

令和4年度卒業論文

SRGANに基づいた敵対的学習による画像からの
ノイズ除去能力の検証

指導教員 椋木 雅之

宮崎大学工学部 情報システム工学科

矢部 優奈

概要

本論文では、画像の超解像を行う SRGAN (Super Resolution Generative Adversarial Network) をもとにノイズ除去画像の生成の検証を行う。画像に含まれるノイズは、画像を操作する際に妨げとなるものである。実際の画像には、レンズによる焦点ボケ、被写体やカメラの動きによるブレ、撮像素子である CCD の特性による画素値の揺らぎやエラーなど、被写体以外に由来する様々な変動(ノイズ)が加わっている。画像上のノイズを取り除き、鮮明な画像にすることは画像データを取り扱う上で重要である。ノイズの例として、画像の拡大により生じるノイズ、インパルスノイズ、ガウシアンノイズなどがある。これらのノイズが画像に加えられた、ノイズ画像からノイズを取り除いた画像を生成するのは、不良設定問題である。不良設定問題とは、解を求めるために必要な情報が一部欠けている問題のことである。ノイズ除去において、ノイズのあるデータからノイズのないデータに変換するためには、ノイズのあるデータには存在しない部分のデータを生成する必要がある。このノイズのあるデータに対するノイズのないデータの生成パターンは無数にあり得る。

低解像度データから高解像度データを復元・生成する技術を超解像という。超解像では、低解像度データから高解像度データに変換するためには、低解像度データには存在しない部分のデータを生成する必要がある。この低解像度データに存在しない部分には無数の生成パターンがあり得る。超解像もノイズ除去同様不良設定問題である。この無数のパターンから適切なパターンを推定する不良設定問題において、超解像では SRGAN という手法が用いられている。ノイズ除去に関しても SRGAN を活用し、ノイズのある画像からノイズのない画像の生成が可能であるか検証する。そこで、本論文では SRGAN をもとにノイズ除去の能力の検証ができるように改良を行った。具体的には、SRGAN の Generator 中の Pixel Shuffler を取り除いた。これにより、入力画像サイズと出力画像サイズを同じにした。さらに、使用する学習データの画素値の正規化を行うことで、生成される画像の品質の向上を図った。評価実験として、画像の拡大によるノイズ、インパルスノイズ、ガウシアンノイズを加えた画像を与え、生成されたノイズ除去画像と、ノイズ画像の比較を行った。結果として、ノイズの種類によってはノイズ除去に対して有効であることが確認できた。

目次

1. はじめに	1
2. 画像からのノイズ除去	3
2.1 ノイズ	3
2.2 ノイズの種類	3
2.2.1 画像の拡大により生じるノイズ	3
2.2.2 インパルスノイズ	5
2.2.3 ガウシアンノイズ	5
2.3 深層学習を利用したノイズ除去	6
2.4 敵対的学習を用いた画像処理	7
2.4.1 敵対的学習 (GAN)	7
2.4.2 GAN を用いたノイズ除去	7
2.4.2.1 Noise Generator [14]	7
2.4.2.2 Image Denoising [15]	8
2.4.3 SRGAN による画像の超解像	8
3. 敵対的学習によるノイズ除去手法	11
3.1 SRGAN に基づくノイズ除去	11
3.2 ネットワーク構造	11
3.3 Generator の変更点	12
4. 実験	14
4.1 実験設定	14
4.2 評価指標	14
4.3 学習データの画素値の正規化	16
4.4 バイキュービック補間によって生じたノイズ除去の検証	17
4.4.1 実験手法	17
4.4.2 実験結果・考察	17
4.5 インパルスノイズ除去の検証	19
4.5.1 実験手法	19
4.5.2 実験結果・考察	19
4.6 ガウシアンノイズ除去の検証	20
4.6.1 実験手法	20
4.6.2 実験結果・考察	21
4.7 実験を通じての考察	22

5. おわりに.....	23
謝辞	24
参考文献	25

1. はじめに

スマートフォンやデジタルカメラの普及に伴い、画像データを取り扱いやすくなった。これに伴い、高画質で鮮明な画像の需要が高まっている。画像データを取り扱いやすくするためにノイズのある画像を鮮明な画像に変換することは重要な課題となっている。ノイズとは信号処理において、本来処理したい対象以外に由来する不要な信号や情報のことである。画像のノイズは画像データ上に現れるランダムな粒状性ノイズや圧縮画像に見られるブロックノイズなどがある。画像上のノイズを取り除き、鮮明な画像にすることは画像データを取り扱う上で重要である。近年、低画質画像からのノイズ除去や高画質化が注目されている。

画像におけるノイズ除去手法として、フィルタ処理を用いた手法がある。この方法では、フィルタの種類によってエッジがぼやけた画像になったり、逆にエッジが強調された画像が生成されたりすることがある。例えば、平滑化フィルタの一種である平均値フィルタは、近傍画素の平均値を取ることで、高周波数領域に現れる細かな濃度変化をノイズとして除去できる。また、ウィナーフィルタではボケやブレによる画像の劣化をモデル化し、その他の加法性ノイズも考慮したうえで、劣化やノイズを除去する。しかし、平均値フィルタでは、ノイズ以外の大きな濃度変化（エッジ）がぼやけた画像になってしまう。また、ウィナーフィルタでは画像劣化のモデル化が適切に行えない場合がある。

一方、近年深層学習技術が進歩し、画像処理の様々な分野で用いられている。ノイズ除去の分野においても、深層学習を用いたノイズ除去や高精細化の研究が盛んに行われている。中でも、敵対的学習（GAN : Generative Adversarial Network）を用いた手法は、人間にとって違和感のない画像を生成できることから、ノイズ除去においても有望な手法となっている。例えば、Linh[1]らはGANによりモデル化したノイズ画像を使って学習することにより、様々なノイズに対してノイズ除去を可能にしている。

