

グレースケール, カラー画像を用いた舌領域抽出の精度向上手法の提案

久家 健斗[†] 荻原 宏是[†] 藤田 悠介[‡] 瀬川 誠^{††} 浜本 義彦[‡] 飯塚 徳男^{‡†}

徳山工業高等専門学校専攻科情報電子工学専攻[†] 山口大学大学院創成科学研究科[†]

山口大学医学部附属病院^{††}

山口総合健診センター[‡]

1. はじめに

舌診とは、舌の粘膜の色、形状、動態、舌粘膜上にある舌苔を診ることで、患者の状態を獲得する方法である。しかし、この舌診は未だ医師の経験によるものが多く、仕組みも過程も科学的に解明されていない。そのため、定量的な評価の確立が求められている[1]。

これまでに、舌診の診断支援システムの研究が行われており、TIAS(Tongue Image Analyze System)[2]や TongueDx[3]の開発がされている。これらのシステムは、舌画像から舌領域を抽出し、舌のみの画像から診断、解析を行っている。また、先行研究では、舌、唇とそれ以外の領域に分割する多クラス U-Net を用いた領域分割法において、舌画像から舌領域の抽出を行っているが、色が類似している舌と唇の領域において誤分類が発生している[4]。舌領域の抽出は診断の精度に大きく影響するため、高精度な領域分割が必要とされる。

本研究では、U-Net を用いた舌、唇とそれ以外の領域に分けた 3 クラス分類モデルを設計し、画像の入力をグレースケール画像のみ、カラー画像のみとその両方を対象とした 3 種類実験を行い、領域抽出精度を比較した。

2. 研究の概要

本研究は、撮影された舌画像を入力データとし、領域分割用深層学習モデルとして U-Net を用いる。

2.1. U-Net

U-Net とは生物医学のために開発されたセマンティックセグメンテーション用のモデルである。図 1 に U-Net の構造を示す。図の左側のエンコーダ部分で入力画像を畳み込み、画像の特徴を抽出する。右側のデコーダ部分でエンコーダによって抽出された特徴を受け取り、逆の畳み込みを行う

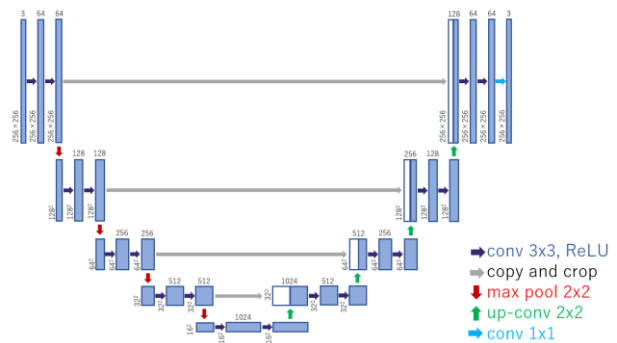


図 1 U-Net の構造

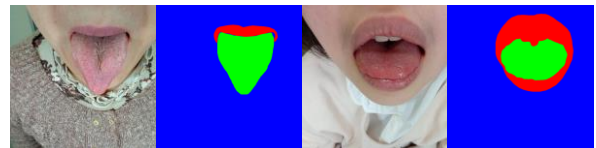


図 2 マスク画像例(左側が原画像, 右側がマスク画像)

ことで、各ピクセルの確率マップを出力する。各層において、エンコーダとデコーダをスキップ接続することで、物体の位置情報をとらえている[5]。

2.2. 実験データとクラス定義

本実験で使用する舌画像は、山口大学医学部附属病院で撮影された画像、125 枚である。舌画像は、すべて同一条件で撮影されている。

U-Net で学習を行うには、領域分割されたマスク画像が必要となる。そのため、舌画像 125 枚に対して、舌領域と唇領域、それ以外の領域に分けたマスク画像を作成した。作成した画像の例を図 2 に示す。

3. 提案手法

3.1. グレースケール, カラー画像を用いた実験

U-Net を用いた舌領域抽出の流れについて説明する。125 枚の舌画像を 256×256 画素のサイズにリサイズし、訓練データ 90 枚、検証データ 20 枚、テストデータを 15 枚ランダムに選び、学習を行う。画像の入力をグレースケール画像のみ、カラー画像のみ、その両方の 3 種類をそれぞれ学習し、判別の違いを観察する。

画像の入力の際、カラーの訓練画像に 1 枚当た

Proposal for Improving the Accuracy of Tongue Area Extraction Using Grayscale and Color Images

[†] Kento KUGE, National Institute of Technology, Tokuyama College

[†] Hiroyuki OGIHARA

[‡] Yusuke HUIJITA, Yamaguchi University

[‡] Yoshihiko HAMAMOTO

^{††} Makoto SEGAWA, Yamaguchi University Hospital

^{‡†} Norio IIZUKA, Yamaguchi Health Examination Center

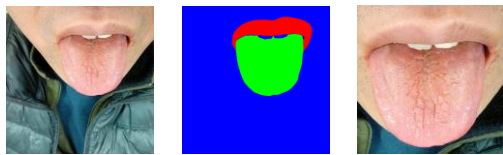


図 3 切り抜き画像例(左側が原画像, 中央がマスク画像, 右側が切り抜き画像)

り 10 枚のデータ拡張を行い, 900 枚生成する. データ拡張のパラメータは, 回転 10 度, 水平移動 5%, 垂直移動 5%, シアー変換 5%, 左右反転ありである. グレースケール画像を入力する場合は, データ拡張されたカラー画像をグレースケールに変換している. グレースケール, カラー両方の入力の際は 5 枚ずつ変換し, 比率を 1:1 としている. 学習時のバッチサイズを 2, エポック数を 100, 最適化アルゴリズムを Adam, 学習率を 0.001 とする.

