

異種医用画像間 DG(Domain Generalization)の精度向上のための入力画像変換 Input Image Conversion for High-Performance Domain Generalization Different types of Medical Images

清水 康生[†]
Koki Shimizu

青木 輝勝[†]
Terumasa Aoki

1. はじめに

近年、医療の現場に AI(Artificial Intelligence)が参入する機会が増えている。厚生労働省が主催する、令和 4 年 4 月に開催した第 12 回保健医療分野 AI 開発加速コンソーシアム内にて、AI の開発や利活用促進に向けての具体策が議論されており、医療 AI の開発は、その社会的意義の大きさから勢いをさらに増していくだろう。しかし、医用画像は患者のプライバシー保護に関する問題や撮影時の内部被ばくの問題、撮影機器の違いによる画像のクオリティの問題などから、自然画像と比べて収集できる枚数に限りがある。加えて、人工知能の学習に利用する場合は、正解ラベルを付与するアノテーション作業が必要であり、この作業は長年勤めている医師や放射線技師などの専門的な知識を有する人が手作業で行っているため、使用できる枚数を増大させることは容易ではない。学習画像の枚数は、人工知能の推論性能に大きな影響を及ぼす要素であるため、この問題は医用画像解析に関わる人工知能の研究開発において大きな障害となっている。

この問題に対して行われている有名な研究分野の 1 つとして、Unsupervised Domain Adaptation(UDA)が広く知られている。この分野は、ラベル付けされた学習画像(ソースドメイン)を用いてラベル付けされていない画像(ターゲットドメイン)に対して推論をできるようにすることを目的としており、これまでに様々な学習方法が提案されてきた。しかし、臨床現場での適用を考えたときに、図 1 のような特定の病理に対して異なる種類の医用画像を用意することが先述の問題によって難しく、状況によってはこの手法が使用できないことがある。

これらの問題を踏まえて、近年 Domain Generalization(DG)という分野が注目を浴びている。この分野ではソースドメイン内で共通の特徴量や確率分布を取得することで、推論における汎用性を高めることを目的とする分野である。これにより、UDA で指摘された問題を解決しつつ、より幅広い画像へ対応することが出来る。しかし、この分野はその難易度などから十分な推論精度をもつモデルの数が、他の分野と比べると少ないことが指摘されており、現在も研究がなされている。

そこで本研究では、Ensemble U-Net と呼ばれるモデルに 2 重正規化(Dual-Normalization)というモジュールを組み込むことで、Domain Generalization の医用画像解析における既存手法よりも精度の高いモデルを提案するとともに、ソースドメインを単一ドメインとしてベジェ曲線を用いた色調変換を前処理として行う事による性能向上を試みた。

上記の達成によって、本研究は次の 3 点について当該分野の貢献に寄与できる。

- Domain Generalization において、新しいモデルを提案したこと

- 異種医用画像間セグメンテーションで既存手法よりも高い精度を達成したこと
- Domain Generalization の有用性を示したことで画像診断 AI の利活用促進につながる可能性があること

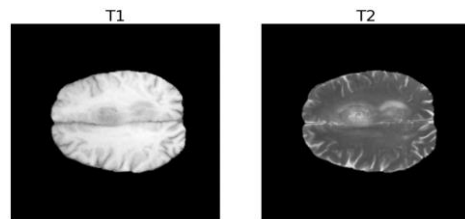


図 1 同じ病理の異なる種類の画像の例

2. 既存研究

2.1 UDA と DG の違い

本節では、UDA と DG の違いについて述べる。両者における学習画像の種類(ソースドメイン)とテスト画像の種類(ターゲットドメイン)の関係性を図に示す。図 2 内では、S をソースドメイン、T をターゲットドメインとし、(a)が UDA、(b)が複数のソースドメインを用いた DG、(c)が本研究の条件である単一ドメインによる DG の様子を表している。どのベン図にも、ソースドメイン同士、およびソースドメインとターゲットドメインが重なっている部分があるが、これは両ドメインに共通する特徴量や画素値の確率分布を表しており、UDA と DG は、ソースドメインとターゲットドメインの共通部分をより広くすることを目的とした研究である。

