

深層学習モデル TransUNet を用いた  
歯科パノラマ X 線画像のセグメンテーション

Segmentation of Dental Panoramic X-ray Images  
Using the Deep Learning Model TransUNet

本田 剛<sup>†</sup>, 鳥井 浩平<sup>†</sup>, 西村 良太<sup>†</sup>, 吉田 稔<sup>†</sup>, 松本 和幸<sup>†</sup>

Tsuyoshi Honda, Kohei Torii, Ryota Nishimura, Minoru Yoshida, Kazuyuki Matsumoto

**あらまし** 歯科パノラマ X 線画像は歯の診療に使われる画像である。歯科パノラマの照合には専門的な知識が必要であり、歯科医師の負担軽減を目的として歯の自動診断システムの研究が行われてきた。その中で、深層学習技術を用いて歯の識別を行う状態分類モデルが開発されている。本研究では状態分類モデルの入力として必要な歯の輪郭情報を作成するため、深層学習技術を用いて歯科パノラマ X 線画像のセグメンテーションを行う。実験では、歯科クリニックで撮影、ラベル付けされたデータを取り扱う。セグメンテーションモデルには、近年医療セグメンテーションタスクで成功を収めている TransUNet モデルを用いて実験を行った。

**キーワード** パノラマ X 線画像, セグメンテーション, 深層学習, TransUNet

## 1. はじめに

歯科パノラマ X 線画像は歯の診療に使われる画像である。歯科パノラマの照合には専門的な知識が必要であり、歯科医師の負担軽減を目的として歯の自動診断システムの研究が行われてきた。他の医療関係の分野では、すでに実用化されている自動診断システムの一つに米 IBM 社の Watson for Genomics (WfG) [1]がある。このシステムは、がんの発症や進行に関与する遺伝子のデータ、抗がん剤の候補を提示することが出来る。このような自動診断システムが歯科画像に対しても研究されている。歯科画像の自動診断システムには、入力として歯の輪郭情報が不可欠である。

本研究では、状態分類モデルの入力として必要な歯の輪郭情報を作成するため、深層学習技術を用いて歯科パノラマ X 線画像のセグメンテーションを行う。

## 2. 提案手法

本研究では、歯科パノラマ X 線画像と歯科診療記録から TransUNet [2]モデルによって学習を行い、歯科画像にセマンティックセグメンテーション処理を行う。歯科診療記録から自作のマスク画像を生成し、歯の領域を一つずつ切り取り、それぞれの画像に対して学習を行うシステムを構築する。

### 2.1 歯科診療記録

歯科診療記録は共同研究先の歯科クリニックで作成されたデータを使用する。表 1 はデータの内容を表したものである。



図 1 歯科パノラマ X 線画像

このデータでは、歯の座標がポリゴン形式で表されており、それぞれの座標をつなぎ合わせることでマスク画像の領域を設定した。

図 2 はマスク画像を生成する順序を表した画像である。

表 1 歯科診療記録  
パラメータ

名称	パラメータ
FILENAME	YCS9601000000.dcm
HEIGHT	1400
WIDTH	2469
POINTS	[ 605.7692307692307, 831.3076923076923 , ... [ 597.8260869565216, 812.3188405797101 ]

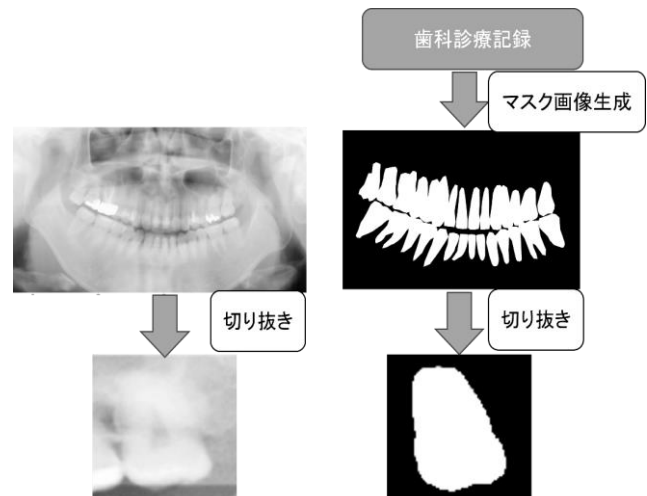


図 2 マスク画像生成の順序

### 3. 実験

実験では、共同研究先の歯科クリニックで撮影されたデジタルX線画像を用いる。入力の際、1872枚の画像を51696枚に分割した。TransUNetの学習用データとして、分割した画像と、生成したマスク画像を使用する。その中の70%を訓練用データとし、残りの30%を検証用データとする。この実験で得られた画像を、入力画像とマスク画像とともに図3に示す。

