーネルを表す特徴量になるという考察のためである。

この劣化カーネル推定器は、図 1 の通り、KB モジュール内で用いられ、1 から t ステージまでで得られた SR 特徴量群、初期特徴量抽出で得られた LR 特徴量 LR_{f0} 、 $t-1$ ステージで得られた推定劣化カーネルの 3 つが入力される。まずは、SR 特徴量群から重要な特徴量を抽出しつつ、 LR_{f0} と解像度を合わせるために、SR 特徴量群をチャンネル方向に連結し、それを数層の畳み込み層に入力する。ここから得られた特徴量 SR'_{ft} と LR_{f0} の差分 (二乗誤差) K_{res} を畳み込み層に入力し、 K'_f を得る。次に、 $t-1$ ステージで得られた推定カーネル K_{t-1} をベクトル化し、空間方向に複製・連結した特徴量 K'_{t-1} と K'_f の二つをチャンネル方向に連結して畳み込み層に入力し、 K_f を得る。 K_f に global average pooling (GAP) と softmax 関数を適用した後、カーネルの形に戻したものを t ステージ目で推定された劣化カーネル K_t とする。

本手法の劣化カーネル推定器は、 K_{t-1} を変形したものを利用して K_t を求めている。これは、 K'_f を出力するまでの畳み込み層の学習を簡単にするためである。

また、4 章以降の実験では、一枚の LR 画像に与えられた劣化カーネルは一つと仮定したため、 K_f に GAP を適用して 1 つの劣化カーネルしか推定しない。しかし、GAP を除けば、入力 LR 画像の 1 画素ごとに与えられた劣化カーネルの推定に、拡張することが可能である。

4. 実験

4.1 モデル構造の詳細

本節では、3 章で紹介したモデルの詳細を述べる。まず、図 1 の KB モジュール中の SFT 層については、図 3 に示

表 1 実験結果. 値は SR 画像と HR 画像との PSNR[db] (Y チャンネルで計算). 使用した劣化カーネルはガウシアンカーネル. 本手法はガウシアンカーネル幅 σ が大きいほど, 劣化除去性能が高い傾向が見られる.

手法	σ	Set5 [17]	Set14 [18]	BSDS100 [19]
SRCNN-CAB [8]	0.2	29.31	26.91	25.51
SRMDNF [9]		31.96	28.35	27.49
SRResNet (劣化カーネルを入力 LR 画像に連結, [12] より)		29.40	26.18	26.34
SRResNet (SFT を使用, [12] より)		31.64	28.23	27.26
IKC [12]		32.39	28.77	27.58
本手法		28.64	26.04	25.58
SRCNN-CAB [8]	1.3	29.50	27.02	25.66
SRMDNF [9]		32.00	28.42	27.53
SRResNet (劣化カーネルを入力 LR 画像に連結, [12] より)		29.11	25.93	26.20
SRResNet (SFT を使用, [12] より)		31.69	28.29	27.29
IKC [12]		32.41	28.82	27.64
本手法		30.83	27.80	26.99
SRCNN-CAB [8]	2.6	28.81	26.15	25.13
SRMDNF [9]		31.77	28.26	27.43
SRResNet (劣化カーネルを入力 LR 画像に連結, [12] より)		28.64	25.63	25.99
SRResNet (SFT を使用, [12] より)		31.35	28.07	27.15
IKC [12]		32.05	28.55	27.47
本手法		30.76	27.76	27.06

表 2 KB モジュールの直後に, IKC のように劣化カーネル自体を特徴量に付与する SFT 層を入れたモデルの実験結果. 評価は表 1 と同様. 表 1 の本手法の結果より, PSNR は向上している.

σ	Set5 [17]	Set14 [18]	BSDS100 [19]
0.2	29.06	26.30	25.83
1.3	31.05	27.88	27.04
2.6	31.03	27.83	27.11

した. SFT 層では最初に, 擬似 LR 画像と入力 LR 画像との差分と, down モジュールからの出力 LR'_{ft} とチャンネル方向に連結させる. これを 2 層の畳み込み層 (2 層ともフィルタサイズ 3) に入力し, その出力にシグモイド関数を適用して γ を得る. これとは別に, 差分がチャンネル方向に連結された特徴量を, 別の 2 層の畳み込み層 (2 層ともフィルタサイズ 3) に入力して, β を得る. 最後に次式に従って LR 特徴量 LR_{ft} を得る. ただし, \odot は要素積を表す.