本研究では、GANを用いてノイズが付加された画像から直接ノイズ除去された画像を生成する手法を調査する。ノイズが付加された画像とそれに対する望ましいノイズ除去画像の組をGANにより学習することで、人間にとって違和感のないノイズ除去画像を生成することを目指す。同様の研究事例としては

Ziyuan ら[2]の研究がある。Ziyuan らは GAN を用いた独自のネットワーク構造を提案し、ノイズ除去を行った。本研究では、画像の超解像処理で有効性が示されている SRGAN[3]のネットワーク構造を利用して、そのノイズ除去への適用可能性について調査する。

以下、2 章では画像からのノイズ除去についての説明を行う。3 章では GAN を用いたノイズ除去について説明した後、提案手法について述べる。4 章では実験、考察について述べる。最後に 5 章では結論を述べる。

2. 画像からのノイズ除去

2.1 ノイズ

画像に含まれるノイズは、画像を操作する際に妨げとなるものである。画像を、撮影したカメラ位置から見た空間的な明るさや色の分布を計測したのと考え、画像には撮影対象となった被写体に由来する情報のみが含まれていることが望ましい。しかし、実際の画像には、レンズによる焦点ボケ、被写体やカメラの動きによるブレ、撮像素子であるCCDの特性による画素値の揺らぎやエラーなど、被写体以外に由来する様々な変動（ノイズ）が加わっている。これらの変動により観測した画像は不鮮明なものとなる。画像上のノイズを取り除き、鮮明な画像にすることは画像データを取り扱う上で重要である[4]。

2.2 ノイズの種類

ノイズには様々な種類がある。本研究では、ノイズ除去の検証実験において以下の3種類のノイズを扱う。

2.2.1 画像の拡大により生じるノイズ

サイズの小さな元画像を大きなサイズの拡大画像に変換するために、画像補間の技術を使うことがある。画像補間においては、拡大画像の画素値を求めるために、元画像の画素と画素の間の画素値を推定する必要がある。そのために、拡大画像の推定対象画素の周囲に当たる元画像の画素値から、推定対象画素の画素値を求める補間法が種々提案されている、しかし、このような補完法では、元画像の画素間隔以上の高周波成分は復元できず、拡大画像にボケのノイズが生じる。

例えば、補間法の一つであるバイキュービック補間[4]では、求めたい拡大画像の画素値 $I(x, y)$ を、まわりの16点の画素値 $f_{11}, f_{12}, \dots, f_{44}$ を用いて、以下の式により求める。

$$I(x, y) = (h(x_1)h(x_2)h(x_3)h(x_4)) \begin{pmatrix} f_{11} & \cdots & f_{14} \\ \vdots & \ddots & \vdots \\ f_{41} & \cdots & f_{44} \end{pmatrix} \begin{pmatrix} h(y_1) \\ h(y_2) \\ h(y_3) \\ h(y_4) \end{pmatrix} \quad (1)$$

ただし、

$$\begin{aligned} x_1 &= 1 + x - [x] & y_1 &= 1 + y - [y] \\ x_2 &= x - [x] & y_2 &= y - [y] \\ x_3 &= [x] + 1 - x & y_3 &= [y] + 1 - y \\ x_4 &= [x] + 2 - x & y_4 &= [y] + 2 - y \end{aligned} \quad (2)$$

また、関数 $h(t)$ は、sinc 関数 $\text{sinc}(t)=\sin(\pi t)/\pi t$ を三次多項式で近似するものであり、一般に以下の式が用いられる。

$$h(t) = \begin{cases} |t|^3 - 2|t|^2 + 1 & (|t| \leq 1) \\ -|t|^3 + 5|t|^2 - 8|t| + 4 & (1 < |t| \leq 2) \\ 0 & (2 < |t|) \end{cases} \quad (3)$$

図 1 に、バイキュービック補間による画像拡大の例を示す。sinc 関数を用いることで、標本化定理に基づき、元画像で表現される周波数以内であれば理論上、正しく補間が行われる。しかし、実際には画像の拡大に、より高周波数の情報が必要であり、拡大画像にはボケが生じる。



図 1 : 左 原画像 右 バイキュービック補間 [5]

2.2.2 インパルスノイズ

インパルスノイズは画像の伝送中のエラーなどにより元画像の画素値が全く別の値に置き換わるノイズである。インパルスノイズは次式でモデル化される[7]。

$$x'_{ij} = \begin{cases} n_{ij} & \text{確率 } p \\ x_{ij} & \text{確率 } (1-p) \end{cases} \quad (4)$$

ここで x_{ij} はノイズを含まない画像の位置 (i,j) における本来の画素値、 x'_{ij} はノイズを含んだ画像の位置 (i,j) における画素値である。また p は画像に含まれるノイズの割合である。 n_{ij} はノイズの画素値であり、下限から上限までの範囲で離散一様分布に従う乱数である。

インパルスノイズの特殊な場合として、次の式でモデル化されるごま塩ノイズがある[7]。

$$n_{ij} = \begin{cases} 0 & \text{確率 } 0.5 \\ 255 & \text{確率 } 0.5 \end{cases} \quad (5)$$