#### 3.1 実験結果

TransUNetによって得られたセグメンテーション結果は図3のPredicted Maskである。ここで、Input Imageは入力画像、Label Maskは生成したマスク画像である。また、Predicted MaskとLabel MaskをDiceLossとIoUスコアによって評価した。表2はその評価数値、図4は学習の推移を示している。

表2 セグメンテーションの評価

DiceLoss	IoU
0.06	0.89

### 4. 結果と考察

今回のセグメンテーション結果は訓練データで評価を行った際、Lossの値は学習が進むごとに減少しているが、検証データの場合は一定の場所からLossの値が横ばいになっている。これは、マスク画像を生成する際にポリゴンデータを使用したことが原因だと考えた。図3のLabel MaskとPredicted Maskを比較するとLabel MaskよりもPredicted Maskの画像がより滑らかな形状をしていることから、マスク画像よりも精密に学習してしまい、セグメンテーション精度が一定の値よりも向上しなかったと考えられる。

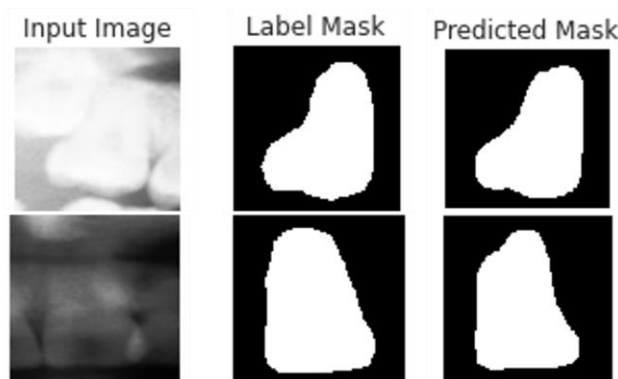


図3 セグメンテーション結果

### 5. おわりに

本稿では、歯科クリニックで撮影された歯科パノラマX線画像から、歯の領域を分割し、深層学習モデルを用いてそれぞれの歯にセグメンテーション処理を施す実験を行った。近年、医療セグメンテーションタスクで成功を収めているTransUNetモデルによってセグメンテーションを行った結果、マスク画像よりも高精度な画像が出力された。今後の課題は、より精度の高いマスク画像の生成手法を提案するか、もしくはすでにマスク画像が存在するデータセットを使用する事である。新たなマスク画像を入手することが出来たら、再度実験と評価を行いたい。

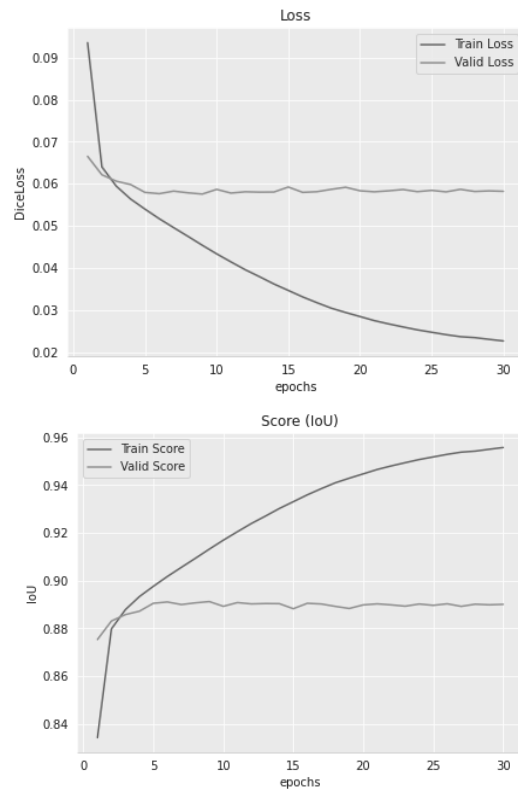


図4 評価結果

#### 参考文献

- [1] D. Ferrucci. Introduction to "this is watson". IBM Journal of Research and Development, Vol. 56, No. 3.4, p. 1, 2012.
- [2] Jieneng Chen, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, YanWang, Le Lu, Alan L. Yuille, Yuyin Zhou. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. CoRR. Mon, 12 Apr 2021.
- [3] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. Attention Is All You Need. 12 Jun 2017.
- [4] Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. 18 May 2015.