

深層学習によるシネ MRI (Magnetic Resonance Imaging) 画像の高フレームレート化 Improvement of frame rate in cine MRI using deep learning

高瀬 唯人[†] 中山 良平[†] 檜作 彰良[†] 市川 泰崇[‡] 石田 正樹[‡] 北川 覚也[‡] 佐久 間肇[‡]
Yuito Takase Ryohei Nakayama Akiyoshi Hizukuri Yasutaka Ichikawa
Masaki Ishida Kakuya Kitagawa Hajime Sakuma

1. はじめに

狭心症や心筋梗塞などの虚血性心疾患は、心筋へ血液を送る冠動脈の血流が悪化し、心筋が酸素不足、栄養不足に陥る疾患である。心疾患は、我が国における死因第 2 位を占めており、その死亡者数は年々増加傾向にある。心疾患は自覚症状が無く、本人の知らないうちに病状が進行してしまうことが、高い死亡率の原因である。しかし、心疾患は早期治療により完治が期待できる疾患であるため、早期発見が非常に重要である。

心臓を対象とした MRI (Magnetic Resonance Imaging) 撮影の一つで、頻繁に用いられる検査法として、シネ MRI [1] がある。シネ MRI は、心電図と同期させて MRI を撮影することにより、1 心拍を 16~40 コマ (フレーム) の動画として観察することが可能である。非侵襲的で正確な診断が出来る MRI 検査であるが、CT (Computed Tomography) 検査に比べて撮影に時間がかかり、また、騒々しく狭い装置内に長時間滞在する必要があるため、患者の体力的負担が大きい。特にシネ MRI は、高いフレームレートの画像を撮影するのに比例して撮影時間が増加する問題がある。フレームレートを下げることにより、撮影時間の短縮は可能であるが、医師が細かな心筋の動きを把握することが困難となり、診断能の低下を引き起こす [2]。

近年、従来のニューラルネットワークに、目的に適した特徴を自動抽出する畳み込み層を導入した畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) が、画像のカテゴリ分類において良好な成果を挙げている。さらに、CNN を用いて解像度を向上させる超解像処理など、CNN を回帰問題に応用する報告もある [3, 4]。超解像処理への応用では、事前に多くの低解像画像とそれに対応する高解像画像を用意し、小領域に分割した低解像パッチと高解像パッチを用いて CNN に信号パターンの関係を学習させる。この高解像度化の概念を時間軸に展開し、CNN により低いフレームレートのシネ MRI 画像を高フレームレート化できれば、撮影時間の大幅な短縮が期待できる。

これまで我々は、超解像処理を行う SRCNN をシネ MRI の高フレームレート化に応用し、良好な結果が得られている。また、最近では U-Net を用いて、低フレームレート画像を高フレームレート化する研究が行われている [5-7]。U-Net は、エンコーダ・デコーダ構造にスキップ接続を加えた手法であり、エンコーダ時の特徴マップをデコーダ時にも利用する技術を導入している。

そこで本研究では、SRCNN, U-Net, また高フレームレート化のために構成した CNN (HFCNN) の比較実験を実施し、低フレームレートのシネ MRI 画像を高フレームレート化するのに適したネットワーク構造を明らかにすることである。

2. 方法

2.1 実験試料

実験試料は、三重大学医学部附属病院の 3T MR 装置 (フィリップス社) で撮像された健常ボランティア 9 名のシネ MRI 画像を用いた。これらのシネ MRI 画像の時間分解能は 20 msec である。

2.2 学習用データの作成

学習用データの作成では、まず、シネ MRI 画像の偶数番目のフレームを間引きし、フレームレートを 1/2 とした低フレームレート画像 (40 msec) を生成した。本研究では低フレームレート画像の各フレーム間に新たなフレームを挿入することにより、高フレームレート画像を生成する。また、挿入画像の前後フレーム画像およびその平均画像、差分画像の 4 枚 1 組をネットワークの入力データとして用い、教師データには挿入位置の間引きした偶数番目のシネ MRI 画像を用いた。図 1 に学習用データの作成例を示す。