ごま塩ノイズは最も単純なインパルスノイズであり、通信路におけるエラーや、ドット欠けなどのモデルとして用いられている。

2.2.3 ガウシアンノイズ

ガウシアンノイズは、撮像素子である CCD が熱雑音などの影響を受け、本来の画素値に微小な変動が加わることで生じるノイズである。ガウシアンノイズは次式でモデル化される[7]。

$$x'_{ij} = x_{ij} + z_{ij} \quad (6)$$

ここで、 z_{ij} は平均 0 分散 σ^2 の正規分布に従うノイズ成分である。ガウシアンノイズは、インパルスノイズと異なり、本来の画素値にノイズ成分が加算される加法性ノイズとして扱われる。また、インパルスノイズがランダムに選択された画素にのみ生じるのに対し、ガウシアンノイズはすべての画素にノイズ成分が加算され

る。インパルスノイズ、ごま塩ノイズ、ガウシアンノイズの例を図 2 に示す。



図 2:(a)Lenna の原画像 (b)インパルスノイズ($p=0.2$)
(c)ごま塩ノイズ($p=0.2$) (d)ガウシアンノイズ($\sigma=20$)[7]

2.3 深層学習を利用したノイズ除去

従来、ノイズ除去には平滑化フィルタやウィナーフィルタなどの信号処理的手法が用いられていた。しかし、信号処理的な手法では、様々な種類のノイズに対応することができなかった。一方で、近年、深層学習技術が著しく向上しており、深層学習をノイズ除去に用いる研究事例も増えている。

深層学習を利用したノイズ除去手法の1つに DnCNN[8]がある。DnCNNは、ノイズを含んだ入力画像とノイズのないクリーン画像(正解)のペアを学習データとして教師あり学習を行う。学習済みの DnCNN にノイズを含んだ入力画像を与えると、ノイズを除去した画像を出力するようになる。この際、学習データの入力画像へのノイズの加え方を様々に行うことで、様々な種類のノイズに対応したノイズ除去を実現できる。しかし、DnCNN のような手法では、ノイズ除去した画像において、細かな部分の鮮明さが失われ、視覚的に劣化したように感

じられることが多い。

2.4 敵対的学習を用いた画像処理

2.4.1 敵対的学習 (GAN)

深層学習を利用した画像処理では、処理結果が視覚的に劣化するという欠点がある。この問題に対し、深層学習の一種である敵対的学習 (GAN: Generative Adversarial Network) [9]による画像生成技術が適用されている。

GAN は Generator(生成器)と Discriminator(識別器)の 2 つのニューラルネットワークで構成されている。Generator はより学習データに近いデータを生成し、Discriminator は入力データが学習データか Generator が生成したデータかを識別する。Generator は Discriminator を騙すことができるように学習し、Discriminator は Generator が生成したデータを見破ることができるように学習する。このような学習により、最終的には Generator が学習データと同じようなデータを生成することが期待される。このような仕組みにより GAN は、学習データと見分けがつかないような、視覚的劣化の少ない新たなデータを生成できるようになる[10]。

GAN を活用した様々な研究が行われている。例えば、DCGAN[11]では実在しない人物の画像を高精細に生成している。また、CycleGAN[12]では画像中の被写体のテクスチャを入れ替えることができ、画像中の馬とシマウマの模様を入れ替える例が示されている。CGAN[13]では、生成する画像の種類(クラス)を指定することができ、指定した数字の画像を生成する例が示されている。

2.4.2 GAN を用いたノイズ除去

ここでは GAN を用いたノイズ除去の既存研究について紹介する。

2.4.2.1 Noise Generator [14]

Linh ら[14]は、実画像に含まれるノイズと同等のノイズを GAN により生成することでノイズ除去の性能向上を行った。Noise Generator では、入力された実画像からノイズのある画像を生成する。Noise Discriminator では、入力された画像が、Noise Generator が生成したノイズ画像か、元々ノイズのある実画像

か識別する。元々ノイズのある画像として Smartphone Image Denoising Dataset(SIDD)を利用するが、対象シーンのバリエーションが少ないため、まず、ノイズモデルにより人工的に生成したノイズ画像で GAN の学習を行った後、SSID を使って改良学習(fine tuning)を行う。これにより、Noise Generator が、任意の画像に実際の場面に近いノイズを付加できるようになる。この Noise Generator を使って学習データを作成し、DnCNN を学習することで、ノイズ除去の性能を向上させている。

2.4.2.2 Image Denoising [15]

Noise Generator では GAN をノイズの生成に利用した。一方で Image Denoising では、ノイズ除去画像生成のために GAN を利用する。Generator ではノイズ画像を入力とし、原画像のようなノイズのない画像を生成する。Discriminator は入力として Generator の生成画像か原画像かのどちらかが与えられ、その画像が Generator の生成画像か原画像かを判断する。これを繰り返すことで、ノイズのある画像からノイズのない画像が生成される。Image Denoising では、この目的のために新たなネットワーク構造を提案している。

2.4.3 SRGAN による画像の超解像

画像の超解像とノイズ除去は、入力画像には含まれない画素の画素値を推定するという共通点を持つ、関連した技術といえる。超解像とは低解像度のデータから高解像度のデータを生成・補完する技術である。画像を拡大した際に、バイキュービック補間などを利用して画素を推定すると画像の劣化、ぼけが生じることがあるが、超解像を行うことで画像劣化を抑えた状態での拡大ができる。

画像の超解像は複数の画像から行う手法と、1枚の画像から行う手法がある。複数の画像から行う手法には、複数の観測画像から位置合わせ処理と再構成処理を同時に行う手法[20]などがある。このような手法は、観測に基づく手法なので、正しい高解像度画像が得られるが、適切な設定で撮影した複数の画像が必要である。

