不確実性適応型損失関数による 頑健な医用画像セグメンテーション

広島大学 大学院先進理工系科学研究科 情報科学プログラム M243422 廣池 友哉

1 はじめに

セマンティックセグメンテーションは，画像中の全てのピクセルに対し物体のクラスのラベルを推定するタスクである．特に医用画像の分野においてはポリープ [1] や頭頸部がん（HNC）放射線治療における危険臓器 [2] などに用いられている．

しかし，医用画像セグメンテーションでは，画像全体に対する関心領域が極小であるクラス不均衡という課題がある．この課題に対処するため，少数クラスを考慮した学習を行うことができる損失関数が考案されてきた．例えば分類タスクで広く使われている Cross-Entropy Loss [3] と比較して，Dice Index と呼ばれる類似度指標に基づいて定義されている Dice Loss [4] は，少数クラスの誤差も全体の平均損失に大きく寄与するため，クラス不均衡下でも検出性能を向上させる効果がある．また、Dice Loss の多くの拡張手法も提案されており，CT 画像 [5, 6] や MRI 画像 [7] において高い性能が報告されている．しかし，これらのアプローチでは検出難易度に関わらず誤差関数の形状が固定されており，画像によって最適な重みづけがされていないことが課題だった．

そこで本研究では，検出難易度を画像毎に算出したものを学習に取り入れ，誤差関数の形状を動的に変化させる手法を提案する．具体的には，学習中に一定 epoch 毎に推論フェーズを挿入し，MC Dropout [8] により各画像毎に複数枚の予測画像を得た後，ピクセル単位で不確実性を計算し，画像全体の不確実性を難易度指標として算出する．この難易度指標を PolyDice-1 Loss の学習に取り入れることで，画像毎に最適な誤差関数を実現する．

本研究の最終的な目的は，画像毎のセグメンテーションの難易度指標を算出したものを学習に取り入れ，誤差関数の形状を適応的に変化させることである．その実現に向けた最初のステップとして，本稿では難易度指標が学習が進むにつれてどのように推移するかを検証する．

2 提案法

図に提案法の概略図を示す．提案法は，まず学習中に一定 epoch 毎に推論フェーズを挿入し，MC Dropout により複数枚の予測画像を得た後，ピクセル単位で不確実性を計算し，画像全体の不確実性を難易度指標として算出する．

2.1 PolyDice-1 Loss の説明

2.1.1 Polynomial Dice Loss の導出

Dice Loss を拡張し多項式展開することを考える．$W \times H$（pixel）の画像に対し，予測画像を $\hat{\mathbf{y}} \in \mathbb{R}^{W\times H}$，その画像に対する正解画像を $\mathbf{y} \in \mathbb{R}^{W\times H}$ とし，$\hat{\mathbf{y}}, \mathbf{y},$ の位置 i, j における画素値をそれぞれ添字 i, j を用いて $\hat{y}_{ij}, y_{ij}$ と表現すると，Dice Loss は下記の式で示される．

$$\mathcal{L}_{\text{Dice}}(\hat{\mathbf{y}}, \mathbf{y}) = 1 - \frac{2\sum_{i=1}^{W}\sum_{j=1}^{H}\hat{y}_{ij}y_{ij}}{\sum_{i=1}^{W}\sum_{j=1}^{H}(\hat{y_{ij}}^2 + y_{ij}^2)} \tag{1}$$

ここで予測画像 $\hat{\mathbf{y}}$ と正解画像 $\mathbf{y}$ をそれぞれ一次元のベクトル $\hat{\mathbf{y}}', \mathbf{y}'$ に平坦化し，L2 ノルム $\|\cdot\|$ と内積 $\langle\cdot,\cdot\rangle$ を用いて Dice Loss は次式のように表現することができる．

$$\begin{aligned}\mathcal{L}_{\text{Dice}} &= 1 - \frac{2\langle\hat{\mathbf{y}}', \mathbf{y}'\rangle}{\|\hat{\mathbf{y}}'\|^2 + \|\mathbf{y}'\|^2} \\ &= 1 - \frac{2\|\hat{\mathbf{y}}'\|\|\mathbf{y}'\|}{\|\hat{\mathbf{y}}'\|^2 + \|\mathbf{y}'\|^2} \times \frac{\langle\hat{\mathbf{y}}', \mathbf{y}'\rangle}{\|\hat{\mathbf{y}}'\|\|\mathbf{y}'\|}\end{aligned} \tag{2}$$