$$LR_{ft} = (LR'_{ft} \odot \gamma) + \beta \quad (2)$$

次に, 図 1 の初期特徴量抽出と初期カーネル推定について述べる. 前者は, 2 層の畳み込み層によって行われており, フィルタサイズはそれぞれ 3 と 1 である. これは, DBPN [7] のものと同じであり, 本手法の SR 部分 (図 1 中の青矢印に関わる部分) が, DBPN と似たモデル構造なためである. 後者は, IKC [12] の初期劣化カーネル推定器と同じ構造をしているが, これは本手法が反復的にカーネルを推定するアルゴリズムとなっており, それが IKC のアルゴリズムと似ているためである. ただし, 劣化カーネルを次元削減表現を出力させる IKC と異なり, 本手法は劣化カーネル自体を出力する必要があるため, 出力されるカー

ネルに対し, softmax 関数を適用して, 劣化カーネルの全要素が $[0, 1]$ となるようにしている.

次に, 図 2 で用いる畳み込み層に関して述べる. SR'_{ft} を出力する畳み込み層は 1 層となっている (フィルタサイズは, 学習する倍率によって異なる. 今回は 8). この畳み込み層の処理の目的は, t ステージ目までで得た SR 特徴量群の次元を, 初期 LR 特徴量 LR_{f_0} に合わせるだけであり, 画像特徴量を大きく変化させないために畳み込み層は少なくしている. K_{res} から K'_f を出力する畳み込み層は 2 層, K_f を出力する畳み込み層は 1 層である. これら 3 層はフィルタサイズは 1 とした.

4.2 実験条件

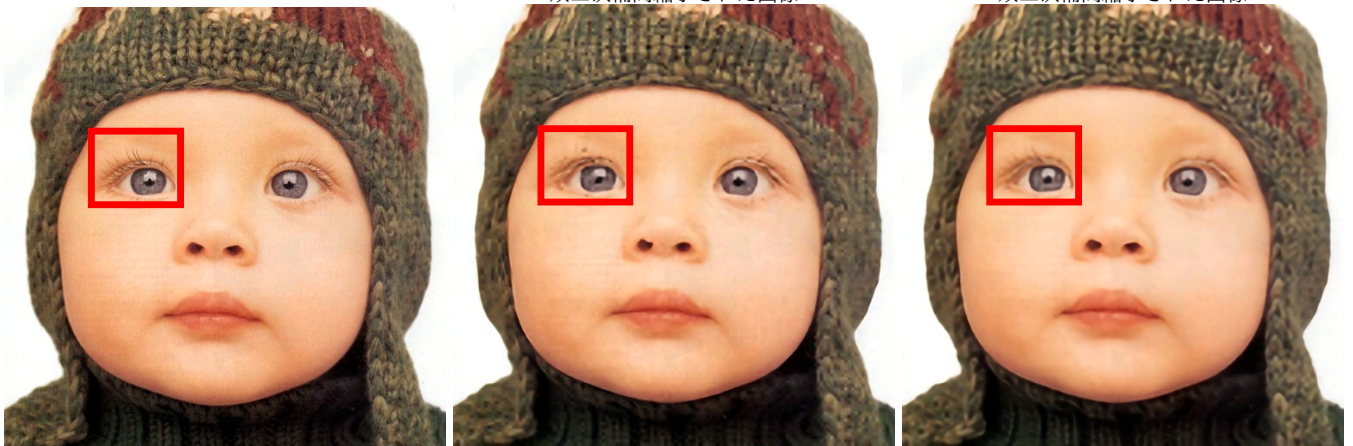
本実験では, KB モジュールを 10 個連結させてできたモデルに対し, LR 画像の辺々 4 倍にする SR を学習させた. また, 式 (1) の縮小過程 \downarrow_s は双三次補間と仮定し, この実験では, 複数の劣化カーネルと双三次補間の組み合わせで, 多様な劣化を表すことにした. 損失関数 L は次式のようにした.

$$L = L_{MSE}(SR, HR) + L_{MSE}(LR_{pseudo}, LR_{input}) + L_{MSE}(SR_{half}, HR) + L_{MSE}(LR_{pseudo_half}, LR_{input}) + L_{L1}(K_{predicted}, K_{GT}) + \|K_{predicted}\|_1 \quad (3)$$

主に, 生成される SR 画像と擬似 LR 画像の復元誤差 (平均二乗誤差) と推定カーネルの推定誤差で構成されている. 第 3, 4 項目は中央 KB モジュールから出力される SR 画像と擬似 LR 画像の復元誤差である. 第 5 項目の L1 損失



(a) $\sigma = 0.2$ のガウシアンカーネルを与えた後、
双三次補間縮小された画像
(b) $\sigma = 2.6$ のガウシアンカーネルを与えた後、
双三次補間縮小された画像



