

## 歯科パノラマ X 線画像を用いた個人識別システムに対する Instance Discrimination 手法の有用性

### The Usefulness of Instance Discrimination Method for Individual Identification Using Panoramic Radiographs

鳥井 浩平\* 高野 栄之\* 桃田 幸弘\* 前田 直樹\* 水頭 英樹\*

北 研二† 誉田 栄一‡ 小林 健一郎‡ 川村 昂輝\* 寺田 賢治\*

Kohei Torii Hideyuki Takano Yukihiko Momota Naoki Maeda Hideki Suito

Kenji Kita Eiichi Honda Kenichiro Kobayashi Koki Kawamura Kenji Terada

#### 1. はじめに

南海トラフ巨大地震では最大で約 298,000 人の死者が出ることが予想されている[1]。遺体は医師や歯科医の協力のもと、警察によって身元確認が行われる。身元確認では歯科所見が主に用いられ、歯科医による歯科所見の照合が行われるが、現状ではほぼ手作業であるため多大な労力と時間を要することが課題となっている。

より迅速な身元確認のために、鳥井らは過去に開発した歯科パノラマ X 線画像から歯科所見を自動的に抽出して電子カルテ入力作業を支援する AI システムを応用することで、デンタルチャート作成作業などの自動化を目指した研究を行っている[2]。しかし、この AI システムを構築するためには、専門医によって歯科所見情報が付与された膨大な数のパノラマ X 線画像データセットを必要とすることが課題となっている。

本研究では歯科所見情報が付与されていない歯科パノラマ X 線画像のみを用いて Instance Discrimination 手法に基づく深層学習モデルの学習を行い、このモデルが生成する特徴ベクトルを用いて個人識別を行うことを目標とする。本論文では、生成された特徴ベクトルに対する次元削減と非階層クラスタリング、中間層における特徴マップの可視化を行い、個人識別に対する有用性について述べる。

#### 2. 手法

Instance Discrimination は画像を個々のインスタンスに分類するタスクである。100 枚の画像がある場合は 100 種類のカテゴリが存在すると仮定し、各画像がどのカテゴリに属するかを識別させる。Instance Discrimination は畳み込みニューラルネットワーク (CNN) の事前学習手法として有用であることが知られており、ラベルが付与されていない大量の画像を CNN に学習させることで、画像を下流タスクに応用可能な低次元の特徴ベクトル (埋め込み表現) に変換するモデルを構築することができる[3][4]。

本研究で提案するモデルを図 1 に示す。入力画像は歯科パノラマ X 線画像の一般的なアスペクト比を考慮し、高さ 256 ピクセル、幅 512 ピクセルの 8 ビット RGB 画像とする。中間層は InceptionV3 のアーキテクチャを基盤とし、最後の Inception モジュールの出力である 2048 チャンネルの特徴マップに対して Global Average Pooling を適用することで得られる 2048 次元の特徴ベクトルに 128 次元の全結合層を

\*徳島大学 Tokushima University

†株式会社 MILAI Technologies MILAI Technologies Co., Ltd.

‡こばやし歯科クリニック Kobayashi Dental Clinic

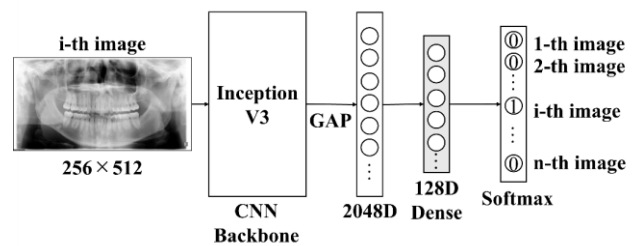


図 1 モデルの概要

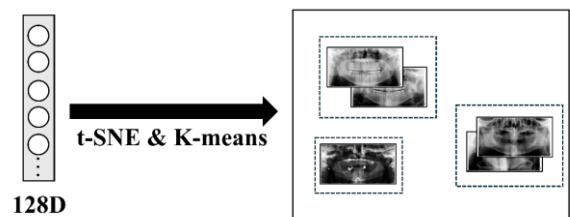


図 2 次元削減と非階層クラスタリング

接続する[5]。この全結合層における 128 次元のベクトルが本研究の対象とする特徴ベクトルである。出力層は活性化関数をソフトマックス関数とした全結合層とする。学習用画像の枚数が  $n$  枚のとき、各画像の学習時に与えられる出力層の正解データは、 $i$  番目の画像であれば  $i$  番目の要素が 1、それ以外の要素が 0 の One-hot ベクトルとする。損失関数は多クラス交差エントロピーとする。

