

修士論文

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

M243422 廣池 友哉

指導教員 古居 彬 准教授

令和 8 年 2 月 3 日

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

Copyright © 2026, Tomoya Hiroike.

# 目次

1	はじめに	2
2	予備知識	4
2.1	PolyDice Loss . . . . .	4
2.2	MC Dropout . . . . .	7
3	提案法	9
3.1	概要 . . . . .	9
3.2	不確実性に基づく画像難易度の定量化 . . . . .	13
3.3	適応的損失形状制御 . . . . .	15
4	実験	17
4.1	実験設定 . . . . .	17
5	結論	21
	謝辞	22
	参考文献	23
	付録 A 何とかの導出	26

# 1 はじめに

医用画像セグメンテーションは、診断支援や治療計画において不可欠な技術であり、正常組織または異常組織の領域を抽出することが求められる。特に大腸ポリープ [1] や頭頸部がん放射線治療における危険臓器 [2] の検出など、臨床応用が進んでいる。

しかしながら、医用画像セグメンテーションには固有の課題が存在する。特に重要な問題として、クラス不均衡が挙げられる。医用画像では背景領域が大部分を占め、対象となる病変は相対的に小さい領域しか占めないことが多い。この状況下では、従来の分類タスクで広く使われている Cross-Entropy Loss [3] は背景領域の学習に偏り、臨床的に重要な小病変や曖昧な境界部分の正確なセグメンテーションが困難となる。

この課題に対処するため、クラス不均衡に対して頑健な Dice Loss [4] やその多くの拡張手法が提案され、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、これらの損失関数は全画像に対して固定的な形状を持つという制約がある。その結果、データセット内に含まれる容易な画像困難な画像を差異を十分に考慮できず、一律の学習戦略しか適用できないという限界がある。

医用画像データセットは、撮影条件や個体差に加え、病変サイズ・形状の著しいばらつき、組織間のコントラストの違い、あるいは境界の不明瞭さなどにより、画像ごとにセグメンテーションの難易度が大きく異なる。このような状況下で形状が固定された損失関数を用いると、学習が容易な典型的なサンプルが勾配の大部分を支配してしまい、本来注力すべき難易度の高い症例や微細な境界領域の学習が十分に進行しないという問題が生じる。したがって、この多様性に対応し、効率的かつ頑健なモデルを構築するためには、画像ごとの難易度に応じて損失関数の挙動を動的かつ適応的に調整することが不可欠である。

本研究では、これら 2 つの要素を組み合わせた適応的学习フレームワークの構築を提案する。損失関数の形状制御には、Dice Loss を多項式展開して得られる PolyDice Loss [8] を用いることで、単一のパラメータを用いて損失関数の形状を柔軟に制御す

る。難易度の定量化には、Monte Carlo Dropout [9]（以下、MC Dropout）による不確実性推定を用いる。MC Dropout は、推論時に Dropout を有効にすることで、モデルの認識的不確実性を効率的に推定することを可能にする。この不確実性はモデルがその画像のセグメンテーションにおいてどの程度の確信を持てないかを反映しており、セグメンテーション難易度の指標として利用可能である。提案手法では、推定された不確実性指標に基づき PolyDice Loss の形状パラメータを動的に制御し、困難な画像には急峻な勾配を、容易な画像には緩やかな勾配を与えることで、効率的かつ頑健な学習の実現を目指す。

本研究の主要な貢献は以下の通りである。

1. **不確実性に基づく画像難易度の動的定量化手法の導入**：MC Dropout を用いて推論時の認識的不確実性を推定し、これを画像単位の「学習難易度」として定量化する手法を提案する。これにより、医用画像の多様なセグメンテーション難易度を、モデル自身の確信度に基づいて客観的に評価することを可能にする。
2. **難易度に応じた損失関数の適応的制御フレームワークの構築**：定量化された難易度指標に基づき、PolyDice Loss の形状パラメータを適応的に制御する学習フレームワークを構築した。提案法は、学習の進行に伴い変化する難易度に応じて損失関数の勾配を自動調整することで、困難な症例への学習の注力と、容易な症例による勾配支配の抑制を同時に実現する。
3. **複数のデータセットを用いた有効性と汎用性の実証**：医用画像データセットを用いた比較実験により、提案手法の有効性を検証した。実験の結果、従来の形状が固定された損失関数と比較してセグメンテーション精度が向上することを示した。

## 2 予備知識

### 2.1 PolyDice Loss

医用画像セグメンテーションで広く使用される Dice Loss は、クラス不均衡に頑健であるが、全画像に対して固定的な形状を持つという制約がある。本研究では、Dice Loss を多項式展開により拡張した PolyDice Loss [8]、特にその実用的な形式である PolyDice-1 Loss を採用する。PolyDice-1 Loss は、単一のパラメータ  $\epsilon$  で損失関数の形状を制御でき、画像の難易度に応じて勾配の急峻さを調整することが可能となる。