(c) HR 画像

(d) 本提案モデルに (a) を入力して得られた
SR 画像 (PSNR:31.51db)

(e) 本提案モデルに (b) を入力して得られた
SR 画像 (PSNR:33.07db)

図 4 各過程で生成された LR 画像 (a), (b) と, (a), (b) を本手法のモデルに入力して得られた SR 画像 (d), (e). (c) は HR 画像. 劣化度合いが小さい (a) の SR 画像 (d) に比べ, 劣化度合いが大きい (b) の SR 画像 (e) の方が, HR 画像との PSNR が高く, また, (e) の方が赤枠内の再現性も高い.

を計算するときは, それぞれのカーネル ($k \times k$ 次元) を PCA によって次元削減したベクトル (k 次元) を使う. これは, IKC [12] に倣い, 低次元空間での劣化カーネルの表現を学習した方が学習が容易かつ十分であると考えたためである. 次元削減行列は, 学習で用いるカーネルからサンプリングされた 500 個のカーネルから得た. また第 6 項目は推定カーネルの正規化項である. また中央 KB モジュールで得られる SR 画像や擬似 LR 画像に対して損失を取ることで, 前半の KB モジュールでも有用な特徴量抽出をする学習をさせるためである.

用いた学習データセットは, 一般に広く使われている, DIV2K [15] の学習データセット 800 枚と Flickr2K [16] の学習データセット 2650 枚の, 合計 3450 枚の 2K 画像を用いた. このデータセットに対し, 上下左右反転のデータ拡張をしている. また, 今回学習に用いた劣化カーネルは, 等方性のガウシアンブレンダーカーネルを選んでいる. ガウシアンカーネルの幅 σ は, [0.2, 4.0] の範囲から, 一様分布

に従って得られたものを用いてガウシアンカーネルを生成し, それを用いた後, 双三次補間で縮小することで LR 画像を生成した. 本実験では, カーネルの大きさは 21×21 で固定している.

テストデータセットは, 広く用いられている Set5 [17], Set14 [18], BSDS100 [19] を使用した. テスト時に使用する劣化カーネルも, 等方性のガウシアンカーネルであり, サイズは学習時と同じで 21×21 で, ガウシアンカーネルの幅 σ は 0.2, 1.3, 2.6 の 3 種類を使用する. 本手法の定量的評価は, 最終的に生成された SR 画像と HR 画像とのピーク対信号雑音比 (PSNR) で行う. PSNR の計算は Y チャネル (輝度) で行い, 評価のときは, 周囲 s ピクセルを省いて評価する (s は拡大倍率値).

5. 実験結果と考察

本実験の定量的評価を, [8, 9], [12] の結果とともに表 1 に示す. これをもとに, 本手法が SR 画像にもたらす効果

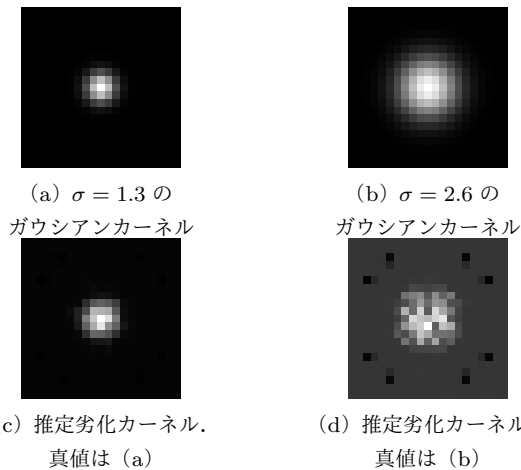


図 5 真値劣化カーネル (a), (b) と推定劣化カーネル (c), (d). (c) の推定精度は比較的高いものの、(d) の推定精度は悪く、カーネルの四隅に余計な値を持ってしまっている。

