

# 学習時の注目領域重視度の動的調整による画像分類精度の改善

船津 朋和<sup>1</sup> 波部 斉<sup>2,3</sup> 永岡 隆<sup>4</sup> 小塚 健倫<sup>5</sup> 根本 充貴<sup>6</sup>

山田 誉大<sup>7</sup> 吉田 久<sup>4</sup> 木村 裕一<sup>3,4</sup> 石井一成<sup>5,7</sup>

近畿大学大学院総合理工学研究科<sup>1</sup> 近畿大学理工学部情報学科<sup>2</sup> 近畿大学情報学研究所<sup>3</sup>

近畿大学生物理工学部生命情報工学科<sup>4</sup> 近畿大学医学部 放射線医学教室 放射線診断学部門<sup>5</sup>

近畿大学生物理工学部医用工学科<sup>6</sup> 近畿大学高度先端総合医療センター PET 分子イメージング部<sup>7</sup>

## 1. はじめに

視覚的説明は深層学習の意思決定を分析するためによく使用されている。特に医用画像を深層学習を用い分類する際、分類結果と注視領域の不一致が起きている場合、分類の信頼性が低下する。そこで分類結果と注視領域の不一致を解決するために任意の注視領域も同時に学習できる Attention Branch Network (ABN) [1]が提案された。ABN のような任意の注視領域を学習させるモデルでは、注視領域の学習と画像分類の学習を同時に行っている。しかし、同時に2種の学習を行っているため、適正な結果が得られずに精度が下がっている場合がある。

そこで本研究では、損失関数の中で注視領域の重要度を定める因子を動的に変化させる手法を提案する。提案手法では学習前半では任意の注視領域を重視しながら学習を進め、学習後半では画像分類の結果を重視した学習を行う。本論文では、CT 画像による COVID-19 性肺炎識別を例にとり、提案手法による学習が効果的に働き、精度が向上することを示す。

## 2. 提案手法

提案手法の概要を図1に示す。手法の流れとして、[1]と同じように分類対象の画像を図2の様な構造である ABN に入力し、出力された結果を元に、式(1)、(2)の様に損失を計算する。 $X_i$ は入力  $i$  番目の画像、 $L_{att}$ は Attention branch の出力に対するクロスエントロピー誤差、 $L_{per}$ は Perception branch の出力に対するクロスエントロピー誤差、 $M(X_i)$ は Attention Map で、 $M'(X_i)$ は任意の注視領域であり、 $L_{map}$ は平均二乗誤差である。

$$L_{abn}(X_i) = L_{att}(X_i) + L_{per}(X_i) \quad (1)$$

$$L_{map}(X_i) = \|M'(X_i) - M(X_i)\|_2 \quad (2)$$

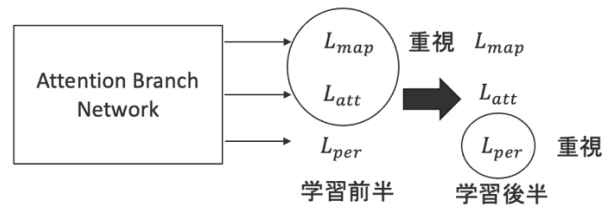


図1 提案手法の概要

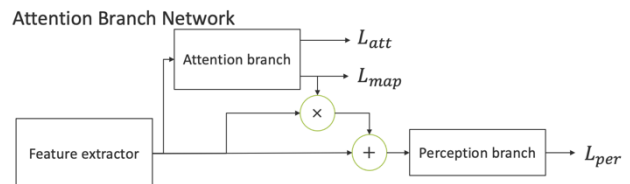


図2 Attention Branch Network

$$L_n(X_i) = L_{abn}(X_i) + \gamma L_{map}(X_i) \quad (3)$$

$$L_d(X_i) = (1 - \alpha)L_{per}(X_i) \quad (4)$$

$$+ \alpha(L_{map}(X_i) + L_{att}(X_i))$$

$$\alpha = \begin{cases} 0.9 & (n < \frac{E}{2}) \\ 0.1 & (n \geq \frac{E}{2}) \end{cases} \quad (n = 1, 2 \dots E - 1) \quad (5)$$

従来手法では、式(3)の様に損失  $L_n(X_i)$  を設定しているが、提案手法では式(4)の様に損失  $L_d(X_i)$  を設定する。式(3)の  $\gamma$  はスケール因子である。式(4)の  $\alpha$  は式(5)の様に 0.9 から 0.1 へ切り替える。式(5)の  $E$  は全体の学習回数であり、 $n$  は現在の学習回数である。このように損失の重み  $\alpha$  を動的に調整することで、図1のように学習初期は Attention Map (注視領域) を重視しながら学習を進め、学習の後半では画像分類の結果を重視した学習ができる。



図3 通常肺炎

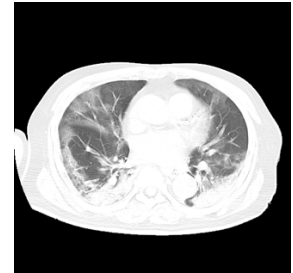


図4 COVID-19 肺炎

Improvement of image classification accuracy by dynamic adjustment of focus area emphasis during learning.