しかし、学習時にそれらをどのように使い、どのようなモデルを作るかという点では大きな違いがある。UDA では、ソースドメインとターゲットドメインが 1 対 1 であり、学習時はソースドメインとターゲットドメインを比較することでどれほど共通部分の範囲が大きくなったかを評価する。したがって、UDA では、単一ドメインの学習を経て任意のドメインに対して特化したモデルが作られる。一方で、DG はソースドメインとターゲットドメインは必ずしも 1 対 1 でないことに加えて、学習時はソースドメインのみしか使用しない。そのため、(b)では、学習時はソースドメインの間で共通部分を大きくすることを目的とし、その結果として複数のソースドメインに対応した汎用的なモデルが完成する。真に汎用的なモデルであれば、学習に使用していないターゲットドメインにも対応できるため、ソースドメインだけを使っても結果としてターゲットドメインとソースドメインの共通部分を広げていることにつながる。さらに、(c)では、単一のソースドメインに対して複数のターゲットドメインを用いた 1 対多の関係になっている。通常このような単一ドメインでの学習では GAN(Generative

Adversarial Networks)などの画像生成や色調変換などを用いてデータ拡張を行う。これにより、学習条件を疑似的に(b)のような複数のソースドメインを用いるものと同様にできるため、ソースドメイン間での共通部分を広げることができる。その結果としてターゲットドメインを含んだ複数のドメインに対して汎用的なモデルが完成する。

2.2 UDA に関する研究

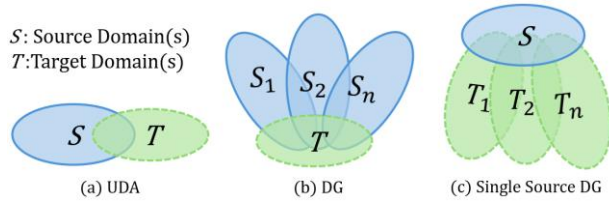


図 2 UDA と DG の違い

UDA とは、ラベル付けされた 1 種類のソースドメインを使って学習したモデルで、ラベル付けされていない他の 1 種類のターゲットドメインを推論させることを目的とする分野である。あるドメインでラベル付きのデータが豊富にあるが、別のドメインでのラベルがほとんどないような状況で用いられる。図 3 は VisDA(The Visual Domain Adaptation)[10]という UDA にて用いられるデータセットを例に説明をする。左の合成画像をソースドメイン、右の実写画像をターゲットドメインとしたとき、この分野のモデルでは、ソースドメインの持つ特徴量の分布や画素値の確率分布をターゲットドメインの持つ特徴量の分布や画素値の確率分布に近づけるための移動量やそれに付随するパラメータを学習する。

Chen ら[11]は前述の両画像の確率分布とは別に、ドメインごとの不変量に注目する学習方法を提案した。この学習方法では、バッチ正規化層(batch normalize layer)をドメインごとに分けることで、それぞれの種類と不変量の区別を、より高い精度で実現している。

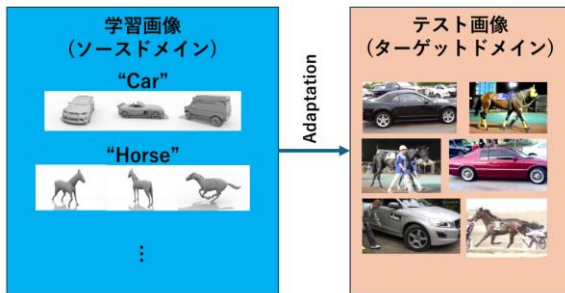


図 3 VisDA データセットの例

2.3 DG に関する研究

DG とは、学習時にソースドメインのみを利用して複数のターゲットドメインに対して推論することを目的とする分野である。これは、ターゲットドメインが複数あり、学習時に使用できない状況であるときに用いられる技術である。図 4 は、PACS(Photo-Art-Cartoon-Sketch)[2]という DG で使用されるデータセットであり、sketch, cartoon, art painting, photo の 4 種類のドメインがある。今回は図中に示している通り、sketch, cartoon, art painting をソースドメイン、

photo をターゲットドメインとする。この分野のモデルでは、ソースドメインの中で共通する特徴量の分布や画素値の確率分布(同時確率分布)を学習する。ターゲットドメインの推論時は、この分布を基に推論を行う。

これらを学習する手法として、Prithivijit ら[3]は、ソースドメイン間の画素値の確率分布を最小にしていけることで、ドメイン間の不変量を取得し、それを用いてターゲットドメインの推論を行うモデルを提案した。また、Segu[4]らは、ソースドメインごとにバッチ正規化層を作るドメイン固有バッチ正規化を提案した。この手法はドメインごとにモデルを作成し、推論時にそのモデルに保存されている統計情報を比較し、最適なモデルを選出するというものである。それと同時に、この研究では 1 種類のソースドメインを用いて学習をするという制約を打破するために、複数のデータ拡張を用いてソースドメインの多様性を確保している。また、Zhou ら[5]はこの手法を単一モデルでの DG が可能になるようなアーキテクチャを考案した。Segu らの研究では、ソースドメインごとにモデルを作成していたが、Zhou らの研究ではモデルの中にソースドメインごとのバッチ正規化層を作成し、それぞれの平均値と分散値を取得する。推論時はターゲットドメインをバッチ正規化層に通すことで平均値と分散値を取得し、それぞれのソースドメインとのユークリッド距離をもって最適なバッチ正規化層を組み込んだモデルでの推論を可能にしている。