SRGAN[3]は、GAN を用いて 1 枚の低解像度画像から高解像度画像を生成する手法である。図 3 は SRGAN のネットワーク構造の概要を示している。また、図 4 は Generator、Discriminator のネットワークの詳細を示している。Generator は低解像度画像とその画像の高解像度画像のペアから、超解像画像を生成する。Generator の Upsample Block (図 4 上部) 中の Pixel Shuffler では、解像度を上げる処理を行っている。従来の逆畳み込みでは、生成される画像に格子状の模様が見られることがあり、これを解決するために Pixel Shuffler を用いて超解像を実現している。Discriminator は入力された画像が生成された画像なのか学習データなのかを識別する。これを繰り返すことで、低解像度画像から高解像度画像を生成することができる。

図 5 に SRGAN による超解像結果とバイキュービック補間での画像拡大の結果を示す。これらは図 5 右の原画像(original)を縦横 1/4 に縮小した画像に対し、それぞれの処理を加えた結果である。SRGAN の超解像では、細部の描写の再現も鮮明なものとなっている。

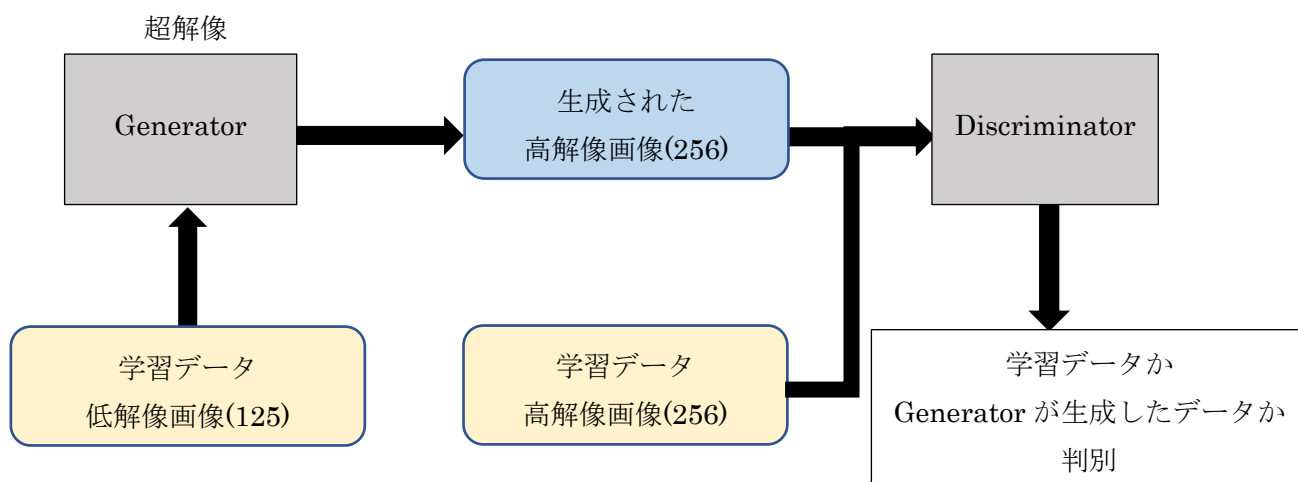


図 3 : SRGAN のネットワーク構造

() 内は画像サイズの一辺

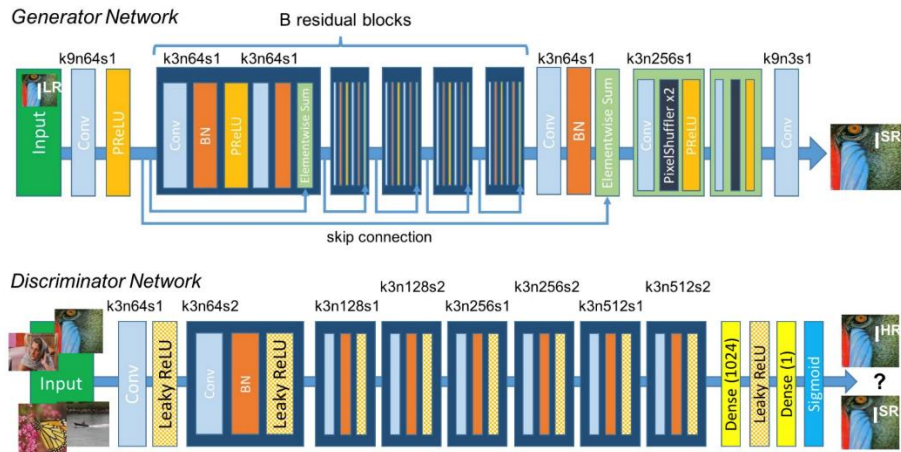


図 4: SRGAN の Generator と Discriminator[3]

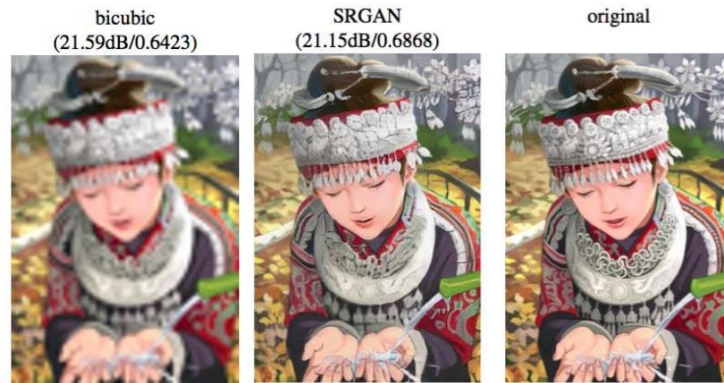


図 5: SRGAN と従来手法の超解像画像[3]

