

ベイズ最適化による医療画像応用 CNN (Convolutional Neural Network) のハイパーパラメータの決定

Determination of hyper-parameters in convolutional neural network for medical image applications using Bayesian optimization method

田中 滉大[†] 中山 良平[†] 檜作 彰良[†]

市川 泰崇[‡] 石田 正樹[‡] 北川 覚也[‡] 佐久間 肇[‡]

Kodai Tanaka Ryohei Nakayama Akiyoshi Hizukuri

Yasutaka Ichikawa Masaki Ishida Kakuya Kitagawa Hajime Sakuma

1. はじめに

近年, CNN (Convolutional Neural Network) は画像の分類や高解像度化に応用され, 良好な成果が数多く報告されている. 画像応用において有用性が期待できる CNN であるが, その構成や学習法など, 事前に手動で決定しなければならないハイパーパラメータが多数あり, それらが不適切であると結果が著しく低下する問題がある. また, CNN で決定すべきハイパーパラメータの組み合わせは無限にあり, 従来最適なハイパーパラメータの探索に用いられてきたグリッドサーチやランダムサーチ[1]では, 膨大な時間を要する問題がある.

このようなハイパーパラメータの最適化問題に対処するため, 機械学習の分野において, ベイズ推定を利用した手法が提案されている[2, 3]. この手法では, 最適化の指標となるブラックボックス関数をガウス過程に従うと仮定し, ベイズ推定により事前分布を事後分布に反映させる. これを繰り返すことにより, 最適なハイパーパラメータの効率的な探索を可能にする.

本研究では, 医療画像を対象とした回帰問題, 分類問題において, CNN のハイパーパラメータの決定におけるベイズ最適化の有用性を明らかにする. まず, CNN により冠動脈 MRA (Magnetic Resonance Angiography) 画像の高解像度化画像を推定する回帰問題を取り扱う. 冠動脈 MRA は放射線被曝が無く, 造影剤を必要としないため患者に優しい非侵襲的検査法である. しかし, MR 装置の機構的な制約により CT (Computed Tomography) に比べ低解像な画像となり, より高解像な画像の取得が望まれている. 近年, 低解像画像から高解像画像を推定する手法として CNN を応用した研究が報告されている[4, 5]. これらの手法では, 参照用の低解像画像とそれに対応する高解像画像の信号パターンの関係を CNN に学習させることにより, 低解像画像から高解像画像を生成する. 冠動脈 MRA 画像の高解像度化に適用する CNN のハイパーパラメータをベイズ最適化で決定することにより, より高精度な高解像度化ができると期待する.

また, CNN により乳房 DCE-MRI (Dynamic Contrast-Enhanced Magnetic Resonance Imaging) 画像の腫瘍病変を良性と悪性に分類する分類問題を取り扱う. 乳房 DCE-MRI は, マンモグラフィに比べ, 高い感度を有するため乳がん検診への応用が期待される[6, 7]. しかし, 乳房 DCE-MRI の特異度はマンモグラフィや超音波検査の特異度と同等以下であるとの報告[8, 9]があり, 特異度の改善が望まれ

ている. 近年, 医用画像の病変の良悪性を分類する手法として CNN を応用した研究が報告されている[10]. これらの手法では, 良性と悪性の病変の信号パターンの違いを CNN に学習させることにより, 病変を良性と悪性に分類する. 乳房 DCE-MRI の腫瘍病変の良悪性分類に適用する CNN のハイパーパラメータをベイズ最適化で決定することにより, より高い正答率で分類できると期待する.

本研究ではベイズ最適化を用いてハイパーパラメータを決定した CNN が, 冠動脈 MRA 画像の高解像度化および乳房 DCE-MRI 画像の腫瘍病変の良悪性分類において, 有用となることを示す.

2. 実験試料

2.1 冠動脈 MRA 画像

実験試料は, 三重大学医学部附属病院 (津市) において, 32 チャンネルコイルを搭載した 1.5 T MR システム (Philips 社製) により撮影された 60 患者の冠動脈 MRA 画像である. これらは, 自由呼吸, ナビゲーターエコー法による呼吸同期, T2 ステップ法による心筋信号抑制, そして SPIR 法による脂肪信号抑制の条件下で撮影された. 冠動脈 MRA 画像の画像サイズは 512×512 pixel, 空間分解能は $0.625 \times 0.625 \times 0.8$ mm³ である.

2.2 乳房 MRA 画像

実験試料は, 北斗病院 (帯広市) において, 3.0 T MR システム (GE MEDICAL SYSTEMS 製) により撮影された 38 患者の 3 次元乳房 DCE-MRI 画像である. これらは造影前, 造影直後, 1 分後, 2 分後, 4 分後の 5 時相で撮影され, 各画像サイズは 512×512 pixel, 空間分解能は 0.66×1.0 mm³ である.