また、近年の DG の研究は自然画像だけではなく医用画像においても同様の研究がなされている。Zhou ら[5]は、ソースドメインが 1 種類である単一ドメインという制約がある中でセグメンテーションを行うことができるモデルの提案をした。この手法ではソースドメインに対して色調変換を行い、U-Net のバッチ正規化層をドメイン固有バッチ正規化に適用することで精度の向上を図っている。

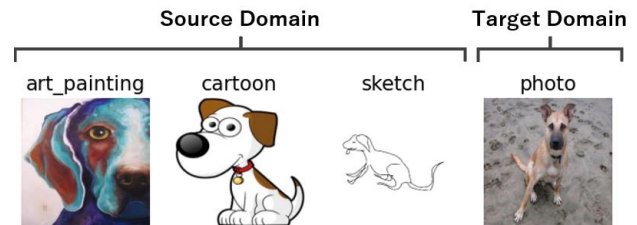


図 4 PACS データセットの例

2.4 ドメイン固有バッチ正規化

ドメイン固有バッチ正規化とは、UDA の際に発生する特徴量の違いを考慮するために、ドメインごとにバッチ正規化(BN)[6]を複数用意するモジュールである。図 5 はバッチ正規化と本モジュールの違いについて図示したものである。通常、このモジュールには $d \in \{S, T\}$ という 2 つのドメインがあり、バッチ正規化同様にアフィンパラメータ γ_d, β_d が用意される。入力画像を x とした時、本モジュールは以下の式で与えられる。

$$\hat{x} = \frac{x_d[i, j, n] - \mu_d}{\sqrt{\sigma_d^2 + \epsilon}} \quad (1)$$

$$\mu_d = \frac{\sum_n \sum_i x_d[i, j, n]}{N \cdot H \cdot W} \quad (2)$$

$$\sigma_a^2 = \frac{\sum_n \sum_{i,j} (x_a[i,j,n])^2}{N \cdot H \cdot W} \quad (3)$$

学習中、本モジュールはドメインごとの平均値と分散値の計算を行い、その値は以下の式で算出される。

$$\bar{\mu}_a^{t+1} = (1 + \alpha)\bar{\mu}_a^t + \alpha\mu_a^t \quad (4)$$

$$(\sigma_a^{t+1})^2 = (1 + \alpha)(\sigma_a^t)^2 + \alpha(\sigma_a^t)^2 \quad (5)$$

これにより、単一のバッチ正規化を行うよりも、ドメインごとの特徴量や確率分布により最適な正規化が可能となる。

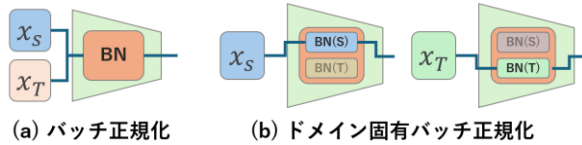


図 5 バッチ正規化とドメイン固有バッチ正規化の違い

2.5 U-Net

U-Net は、2015 年に Ronneberger ら[7]が提案した医用画像のセグメンテーション用のモデルである(図 6)。このモデルは従来の CNN(Convolutional Neural Network)モデル同様、畳み込みを複数回おこなう構造であり、contracting path(bottom-up-path)、expansive path(bottom-down-path)、skip connection という 3 つの特徴がある。contracting path は、1 つの層で 3×3 の畳み込みを 2 回行う。この 2 回の畳み込み処理の後、ReLU 関数と 2×2 の maxpooling を行い、stride が 2 の畳み込みを行うことでダウンサンプリングを実現する。expansive path も同様に、2×2 の畳み込みを 2 回行うのだが、畳み込みの前に同じ階層の contracting path の特徴マップを結合する(これを skip connection と呼ぶ)。これにより、畳み込みによって失われる縁のピクセルを補うことができる。

U-Net は現在でも様々な研究がおこなわれており、特徴量の消失を防ぐために residual block を組み込んだモデル[8]や、3D 画像への応用[9]などがある。