†Tomokazu Funatsu, Kindai University.

‡Hitoshi Habe, Takashi Nagaoka, Takanori Koduka, Mitsutaka Nemoto, Takahiro Yamada, Hisashi Yoshida, Yuuichi Kimura, Ishi Kazunari, Kindai University.

### 3. 実験

#### 3.1. データセット

実験に使用するデータセットとして、近畿大学病院で撮影された、図 3, 4 の様な通常肺炎と、COVID-19 肺炎の 2 種のラベルを持つ CT 画像を使用する。それぞれ使用する枚数は、COVID-19 肺炎が 6621 枚、通常肺炎が 7092 枚である。Attention Map の教師データとして、肺のみを切り出したデータを使用する。これは Hofmanninger らが作成した U-Net のパラメータ [3] を使用して生成した。例を図 5, 図 6 に示す。

本研究は近畿大学医学部、生物理工学部、総合理工学研究科の倫理委員会の承認の下で実施した。



図 5 入力データ



図 6 教師データ

#### 3.2. 実験方法

実験では従来手法の「損失関数の調整なし」と提案手法の「損失関数の調整あり」を比較する。ABN のバックボーンネットワークとして、ResNet18 [3] を使用する。今回使用するデータセットは小規模であり、一通りの評価では正しい評価ができない可能性がある。そのため 3 分割交差検証を実施することで正確な評価をできる様にする。

性能の評価は、スライス単位の指標と、患者単位の指標を使用する。スライス単位の指標は正診率、感度、特異度、AUC を使用し、患者単位での指標は、1 患者全てのスライス画像で出力された softmax の平均を元にした診断結果、感度、特異度を使用する。

### 4. 結果と考察

実験の結果は表 1 の様になった。損失関数の調整ありの方が全体的に精度が高く出ている。

表 1 実験結果

	正診率	感度	特異度	AUC
従来(調整無)	83.1%	<b>88.9%</b>	77.8%	0.914
提案(調整有)	<b>84.0%</b>	81.3%	<b>86.5%</b>	<b>0.923</b>

AUC と患者ごとの診断結果は表 2 の様になった。損失関数の調整ありの方が指標は高く出ている。

表 2 実験結果

	診断結果	感度	特異度
従来(調整無)	<u>299</u> <u>346</u>	<u>115</u> <u>119</u>	<u>184</u> <u>227</u>
提案(調整有)	<u>309</u> <u>346</u>	<u>104</u> <u>119</u>	<u>205</u> <u>227</u>

表 1 を見ると、損失関数の調整による学習が効果的に働いているため、精度が向上している。感度は調整なしの方が高くなっているが、特異度が大きく落ちていることから、COVID-19 肺炎へ偏った振り分けを行なったため感度が向上したと考えることができる。

患者ごとの診断結果でも同じ傾向が出ており、本手法はスライス一枚に対しそれぞれのラベルへ確信度が高く分類できていることがわかる。このことから、提案手法による学習は効果的であったため精度が向上したといえる。

### 5. まとめ

本項では、損失関数の中で注視領域の重要度を定める因子を動的に変化させる手法を提案し、COVID-19 と通常肺炎の診断では通常の学習より良い結果を示した。今後は Microsoft COCO [4] 等のデータセットを使い、手法の汎化性能を確認していきたい。

本研究の一部は、“オール近大”新型コロナウイルス感染症対策支援プロジェクトおよび科研費 JP21H05302 の助成を受けた。

### 参考文献

- [1] 三津原将弘, 福井宏, et al, “Attention map を介した Deep Neural Network への人の知見の組み込み,” 電子情報通信学会論文誌, Vol. J104-D, No. 11, pp. 796-807, 2021.
- [2] Hofmanninger, J., Prayer, F., Pan, J. et al. “Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem.” Eur Radiol Exp 4, 50 (2020). <https://doi.org/10.1186/s41747-020-00173-2>
- [3] Yu, Xin and Yu, Zhiding, et al. “Learning Strict Identity Mappings in Deep Residual Networks”, CVPR 2016
- [4] Tsung-Yi Lin, Michael Maire et al. “Microsoft COCO: Common Objects in Context”, 2015. <https://cocodataset.org/>