

## アンサンブル学習に基づく超解像画像の生成 Generation of Super-Resolution Images Based on Ensemble Machine Learning

筈見 晃清<sup>1)</sup> 戸崎 哲也<sup>2)</sup>  
Kousei Tomami Tetsuya Tozaki

### 1. はじめに

本研究室では、医用画像生成の一環として医用画像からの疑似解剖画像生成[1]に取り組んでいる。医用画像は、医師や研究者にとって診断や治療の重要な情報源となっている。しかし、既存の画像生成技術では十分な量と品質の画像を入手することは困難である。そこで、生成された画像の品質を改善するために、超解像技術の使用が有効と考えられる。超解像技術を用いることで、低解像度の画像から高解像度の画像を生成することができる。超解像技術の研究は、深層学習を用いたアルゴリズムの開発により進歩している。これを用いることで、高解像度の画像を生成するために必要な情報を学習することができる。

本論文では、既存の画像に対して提案した超解像技術を使用し、高品質な生成画像の作成における有効性を評価する。具体的には、深層学習に基づくアンサンブル学習の手法の検討と、本手法により生成された画像の評価を示す。また、従来の超解像手法との比較も行うことで、提案手法の有効性・改善点についても考察する。

### 2. 手法

#### 2.1 データセット

本実験に使用するデータセットには、超解像タスクにおいてベンチマークとして広く利用される 91-image, Set5 に対して前処理を施し低解像度画像・高解像度画像の 2 つを作成する。その後、それらを HDF5 形式に変換し使用する。以下に、91-image, Set5 の画像データの一部を図 1 に示す。



図 1 91-image, Set5

#### 2.2 アンサンブル学習

本実験では、複数のモデルを用いたアンサンブル学習による超解像手法を使用する。使用するモデルの数は 10 とし、それぞれのモデルに対して学習データとして 91-image, 評価データとして Set5 を使用する。モデルのネットワークには、SRCNN[2]を使用する。SRCNN は 2014 年に Chao Dong らによって提案された手法であり、CNN を用いた超解像モデルである。また、それぞれのモデルで得る結果はごくわずかに異なるようにするために、各モデルの学習率を変更した。それぞれのモデルの学習率をまとめたものを表 1 に示す。

- 1) 神戸市立工業高等専門学校 専攻科 電気電子工学専攻
- 2) 神戸市立工業高等専門学校 電子工学科

表 1 各モデルの学習率

モデル名	学習率
model1	0.0001
model2	0.0002
model3	0.0003
model4	0.0004
model5	0.0005
model6	0.0006
model7	0.0007
model8	0.0008
model9	0.0009
model10	0.001

#### 2.3 画像の再構成

ここではアンサンブル学習により、それぞれのモデルから生成した超解像画像の重ね合わせを行うことで、1 枚の超解像画像を得る。画像を重ね合わせる際には、それぞれのモデルから得た画像の全画素の画素値の総和を  $x_1, x_2, \dots, x_{10}$  とする。はじめにそれぞれの画像に対して、全画素値の総和に対する割合(重み)を求める。  $i$  番目の重みを  $w_i$  とすると、  $w_i$  は次の式(1)で求めることができる。

$$w_i = \frac{x_i}{\sum_{k=1}^{10} x_k} \quad (1)$$

重ね合わせる際には、重み  $w_i$  と画素値  $x_i$  の加重平均をとる。すなわち、再構成された超解像画像  $x_T$  は、それぞれのモデルから生成された画像を  $x'_1, x'_2, \dots, x'_{10}$  とすると、次の式(2)で求めることができる。

$$x_T = \sum_{i=1}^{10} w_i x'_i \quad (2)$$

これにより 1 枚の超解像画像を得る。2.1~2.3 の内容をまとめて 1 つのモデルとして図にしたものを、図 2 に示す。

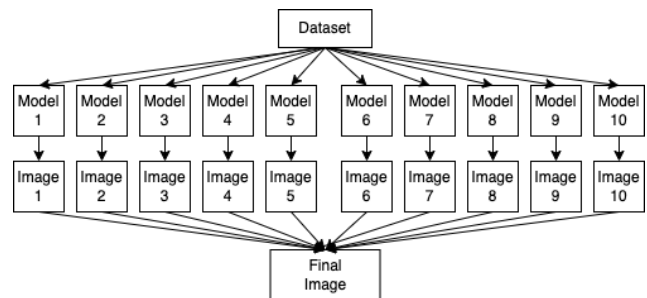


図 2 考案したモデル

## 2.4 評価

生成した画像は、PSNR[3](Peak Signal-to-Noise Ratio, 以下 PSNR と呼ぶ)と SSIM[4](Structural Similarity, 以下 SSIM と呼ぶ)の2つの一般的な画像評価指標により評価を行う。