3. 敵対的学習によるノイズ除去手法

3.1 SRGAN に基づくノイズ除去

SRGAN を活用することで、超解像画像が生成できる。超解像、ノイズ除去この二つはともに不良設定問題である。不良設定問題とは、解を求めるために必要な情報が一部欠けている問題のことである。超解像では、低解像度データから高解像度データに変換するためには、低解像度データには存在しない部分のデータを生成する必要がある。この低解像度データに存在しない部分には無数の生成パターンがあり得る。また、ノイズ除去では、ノイズのあるデータからノイズのないデータに変換するためには、ノイズのあるデータには存在しない部分のデータを生成する必要がある。このノイズのあるデータに対するノイズのないデータの生成パターンは無数にあり得る。この無数のパターンから適切なパターンを推定する不良設定問題において、超解像では SRGAN という手法が用いられている。ノイズ除去に関しても SRGAN を活用し、ノイズのある画像からノイズのない画像の生成が可能であるか検証する。

3.2 ネットワーク構造

ここでは提案手法の大まかな構成を述べる。提案手法は SRGAN をノイズ除去ができるか検証するために変更したものである。提案手法では Generator と Discriminator の構造は基本的に SRGAN と同様とし、入力データに対して出力データの画素値の値をそのまま保持できるように変更を加えている。学習データには、ノイズを付与した画像とノイズを付与していない画像のペアを使用する。

図 6 に提案手法のネットワーク構造を示す。提案手法は、Generator と Discriminator の 2 つのニューラルネットワークで構成されている。図 6 中の Generator の処理を経る前後で画像サイズは $256 \times 256\text{px}$ のままである。SRGAN の場合には Pixel Shuffler を介しており、画像サイズが $125 \times 125\text{px}$ から $256 \times 256\text{px}$ になっており、超解像の処理を行っている (図 3) が、図 6 の提案手法中の Generator の処理を経る前後で画像サイズは $256 \times 256\text{px}$ のままで

ある。

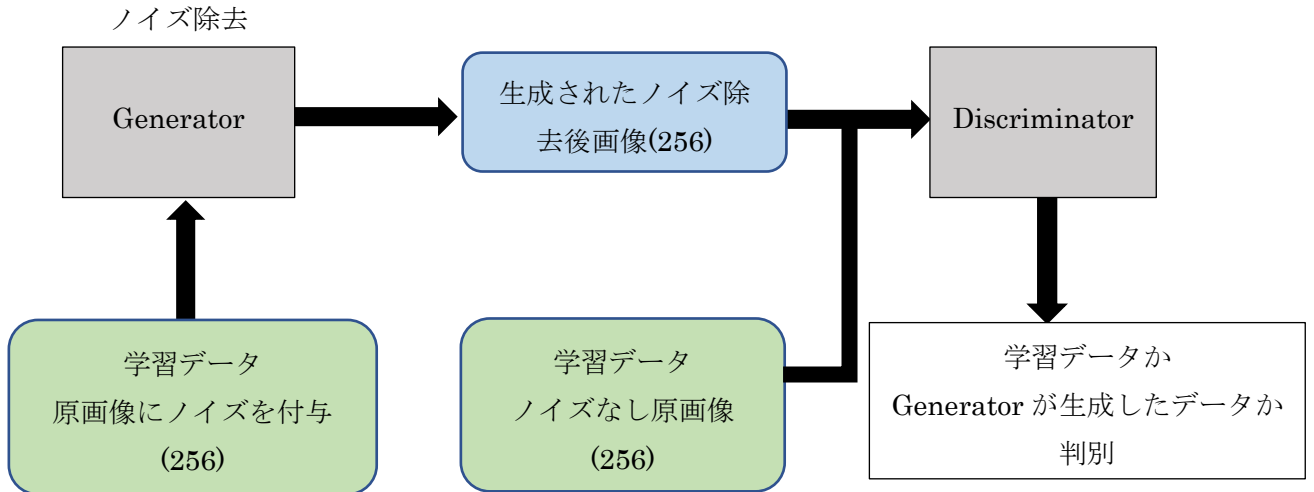


図 6 : 提案手法のネットワーク構造

3.3 Generator の変更点

図 7 に提案手法の Generator のネットワーク構造を示す。Generator の構造も基本的に SRGAN と同様であるが、ノイズ除去能力を検証するために Generator 中の Pixel Shuffler を除いた構造とした。SRGAN では Pixel Shuffler を使うことによって解像度を上げている (図 4)。Pixel Shuffler では入力マップのピクセルを並び替えるという方法で高解像度の画像を生成する。提案手法では、この Pixel Shuffler を取り除くことで入力画像と出力画像の解像度を同じにした。Generator の処理を経ることでノイズのある画像からノイズのない画像を生成する。

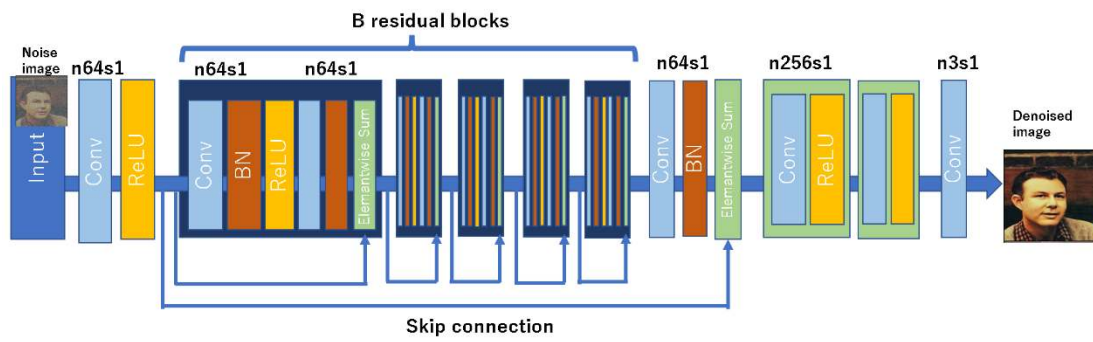


図 7 : 提案手法の Generator

4. 実験

本実験では、SRGAN を用いた画像からのノイズ除去能力の検証を行う。本実験では、拡大に伴い生じるノイズ、ごま塩ノイズ、ガウシアンノイズの 3 つのノイズモデルに対して実験を行う。