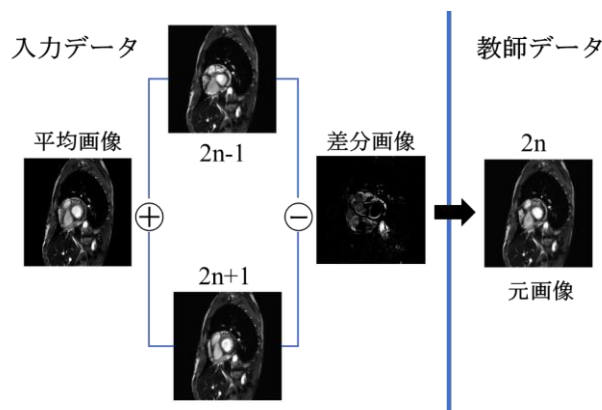


図 1 学習用データの作成

[†] 立命館大学 Ritsumeikan University

[‡] 三重大学 Mie University

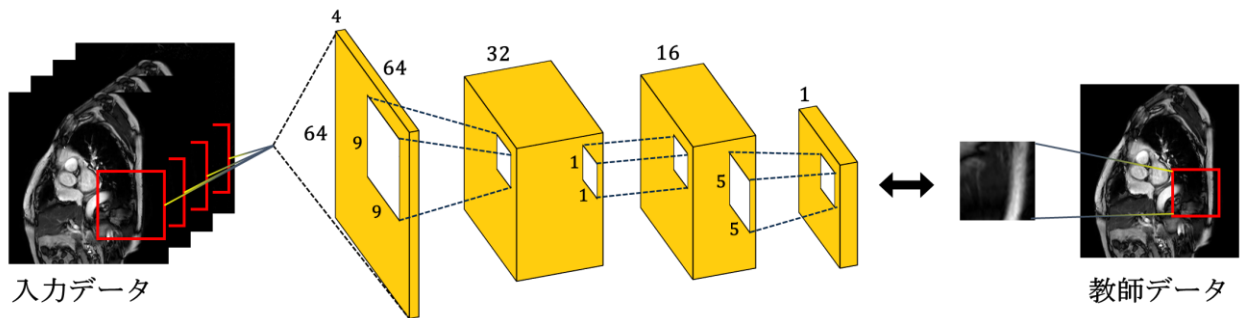


図2 SRCNNを用いたシネMRIの高フレームレート化

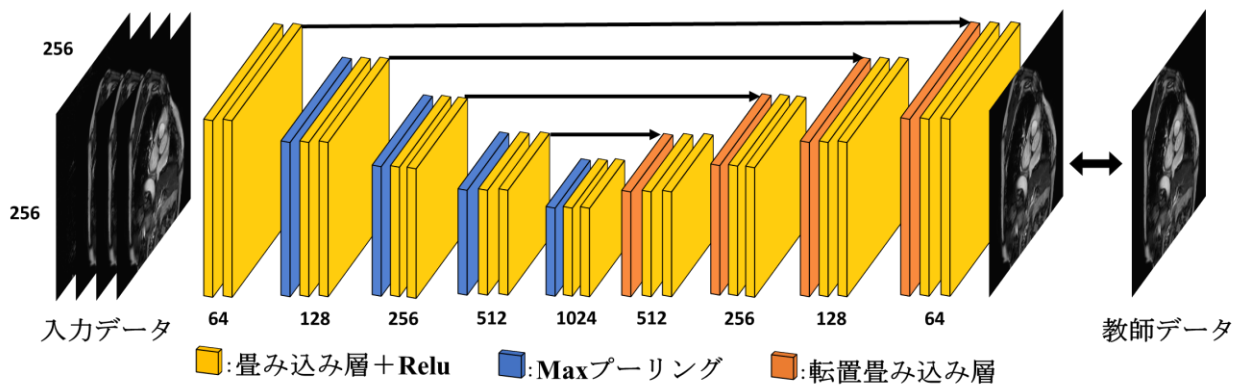


図3 U-Netを用いたシネMRIの高フレームレート化

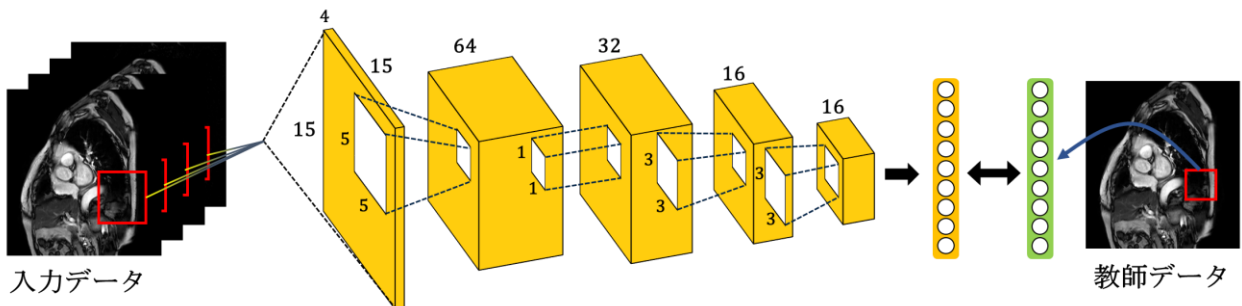


図4 HFCNNを用いたシネMRIの高フレームレート化

2.3 各ネットワークによる高フレームレート化

本研究では、SRCNN、U-Net、高フレームレート化のためのHFCNNを用いて、シネMRI画像の高フレームレート化を行った。

SRCNNの構成を以下に示す。

- 入力層
- 畳み込み層 (9×9のフィルタ 32枚) + ReLU
- 畳み込み層 (1×1のフィルタ 16枚) + ReLU
- 畳み込み層 (5×5のフィルタ 1枚) + ReLU
- 出力層

SRCNNの学習では、入力データを64×64のパッチに分割し、SRCNNの入力とした。また、教師データとして、間引きした偶数番目のシネMRI画像の入力パッチに対応する位置の52×52パッチを与えた。そして、SRCNNの出力値と教師データとのMSE (Mean Square Error) が最小となるようにSRCNN内の重み