この式において，$s = \frac{2\langle\hat{\mathbf{y}}', \mathbf{y}'\rangle}{\|\hat{\mathbf{y}}'\|^2 + \|\mathbf{y}'\|^2}$，$\theta = \frac{\langle\hat{\mathbf{y}}', \mathbf{y}'\rangle}{\|\hat{\mathbf{y}}'\|\|\mathbf{y}'\|}$ とおくと，$\mathcal{L}_{\text{Dice}} = 1 - s\cos\theta$ となり，Dice Loss は 2 つの成分，すなわちベクトル間の角度 θ に基づく方向成分（コサイン類似度）と，ベクトルの大きさに基づくスケール成分 s から構成されていることがわる．両方のベクトルが単位ノルムである場合、スケール成分 $s = 1$ となり，Dice Loss はコサイン類似度そのものと一致する [7]．ここで，θ は予測ベクトルと正解ベクトルのなす角であり，予測と正解が大きく異ならない，つまり $\theta \approx 0$ と仮定し，$\cos\theta$ を $\theta = 0$ まわりでテイラー展開すると以下のように近似できる．

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots \tag{3}$$

スケール成分 s を一定と仮定し，この近似式を Dice Loss に代入して整理すると，以下のようになる．

$$\begin{aligned}\mathcal{L}_{\text{Dice}} &= 1 - s\left(1 + \frac{1}{2!}\theta^2 - \frac{1}{4!}\theta^4 + \cdots\right) \\ &= (1-s) + s\left[\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(2k)!}\theta^{2k}\right]\end{aligned} \tag{4}$$

ここで，各テイラー項の係数 α_k を $\alpha_k = \frac{(-1)^{k-1}}{(2k)!}$ とすると，Dice Loss を多項式展開した Polynomial Dice Loss は以下のように表現できる．

$$\mathcal{L}_{\text{PolyDice}} = (1-s) + s\sum_{k=1}^{\infty}\alpha_k\theta^{2k} \tag{5}$$

この損失関数を計算するために、角度 θ は以下のようにコサインの逆関数を用いて算出される．

$$\theta = \arccos\left(\frac{\langle\hat{\mathbf{y}}', \mathbf{y}'\rangle}{\|\hat{\mathbf{y}}'\|\|\mathbf{y}'\|}\right) \tag{6}$$

このようにして，Dice Loss を幾何学的に解釈したうえでテイラー展開を用いることで Polynomial Dice Loss が導かれる．

2.1.2 PolyDice-1 Loss の導出

PolyLoss [9] のアプローチに倣い，多項式の次数に基づいて Polynomial Dice Loss のバリエーションを検討する．まず，テイラー展開で得られた無限級数を第 K 項までで打ち切ることで，以下のような損失関数を定義する：

$$\mathcal{L}_{\text{PolyDice-K}} = \mathcal{L}_{\text{PolyDice}} + s\sum_{k=1}^{K}\varepsilon_k\theta^{2k} \tag{7}$$

Leng ら [9] は，多項式展開の最初の K 項に対してパラメータ $\epsilon_i (i = 1, \cdots, K)$ を用い，高次の項を省略し，低次の項を強調する損失関数を提案した．具体的には次のように定義される：

$$\mathcal{L}_{\text{PolyDice-K}} = (1-s) + s\sum_{k=1}^{K}(\alpha_k + \varepsilon_k)\theta^{2k} \tag{8}$$

しかし，これらのハイパーパラメータ ϵ_i を調整するのは難しく，特に医用画像セグメンテーションでは交差検証による検証が用いられるため，このような複雑なチューニングは実用的ではない．Leng ら [9] は，第 1 項（θ^2）のみを調整するだけで十分な性能向上が得られることを示し，よりチューニングが容易なケース $K = 1$ の損失関数を提案している．これに従い，以下のように PolyDice-1 Loss を定義する．

$$\mathcal{L}_{\text{PolyDice-1}} = (1-s) + s(\alpha_1 + \epsilon_1)\theta^2 \tag{9}$$

2.2 MC Dropout による複数回の推論

前節で導入した PolyDice-1 Loss を用いて学習を行う際に，一定 epoch 毎に推論フェーズを挿入し，MC Dropout により複数枚の予測画像を得る．MC Dropout は，ニューラルネットワークの一部のユニットを確率 p でランダムに不活化する Dropout 処理を学習時のみでなく推論時にも用いることで，予測を分布として計算する手法である．出力を落とすノードを N 回変えることで，出力される値は N 回分の推論結果の分布として得られ，この集合は近似的に予測事後分布の標本集合として解釈できる．提案法では，学習段階で M（epoch）毎に推論フェーズを挿入して Dropout 確率 p で MC Dropout を N 回実行することで，各学習段階で N 枚の複数の予測画像を得る．