#### 2.1.1 Dice Loss の定義

画像サイズを  $H \times W$  とし、ピクセル位置を  $(i, j)$  で表す ( $i \in \{1, \dots, H\}, j \in \{1, \dots, W\}$ )。セグメンテーションタスクにおいて、モデルの予測確率マップを  $\hat{\mathbf{Y}} = \{\hat{y}_{i,j}\}_{i,j} \in \mathbb{R}^{H \times W}$ 、その画像に対する正解マスクを  $\mathbf{Y} = \{y_{i,j}\}_{i,j} \in \mathbb{R}^{H \times W}$  とすると、Dice Loss は次式で定義される。

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

#### 2.1.2 幾何学的解釈と多項式展開

予測確率マップ  $\hat{\mathbf{Y}}$  と正解マスク  $\mathbf{Y}$  をそれぞれ長さ  $HW$  のベクトル  $\hat{\mathbf{y}}, \mathbf{y}$  として平坦化すると、Dice Loss は以下のように分解できる。

$$\mathcal{L}_{\text{Dice}} = 1 - s \cos \theta \quad (2)$$

ここで、 $s = \frac{2\langle \hat{\mathbf{y}}, \mathbf{y} \rangle}{\|\hat{\mathbf{y}}\|^2 + \|\mathbf{y}\|^2}$  はスケール成分、 $\theta = \arccos \frac{\langle \hat{\mathbf{y}}, \mathbf{y} \rangle}{\|\hat{\mathbf{y}}\| \|\mathbf{y}\|}$  は 2 つのベクトル間の角度を表す。この分解により、Dice Loss はスケール成分  $s$  と  $\cos \theta$  の積として理解できる。

方向成分  $\cos \theta$  に対して Taylor 展開を適用することで、PolyDice Loss の多項式表

現を導出する。つまり学習が進むにつれて予測  $\hat{\mathbf{y}}$  は正解  $\mathbf{y}$  に近づくため、両ベクトルのなす角  $\theta$  は  $0$  に近づく。この性質を利用し、 $\cos \theta$  を  $\theta = 0$  まわりで泰イラー展開すると以下のように近似できる。

$$\cos \theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots \quad (3)$$

これを Dice Loss に代入し、整理すると PolyDice の一般形が得られる：

$$\mathcal{L}_{\text{PolyDice}} = 1 - s \left( 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots \right) \quad (4)$$

$$= (1 - s) + s \sum_{k=1}^{\infty} \alpha_k \theta^{2k} \quad (5)$$

ここで、 $\alpha_k = \frac{(-1)^{k-1}}{(2k)!}$  は各 Taylor 項の符号係数である。

### 2.1.3 PolyDice-1 Loss

PolyLoss [10] は分類タスクにおいて Cross-Entropy Loss を多項式展開し、第 1 項のみを調整可能とすることで実用的な性能向上を達成した。PolyDice Loss [8] はこのアプローチを Dice Loss に適用したものであり、本研究では第 1 項のみを調整する PolyDice-1 Loss を採用する。

$$\mathcal{L}_{\text{PolyDice-1}} = (1 - s) + s \left( \frac{1}{2} + \epsilon \right) \theta^2 \quad (6)$$

ここで、 $\epsilon \in \mathbb{R}$  は損失関数の形状を制御するハイパーパラメータである。図 1 に、 $\epsilon$  に応じた PolyDice-1 Loss の形状変化を示す。 $\epsilon > 0$  では予測誤差に対するペナルティが強化され、 $\epsilon < 0$  では緩和される。この柔軟な形状制御が可能な特性は、本研究の適応的学習フレームワークにおいて重要な役割を果たす。後述する手法では、この  $\epsilon$  を動的に調整することで、個々の画像の難易度に応じた勾配制御を実現し、学習戦略をサンプル単位で最適化することを可能にしている。