4.1 実験設定

学習、評価に用いる画像として、データセット CelebA[17]のデータを使用する。学習には 502 枚分の画像を使用した。学習時にデータはピクセル数が 256×256 のサイズになるようリサイズする。学習はバッチサイズが 2、epoch 数が 200 回とした。

評価に使用したデータは学習の最後に出力されたものから 12 枚の画像を取得し、利用している。

4.2 評価指標

評価指標として、MSE、PSNR、SSIM、GMSD、Perceptual Index を用いて定量評価を行った。各指標は式(6)～(10)で算出する。

MSE (Mean Squared Error) は Generator が生成した画像と学習データの画像のピクセル値の平均二乗誤差を表している。 y_i が実際の値、 \hat{y}_i が予測値である。MSE は値が小さいほど誤差の少ないモデルである。

PSNR(Peak signal-to-noise ratio)は 2 画像の同位置の画素同士の差を参照し、画素値がどのくらい変わったのかを示している。PSNR は大きいほど画像品質が良い。

SSIM(structural similarity)は二つの画像の間の「輝度」「コントラスト」「構造」の 3 つの要素を比較し、周囲のピクセルとの相関を取り込んだ指標で、PSNR よりも人間の視覚的評価を反映できるようになっている。 μ_x は画像 X の小領域内の平均画素値、 μ_y は画像 Y の小領域内の平均画素値、 σ_x は画像 X の小領域内の画素値の標準誤差、 σ_y は画像 Y の小領域内の画素値の標準誤差、 σ_{xy} は画像 X と画像 Y 画素値の共分散、 c_1, c_2 は分母の値が小さくなった際に出力値を安定させるための定数である。SSIM は 0 のときに画質最低、1 のときに画質最高を示す。

GMSD(Gradient Magnitude Similarity Deviation)は画像の勾配の大きさの類

似度を比較し、評価を行う。 r はコントラスト類似度と構造類似度の類似性、 d は画像のゆがみを表している。 $GMSD$ は値が小さいほど画像品質が良い。

Perceptual Index(PI) は PIRM(Perceptual Image Restoration and Manipulation)2018 という画像の復元、操作に関連する技術を競う大会において画像品質を評価する際に用いられたものである。Ma et al[21]と NIQE (Naturalness Image Quality Evaluator)[19]を利用し、人間の判断に近い評価ができる。PIは値が小さいほど、画像品質が良い。

$$MSE = \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2 \quad (6)$$

$$PSNR = 10 \cdot \log_{10} \frac{255^2}{MSE} \quad (7)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (8)$$

$$GMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (GMS(i) - GSM)^2} \quad (9)$$

ただし、

$$\mathbf{h}_x = \begin{bmatrix} 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \end{bmatrix}, \mathbf{h}_y = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 \\ -1/3 & -1/3 & -1/3 \end{bmatrix} \quad (9.1)$$

$$\mathbf{m}_r(i) = \sqrt{(\mathbf{r} \otimes \mathbf{h}_x)^2(i) + (\mathbf{r} \otimes \mathbf{h}_y)^2(i)}$$

$$\mathbf{m}_d(i) = \sqrt{(\mathbf{d} \otimes \mathbf{h}_x)^2(i) + (\mathbf{d} \otimes \mathbf{h}_y)^2(i)} \quad (9.2)$$

$$GMSM = \frac{1}{N} \sum_{i=1}^N GMS(i) \quad (9.3)$$

$$GMS(i) = \frac{2\mathbf{m}_r(i)\mathbf{m}_d(i) + c}{\mathbf{m}_r^2(i) + \mathbf{m}_d^2(i) + c} \quad (9.4)$$

$$PI = \frac{1}{2}((10 - Ma) + NIQE) \quad (10)$$

4.3 学習データの画素値の正規化

提案手法を用いて、試験的にノイズ除去をするための学習を行った際に、生成される画像と元画像の彩度が異なる結果が見られた。使用した SRGAN の実装[18]では、入力画像に対して平均・分散を仮定して、その値を基準に画素値の正規化を行っている。このとき、想定していたデータセットが、本研究で用いるデータセット(CelebA)と異なるため、平均・分散が大きく変化してしまうことが原因であった。これを改善するために使用している学習データ CelebA の全画像の平均分散を求め、画素値の正規化の基準とした。変更後の結果を図 8 に示す。



図 8 : 平均・分散の補正の効果

4.4. バイキュービック補間によって生じたノイズ除去の検証

この実験では **SRGAN** を用いたノイズ除去の有効性について検証する。

4.4.1 実験手法

学習画像に対し、一度画像を縦横 $1/4(64 \times 64)$ の大きさに縮小させた後、バイキュービック補間を用い再度元の大きさ (256×256) に戻す処理を施す。これにより 2.2.1 節で述べたような画像の拡大に伴うボケのノイズが生じる。このノイズ画像、元画像を用いて学習を行った。

4.4.2 実験結果・考察

図 9 にノイズ除去結果を示す。また、表 1 にノイズ除去結果の評価指標の値を示す。表は左の列に各評価手法、中央の列に元画像とノイズのある画像を比較した際の評価値、右の列に元画像と生成画像を比較した際の評価値を記している。太字のほうの評価値が優れていることを示している。