3. 実験方法

3.1 ベイズ最適化によるハイパーパラメータの決定

ベイズ最適化では, ハイパーパラメータと目的関数の関係をブラックボックス関数の入出力とみなす. そして, ブラックボックス関数をガウス過程に従うと仮定し, 目

[†] 立命館大学 Ritsumeikan University

[‡] 三重大学 Mie University

的関数を最小化 (最大化) するハイパーパラメータの組み合わせを探索する。

まず, 学習データからランダムに選択した標本点 (ハイパーパラメータの組み合わせ) に対し, ガウス過程回帰モデルに基づいて事前分布を予測する。ただし, 初期の標本点に関しては, ハイパーパラメータの探索範囲内でランダムに 4 点を選択し, 各点で評価を行う。もし評価を失敗する標本点がある場合は, 評価ができるまで標本点の選択を繰り返す。そして, 以下の手順を繰り返し行うことにより, ハイパーパラメータの最適化を行う。

- ① ガウス過程回帰モデルの更新
- ② ベイズ推定により事前分布から事後分布の予測
- ③ 事後分布に基づき獲得関数を最大化する新しい標本点の計算

上記①～③の処理を 40 回繰り返す, 適切なハイパーパラメータの組み合わせを用いて CNN を学習する。そして, 目的関数を最大化する組み合わせを最適なハイパーパラメータとして決定する。冠動脈 MRA 画像の高解像度化では, 目的関数として, CNN により低解像度画像から生成した高解像度化画像の元画像に対する RMSE (Root Mean Squared Error : 平均平方誤差) を用いた。また, 乳房 DCE-MRI 画像の腫瘍病変の良悪性鑑別では, 正答率を用いた。

本研究では, 探索範囲内全てのハイパーパラメータの組み合わせに対する目的関数の出力分布の推定に, ガウス過程回帰モデル[11]を用いる。ガウス過程回帰モデルは, ノンパラメトリックでカーネルを軸にした確率モデルである。任意の n 個の入力データで, 部分集合 $\{x_1, x_2, \dots, x_n\} \subset X$ に対して $(f(x_1), f(x_2), \dots, f(x_n)) \in \mathbb{R}^n$ が, n 次元ガウス過程回帰モデルに従うと仮定し, 平均関数 $m(x)$ と共分散関数 K_θ により定義される。

$$f \sim \mathcal{GP}(f|m, K_\theta)$$

$$m(x_{1:n}) = (m(x_1), m(x_2), \dots, m(x_n))^T$$

$$K_\theta(x_{1:n}) = \begin{bmatrix} k_\theta(x_1, x_1) & k_\theta(x_1, x_2) & \dots & k_\theta(x_1, x_n) \\ k_\theta(x_2, x_1) & k_\theta(x_2, x_2) & \dots & k_\theta(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k_\theta(x_n, x_1) & k_\theta(x_n, x_2) & \dots & k_\theta(x_n, x_n) \end{bmatrix}$$

共分散関数 K_θ はガウス過程回帰モデルの出力に大きな影響を与える。このとき, 共分散関数として各予測子について異なる特徴スケール σ_m を持つ ARD Matern 5/2 kernel を用いた。これにより, 標本点 x_i から近い標本点同士は出力値 y_i が近くなると期待できる。また共分散関数は, 標本点の類似度を測る指標となり [11, 12], 共分散関数を用いることで高次元の分布を柔軟に表現することが可能となる。ここで, x_i と x_j は d 行 1 列のベクトルであり, ベクトル θ によって $k(x_i, x_j|\theta)$ と表される。ARD Matern 5/2 kernel は次式で定義される。

$$k(x_i, x_j|\theta) = \sigma_f^2 \left(1 + \sqrt{5}r + \frac{5}{3}r^2\right) \exp(-\sqrt{5}r)$$

$$r = \sqrt{\sum_{m=1}^d \frac{(x_{im} - x_{jm})^2}{\sigma_f^2}}$$

$$i \neq j \text{ および } i = 1, 2, \dots, n$$

$$\theta_{d+1} = \log \sigma_f$$

$$\theta_m = \log \sigma_m (m = 1, 2, \dots, d)$$

本研究では獲得関数として, EI (Expected Improvement) [13]を用いる。 EI は事前に予測された値の近傍だけでなく, 大域的に標本点を探索することが可能であるため, 局所解に陥りにくい利点がある。 EI は次式で定義される。

$$EI(x, Q) = E_Q \left[\max \left(0, \mu_Q(x_{best}) - f(x) \right) \right]$$

標本点: x , 事後分布: Q
 最小の事後平均の位置: x_{best}
 事後平均の最小値: $\mu_Q(x_{best})$
 目的関数: $f(x)$