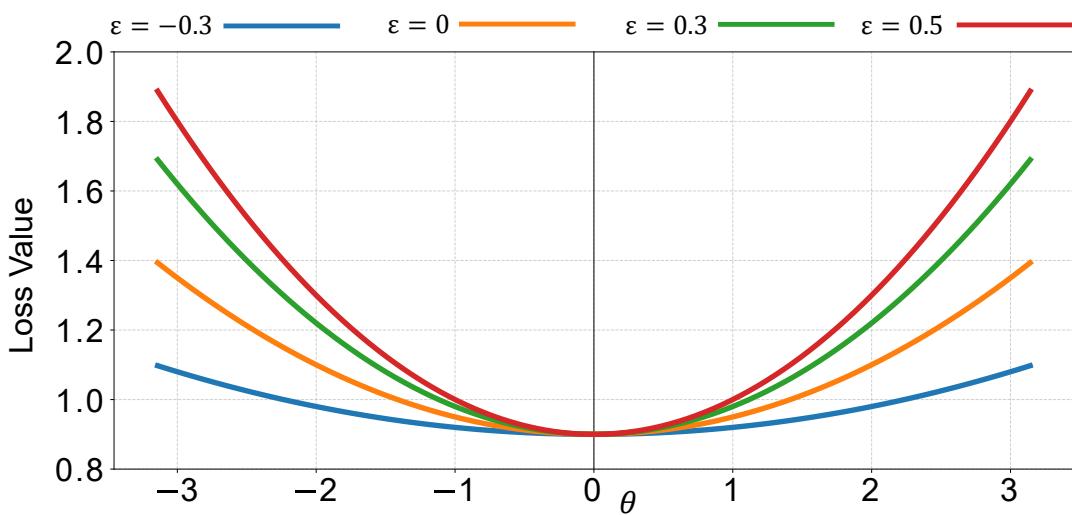


Fig. 1. Plot of PolyDice-1 Loss( $s = 0.1$ )

## 2.2 MC Dropout

### 2.2.1 認識的不確実性と偶然的不確実性

不確実性は、その要因の違いから偶然的不確実性と認識的不確実性に大別される [11].

- **偶然的不確実性 (Aleatoric Uncertainty)**

撮像装置に起因するノイズや低解像度、あるいは組織境界の物理的な曖昧さなど、データそのものに内在する情報不足に起因する。この不確実性はデータの本質的な統計的性質であるため、同一ドメインの訓練データを追加しても解消されない特性を持つ。

- **認識的不確実性 (Epistemic Uncertainty)**

モデルが未学習のパターンや、訓練データの不足による知識の欠如に起因する。これは適切な学習データを追加し、モデルが対象の分布をより詳細に記述することで低減が可能である。

モデルの予測に付随する不確実性を評価し、その依拠する要因を分離する指標として、相互情報量が広く用いられている。全不確実性を示す予測エントロピーから、偶然的不確実性を示す期待エントロピーを減ずることで、以下の通り認識的不確実性を定量化できる。

$$I(y_{i,j}, \omega | \mathbf{X}) = \underbrace{H[\mathbb{E}_{\omega \sim p(\omega|\mathcal{D})}[p(y_{i,j} | \mathbf{X}, \omega)]]}_{\text{Predictive Entropy}} - \underbrace{\mathbb{E}_{\omega \sim p(\omega|\mathcal{D})}[H[p(y_{i,j} | \mathbf{X}, \omega)]]}_{\text{Expected Entropy}} \quad (7)$$

ここで、各記号の定義は以下の通りである。

- $\mathbf{X}$ : 入力画像
- $y_{i,j}$ : ピクセル位置  $(i, j)$  における予測クラスラベル
- $\omega$ : モデルのパラメータ

- $p(\omega | \mathcal{D})$ : 訓練データ  $\mathcal{D}$  に基づくパラメータの事後分布
- $H[\cdot]$ : シャノン・エントロピー
- $\mathbb{E}_{\omega \sim p(\omega | \mathcal{D})}[\cdot]$ : パラメータ分布に関する期待値

### 2.2.2 MC Dropout の原理と応用

Dropout は元々、ニューラルネットワークの過学習を防ぐ正則化手法として提案された [12]. 訓練時に各層のニューロンを確率  $p$  でランダムに不活性化することでモデルの汎化性能を向上させる. 通常、推論時には Dropout は無効化され、全ニューロンが活性化される.

MC Dropout は、学習時のみだけでなく、推論時にも Dropout を有効にすることで、モデルの認識的不確実性を推定する手法である. 通常、推論時の出力は決定論的であるが、Dropout を有効にすることで、異なるサブネットワークが形成され、確率的な出力が得られる. 具体的には、各ニューロンの脱落を決定する Dropout マスクは、重みそのものではなく、重みに乗算されるベルヌーイ分布に従う確率変数として表現できる. このような確率的マスクを導入した学習目的関数を最小化することは、深層ガウス過程に対する変分推論の枠組みにおいて、真の事後分布を特定の変分分布で近似し、その変分下界を最適化することに対応すると解釈できる [9]. したがって、同一入力に対して確率的な推論を複数回実行し、その結果を平均化する操作は、ベイズ推論における予測分布の周辺化をモンテカルロ法で近似していると解釈できる.