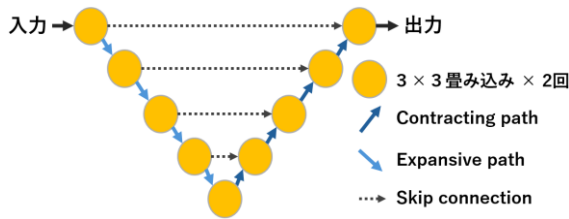


図 6 U-Net のネットワーク図

2.6 Ensemble U-Net

U-Net は、その構造から目的に合わせてユーザーの任意の層の深さでモデルを構築することができる(図 7)。そのモデルの柔軟さから、これまで数々のモデルが提案されてきた。その中の 1 つが Ensemble U-Net である。モデルの概要を図 8 に示す。図 8 中の下向き矢印はダウンサンプリング、上向き矢印はアップサンプリング、点線の矢印はスキップコネクションを表している。このモデルの特徴は 2 つある。

1 点目は複数の U-Net が 1 つのモデルに組み込まれている点である。図 8 中では、1 から 5 層の U-Net が 1 つのモデ

ルの中に組み込まれていることが確認できる。複数のモデルを組み込むメリットとしては、浅い層と深い層のそれぞれで得ることができる特徴量を考慮して推論ができることである。これにより、より確実な推論が可能になる。

2 点目はエンコーダを共有しているという点である。このモデルは 1 つのモデルに複数の U-Net を組み込んでいるため、特徴量抽出について大きな性能差が生まれる。もしそれぞれを独立して学習させる場合、同じ学習回数でも浅いものは過学習を起し、深いものは学習が収束する前に推論が始まってしまう。そこで、特徴量を抽出するモジュールであるエンコーダを共有することによってそれらの問題を解決している。

また、このモデルは推論が組み込まれている U-Net すべてから行われる。図 8 を例に挙げると、モデルの中には 4 つの U-Net が組み込まれているため、推論結果が 4 種類出てくることになる。Ensemble U-Net では、それらの推論結果の平均値をとり、sigmoid 関数へ通すことで推論結果を出している。

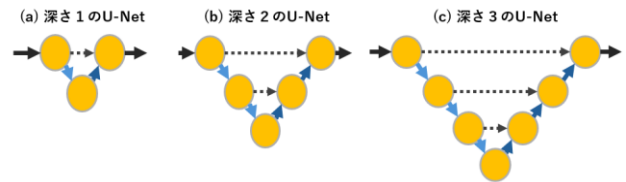


図 7 異なる深さの U-Net

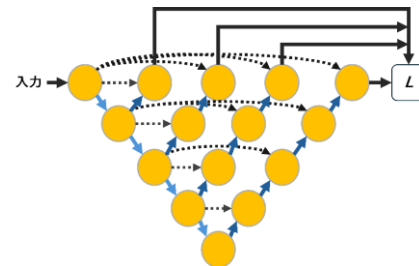


図 8 Ensemble U-Net のネットワーク図

3. 提案手法

3.1 色調変換

本研究は単一ドメインであるため、ソースドメインの多様性を考慮してデータ拡張を行う。一般的に、医用画像はグレースケールであるが、使用する機器の種類によって写り方が全く異なる。その様子を図 9 に示す。両者ともに MR(Magnetic Resonance)画像であるが、左は T2、右は T1 という投影するときのプロトコルが異なった画像である。赤く囲まれているのが脳腫瘍であり、今回のセグメントの対象である。しかし、T2 の画像では脳腫瘍は周りよりも明るく、T1 の画像では周りよりも暗く写っており、両者が正反対の性質を持つことが確認できる。本研究ではこのような写り方が正反対の画像に対応するとともに、単一ドメインであるという事を考慮し、データの多様性と過剰な色調変換を防ぐためにベジェ曲線を用いた色調変換を行う。

ベジェ曲線は、始点と終点(P_0 と P_3)、2 つの制御点(P_1 と P_2)の計 4 点によって生成され、式(6)によって定義される。

$$B(t) = \sum_{i=0}^n \binom{n}{i} P_i (1-t)^{n-i} t^i, n=3, t \in [0,1] \quad (6)$$

ここで、 t は0から1までの小数を表す。また、本研究ではこの関数の値域は-1から1とした。

色調変換の様子を図10に示す。ここではT2というドメインの画像を用いて説明を行う。この曲線は、始点と終点を $P_0(-1,-1), P_3(1,1)$ とした場合は増加傾向、両者を逆にした場合は減少傾向の曲線を得ることができる。また、 $P_0 = P_1, P_2 = P_3$ とした場合は線形関数を得ることができる(図10中、2列目と5列目)。加えて、本研究では0から1の乱数 v を2種類生成し、それぞれを $P_1(-v, v), P_2(v, -v)$ とした。これにより、2種類の制御点と増加傾向(3, 4列目)と減少傾向(6, 7列目)を組み合わせることで、本研究では6種類の色調変換を行った。