主観評価としては拡大によるボケは低減できているが、元画像のような繊細な描写はできてないという評価である。

表 1 の結果では **PSNR** と **MSE**、**SSIM** においてノイズ除去前のほうが評価値が優れていた。**PSNR**, **MSE** の二つの指標は評価値を算出する際に両画像の同じ位置の画素値を比較している。4.3 節で述べたように、画素値の正規化による色相や彩度の変化に対処したが、図 9 右のノイズ除去後の画像は元画像よりも彩度が高くなっている。このため、単純な画素値の比較では、除去前の方が良い結果になったと考えられる。

拡大によるボケのノイズ除去は、**SRGAN** が本来適用されていた超解像に近い処理であり、ノイズ除去が精度よく行えると予想していた。しかし、補間により生じたボケた画素値が、超解像と同等の未知の画素値の推定をむしろ妨げていると言える。**SRGAN** による超解像の結果と、本研究における拡大によるボケのノイズ除去の定量比較は今後の課題とする。



図 9: 画像の拡大によるノイズ
 (左: 元画像 中央: ノイズ画像 右: ノイズ除去後)

表 1: 画像拡大によるノイズ画像とノイズ除去後の画像に対する定量評価 (12 枚平均)

	ノイズ除去前	ノイズ除去後
MSE ↓	69.61495884	430.4116997
PSNR ↑	30.20821791	22.16417047
SSIM ↑	0.8938378601	0.8160060595
GMSD ↓	0.07791627819	0.07045733962
Perceptual Index ↓	7.609141667	5.220491667

4.5. インパルスノイズ除去の検証

4.5.1 実験手法

学習画像に対し、2.2.2 節で述べたインパルスノイズであるごま塩ノイズを学習画像に与え、ノイズ除去能力を検証する。ごま塩ノイズの発生確率は $p=0.025$ とした。生成したノイズ画像を用いて学習を行った。

4.5.2 実験結果・考察

図 10 右に画像のごま塩ノイズに対するノイズ除去結果を示す。また、表 2 に、ノイズ除去結果の評価指標の値を示す。

主観評価としては、図 10 から GAN を用いたノイズ除去によってごま塩ノイズが取り除かれていることがわかる。ノイズ画像でノイズがあった画素も生成画像では自然に復元できており、ノイズ除去性能は高いと言える。

表 2 の結果から、すべての指標において、ノイズ除去後の方が、評価値が優れていた。

ごま塩ノイズでは、ノイズの画素値に置き換えられた画像の画素値を復元する必要がある。復元が必要な画素の位置は不定であるが、ノイズの発生確率を $p=0.025$ としているためその頻度は超解像に比べても少ない。そのため、本来 SRGAN が対象とする超解像よりも復元問題としては容易であり、高いノイズ除去性能を示したと言える。



図 10:ごま塩ノイズ
(左 : 元画像 中央 : ノイズ画像 右 : ノイズ除去後)

表 2 : ごま塩ノイズ画像とノイズ除去後の画像に対する定量評価 (12 枚平均)

	ノイズ除去前	ノイズ除去後
MSE ↓	685.2363798	670.3114794
PSNR ↑	19.87264412	20.21245139
SSIM ↑	0.4448959278	0.842578498
GMSD ↓	0.2133628508	0.06927462612
Perceptual Index ↓	15.23950833	3.543283333

4.6. ガウシアンノイズ除去の検証

4.6.1 実験手法

学習画像に対し、2.2.3 節で述べたガウシアンノイズを学習画像に与え、ノイズ除去能力を検証する。ガウシアンノイズの分散は $\sigma^2=128$ とした。生成したノイズ画像を用いて学習を行った。

4.6.2 実験結果・考察

図 11 右に画像のガウシアンノイズに対するノイズ除去結果を示す。また、表 3 にノイズ除去結果の評価指標の値を示す。

主観評価としては、図 11 からガウシアンノイズを提案手法を用いることで取り除くことができている。曖昧な背景部分においても元画像に近い画像が生成できている。

表 3 の結果では、すべての指標において、ノイズ除去画像の方が画像品質が優れていた。

本実験では、ガウシアンノイズの分散を $\sigma^2=128$ とかなり大きく設定しており、入力となるノイズ画像は原画像に比べて大きく劣化している。それにも関わらず、ノイズ除去画像は視覚的に良い品質となっており、評価指標による数値評価も改善している。SRGAN の持つ、細部まで復元する能力が本実験でのノイズ除去にも効果があったと言える。



図 11:ガウシアンノイズ
(左：元画像 中央：ノイズ画像 右：ノイズ除去後)

表 3 : ガウシアンノイズ画像とノイズ除去後の画像に対する定量評価 (12 枚平均)

	ノイズ除去前	ノイズ除去後
MSE ↓	2699.818602	347.341167
PSNR ↑	14.01129642	22.95230195
SSIM ↑	0.184836697	0.8452572904
GMSD ↓	0.1880677283	0.08583008405
Perceptual Index ↓	10.30899091	4.166854545

4.7 実験を通じての考察

これらの結果から、主観評価としてはノイズのある画像からノイズを除去した画像を生成することができるという評価となった。また、4.4 節では提案手法より生成されたより、ノイズ画像の方が優れていた評価指標が多い一方で、4.5 節、4.6 節ではノイズ画像より提案手法より生成画像の方が全体的に評価指標が優れていた。この結果から、提案手法では拡大によるノイズに対するノイズ除去能力はあまりなく、ごま塩ノイズやガウシアンノイズなどのノイズに対しては有効であると言える。4.3 節では学習データの画素値の正規化を行ったが、4.4 節の結果を見ると、さらに改善が必要であると考えられる。