提案法では、MC Dropout を用いてモデルがその画像のセグメンテーションにおいてどの程度の確信を持てないかを定量化し、各画像の学習難易度として解釈することで、後述する損失関数の適応的制御に利用する.

### 3 提案法

#### 3.1 概要

訓練データセットを  $\mathcal{D} = \{(\mathbf{X}_n, \mathbf{Y}_n)\}_{n=1}^N$  とする。ここで、 $N$  は訓練画像の総数、 $\mathbf{X}_n \in \mathbb{R}^{H \times W \times C}$  は  $n$  番目の入力画像 ( $C$  はチャンネル数)、 $\mathbf{Y}_n = \{y_{n,i,j}\}_{i,j} \in \mathbb{R}^{H \times W}$  は対応する正解マスクである。MC Dropout による不確実性推定では、各画像に対して  $T$  回の確率的推論を行う。 $t$  回目の推論 ( $t \in \{1, \dots, T\}$ ) における予測確率マップを  $\hat{\mathbf{Y}}_n^{(t)} = \{\hat{y}_{n,i,j}^{(t)}\}_{i,j}$  と表記する。

図 2 に提案手法の概要を、アルゴリズム 1 に詳細なアルゴリズムを示す。本手法の設計は、主に以下の 2 つの観点に基づいている。

第一に、学習サンプルに対する難易度評価を、学習プロセスの中で動的に更新する必要がある点である。画像の「難しさ」は絶対的なものではなく、モデルの学習進捗によって変化する相対的なものである。学習初期には困難であった画像も、学習が進むにつれてモデルが特徴を捉え、容易な画像へと変化しうる。したがって、学習開始前に固定的に難易度を決定するのではなく、現在のモデルの状態に基づいて難易度を逐次再評価することで、その時点でのモデルにとって確信を持てない画像に学習を集中できる。

第二に、難易度の定量化指標として、認識的不確実性が適している点である。単なる予測誤差の大きさは、画像のノイズや偶然的不確実性によっても増大するため、必ずしも「学習すれば改善できる難しさ」を表さない。一方で、認識的不確実性は「モデルの知識不足」に起因するため、モデルが未学習のパターンや判断に迷っている領域を直接的に反映する。この指標を用いることで、ノイズに影響されることなく、学習によって改善可能な「モデルにとってどれくらい確信を持てないか」を適切に定量化することが可能となる。

提案法では、学習中各画像に対して  $\tau$  エポックごとに MC Dropout を用いて画像毎に複数枚推論を行い、予測のばらつきから不確実性を定量化する。この不確実性情

報は、モデルがその画像のセグメンテーションにおいてどの程度の確信を持てないかを反映する。その後、この不確実性情報を画像単位に集約し、PolyDice Loss の形状を動的に制御することで、難しい画像には急峻な勾配を、簡単な画像には緩やかな勾配を与える。更新された  $\epsilon$  を次の  $\tau$  エポック間の学習に適用することで、学習の進行度に応じて最適化の重み付けを動的に変化させる適応的学習を実現する。