また、変換された画像において、増加傾向の曲線で変換された画像はソースドメインと性質が変わらないことから、本研究では source similar domain(D_{ss})と呼び、減少傾向の曲線で変換された画像はソースドメインと性質が正反対であることから、source dissimilar domain(D_{sd})と呼ぶ。これらの操作により、ソースドメインの多様性を確保したとともに、学習データの総量を増やすことができる。

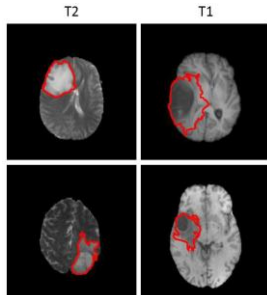


図9 MR画像のT2(左)とT1(右)における写り方の違い

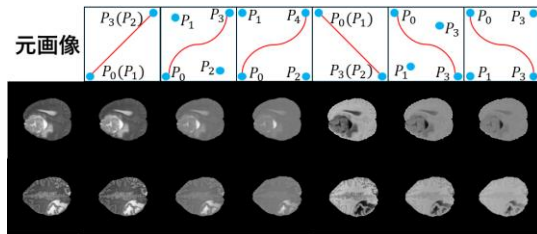


図10 T2画像の色調変換のようす

3.2.2 重正規化

2重正規化は、Zhou[5]によって提案されたアーキテクチャである(図11)。これは、モデルの学習時に D_{ss} と D_{sd} それぞれでバッチ正規化層を分けて学習させることを目的としており、式(7)で定義される。

$$DN(z; d) = \gamma \frac{z - \mu_d}{\sqrt{\sigma_d^2 + \epsilon}} + \beta_d \quad (7)$$

このとき、 z は入力画像、 d はドメインの種類(domain label)、 μ_d は平均値、 σ_d^2 は分散値、 ϵ はゼロで除算されることを防ぐための小さな値、 γ, β_d はアフィンパラメータを表す。

バッチ正規化は、学習時に入力画像の平均値と分散値を取得し、式(7)を更新していく。更新の際は、以下の式を用いて計算される。

$$\mu_d^{t+1} = (1 + \alpha)\mu_d^t + \alpha\mu_d^t \quad (8)$$

$$(\sigma_d^{t+1})^2 = (1 + \alpha)(\sigma_d^t)^2 + \alpha(\sigma_d^t)^2 \quad (9)$$

このとき、 t は直前の対象物を指し、 μ_d^t と $(\sigma_d^t)^2$ はドメイン d における直前の平均値と分散値を示す。これを行う事によってドメインごとに統計情報を保持することができる。

学習が終了すると、各バッチ正規化層の中には、 $\mu_d, \sigma_d^2, \gamma, \beta_d$ の4種類の統計情報が保持される。ここでは、ソースドメインとターゲットドメインの平均値と分散値を比較することで、適切なバッチ正規化層を選択することができる。ここで、この統計情報を e としたとき、以下のようにならされる。

$$e_d = [e_d^1, e_d^2, \dots, e_d^L] \\ = [(\mu_d^1, (\sigma_d^1)^2), (\mu_d^2, (\sigma_d^2)^2), \dots, (\mu_d^L, (\sigma_d^L)^2)] \quad (10)$$

このとき、 d は D_{ss}, D_{sd} のような domain label、 L は L 番目のバッチ正規化層を表す。ターゲットドメインにおいても同様に、以下のようにならされる。

$$e_t = [(\mu_t^1, (\sigma_t^1)^2), (\mu_t^2, (\sigma_t^2)^2), \dots, (\mu_t^L, (\sigma_t^L)^2)] \quad (11)$$

これらの統計量をもって、適切なバッチ正規化層を選択する。算出の際は、ユークリッド距離を用いて D_{ss}, D_{sd} 、ターゲットドメインのそれぞれの距離を求める。 l 番目のユークリッド距離を算出する場合は、以下のようにならされる。

$$W(e_t^l, e_d^l) = \|\mu_t^l - \mu_d^l\|_2^2 + \|\sigma_t^l - \sigma_d^l\|_2^2 \quad (12)$$

この計算を全ての層で行い、ターゲットドメインとの距離が近かったソースドメインのバッチ正規化層を組み込んだモデルで推論を行う事で、単一のモデルでターゲットドメインの推論が可能になる。(図12)