3.2 CNN による冠動脈 MRA 画像高解像度化

冠動脈 MRA 画像を高解像度化する CNN に, 高解像画像と低解像画像の対応する位置の信号パターンを学習させる。本研究では, 冠動脈 MRA 画像を 1 回ダウンサンプリングした 256×256 画像を高解像画像として定義した。一方, 低解像画像は高解像画像をダウンサンプリングすることにより 128×128 の縮小画像を生成し, それを Bicubic 法により 256×256 に拡大した画像で定義した。そして, 低解像画像を 11×11 の大きさのパッチ (小領域) に分割し, CNN の入力とした。また, 入力パッチに対する教師データとして, 低解像パッチに対応する位置の高解像パッチを与えた。そして, CNN の出力値と教師データ間の MSE が最小となるように CNN の学習を行った。CNN の学習には, 確率的勾配降下法 (SGD: Stochastic Gradient Descent), ミニバッチサイズ 128 とエポック数 100 を用いた。

ベイズ最適化で探索する CNN の構成に関するハイパーパラメータの範囲は, 畳み込み層の数: 1~5, 各畳み込み層のフィルタサイズ: 1, 3, 5, 7, 各畳み込み層のフィルタ枚数 16, 32, 64 とした。また, CNN の学習に関するハイパーパラメータの範囲は, 学習率: $1e-9 \sim 1e-4$, Momentum: 0.9~0.95, L2Regularization: 0.001~0.05 とした。

低解像 MRA 画像から生成された高解像度化画像の正確さを評価するため, 生成画像の MRA 元画像に対する RMSE, PSNR (Peak Signal-to-Noise Ratio: ピーク信号対雑音比) および SNR (Signal-to-Noise-Ratio) を用いた。RMSE は元画像と生成画像間の画素値の差を表す指標であり, 0 に近いほど誤差が少なく, 元画像に近い画像が得られたことを示す。

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

N : 画像数

I : 高解像 MRA 画像

I' : 高解像度化 MRA 画像

また, PSNR は高解像度化画像が MRA 元画像と比較して, どの程度劣化したかを表す指標である。PSNR は次式で定義され, PSNR の値が大きいほど, 劣化が少ないことを示す。

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

MAX_I : 高解像 MRA 画像の最大画素値

SNRは信号レベルに対する雑音レベルの比であり、雑音の大きさ N に対する信号の大きさ S の比で定義される。

$$SNR = \frac{S}{N}$$

S : 冠動脈内の平均画素値

N : 冠動脈内の画素値の標準偏差

3.3 CNNによる腫瘍病変の良悪性鑑別

DCE-MRI 画像から腫瘍径が最大となるスライスを選択し、腫瘍を中心とした 30×30 の関心領域を 5 つの時相から抽出する。それら 5 関心領域を時相順に連結し、CNN の入力画像 (30×150) とした。また、各関心領域を 90 度、 180 度、 270 度回転し、それらを連結することにより、各症例あたり 4 学習データを生成した。本研究では、CNN の学習と評価には leave one patient out テスト法を用いた。この手法では評価用として 1 患者データを除き、それ以外の 37 患者分の学習データを用いて CNN を学習する。そして、学習した CNN で評価用 1 患者データを評価した。この処理を全ての患者データが評価データとして用いられるまで繰り返した。

ベイズ最適化で探索する CNN の構成に関するハイパーパラメータの範囲は、畳み込み層の数 : 1~5, 各畳み込み層のフィルタサイズ : 1, 3, 5, 7, 各畳み込み層のフィルタ枚数 : 16, 32, 64, また、学習に関する範囲は、学習率 : $1e-1 \sim 1e-3$, Momentum : $0.9 \sim 0.95$, L2Regularization : $0.001 \sim 0.05$ とした。

良悪性分類の評価指標として、正答率(Classification accuracy), 感度(Sensitivity), 特異度(Specificity), 陽性予測度(PPV: Positive predict value), 陰性予測度(Negative predict value)の 5 指標を用いた。ここで、真陽性 : TP(True Positive), 偽陽性 : FP(False Positive), 真陰性 : TN(True Negative)としたときの各指標は以下で定義される。

$$Classification\ accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$PPV = \frac{TP}{TP + FP}$$

$$NPV = \frac{TN}{TN + FN}$$

4. 結果と考察

ベイズ最適化により、冠動脈 MRA 画像を高解像度化する CNN の構成は、 7×7 のパッチを入力とする入力層、 3×3 のフィルタを 32 枚有する畳み込み層、 1×1 のフィルタを 32 枚有する畳み込み層、そして、全結合層と 3×3 の出力層と決定された。また、各畳み込み層の後には、活性化関数として、ReLU (Rectified Linear Unit) 関数が用いられた。ベイズ最適化で決定された高解像度化 CNN の構成を図 1 に示す。