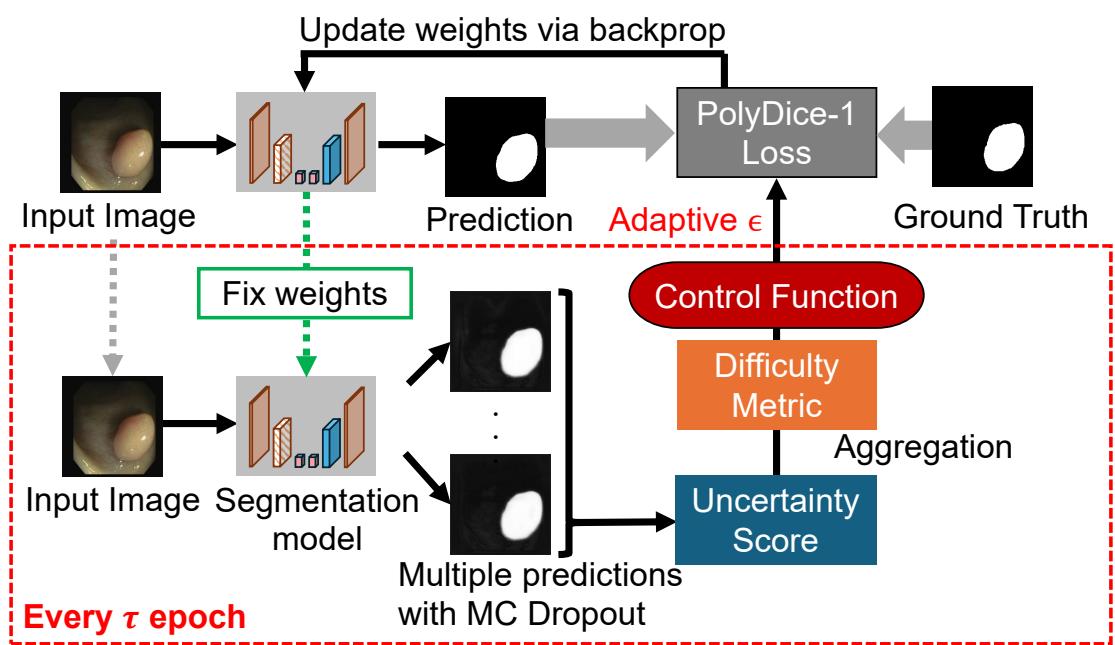


Fig. 2. Overview of the proposed adaptive learning framework

---

**Algorithm 1** Uncertainty-based Adaptive PolyDice-1 Loss Learning

---

**Require:** Training dataset  $\mathcal{D}$ , Max epochs  $E$ , Model  $f_\theta$

**Require: Hyperparameters:** Adaptive start epoch  $E_0$ , Interval  $\tau$ , MC iterations  $N$ , Slope  $k$ , Range  $[\epsilon_{\min}, \epsilon_{\max}]$

```
1: Initialize model parameters  $\theta$ 
2: Initialize loss parameter  $\epsilon_i \leftarrow 0$  for all images  $x_i \in \mathcal{D}$ 
3: for  $e = 1$  to  $E$  do                                 $\triangleright$  Step 1: Adaptive Control Phase
4:   if  $e \geq E_0$  and  $(e - E_0) \pmod{\tau} = 0$  then
5:      $List_D \leftarrow \emptyset$ 
6:     for each image  $x_i \in \mathcal{D}$  do
7:       Perform  $N$  stochastic inferences:  $\{\hat{y}^{(n)}\}_{n=1}^N \leftarrow \text{MC\_Dropout}(x_i)$ 
8:       Update  $\epsilon_i \leftarrow \dots$ 
9:     end for
10:   else
11:      $\triangleright$  Keep  $\epsilon_i$  fixed (Initial Training / Interval)
12:   end if
13:    $\triangleright$  Step 2: Training Phase
14:   for each batch  $(\mathbf{x}, \mathbf{y}) \in \mathcal{D}$  do
15:     Update parameters:  $\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}$ 
16:   end for
17: end for
18: return Trained parameters  $\theta$ 
```

---

## 3.2 不確実性に基づく画像難易度の定量化

### 3.2.1 学習中の MC Dropout 推論

提案法では、学習プロセスを初期学習期間と適応的学習期間の 2 段階に分割する。エポック 1 から  $E_0 - 1$  までの期間は初期学習期間とし、損失形状パラメータ  $\epsilon$  を 0 に固定して学習を行う。この期間を設ける理由は、学習初期のモデルは特徴表現が未成熟であり、この段階での不確実性は画像の本質的な難易度よりもモデルの初期化に依存するためである。 $E_0$  はモデルが基礎的なセグメンテーション能力を獲得するのに十分なエポック数として設定する。

適応的学習期間 ( $e \geq E_0$ ) においては、周期  $\tau$  ごとに不確実性の再評価と  $\epsilon$  の更新を行う。すなわち、更新は  $e \in \{E_0, E_0 + \tau, E_0 + 2\tau, \dots\}$  を満たすエポックの学習開始前に実行される。更新の際は、その時点のモデルパラメータ  $\mathbf{W}$  を固定し、訓練データの各画像  $x \in \mathcal{X}$  に対して Dropout 率  $p \in (0, 1)$  で  $N$  回の確率的推論を行う。得られる予測集合を  $\{\hat{\mathbf{Y}}^{(n)}\}_{n=1}^N$  とする：

$$\hat{\mathbf{Y}}^{(n)} = f_{\mathbf{W}}(x; \mathbf{z}^{(n)}), \quad \mathbf{z}^{(n)} \sim \text{Bernoulli}(1 - p) \quad (8)$$

ここで、 $\mathbf{z}^{(n)}$  は  $n$  回目の推論における Dropout マスクであり、 $\hat{\mathbf{Y}}^{(n)}$  はその予測確率マップである。この確率的推論により得られる予測のばらつきから、モデルの認識的不確実性を定量化する。

### 3.2.2 ピクセル単位の不確実性指標の計算

得られた  $N$  枚の予測画像に対し、ピクセル単位の不確実性指標として認識的不確実性を直接捉えることができる相互情報量  $I_{i,j}$  を計算する。

$$I_{i,j} = H(\bar{y}_{i,j}) - \frac{1}{N} \sum_{n=1}^N H(\hat{y}_{i,j}^{(n)}) \quad (9)$$