128 次元の特徴ベクトルに対する次元削減は t-SNE を用いる[6]。t-SNE を用いて 128 次元の特徴ベクトルを 2 次元に次元削減した後、K-means による非階層クラスタリングを行う[7]。クラスタリング結果に基づいてテスト用画像のグループ分けを行い、各グループにおいて共通する画像特徴や歯科所見などを確認する (図 2)。

また、入力層から数えて 1 番目の Inception モジュール (mixed1)、5 番目の Inception モジュール (mixed5)、10 番目の Inception モジュール (mixed10) が出力する特徴マップを対象にそれぞれ可視化を行う。可視化の方法として、まず特徴マップをチャンネル方向に加算し、各値が 0 から 255 までの値となるように正規化を行う。次に中心化を行い、特徴マップを原画像と同じサイズ (256×512) にリサイズする。ただし、中心化によって 0 未満になる値は 0 に置換する。最後に原画像と特徴マップを合成することで、モデルが画像上のどこに注目しているかを示すヒートマップ画像を作成する。

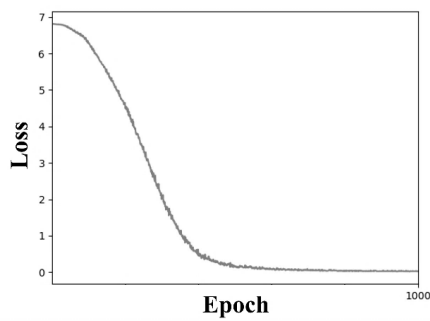


図 3 損失の推移

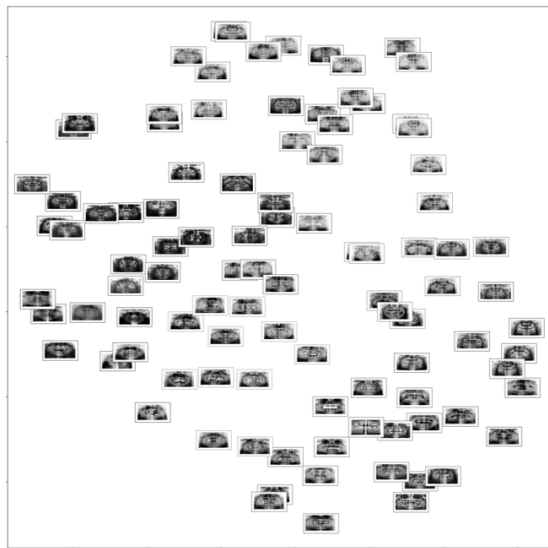


図 4 テスト用画像の特徴ベクトルの分布

### 3. 実験

パノラマ X 画像は Panetta らの Tufts Dental Database から 1,000 枚収集し、ランダムに選択した 900 枚を学習用、残り 100 枚をテスト用とした[8]。また、学習時には各画像に対して、水平方向反転、回転、平行移動、拡大縮小、せん断、ガウシアンぼかし、ガウシアンノイズ付加をランダムに適用し、明るさとコントラストをランダムに変化させた。

パラメータの最適化には学習率 0.01、バッチサイズ 32 の勾配降下法を用いた。学習回数は 1,000 回とした。また、t-SNE における Perplexity は 30 とし、K-means におけるクラスタ数は 10 とした。

### 4. 結果と考察

学習曲線は図 3 のようになり、損失が収束したことを確認した。次元削減後のテスト用画像の特徴ベクトルの分布にしたがって画像をプロットした結果は図 4 のようになった。