5. おわりに

本研究では、SRGAN を画像からのノイズ除去に適用して、その性能を調査した。SRGAN では超解像の際に Pixel Shuffler を用いて解像度を上げていたが、この Pixel Shuffler を除去し、解像度をそのままに出力をするよう修正した。さらに、ノイズ除去を行った際の出力画像の彩度の変化を抑えるために、画素値の正規化のための平均・分散を再設定した。その後、複数の種類のノイズ画像を生成し、生成されたノイズ除去された画像とノイズ画像を比較し、SRGAN を用いたノイズ除去がどの程度有効であるかを評価した。結果として、各ノイズ画像からノイズを除去した画像の生成が可能であることが確認できた。また、今回の実験では、提案手法がガウシアンノイズ、ごま塩ノイズに対しては有効であるが、画像の拡大によるノイズにはあまり効果的でないということが分かった。

今後の課題としては、今回は拡大によるノイズ、ごま塩ノイズ、ガウシアンノイズのみで実験を行ったが、他のノイズモデルの比較を行うことが挙げられる。今回使用したノイズは人工的に作成したものであるので、任意の画像に実際の場面に近いノイズを付加した検証を行うことが必要であると考え。また、図 9,10,11 から、出力画像の彩度の違いを低減させる必要があると考え。この彩度の差がなくなることで評価値はさらに良くなると考える。他には、今回は提案手法に対する評価を行ってきたが、他のノイズ除去を行う手法との比較が必要であると感じた。本研究は、SRGAN の最低限の拡張しかしていない。そのため、他手法との比較、検討をもとにノイズ除去のための Discriminator や Generator のパラメータ、構造の変更をする必要があると考え。

謝辞

本研究を行うにあたって、椋木研究室の椋木教授をはじめとする研究室の皆様には非常に世話になりました。椋木教授におかれましては、論文執筆の際には何度も添削をしてくださったり、研究で行き詰った際には適切な助言を頂いたり、大変感謝しております。

また、先輩方に置かれましては、研究に関することから、学校生活、進路に関することまで、様々なためになる助言をいただきました。ありがとうございました。

最後になりますが、執筆にあたって関わってくださったすべての方にお礼申し上げます。ありがとうございました。

参考文献

- [1] Linh Duy Tran, Son Minh Nguyen, Masayuki Arai: “GAN-based Noise Model for Denoising Real Images”, ACCV, LNIP 12625, pp.560-572 (2020)
- [2] Ziyuan Wang, Lidan Wang, Shukai Duan, Yunfei Li : “An Image Denoising Method Based on Deep Residual GAN”, Journal of Physics Conference Series, vol.1550 no.3(2020)
- [3] Christian Ledig, et.al: “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”, CVPR, pp. 4681-4690 (2016)
- [4] 松坂喜幸: “デジタル画像処理[改定新版]”, オーム社, P47 (2015)
- [5] 岩本祐太郎, 韓 先花, 健山 智子, 大橋 基範, 笹谷 聡, 陳 延偉: “勾配情報に基づくエッジ保存補間と超解像技術への応用”, 電気学会論文誌 C, vol.131, no.11, pp.1901-1906 (2011)
- [6] 山口拓郎: “ノイズを含む画像の任意倍率拡大に関する研究”, 慶應義塾大学大学院理工学研究科博士論文(2018)
- [7] 佐藤正平, 狩野均: “セルオートマトンによる多値画像のノイズ 低減手法 (評価方法の検討)”, 情報処理学会研究報告 vol. 2009-BIO-19, no.33 pp.1-8(2009)
- [8] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang: “Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising”, IEEE Transactions on Image Processing(2016)
- [9] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio: “Generative adversarial nets”, Advances in Neural Information Processing Systems (NIPS), pp. 2672–2680 (2014)
- [10] 岡 和寿: “SRGAN の 3 次元モデル超解像への拡張”, 宮崎大学大学院工学研究科修士論文(2018)
- [11] Radford, Alec, Luke Metz, Soumith Chintala: “Unsupervised representation learning with deep convolutional generative adversarial networks”, arXiv preprint arXiv:1511.06434 (2015)
- [12] ZHU, Jun-Yan, et al.: “ Unpaired image-to-image translation using cycle-consistent adversarial networks”, ICCV, pp. 2223-2232,(2017)
- [13] Mirza, Mehdi, Simon Osindero: “Conditional generative adversarial nets”, arXiv preprint arXiv:1411.1784 (2014)
- [14] Linh Duy Tran, Son Minh Nguyen, Masayuki Arai: “GAN-based Noise Model for Denoising Real Images”, ACCV, pp.560–572(2020)
- [15] WANG, Ziyuan, et al.: “An image denoising method based on deep residual GAN”, Journal of Physics, Conference Series. IOP Publishing, p. 032127(2020)
- [16] 金田 健太郎: “3 次元モデルの超解像における 敵対学習の有効性”, 宮崎大学工学部

情報システム工学科卒業論文(2020)

[17] Ziwei Liu, Ping Luo, Xiaogang Wang, Xiaoou Tang: “Large-scale CelebFaces Attributes (CelebA) Dataset”, CelebA

<https://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>

[18] eriklindernoren : “eriklindernoren/PyTorch-GAN”,github,
<https://github.com/eriklindernoren/PyTorch-GAN>

[19] Mittal A., Soundararajan R., Bovik A.C.: “Making a ‘completely blind’ image quality analyzer”, IEEE Signal Processing Letter. vol20 no.3, pp.209–212 (2013)

[20] 田中正行, 奥富正敏: “画素数の壁を打ち破る 複数画像からの超解像技術”, 映像情報メディア学会誌, vol. 62, no. 3, pp. 337-342 (2008)

[21] Ma C., Yang C.Y. Yang, X. Yang, M.H.: “Learning a no-reference quality metric for single-image super-resolution”, Computer Vision and Image Understanding vol158, pp. 1–16 (2017)