ここで  $H(\cdot)$  はエントロピー関数、 $\bar{y}_{i,j}$  は予測の平均を表す。この相互情報量はモデルパラメータの不確実性を反映しており、値が高い領域はモデルが十分に学習できてい

いないことを示唆する。

### 3.2.3 画像単位への集約

医用画像セグメンテーションでは、背景領域が画像の大部分を占める一方で、関心領域である病変部は極めて小さいクラス不均衡が存在する。背景領域は一般に推論が容易であり、その不確実性は極めて低い値をとる傾向がある。そのため、画像全体で不確実性の平均を算出すると、大量の背景画素による低い値が難易度指標全体を支配してしまい、本来捉えるべき病変部の局所的な難易度を適切に定量化できない可能性がある。したがって、病変検出の難易度を鋭敏に反映させるため、本手法では正解画像における病変領域に限定して相互情報量の平均を算出する。

画像領域全体を  $\Omega$ 、正解マスクにおける陽性領域の画素集合を  $\mathcal{P} = \{(i, j) \in \Omega \mid y_{i,j} = 1\}$  とする。ここで、算出される相互情報量には突発的なノイズや極端な外れ値が含まれる可能性があり、これらが難易度指標の定量化を不安定にさせる要因となる。そのため、スコアの算出に先立ち、統計的な外れ値除去を行う。具体的には、領域  $\mathcal{P}$  内の相互情報量の平均を  $\mu_{\mathcal{P}}$ 、標準偏差を  $\sigma_{\mathcal{P}}$  とし、有効な画素集合  $\mathcal{P}'$  を以下のように定義する。

$$\mathcal{P}' = (i, j) \in \mathcal{P} \mid \mu_{\mathcal{P}} - 2\sigma_{\mathcal{P}} \leq I_{i,j} \leq \mu_{\mathcal{P}} + 2\sigma_{\mathcal{P}} \quad (10)$$

この有効集合  $\mathcal{P}'$  を用いて、画像全体の難易度スコア  $D$  は次式で算出される。

$$D = \frac{1}{|\mathcal{P}'|} \sum_{(i,j) \in \mathcal{P}'} I_{i,j} \quad (11)$$

ここで、 $|\mathcal{P}'|$  は外れ値除去後の陽性領域の画素数を表す。

次に、各サンプルの相対的な難易度を決定するため、データセット全体の難易度スコア分布に基づいて正規化を行う。ここでの目的は、画像毎の数値的なスケールの差異を吸収し、各画像が分布全体の中で相対的にどの程度難しいかを評価することである。算出された全画像のスコア集合における  $q$  パーセンタイル値を  $D_q$ 、標準偏差を  $\sigma_D$  とすると、正規化されたスコア  $D_{\text{norm}}$  は次のように計算される。

$$D_{\text{norm}} = \frac{D - D_q}{\sigma_D + \delta} \quad (12)$$

ここで、式中の  $\delta$  は数値的安定性のための微小定数である。 $D_q$  による減算は、後述する制御関数の中心点を入力を合わせるための中心化である。分布の平均値ではなく  $q$  パーセンタイル値を用いることで、容易なサンプルが多数を占める分布においても、難易度の基準点を柔軟に設定し、外れ値の影響を受けずに困難な画像の難易度を定量化可能である。 $\sigma_D$  による除算は尺度の統一であり、難易度の変動幅を統一することで、制御関数の感度がデータセットごとの不確実性のスケールに依存しないように調整する役割を持つ。

### 3.3 適応的損失形状制御

#### 3.3.1 制御関数の設計

得られた難易度指標  $D_{\text{norm}}$  に基づき、PolyDice-1 Loss の形状パラメータ  $\epsilon$  を動的に更新する。更新式には以下のシグモイドベースの制御関数を用いる。

$$\sigma(x) = \frac{1}{1 + \exp(-kx)} \quad (13)$$

$$\epsilon = \epsilon_{\min} + (\epsilon_{\max} - \epsilon_{\min})\sigma(D_{\text{norm}}) \quad (14)$$

ここで、 $k$  は関数の感度を制御するハイパーパラメータであり、 $\epsilon_{\min}, \epsilon_{\max}$  は  $\epsilon$  の変動範囲である。本手法でシグモイド関数を採用する理由は、有界性と滑らかさにある。ステップ関数のような急激な切り替えは学習の安定性を損なう恐れがあり、線形関数ではパラメータ  $\epsilon$  が適切な範囲  $[\epsilon_{\min}, \epsilon_{\max}]$  を逸脱する可能性がある。シグモイド関数を用いることで、難易度が低い領域から高い領域への遷移を滑らかに行いつつ、出力値を常に所定の範囲内に厳密に制約することが可能となる。