PSNR はピーク信号対雑音比のことを指し、以下の式(3)のように求められる。

$$PSNR = 10 \log_{10} \left( \frac{MAX^2}{MSE} \right) \quad (3)$$

式(3)において、MSE(Mean Squared Error)は画像の平均 2 乗誤差、MAX は画像の最大輝度値を表す。PSNR では名前の通り最大輝度と雑音の比率から評価値を算出し、大きな値になるほど高い評価であることを表す。

しかし PSNR では、画像全体が少しずつ劣化している時と、局所的に大きく劣化している場合の評価値が等しくなり得るため、人間の主観による評価とは一致しないことや、一貫性のある評価が難しいことが問題となる。そこで、異なる評価指標である SSIM を使用する。

SSIM は画像構造の類似度の評価を数値化できるとされる評価指標である。その評価値は輝度、コントラスト、構造の比較項の乗算からなる以下の式(4)で求められる。

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (4)$$

式(4)において、 $\mu_x$ ,  $\mu_y$  は輝度  $x, y$  の平均値、 $\sigma_x$ ,  $\sigma_y$  は  $x, y$  の標準偏差、 $\sigma_{xy}$  は  $x, y$  の共分散、 $C_1$ ,  $C_2$  は定数である。SSIM で算出された値は 0~1 の範囲に収まり、1 に近づくほど画質が良いことを示す。実際には、1 枚の画像を小領域に分けて計算を行い、それらの総和を範囲で割ることで得た平均値である MSSIM(Mean Structural Similarity)を使用する。MSSIM の定義式を式(5)として以下に示す。

$$MSSIM = \frac{1}{M} \sum_{j=1}^M SSIM(x_j, y_j) \quad (5)$$

式(5)において、 $x_j, y_j$  は小領域の元画像、復号画像を表す。本手法では最終的に RGB3 チャンネル画像を生成するため、それぞれのチャンネルについて MSSIM を算出した後、それらの算術平均をとったものを最終的な SSIM の値とする。

## 3. 結果

考案した学習モデルおよび画像生成モデルから、超解像画像を生成した。はじめに、生成した画像の一部を図 3 に示す。左が従来法である SRCNN で、右が提案手法により生成した画像である。本手法で生成した画像では、少しコントラストが強調されたようにも感じられる。

次に、評価指標から SRCNN による超解像と、本手法を使用した場合で値の比較を行った。算出した PSNR, SSIM の値をまとめた表を表 2, 表 3 に示す。表 2, 表 3 より、PSNR および SSIM の値どちらも、本手法で生成した画像の方が高い数値を記録した。

また、算術平均による画像の生成も行なった。加重平均の場合と比較して画像が少し粗い、明度が低い画像が生成された。加えて PSNR, SSIM の値ともに、上の 2 つの場合に比べて小さいという結果だった。



図 3 生成した画像

表 2 SRCNN での評価値

画像名	PSNR	SSIM
butterfly_GT	28.14	0.9163
Zebra	29.09	0.8676
PPT3	26.57	0.9258

表 3 本手法での評価値

画像名	PSNR	SSIM
butterfly_GT	28.52	0.9238
Zebra	29.36	0.8729
PPT3	26.77	0.9304

## 4. 考察

実験結果より、本手法での生成画像の方が PSNR, SSIM ともに高い数値を記録した。理由としては、実験条件として設定した各モデルにおける学習率の違いが影響していると考えられる。学習率を異なる値に設定することで、わずかながら互いに異なる特徴を持つモデルが作成される。そして 1 つのモデルだけでは抽出できなかった特徴を、複数モデルにすることで捉える事が可能になったと考えられる。これにより、PSNR では信号に対するノイズの抑制の向上、SSIM では元画像との類似性が向上したと考えられる。

また算術平均の場合では、PSNR, SSIM ともに最も低い値を記録した。これは、個々の画像に含まれるノイズ成分や特徴構造のずれが平均化され、画像の鮮明さやコントラストが低下したことが原因と考えられる。

## 5. まとめ

本実験を通じて、提案手法の有効性を確認することができた。今回は超解像モデルとして SRCNN を使用した。今後は、VDSR や SR3 などの他の超解像モデルや、画像生成モデルによる実装、パターン認識やコンピュータビジョンの分野で用いられるフレーム補間手法などの導入も検討する。そして超解像の精度をより向上させ、高品質な画像の生成を目指す。

### 参考文献

- [1] 皆見 晃清(2023), "GAN に基づく医用画像からの疑似解剖画像生成とその評価に関する研究", 神戸市立工業高等専門学校専攻科, 特別研究論文
- [2] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution", European Conference on Computer Vision, Vol.3, (2014).
- [3] Q. Huynh-Thu, and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," IET, vol. 44, Issue 13, pp. 800 - 801, (2014).
- [4] Zhou Wang, "Image quality assessment: from error visibility to structural similarity," IEEE, pp. 600-612, (2004)