と、他手法との比較をそれぞれ 5.1 節, 5.2 節で述べる。

5.1 本手法の効果

表 1 で本手法の結果に注目すると、ガウシアンカーネルの幅 σ が大きくなるほど、劣化の除去精度が上がっているように見られる。一般に、ガウシアンカーネルの幅が小さい場合、除去すべき画像劣化は小さいため、PSNR が大きくなり、逆にガウシアンカーネル幅が大きい場合は、除去すべき劣化が大きいため、PSNR が下がりやすくなる。しかし、本手法の結果を見ると、 $\sigma = 1.3$ と $\sigma = 2.6$ の PSNR にほとんど差はなく、逆に $\sigma = 0.2$ でのテスト結果の PSNR は、 $\sigma = 2.6$ よりも悪い結果となっている。このことから、本手法の大きな特徴である擬似 LR 画像と入力 LR 画像の差分をモデル内で取り、特徴量へ付与することは、劣化度合いが大きいほど劣化除去力が大きくなると言える。 σ が大きいほど、入力 LR 画像のボケの度合いが大きくなるため、前半の KB モジュールで得られる差分は大きくなりやすくなると考えられる。その場合、SFT 層での処理によって LR 特徴量は大きく変化する。このような処理が、より有効な特徴量を抽出するのに有用に働いたのだと考えられる。一方で、 σ が小さい場合は、前半の KB モジュール時点で、擬似 LR 画像と入力 LR 画像の差分は小さくなりやすくなるため、SFT 層を用いても LR 特徴量の変化量は小さくなる。この場合は本モデルは DBPN と酷似した処理しかできなくなる。そうなる縮小過程によるボケ以外の劣化に対応できなくなってしまう、結果として SR の性能が悪くなったと考える。

また、図 4 で、定性的に本手法の結果を見てみると、図 4 の (d), (e) について、特に赤枠内で、(d) では睫毛の先が、HR 画像には見られないような黒点になってしまっているが、(e) にはそれが見られないのが分かる。

5.2 他手法との比較

次に表 1 で他手法と比較してみると、本手法は SRMD や IKC よりも悪い結果となってしまった。SRMD や IKC と本手法の大きな違いは、劣化カーネル自体を、画像特徴量に付与しているかどうかであり、今回の結果から、劣化カーネル自体を特徴量に付与したほうが良い、ということが分かる。よって、3.1 節で述べた拡張は行うべきだと分かった。また、劣化カーネル自体を特徴量に付与することで、劣化度合いが小さい場合でも特徴量を変化させることができるため、さらなる SR 精度向上が見込める。そこで、IKC のように SFT 層を使って、特徴量 LR_{ft} に推定劣化カーネル K_t を付与した結果を表 2 に示し、表 1 の本手法との結果を比べてみる。このときに使用した SFT 層の構造は、図 3 と同じであるが、入力を変えており、図 3 の LR'_{ft} は KB モジュールの出力 LR_{ft} に、擬似 LR 画像と入力 LR 画像の差分は、劣化カーネル K_t を図 2 の K'_{t-1} と同じように変形した特徴量にしている。この SFT 層から出力された特徴量を新しい LR 特徴量 LR_{ft} として、次の KB モジュールに入力させた。すると、すべてのテスト条件で、表 2 の結果のほうが、表 1 以上の結果を出すことができた。この結果からも、劣化カーネル自体を特徴量に付与することは有効であると分かった。

5.3 今後の課題

今回の実験を通して、擬似 LR 画像と入力 LR 画像との差分を用いることは、程度の大きい画像劣化の除去に有効だということが分かった。しかし、SR と劣化除去を同時に学習するような本手法で扱うときには注意が必要なものであると考えられる。なぜなら、同時に学習を行う場合、擬似 LR 画像と入力 LR 画像の差分が小さくなったとしても、必ずしも SR 画像が HR 画像に近くなる、あるいは推定カーネルが真値カーネルと近い値になるとは限らないためである。このため、例えば途中の KB モジュールで生成される SR 画像 SR_t が HR 画像とかけ離れていても、推定されたカーネルとの演算結果が入力 LR 画像と近ければ、SFT 層は上手く機能せず、また、擬似 LR 画像と入力 LR 画像との損失も小さくなるため学習が進まなくなってしまう。また、同時学習の場合では、擬似 LR 画像と入力 LR 画像との差分が大きい場合、SR と劣化推定のどちらが悪いかを一意に特定できないため、学習が上手く進まなくなるとも考えられる。よって、SR と劣化推定の学習を最初は個別で行い、後にモデル全体を再学習するといった工夫をしなければならないと考える。例えば、SR を個別で学習させる場合、使用する劣化カーネルは真値を使い、劣化推定を個別で学習させる場合は真値 HR 画像を使うことになる。これにより、擬似 LR 画像と入力 LR 画像の差分が大きい場合は、必ず一方の推定に原因があるため正しい学習ができると考えられる。

また、現時点での本手法の劣化推定精度は、図5に示す通りであり、未だ改善の余地がある。図2の(c)は真値である(a)と似た推定劣化カーネルであるが、(d)は真値(b)とは差があり、またカーネルの四隅にも値を持ってしまっている。この劣化推定性能を向上させるための、本手法のさらなる改善案として、擬似LR画像と入力LR画像の差分を、次のKBモジュール内の劣化カーネル推定にも用いること、が挙げられる。例えば、差分を図2の K'_{t-1} に乗算したり、SFT層を使って付与するなどして、 $t-1$ ステージ目の劣化推定の結果を考慮させれば、劣化推定性能の向上が見込めると考える。