この制御により、難しい画像に対して大きな  $\epsilon$  を割り当てることで、損失関数の勾配が急峻になる。これは、同じ予測誤差に対してより大きな損失値と勾配を与えることを意味し、結果として難しい画像からの学習信号が相対的に強化される。一方、すで

に十分に学習できている簡単な画像には小さな  $\epsilon$  を割り当て、過学習を防ぎつつ学習リソースを難しい画像に集中させる。更新された  $\epsilon$  は次の  $\tau$  エポック間の学習に適用され、これによりモデルは困難な画像の学習を重点的に行うことが可能となる。

### 3.3.2 学習アルゴリズム

ああ

## 4 実験

提案法の有効性を検証するために、医用画像データセットを用いた実験を行った。

### 4.1 実験設定

#### 4.1.1 データセットおよび実装の詳細

CVC-ClinicDB データセット [13] データセットおよび Kvasir-SEG データセット [14] を用いて実験を行った。CVC-ClinicDB データセットは 612 枚の大腸内視鏡画像 ( $384 \times 288$  pixel), Kvasir-SEG データセットは 1000 枚の大腸内視鏡画像で、いずれも大腸内視鏡画像とそれに対応するポリープの正解マスクから構成される。

セグメンテーションモデルには表 1 に示される構造の U-Net [15] を採用した。学習には Adam optimizer [16] を使用し、バッチサイズ 32, 学習率  $10^{-3}$  に設定した。損失関数には PolyDice-1 Loss ( $\epsilon = 0$ ) を用いた。前処理として全画像を  $W = 224$  pixel,  $H = 224$  pixel にリサイズし、訓練時には 50% の確率で上下左右反転および明度・コントラストの変更を適用した。最大エポック数  $E$  は 200 に設定した。

MC Dropout による不確実評価は、 $\tau = 10$  エポックごとに実施した。Dropout 層はエンコーダの最終ブロックとデコーダの最終ブロックに配置し、各評価時には、 $p = 0.5$  で  $N = 10$  回の確率的推論を行った。またバッチ内の難易度指標の正規化時のペーセンタイルは  $p = 25$  に設定し、適応的学习の開始エポックは  $E_{\text{start}} = 0$  とし、 $\epsilon_{\min} = 0$ ,  $\epsilon_{\max} = 0.5$ ,  $k = 2$  とした。

#### 4.1.2 比較条件および比較指標

適応的学习の有効性を評価するため、複数の条件下で比較実験を行った。まず、医用画像セグメンテーションにおける標準的なベースラインとして Dice Loss を検討し、併せてクラス不均衡問題に対処する代表的手法である Focal Loss [17] とも比較を行った。また易しいサンプルの重みを下げる割合を調整するパラメータは  $\gamma = 2$  と設定し

た。次に, PolyDice-1 Loss におけるパラメータ  $\epsilon$  を動的に制御することの有効性を検証するため,  $\epsilon$  を固定値とした場合との比較を行った。これには, Dice Loss を近似した  $\epsilon = 0$  の設定に加え, 予備実験により探索されたデータセットごとの最適な固定値を用いた PolyDice-1 Loss が含まれる。これらの比較手法と, 不確実性指標に基づいて  $\epsilon$  を適応的に更新する提案法のセグメンテーション精度を比較することで, 提案手法の有効性を検証する。

またセグメンテーションの性能の評価には, Dice 係数, IoU, Precision, Recall を用いた。

Table 1. Overview of the U-Net Architecture

<b>Layer</b>	<b>Output Size</b>
— <i>Encoder</i> —	
Input	$224 \times 224 \times 3$
inc (DoubleConv)	$224 \times 224 \times 64$
down1 (MaxPool + DoubleConv)	$112 \times 112 \times 128$
down2 (MaxPool + DoubleConv)	$56 \times 56 \times 256$
down3 (MaxPool + DoubleConv)	$28 \times 28 \times 512$
down4 (MaxPool + DoubleConv)	$14 \times 14 \times 512$
— <i>Decoder</i> —	
up1 (Upsample + DoubleConv)	$28 \times 28 \times 256$
up2 (Upsample + DoubleConv)	$56 \times 56 \times 128$
up3 (Upsample + DoubleConv)	$112 \times 112 \times 64$
up4 (Upsample + DoubleConv)	$224 \times 224 \times 64$
outc (Conv2d)	$224 \times 224 \times 2$

Table 2. Performance Comparison with Existing Loss Functions on Kvasir-SEG Dataset