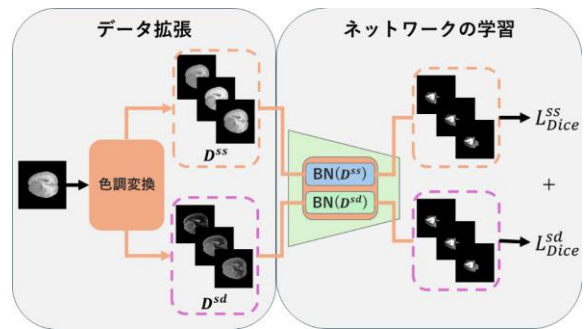


図11 2重正規化の学習時のパイプライン

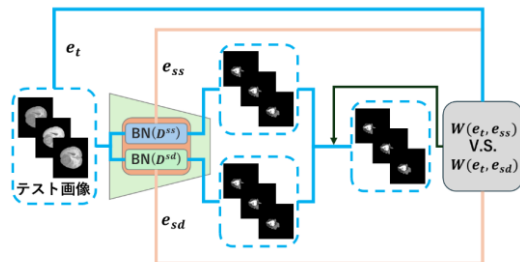


図12 2重正規化の推論時のパイプライン

3.3 Ensemble U-Net with Dual-Normalization

提案手法のネットワークは図 8 に示したとおりである。このモデルの特徴としては、各畳み込みブロック (convolution block) のバッチ正規化層を 2 重正規化が適用している点である。

本研究の根幹となる 2 重正規化を各バッチ正規化層に適用することで、ソースドメイン内のそれぞれのドメインごとに学習が可能になるため、ターゲットドメインに対してより高い精度の推論が可能となる。

4. 評価

4.1 評価方法と実験条件

本節では、初めに使用する評価指標を述べ、続いて実験条件について記述する。

始めに、評価指標についてである。本研究では Dice と Harsdorf Distance (HD) の 2 つの指標を用いる。Dice の算出式を式(13)に示す。ここで、 D は推論結果と正解画像の一致する面積の比率、 N はテスト画像の総数、 i は画像のインデックス番号、 P は推論結果、 G は正解ラベルを表す。

$$D = \frac{2\sum_i^N P_i G_i}{\sum_i^N P_i^2 + \sum_i^N G_i^2} \quad (13)$$

式(13)より、Dice は正解画像と推論結果の一致する面積の比を表しており、推論結果 P が正解ラベル G に近づくことで数値が上がるのがわかる。

続いて、Harsdorf Distance の算出式を式(14)に示す。ここで、 $d_H(P, G)$ は推論結果 P と正解ラベル G のピクセル間の距離を表し、 p は推論結果 P のピクセルの座標、 g は正解ラベル G の座標を表す。

$$d_H(P, G) = \max\{sup_{p \in P} d(p, G), sup_{g \in G} d(g, P)\} \quad (14)$$

式(14)より、HD は正解画像と推論結果のすべてのピクセル間の距離を算出し、その値で最も大きいものを表しており、推論結果 P が正解ラベル G に近づくことで数値が下がることが分かる。

上記の 2 つの評価指標を用いることで、推論結果と正解ラベルの一致する面積を求めるとともに、推論が正しい位置に行われているかを数値化する。

4.1.1 また、本研究では The Cross-Modality Brain Tumor Segmentation Challenge 2018 dataset[13] を用いて実験を行い、80% を学習データ、20% をテストデータとして使用した。各画像は $[-1, 1]$ の範囲で正規化し、画像サイズは 64×64 にリサイズして使用した。比較対象として、既存手法である U-Net with Dual-Normalization を使用した。両ネットワークはバッチサイズ 64、学習率 1.0×10^{-3} で設定をし、NVIDIA-TITAN Xp(12GB) を用いて実験を行った。

4.2 既存手法との比較

本稿では、比較実験として T1 と T2 のそれぞれをソースドメインとして学習したモデルを用いる。それらの推論精度と出力画像を表 1、表 3、図 13、図 14 に示す。また、表 1 と表 3 の Dice と DH の平均値を表 2 と表 4 に示す。図 12、図 13 はどちらも、1 行目画像が入力画像、1 行目画像が正解ラベル、3 行目が提案手法の推論結果、4 行目の画像が既存手法の推論結果である。また、表 1、表 2 には、一番上

の行にソースドメインが書いてあり、Dice、HD の数値をモダリティごとにプロットした。モダリティごとに既存手法と提案手法を比べ、数値が優れている方を太字にしている。以降は、それぞれの表と可視化した情報を基に既存研究との精度比較を行った結果を示す。