図 2 に MRA 元画像および Bicubic 法、ベイズ最適化なし CNN, ベイズ最適化あり CNN による生成画像の比較を示す。ここで、ベイズ最適化なし CNN は 128 のハイパーパラメータの組み合わせの中で最も RMSE が高かった組み合わせで決定された CNN である。ベイズ最適化あり/なし CNN による生成画像は、MRA 元画像に比べノイズが低減することが確認できた。また、Bicubic 法、ベイズ最適化なし CNN による生成画像は、ベイズ最適化あり CNN による生成画像に比べ、少しボケた画像となった。表 1 に Bicubic 法、ベイズ最適化なし CNN, ベイズ最適化あり CNN による生成画像の MRA 元画像に対する忠実度を示す。ベイズ最適化あり CNN による生成画像が他の手法による生成画像よりも、全ての指標において、忠実度が高い結果となり、その有用性が示された。

また、ベイズ最適化により、乳房 DCE-MRI 画像の腫瘍病変の良悪性を分類する CNN の構成は、 30×150 の入力層、 5×5 のフィルタを 16 枚有する畳み込み層、 5×5 のフ

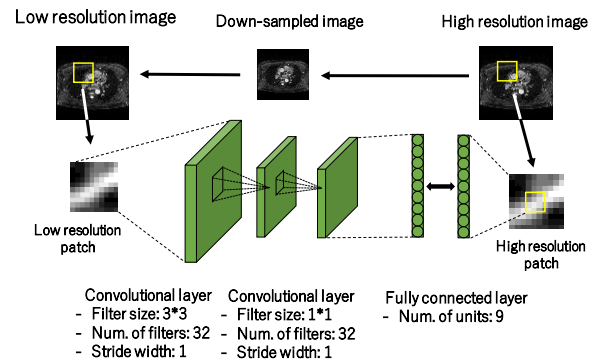
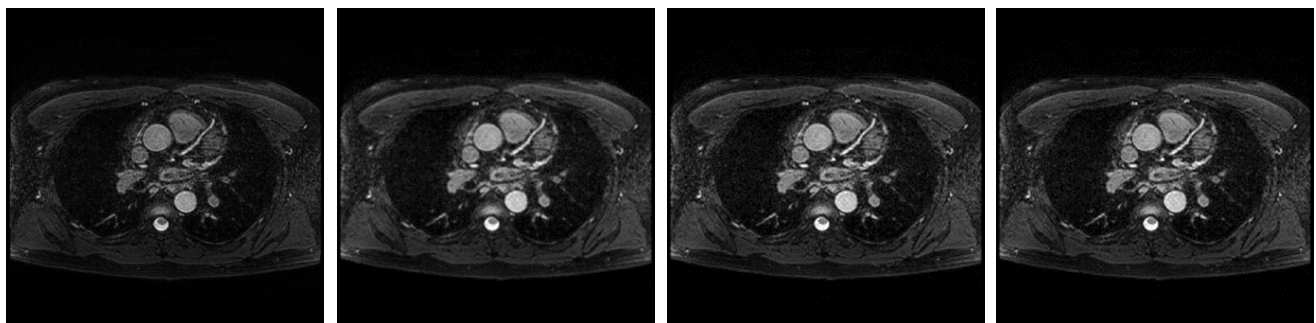


図1 冠動脈MRA画像の高解像度化におけるCNN構造



MRA元画像

Bicubic法による生成画像

ベイズ最適化なしCNN
による生成画像

ベイズ最適化ありCNN
による生成画像

図2 MRA元画像と生成画像の比較

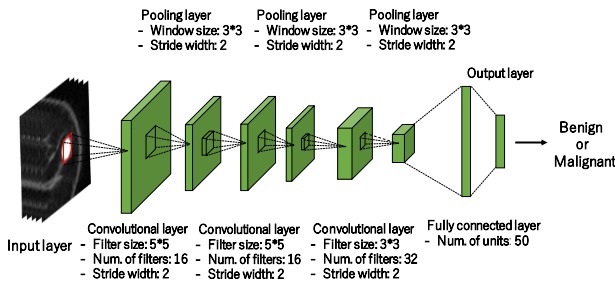


図3 腫瘍病変の良悪性自動分類におけるCNN構造

フィルタを16枚有する畳み込み層、 3×3 のフィルタを32枚有する畳み込み層、そして、全結合層と2つの出力を有する出力層と決定された。各畳み込み層の後には、ReLU関数とプーリング層が用いられた。ベイズ最適化で決定された良悪性分類CNNの構成を図3に示す。