Method	Dice	IoU	Precision	Recall
Dice Loss	0.7895	0.7021	0.8281	0.8154
Focal Loss ( $\gamma = 2$ )	0.7192	0.6082	0.8634	0.6769
PolyDice-1 ( $\epsilon = 0$ )	0.8095	0.7198	0.8461	0.8268
PolyDice-1 (Opt. $\epsilon$ )	0.8095	0.7198	0.8461	0.8268
Adaptive PolyDice-1	<b>0.8272</b>	<b>0.7440</b>	<b>0.8707</b>	<b>0.8397</b>

## 5 結論

本論では、何とか法に基づく何とかかんとかについて述べた。特に、何とかかんとかすることで、何とかが可能である。

実験では、何とかかんとかを目的に何とかを行った。その結果、何とかであったことから、何とかである可能性が示された。

今後は、何とかかんとかを改善していく必要がある。

## 謝辞

本研究を遂行するにあたり、終始熱心なご指導と多大なるご鞭撻を賜りました、指導教員の古居彬准教授に心より感謝申し上げます。古居准教授には、研究の方向性に関する議論から、論文執筆、プレゼンテーション技術に至るまで、細部にわたり丁寧にご指導いただきました。配属当初は期待と不安が入り混じっておりましたが、先生の時に親身で、時に厳格なご助言のおかげで、迷うことなく研究に邁進することができました。

本論文の査読の労をお取りいただきました向谷博明教授、曾智准教授に厚く御礼申し上げます。先生方から頂いた鋭いご指摘は、本研究の質を高め、議論を深める上で極めて重要な指針となりました。

また、共同研究において多角的な視点から貴重なご意見をいただきました相澤宏旭助教に深く感謝致します。学部時代の講義で得た基礎知識が、最先端の研究課題といかに結びつくかをご教示いただき、学問の奥深さと面白さを改めて実感する得難い機会を与えてくださいました。

日々の研究生活において、知能生体情報学研究室の皆様の存在は、私にとって大きな心の支えでした。全体ゼミでの活発な議論はもとより、研究の合間に食事を共にし、日々の他愛もない会話で笑い合った時間は、私にとってかけがえのない思い出です。

最後に、私の大学院進学を快く認め、在学期間中、経済的・精神的に支え続けてくれた両親に、心からの感謝を捧げます。何不自由なく研究に打ち込める恵まれた環境は、家族の献身的な支えがあったからこそ実現したものです。

ここに記して、深甚なる感謝の意を表します。

## 参考文献

- [1] G.-P. Ji, G. Xiao, Y.-C. Chou, D.-P. Fan, K. Zhao, G. Chen, and L. Van Gool, “Video polyp segmentation: A deep learning perspective,” *Machine Intelligence Research*, vol. 19, no. 6, pp. 531–549, 2022.
- [2] F. Maleki, W. T. Le, T. Sananmuang, S. Kadoury, and R. Forghani, “Machine learning applications for head and neck imaging,” *Neuroimaging Clinics*, vol. 30, no. 4, pp. 517–529, 2020.
- [3] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, 2015.
- [4] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-net: Fully convolutional neural networks for volumetric medical image segmentation,” in *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571. Ieee, 2016.
- [5] W. Zhu, Y. Huang, L. Zeng, X. Chen, Y. Liu, Z. Qian, N. Du, W. Fan, and X. Xie, “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, X. Liu, C. Li, Z. Xu, J. Ruan, H. Zhu, T. Meng, K. Li, N. Huang, and S. Zhang, “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, pp. 107695, 2024.
- [8] H. Aizawa, “Polynomial dice loss for medical image segmenta-

tion(submitted),” .

- [9] 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*, vol. 48, pp. 1050–1059, 20–22 Jun 2016.
- [10] Z. Leng, M. Tan, C. Liu, E. D. Cubuk, J. Shi, S. Cheng, and D. Anguelov, “Polyloss: A polynomial expansion perspective of classification loss functions,” in *International Conference on Learning Representations*, 2022.
- [11] L. Smith and Y. Gal, “Understanding measures of uncertainty for adversarial example detection,” *arXiv preprint arXiv:1803.08533*, 2018.
- [12] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, no. 56, pp. 1929–1958, 2014.
- [13] J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño, “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.
- [14] D. Jha, P. H. Smedsrød, M. A. Riegler, P. Halvorsen, T. d. Lange, D. Johansen, and H. D. Johansen, “Kvasir-seg: A segmented polyp dataset,” in *Proceedings of the 26th International Conference on Multimedia Modeling (MMM 2020), Part II*, vol. 11962 of *Lecture Notes in Computer Science*, pp. 451–462, Springer. 2020.
- [15] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241. Springer, 2015.

- [16] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.
- [17] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal loss for dense object detection,” in *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988, 2017.

## 付録 A 何とかの導出

提案法における何とかを導出する。