3 セグメンテーションの難易度指標の算出

3.1 ピクセル単位の不確実性の計算

前節で画像毎に得られた N 枚の予測画像に対し，画素単位で不確実性を計算する．具体的には，各画素における予測値の分散を不確実性として採用する．予測値の分散を用いることで，各画素における予測のばらつきを評価できると考えられる．

3.2 画像全体の難易度指標の算出

前節で得られたピクセル単位の不確実性を用いて，画像全体の難易度指標を算出する．具体的には，画像全体の不確実性の○○を難易度指標として採用する．○○を用いることで，画像全体の××を評価できると考えられる．

4 実験

4.1 実験条件

提案法の有効性を検証するために，医用画像セグメンテーションデータを用いて実験を行った．データセットはCVC-ClinicDB [10] を用い，5-fold 交差検証に基づき訓練データとテストデータに分割された．なお，訓練データのうち10%をランダムに抽出し，検証データとした．各画像は$W = 224$ pixel, $H = 224$ pixelでリサイズを行い，過学習抑制のため，訓練データに対する空間的データ拡張として50%の確率で上下左右反転及び画像の明るさ・コントラストの変更を施した．

セグメンテーションモデルはU-Net [11] を用い，学習には，バッチサイズ 32, 学習率 10^{-3} の Adaptive moment estimation（以下, Adam）[12] を用いた．最大エポック数は200に設定し，検証データに対する損失が最小のエポックにおけるモデルインスタンスを，最終的なテストデータの評価に用いた．

また，提案法におけるMC Dropoutのパラメータとして，Dropoutを行う確率$p = 0.5$，推論回数$N = 10$，推論フェーズの挿入間隔$M = 10$とした．

4.2 結果と考察

未定

5 まとめと今後の課題

本稿では，医用画像セグメンテーションにおけるクラス不均衡の課題に対処するため，検出難易度を画像毎に算出したものを学習に取り入れ，誤差関数の形状を動的に変化させる手法を提案した．提案法では，学習中に一定epoch毎に推論フェーズを挿入し，MC Dropoutにより複数枚の予測画像を得た後，ピクセル単位で不確実性を計算し，画像全体の不確実性を難易度指標として算出した．医用画像セグメンテーションデータを用いた実験の結果，○○が明らかになった．今後は××および学習に不確実性を取り入れることで，動的な損失関数を実現し，提案法の有効性を検証する予定である．

参考文献

[1] G.-P. Ji *et al.*, “Video polyp segmentation: A deep learning perspective,” *Machine Intelligence Research*, vol. 19, no. 6, pp. 531–549, 2022.

[2] F. Maleki *et al.*, “Machine learning applications for head and neck imaging,” *Neuroimaging Clinics*, vol. 30, no. 4, pp. 517–529, 2020.

[3] J. Long *et al.*, “Fully convolutional networks for semantic segmentation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[4] F. Milletari *et al.*, “V-net: Fully convolutional neural networks for volumetric medical image segmentation,” in *2016 fourth international conference on 3D vision (3DV)*. Ieee, 2016, pp. 565–571.

[5] W. Zhu *et al.*, “Anatomynet: deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy,” *Medical physics*, vol. 46, no. 2, pp. 576–589, 2019.

[6] G. Wang *et al.*, “A noise-robust framework for automatic segmentation of covid-19 pneumonia lesions from ct images,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2653–2663, 2020.

[7] S. Kato and K. Hotta, “Adaptive t-vmf dice loss: An effective expansion of dice loss for medical image segmentation,” *Computers in Biology and Medicine*, vol. 168, p. 107695, 2024.

[8] Y. Gal and Z. Ghahramani, “Dropout as a bayesian approximation: Representing model uncertainty in deep learning,” in *Proceedings of The 33rd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, M. F. Balcan and K. Q. Weinberger, Eds., vol. 48. New York, New York, USA: PMLR, 20–22 Jun 2016, pp. 1050–1059.

[9] Z. Leng *et al.*, “Polyloss: A polynomial expansion perspective of classification loss functions,” in *International Conference on Learning Representations*, 2022.

[10] J. Bernal *et al.*, “Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians,” *Computerized Medical Imaging and Graphics*, vol. 43, pp. 99–111, 2015.

[11] O. Ronneberger *et al.*, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[12] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.