係数を更新した。ここで、SRCNNの最適化にはSGD (Stochastic Gradient Descent) を用いた。また、初期学習率は 1.0×10^{-7} 、ミニバッチサイズは400である。図2にSRCNNの構成を示す。

U-Netは、4つの最大プーリング層と4つの転置畳み込み層から構成されており、それぞれに2つの畳み込み層を配置している。図3に示すように、解像度が低くなるにつれ、畳み込み層のフィルタ枚数が増える構成となっている。ここで、全ての畳み込み層におけるフィルタサイズは、3×3である。また、最大プーリング層でのフィルタサイズは、2×2である。U-Netの学習では、入力データから、心臓を中心とした256×256の大きさに切り取り、これをU-Netの入力として与えた。また、教師データとして、間引きした偶数番目のシネMRI画像において、入力データと同様、256×256の大きさに切り出した領域を与えた。そして、U-Netの出力値と教師データとのMSEが最小となるようにU-Net内の重み係数を更新した。ここで、U-Netの最適化にはAdamを

用いた。初期学習率は 1.0×10^{-5} 、ミニバッチサイズは 5 である。

本研究で、高フレームレート化のための HFCNN の構成を以下に示す。

- 入力層
- 畳み込み層 (5×5 のフィルタ 64 枚) + ReLU
- 畳み込み層 (1×1 のフィルタ 32 枚) + ReLU
- 畳み込み層 (3×3 のフィルタ 16 枚) + ReLU
- 畳み込み層 (3×3 のフィルタ 16 枚) + ReLU
- 全結合層
- 出力層