K-means により分割されたテスト用画像のグループの中で、歯が著しく欠損している画像が集まったグループがあった(図 5)。また、クラウンやインプラントなどの修復歯が多数確認できる画像が集まったグループや、乳歯を含む画像が集まったグループがあることも確認した。これらの結果から、大域的な歯科所見の情報を反映した特徴ベクトルが得られたと考えられる。特徴マップを可視化した画像からも、モデルが修復歯や歯の領域に注目していること

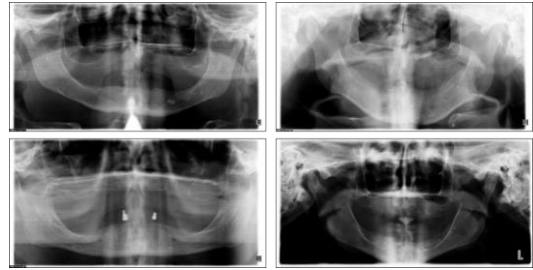


図 5 歯が著しく欠損している画像が集まったグループの画像

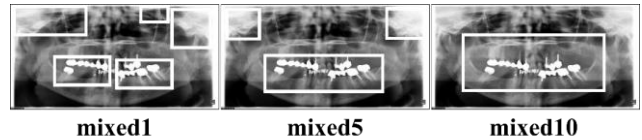


図 6 モデルの注目領域

を確認した(図 6)。しかし、図 4 と図 6 に示す通り、撮影機種に依存した画像特徴が特徴ベクトルに強く反映されていることも確認した。また、Instance Discrimination 手法を用いると、同じ歯科所見情報を持つ画像であっても異なるカテゴリの画像として学習してしまうため、学習が望む方向に進まず、機種依存性をより強めるような学習となった可能性も考えられる。より歯科所見のみに着目した特徴ベクトルを得るには、ドメイン汎化手法などを新たに導入することで機種依存性を取り除く必要があると考えられる。

### 5. まとめ

Instance Discrimination 手法より、明示的なラベルデータを必要とせずに、個人識別に有用な歯科パノラマ X 線画像の特徴ベクトルを生成できる可能性が示唆された。今後は機種依存性を取り除く工夫を導入し、より大規模な実データを用いて実験および定量的評価を行う。

#### 参考文献

- [1] 内閣府, “南海トラフ巨大地震 最大クラス地震における被害想定について”, 内閣府 防災情報のページ (2025). [https://www.bousai.go.jp/jishin/nankai/taisaku\\_wg\\_02/pdf/saidai\\_01.pdf](https://www.bousai.go.jp/jishin/nankai/taisaku_wg_02/pdf/saidai_01.pdf)
- [2] 鳥井 浩平, 菅田 栄一, 北 研二, “AI を用いたパノラマ X 線画像からのカルテ入力支援システムの開発”, 歯科放射線, Vol.62, No.1 (2022).
- [3] Wu, Z., Xiong, Y., Yu, S. X., Lin, D., “Unsupervised Feature Learning via Non-Parametric Instance Discrimination,” IEEE Conference on Computer Vision and Pattern Recognition (2018).
- [4] Tao, Y., Takagi, K., Nakata, K., “Clustering-friendly representation learning via instance discrimination and feature decorrelation,” International Conference on Learning Representations (2021).
- [5] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., “Rethinking the Inception Architecture for Computer Vision,” IEEE Conference on Computer Vision and Pattern Recognition (2016).
- [6] Van der Maaten, L., Hinton, G., “Visualizing data using t-SNE,” Journal of Machine Learning Research,” Vol.9, No.11 (2008).
- [7] Lloyd, S. P., “Least squares quantization in PCM”, Technical Report RR-5497, Bell Lab (1957).
- [8] Panetta, K., Rajendran, R., Ramesh, A., Rao, S. P., Agaian, S., “Tufts Dental Database: A Multimodal Panoramic X-Ray Dataset for Benchmarking Diagnostic Systems,” IEEE Journal of Biomedical and Health Informatics, Vol.26, No.4 (2021).