3.2. 切り抜き画像を用いた実験

さらなる精度向上のために, 切り抜き画像を用いた舌領域抽出を行う. まず, 図 3 のような唇と舌の部分を取り抜いた画像を用意し, 学習を行う. 学習の条件は, グレースケール, カラー両方の入力した時と同じである. 次に, 3.1 で作成した, グレースケール, カラー画像での学習モデルを使用し, 出力された結果をもとに, 舌と唇部分のみを取りぬく. 最後に, 切り抜いた画像をテスト画像とし, 切り抜き画像を用いた学習モデルに入力し, 結果を出力させる.

3.3. 評価方法

モデルから出力された舌領域抽出の評価には IoU(Intersection over Union)を用いる. IoU とは物体検出での評価指標に使われ, 画像の重なりを示している. モデルが予測した舌領域と正解の舌領域との重なりが大きいほど値が 1 に近づく. 3 種類の試行を 10 回行い, 全データの IoU の平均値を求める.

4. 結果

結果を表 1 と, 図 4 に示す. 表 1 よりグレースケールとカラーの両方を入力した場合が, 単体で入力した場合よりも, 精度が良いことが確認できた. また, 図 3 より, グレースケールでの学習は形を中心に判別し, カラーでの学習は, 色を中心に判別していることが確認できる. 両方を使った学習では, 色と形をどちらも重視するようになったことが要因だと考えられる.

切り抜き画像を使った学習では, IoU が 0.967 と高精度の結果が出た. 図 4 の中央の画像のように, 少し舌を誤分類しているが, ほとんどの舌画像において正確に舌部分のみを抽出できていることが確認できた. 図 4, 5 の右 2 つのように唇が舌を

表 1 画像の入力と IoU における比較

画像の入力	IoU
グレースケールのみ	0.902
カラーのみ	0.941
両方	0.952
両方 + 切り抜き	0.967

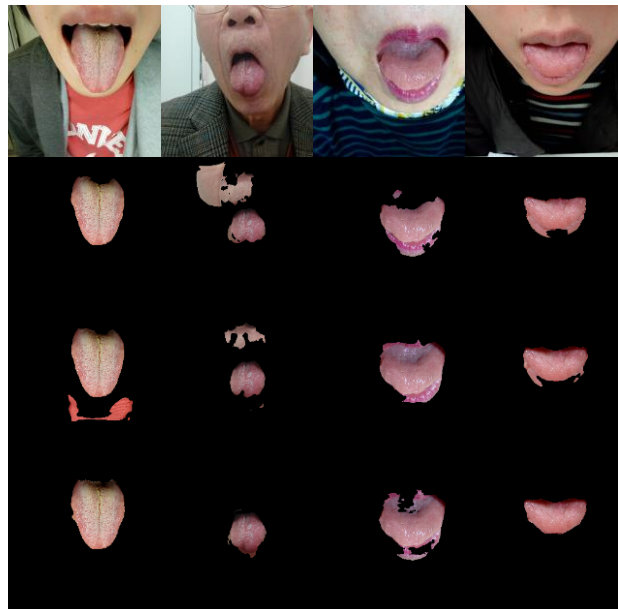


図 4 出力結果例(上から原画像, グレースケールのみ, カラーのみ, 両方)

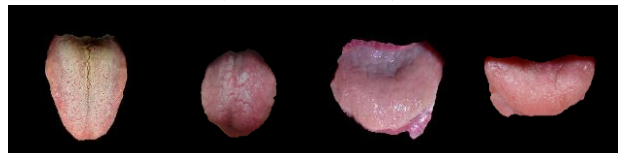


図 5 切り抜き画像を用いた学習結果例

囲んでいる画像は, 正確に抽出できていない. さらなる精度向上には, このような画像データを増やすことが必要だと考えられる.

参考文献

- [1] 安振宇, 呉韜, 中田充, 葛崎偉. “舌画像の深層学習による臓腑健康状態の自動診断.” *Journal of Signal Processing* 27.5, pp133-143.2023
- [2] T. Nakaguchi, S. Namiki, N. Tsumura, Y. Kasahara, T. Namiki, Y. Miyake, K. Terasawa, “Regional Image Analysis of the Tongue Color Spectrum”, *International Journal of Computer Assisted Radiology and Surgery*, Vol. 6, No. 1, pp.143-152, 2011.
- [3] 劉維尼, 椎尾一郎, “TongueDx: スマートフォンのための舌診システム”, 情報処理学会インタラクシオン, 2014.
- [4] 小倉成志朗, 藤田悠介, 荻原宏是, 瀬川誠, 浜本義彦, 飯塚徳男. “多クラス U-Net を用いた舌診断のための領域分割の高精度化”. 動的画像処理実利用化ワークショップ DIA2022. 2023
- [5] O.Ronneberger, P. Fischer, T. Brox, “U-net: Convolution networks for biomedical image segmentation” *International Conference on Medical image computing and computer-assisted intervention*, Springer, pp.234-241. 2015