HFCNN の学習では、入力データを 15×15 のパッチに分割し、HFCNN の入力とした。また、教師データとして、間引きした偶数番目のシネ MRI 画像の入力パッチに対応する位置の 3×3 パッチを与えた。そして、HFCNN の出力値と教師データとの MSE が最小となるように HFCNN 内の重み係数を更新した。ここで、HFCNN の最適化には SGD を用いた。また、初期学習率は 1.0×10^{-7} 、ミニバッチサイズは 400 である。図 4 に、高フレームレート化のための HFCNN の構成を示す。

2.4 評価指標

本研究では、低フレームレートの画像から生成された画像のシネ MRI 画像に対する Root MSE, PSNR (Peak Signal to Noise Ratio), SSIM (Structural Similarity) の 3 つの指標を用いて、提案手法の評価を行った。

RMSE は、シネ MRI 画像と高フレームレート化画像間の画素値の差を表す指標である。RMSE は次式で定義され、RMSE の値が、0 に近いほど信号値の誤差が少なく、元のシネ MRI 画像に近い画像が得られたことを示す。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (I - I')^2}$$

N : 画像の画素数

I : シネ MRI 画像

I' : 生成画像

PSNR は高フレームレート化画像がシネ MRI 画像と比較して、どの程度劣化したのかを表す指標であり、値が大きいほど劣化が少ない画像であることを示す。

$$PSNR = 20 * \log_{10} \left(\frac{MAX_I}{RMSE} \right)$$

MAX_I : シネ MRI 画像の最大画素値

SSIM は、画素値、コントラスト、構造の 3 つの要素がどの程度変化したかを統合的に評価する指標であり、値が 1 に近いほど忠実度が高いことを示す。

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

$\mu_x, \mu_y, \sigma_x, \sigma_y$, そして、 σ_{xy} はシネ MRI 画像と高フレームレート化画像の平均画素値、標準偏差、そして共分散値を示す。 C_1, C_2 は定数であり、 $C_1 = (K_1 L)^2$, $C_2 = (K_2 L)^2$ で与えられる。ここで、 K_1, K_2 は定数、 L が表示される画素値の最大値である。

3. 結果と考察

実験に使用したコンピュータの CPU は、Intel® Core i7-78200K CPU@3.60GHz, GPU は NVIDIA GeForce GTX 1080Ti (11GB), メモリは 128GB である。また、プログラミング言語は MATLAB2018a (MathWorks 社) を使用した。

図 5 にシネ MRI 画像 (元画像), 低フレームレート画像を SRCNN, U-Net, HFCNN により高フレームレート化した画像を示す。SRCNN, U-Net による高フレームレート化画像では、心臓の収縮の影響で心壁部が二重になりボケていることが確認できた。一方、HFCNN による高フレームレート化画像では、心壁部にボケが見られず、心臓の動きを忠実に再現することができた。

表 1 に SRCNN, U-Net, HFCNN による高フレームレート化画像の元のシネ MRI 画像に対する忠実度の比較を示す。HFCNN による高フレームレート化画像が、