6. まとめ

本研究では、SRと、劣化推定からの除去を二つの深層学習モデルを使って実現しようとした。これら二つのモデルを学習するときに、推定される劣化カーネルとSR画像を使って、擬似LR画像を生成し、入力LR画像との差分を特徴量に利用し、さらに損失関数の一つとして扱うことで多様な劣化に対応するSRの性能向上を狙った。本実験を通して、擬似LR画像と入力LR画像との差分の利用は、程度が大きい劣化の除去に適しているということが分かった。また、この有用性から、差分の利用は劣化カーネル推定にも利用することで、劣化推定性能も向上させることができ、手法全体の性能向上につながると思われる。しかし、本手法のようなSRと劣化推定を同時に学習する場合は、差分が小さくなっただけでは、SRや劣化推定性能が上がったとは言えず、逆に間違った学習をしてしまう可能性がある。このため、SRと劣化推定を最初から同時学習させるのではなく、個別に学習をした後に同時学習をするなどの工夫が必要である。

参考文献

- [1] Dong, C., Loy, C., He, K. and Tang, X.: *Image super-resolution using deep convolutional networks*, IEEE transactions on pattern analysis and machine intelligence, 2016
- [2] Kim, J., Kwon Lee, J. and Mu Lee, K.: *Accurate image super-resolution using very deep convolutional networks*, In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016
- [3] Goodfellow, J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: *Generative adversarial nets*, Advances in Neural Information Processing Systems, 2014
- [4] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. and W., Shi: *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*, In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017
- [5] Johnson, J., Alahi, A. and Fei-Fei, L.: *Perceptual losses for real-time style transfer and super-resolution*, In European conference on computer vision, 2016
- [6] Sajjadi, M. S., Scholkopf, B. and Hirsch, M.: *Enhancenet: Single image super-resolution through automated texture synthesis*, In Proceedings of the IEEE International Conference on Computer Vision, 2017
- [7] Haris, M., Shakhnarovich, G. and Ukita, N.: *Deep back-projection networks for super-resolution*, In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018
- [8] Riegler, G., Schuler, S., Ruther, M., and Bischof, H.: *Conditioned regression models for non-blind single image super-resolution*, In Proceedings of the IEEE International Conference on Computer Vision, 2015
- [9] Zhang, K., Zuo, W. and Zhang, L.: *Learning a single convolutional super-resolution network for multiple degradations*, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018
- [10] Cornillere, V., Djelouah, A., Yifan, W., Sorkine-Hornung, O. and Schroers, C.: *Blind image super-resolution with spatially variant degradations*, ACM Transactions on Graphics, 2019
- [11] Wang, X., Yu, K., Dong, C. and Change Loy, C.: *Recovering realistic texture in image super-resolution by deep spatial feature transform*, In Proceedings of the IEEE conference on computer vision and pattern recognition, 2018
- [12] Gu, J., Lu, H., Zuo, W. and Dong, C.: *Blind super-resolution with iterative kernel correction*, In Proceedings of the IEEE conference on computer vision and pattern recognition, 2019
- [13] Efrat, N., Glasner, D., Apartsin, A., Nadler, B. and Levin, A.: *Accurate blur models vs. image priors in single image super-resolution*, In Proceedings of the IEEE International Conference on Computer Vision, 2013
- [14] Yang, C. Y., Ma, C. and Yang, M. H.: *Single-image super-resolution: A benchmark*, In European Conference on Computer Vision, 2014
- [15] Agustsson, E. and Timofte, R.: *Ntire 2017 challenge on single image super-resolution: Dataset and study*, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017
- [16] Timofte, R., Agustsson, E., Van Gool, L., Yang, M. H. and Zhang, L.: *Ntire 2017 challenge on single image super-resolution: Methods and results*, In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2017
- [17] Bevilacqua, M., Roumy, A., Guillemot, C. and Alberi-Morel, M. L.: *Low-complexity single-image super-resolution based on nonnegative neighbor embedding*, 2012
- [18] Zeyde, R., Elad, M. and Protter, M.: *On single image scale-up using sparse-representations*, In International conference on curves and surfaces, Springer, Berlin, Heidelberg, 2010
- [19] Martin, D., Fowlkes, C., Tal, D. and Malik, J.: *A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics*, In Proceedings Eighth IEEE International Conference on Computer Vision, 2001