はじめにソースドメインを T1 にしたモデルについて考察。表 1 の結果より、このモデルでは、T1ce の Dice、HD 共に既存手法を上回ることができなかったが、T2 は Dice が 57.93%、DH は 4.26mm 既存手法を上回り、大幅な精度向上を達成した。flair に関しても、Dice は 23.68%、DH は 0.74mm 既存手法を上回ることができ、精度の向上を確認した。この推論結果の 1 つを可視化した図 12 を見ると、T1ce は既存手法と比べて予測している場所と位置が既存研究と比べて大きくかけ離れていることが分かる。一方で、T2、Flair に関しては、既存研究よりも正解ラベルに近い位置をセグメントできていることが分かる。

続いて、T2 をソースドメインとして学習したモデルについてである。表 2 の結果より、T1 は Dice と DH がそれぞれ 1.25%、0.65mm の改善、T1ce は 3.77%、0.11mm の改善、flair は 0.11%、0.28mm の改善と、T1 をソースドメインにした時と比べて大幅に上昇したものはないが、すべてのカテゴリで精度の向上が確認できた。この結果を可視化したものが図 13 である。数値でも確認した通り、大幅な改善ではなかったため視覚的に分かりづらい点があるが、flair に注目すると既存手法よりも近い位置をセグメントしているこ

表 1 T1 を用いた学習による推論の数値結果

Source Domain: T1						
評価指標	Dice(%) ↑			HD(mm) ↓		
モダリティ	T2	T1ce	flair	T2	T1ce	flair
既存手法	23.38	9.34	12.24	9.02	8.91	9.2
提案手法	2.68	67.27	35.92	10.31	4.05	8.46

表 2 表 1 における評価指標ごとの平均値

Source Domain: T1		
評価指標	Average Dice(%) ↑	Average HD(mm) ↓
既存手法	14.99	9.04
提案手法	35.29	7.81

表 3 T2 を用いた学習による推論の数値結果

Source domain: T2						
評価指標	Dice(%) ↑			HD(mm) ↓		
モダリティ	T1	T1ce	flair	T1	T1ce	flair
既存手法	45.56	28.79	0.31	7.61	8.51	9.31
提案手法	46.81	32.56	0.42	6.96	8.4	9.03

とから、精度が改善していることが確認できる。

表 4 表 3 の評価指標ごとの平均値

Source Domain: T2		
評価指標	Average Dice(%) ↑	Average HD(mm) ↓
既存手法	24.89	8.48
提案手法	26.59	8.13

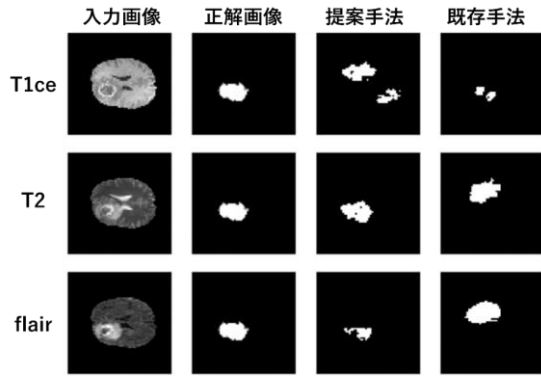


図 13 表 1 の推論結果を可視化

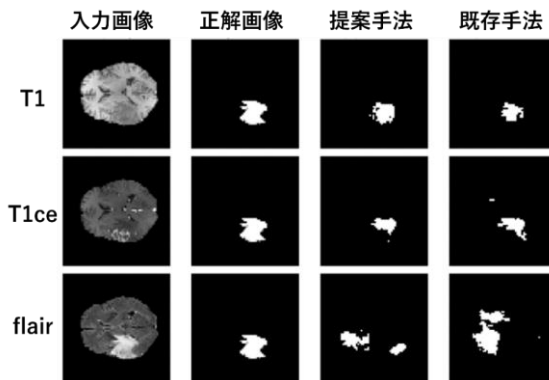


図 14 表 2 の推論結果を可視化

5. まとめと今後の課題

5.1 まとめ

本稿では、DG における異種医用画像間のセグメンテーションに有効なモデルとして、Ensemble U-Net with Dual-Normalization を提案した。この手法により、従来の研究で問題となったソースドメインの多様性に関する問題を解決しつつ、既存手法よりも高い精度を達成することができる。Dice と Harsdorf Distance を用いた定量的評価では、既存手法の精度を Dice は最大で 57.93 ポイント、Harsdorf Distance では最大で 4.26 ポイントの大幅な改善が実現できることを示した。