表2にベイズ最適化なしCNN、ベイズ最適化ありCNNによる腫瘍病変の良悪性鑑別結果を示す。ここで、ベイズ最適化なしCNNは64のハイパーパラメータの組み合わせの中で最も正答率が高かった組み合わせで決定されたCNNである。すべての指標において、ベイズ最適化ありCNNによる良悪性分類結果は、ベイズ最適化なしCNNよりも高い結果が得られた。

表1 MRA元画像に対する忠実度

	Bicubic 法	ベイズ最適化無 CNN	ベイズ最適化有 CNN
RMSE	9.21	8.01	7.66
PSNR	36.11	37.90	37.96
SNR	69.22	70.71	70.79

表2 乳房DCE-MRI画像における良悪性分類の結果比較

	ベイズ最適化無 CNN	ベイズ最適化有 CNN
正答率	78.2%	89.1%
感度	73.3%	90.0%
特異度	84.0%	88.0%
陽性予測度	84.6%	90.0%
陰性予測度	72.4%	88.0%

5. まとめ

本研究では、ベイズ最適化を用いてハイパーパラメータを決定したCNNにより、冠動脈MRA画像の高解像度化および乳房DCE-MRI画像における腫瘍病変の良悪性分類を行った。高解像度化の回帰問題、良悪性分類の分類問題ともに、ベイズ最適化ありCNNが、64~128の適当なハイパーパラメータを組み合わせたCNNより、有意に高い結果を示した。ベイズ最適化は効率的に高性能なCNNを構成することが可能であり、幅広い応用が期待できる。

謝辞

本研究に際し、協力いただいた立命館大学大学院先進医用画像解析処理研究室諸氏、並びに実験試料の提供および臨床的助言を頂いた三重大学医学部附属病院の医師の皆様に深謝する。

参考文献

- [1] James B, Yoshua B: Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, 13 pp.281-305, 2012.
- [2] Mockus J, Tiesis V, Zilinskas A: The application of Bayesian methods for seeking the extremum. *Towards Global Optimization 2*, pp.117-129, 1978.
- [3] Snoek J, Larochelle H, Adams RP: Practical Bayesian optimization of machine learning algorithms. *Proceedings of Advances in Neural Information Processing Systems 25 (NIPS 2012)*, pp. 2960-2968, 2012.
- [4] Dong C, Loy CC, He K, Tang X: Learning a Deep Convolutional Network for Image Super-Resolution, *European Conference on Computer Vision (ECCV)*. pp.184-199 2014.
- [5] Dong C, Loy CC, He K, Tang X: Image Super-Resolution Using Deep Convolutional Networks, *IEEE*. 6 pp.295-307 2015.
- [6] Preda P, Turetschek K, Daldrup H: The choice of region of interest measures in contrast-enhanced magnetic resonance image characterization of experimental breast tumors, *Invest Radiol* 40 349-54, 2005.
- [7] Warner E, Messersmith H, Causer P, et al: Systematic review: using magnetic resonance imaging to screen women at high risk for breast cancer, *Ann. Intern. Med.*, 148(9), pp.671-9, 2008.
- [8] Leach MO, Boggis CR, Dixon AK, et al: Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer, a prospective multicentre cohort study (MARIBS). *Lancet* 365, pp.1769-1778, 2005.
- [9] Kuhl CK, Schrading S, Leutner CC, et al: Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer, *J.Clin.Oncol.*, 23 pp.8469-8476, 2005.
- [10] Kunishige M, Kengo Y, Seiji K, et al: Usefulness of Artificial Neural Network for Differential Diagnosis of Hepatic Masses on CT Images, *Academic Radiology*. Volume13, Issue8, pp.951-962, 2006.
- [11] Rasmussen CE and Williams CKI: *Gaussian Processes for Machine Learning*, MIT Press. Cambridge, Massachusetts, 2006.
- [12] Neal RM: *Bayesian Learning for Neural Networks*. Springer, New York. *Lecture Notes in Statistics*, 118, 1996.
- [13] Mockus J: On Bayesian methods for seeking the extremum, *Optimization Techniques IFIP Technical Conference Novosibirsk*, July 1-7, 1974.