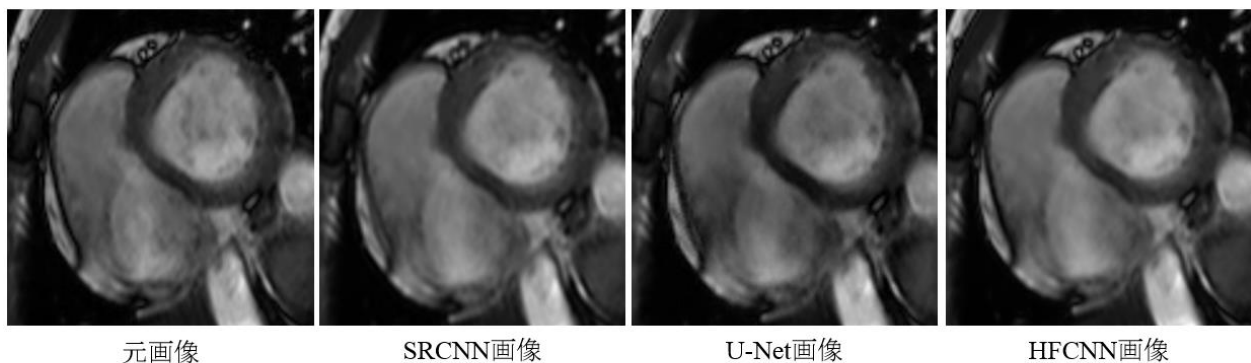


図 5 元画像, SRCNNによる高フレームレート化画像, U-Netによる高フレームレート化画像およびHFCNNによる高フレームレート化画像

他の手法に比べて全ての指標において最も忠実度が高い結果となった。

時間分解能 40 msec の低フレームレート画像では、フレームごとの心臓の変化量が少なかったため、結果画像全体に大きな差は見られなかった。しかし、変化量の多い心壁部では、HFCNN による高フレームレート化画像が、SRCNN, U-Net と比較して、最も忠実に再現できた結果となり、臨床における有用性が示唆された。

また、SRCNN や U-Net に比べて、HFCNN の方が高い忠実度を示した理由として、 15×15 のパッチを学習に用いたことにより、心臓の小さな動きに対応することができたと考えられる。

表 1 元のシネ MRI 画像に対する忠実度

	SRCNN	U-Net	HFCNN
RMSE	8.21	12.79	6.76
PSNR	37.16	33.96	38.95
SSIM	0.996	0.992	0.997

4. まとめ

本研究では深層学習を用いて、シネ MRI 画像の高フレームレート化を行った。SRCNN, U-Net より、本研究で提案した HFCNN による高フレームレート画像が、元のシネ MRI 画像に対する忠実度が高い結果を示し、その有用性が示唆された。

謝辞

研究遂行にあたり貴重なご助言を賜った三重大学医学部附属病院放射線診断科のスタッフの方々に深謝いたします。

参考文献

- [1] Osamu Oshiro, Ayumu Matani, Knihiro Chihara, Taisei mikami & Akira Kitabatake “3D visualization of myocardial motion and blood flow using cine-MR images.” Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. (1996)
- [2] Yudai Suzuki, Keigo Kawaji, Amit R. Patel, Satoshi Tamura & Satoru Hayamizu. “Toward effective noise reduction for sub-Nyquist high-frame-rate MRI techniques with deep learning.” Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. Page.1136 – 1139 (2017)
- [3] Kathiravan Srinivasan, Avinash Ankur & Anant Sharma. “Super-resolution of Magnetic Resonance Images using deep Convolutional Neural Networks.” IEEE International Conference on Consumer Electronics. (2017)
- [4] Chao Dong, Chen Change Loy, Kaiming He & Xiaoou Tang. “Image super-resolution using deep convolutional networks.” IEEE transactions on pattern analysis and machine intelligence. Vol.38 Page.295 – 307 (2015)
- [5] Huaizu Jiang, Deqing Sun, Varan Jampani, Ming-Hsuan Yang, Erik Learned-Miller, & Jan Kautz. “Super sloMo: High quality estimation of multiple intermediate frames for video interpolation.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Page.9000 – 9008 (2018)
- [6] Ziwei Liu, Raymond A. Yen, Xiaoou Tang, Yiming Liu & Aseem Agarwala. “Video frame synthesis using deep voxel flow.” Proceedings of the IEEE International Conference on Computer Vision. Page.4463 – 4471 (2017)
- [7] Simon Niklaus, Long Mai & Feng Liu. “Video frame interpolation via adaptive separable convolution.” Proceedings of the IEEE International Conference on Computer Vision. Page.261 – 270 (2017)