上記の研究結果より、本研究では以下の 3 点について当該分野に寄与するものと考えられる。

- 単一ドメインのセグメンテーションにおいて多様性を確保しつつ推論ができるフレームワークを提案した
- Domain Generalization において有効なモデルの提案をした
- 既存手法を上回るモデルの提案をした

5.2 今後の課題

本研究では、Ensemble U-Net に Dual-Normalization を適用することで、既存手法よりも優れた精度を達成したが、以下の課題が残った。

- ソースドメインによってターゲットドメインごとの推論精度への影響が顕著に出てしまう

- 色調変換の時に使用する関数の性質上、 D_{SS} と D_{SD} の中間にあたるような画像を作り出せず、ソースドメインの多様性を欠いてしまった

今後は、DG においてソースドメインの多様性を確保できるような変換手法の提案や、ソースドメインに縛られない学習が可能になるような学習方法の提案などを行う予定である。

参考文献

- [1] 厚生労働省, 第 12 回 保健医療分野 AI 開発促進コンソーシアム資料, https://www.mhlw.go.jp/stf/newpage_25103.html, Accessed on 18/01/2024
- [2] Li, Da and Yang, Yongxin and Song, Yi-Zhe and Hospedales, Timothy M. "Deeper, broader and artier domain generalization", Proceedings of the IEEE international conference on computer vision, pp. 5542-5550, 2019
- [3] Yingjun Du, Jun Xu, Huan Xiong, Qiang Qiu, Xiantong Zhen, Cees GM Snoek, and Ling Shao. Learning to learn with variational information bottleneck for domain generalization. In Computer Vision-ECCV 2020: 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part X 16, pages 200-216. Springer, 2020.
- [4] Mattia Segu, Alessio Tonioni, and Federico Tombari. Batch normalization embeddings for deep domain generalization. Pattern Recognition, 135:109115, 2023.
- [5] Ziqi Zhou, Lei Qi, Xin Yang, Dong Ni, and Yinghuan Shi. Generalizable cross-modality medical image segmentation via style augmentation and dual normalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 20856-20865, 2022.
- [6] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning, pp. 448-456. pmlr, 2015.
- [7] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, pp. 234-241. Springer, 2015.
- [8] Zhengxin Zhang, Qingjie Liu, and Yunhong Wang. Road extraction by deep residual u-net. IEEE Geoscience and Remote Sensing Letters, Vol. 15, No. 5, pp. 749-753, 2018.
- [9] Ozgün C, ipek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19, pp. 424-432. Springer, 2016.
- [10] Xingchao Peng, Ben Usman, Neela Kaushik, Judy Hoffman, Dequan Wang, and Kate Saenko. Visda: The visual domain adaptation challenge. arXiv preprint arXiv:1710.06924, 2017.
- [11] Woong-Gi Chang, Tackgeun You, Seonguk Seo, Suha Kwak, and Bohyung Han. Domain-specific batch normalization for unsupervised domain adaptation. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7346-7354, 2019.
- [12] Bjoern H. Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, Levente Lenczi, Elizabeth Gerstner, Marc-André Weber, Tal Arbel, Brian B. Avants, Nicholas Ayache, Patricia Buendia, D. Louis Collins, Nicolas Cordier, Jason J. Corso, Antonio Criminisi, Tilak Das, Hervé Delingette, C, a gatay Demiralp,
- [13] Christopher R. Durst, Michel Dojat, Senan Doyle, Joana Festa, Florence Forbes, Ezequiel Geremia, Ben Glocker, Polina Golland, Xiaotao Guo, Andac Hamamci Khan M. Iftekharuddin, Raj Jena, Nigel M. John, Ender Konukoglu, Danial Lashkari, Jos'e António Mariz, Raphael Meier, Sérgio Pereira, Doina Precup, Stephen J. Price, Tammy Riklin Raviv, Syed M. S. Reza, Michael Ryan, Duygu Sarikaya, Lawrence Schwartz, Hoo-Chang Shin, Jamie Shotton, Carlos A. Silva, Nuno Sousa, Nagesh K. Subbanna, Gabor Szekely, Thomas J. Taylor, Owen M. Thomas, Nicholas J. Tustison, Gozde Unal, Flor Vasseur, Max Wintermark, Dong Hye Ye, Liang Zhao, Binsheng Zhao, Darko Zikic, Marcel Prastawa, Mauricio Reyes, and Koen Van Leemput. The multimodal brain tumor image segmentation benchmark (brats). IEEE Transactions on Medical Imaging, Vol. 34, No. 10, pp. 1993-2024, 2015.