

距離学習とアンサンブル学習を併用した歯根吸収の判別法

A Method for Discrimination of Root Resorption Using Metric Learning and Ensemble Learning

田村 晃聖*, 神谷 亨*, 小田 昌史**, 森本 泰宏**

*九州工業大学, **九州歯科大学

Kosei Tamura*, Tohru Kamiya*, Masafumi Oda**, Yasuhiro Morimoto***

*Kyushu Institute of Technology, **Kyushu Dental University

E-mail address: tamura.kosei905@mail.kyutech.jp

1. はじめに

歯根吸収は主に、外圧による歯の緩衝、副甲状腺ホルモンの過剰分泌、細菌感染によるものなどがある。歯科矯正治療の中で起こることも多く、矯正治療中は定期的な検査が必要とされている。歯根吸収は歯科矯正治療を受けていない人も約 5~10%罹っているとされている [1]。歯根吸収は進行具合により治療の結果に影響を与えるため、早期発見・早期治療が必要である。早期発見のための歯科器具として歯科用 CT 撮影装置とパノラマエックス線撮影装置がある。パノラマエックス線画像から歯根吸収を診断することは困難であるため [2]、歯根吸収の判別は一般に歯科用 CT 撮影装置を必要とする。診断における現在の問題点として歯科用 CT 撮影装置の所有率は低いことが挙げられる。そのため、パノラマエックス線画像から歯根吸収を判別する必要がある。

一方、深層学習の分野では人間の精度を超える検出能を有する AI が話題になっている。特に異常検知の分野では、少量の画像データを用いた実験で深層距離学習を用いた手法が高い精度を得ている場合もある。しかし、一般画像分類ではアンサンブル学習という手法も広く用いられている。そこで本稿では、この二つの手法を併用することにより、少数のデータを用いた歯根吸収の自動分類のための画像診断支援法を提案し、実画像を用いた識別結果について述べる。

2. 解析法

全体の流れは図 1 のとおりである。

まず、前処理ではパノラマエックス線画像から関心領域の抽出を行う。その後、画像のサイズを変更した後、画像の輝度値の標準化を行う。次に、特徴量抽出を行う。これにより、前処理で得られた画像を数値データに変換する。さらにアンサンブル学習にこれらの数値データを与え、それぞれの弱学習器による予測を行う。最後に予測結果をもとに最終的なネットワークの歯根吸収の予測結果を得る。以下にその詳細を示す。

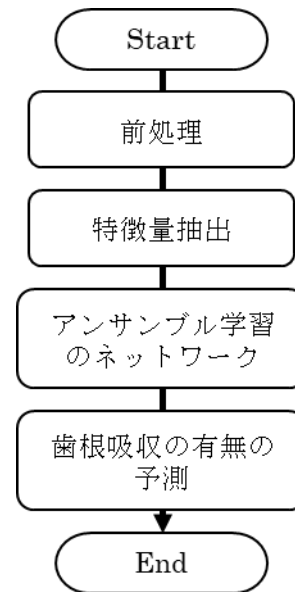


図 1 全体の流れ

2.1 前処理

歯根吸収は下顎の第三大臼歯と第二大臼歯の間で起こる症例が多い。本論文での処理対象は、下顎の第三大臼歯と第二大臼歯の間の歯根吸収とした。歯根吸収の有無の識別を行うため、パノラマエックス線画像から処理対象である下顎の第二大臼歯の領域を歯科医の指導の下、手動で抽出を行う。歯根吸収のある症例のパノラマエックス線画像の例を図 2 に、その画像を抽出して得られた左右の下顎第二大臼歯の画像を図 3 にそれぞれ示す。抽出後の画像枚数は 363 枚(うち歯根吸収あり 69 枚、歯根吸収なし 294 枚)である。

関心領域抽出後の画像をもとにラベル付けを行う。その後、画像幅 224、画像の高さ 224 に線形補間により変形する。変形後の画像は輝度値にばらつきがあるため、式(1)を用い標準化を行う。

$$\text{image}(x, y) = \frac{\text{image}(x, y) - \mu}{\sigma} \quad (1)$$

ただし、 μ 、 σ はそれぞれ画像の輝度値平均と輝度値の分散を表し、本論文では $\mu = 0$ 、 $\sigma = 1$ を用いた。



図2 パノラマエックス線画像

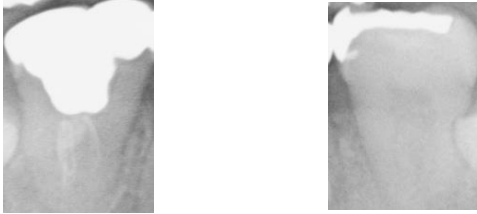


図3 関心領域抽出後の画像

2.2 特徴量抽出

前処理により得られた画像を特徴量抽出器として本稿で採用した VGG16 [3] に入力する. VGG16 は 16 層の畳み込みネットワークであり, その単純な構造から採用した. VGG16 は一般画像分類 ImageNet [4] を用い, 事前に学習したモデルを用いた. この VGG16 の出力は 1000 個の数値データである.

2.3 距離学習

得られた 1000 個の特徴量をネットワークに入力として与える. 表 1 に示すように, これらの 1000 個の数値データを入力とする全結合層 3 層と Softmax 層 1 層のネットワークを用いる. 出力は入力データに対する歯根吸収とする確率と歯根吸収ではないとする確率となる. ネットワークの全結合層の 1 層経たのち距離学習の損失関数を用いることによる特徴量の学習を行う.

距離学習の損失関数として Center Loss [5], Triplet Loss [6] を用いる.

Center Loss は式(2)で定義される損失関数である.

$$L_{Center Loss} = \|x_i - c_{y_i}\|_2^2 \quad (2)$$

ただし, x は特徴量, c_y は正解ラベル y である中心である. 学習時, c_i ($i = 1, 2$) は式(3)を用い, 更新を行う.

$$c_i = c_i - \alpha \times \Delta c_i \quad (3)$$

ただし,

$$\Delta c_i = \frac{\sum_{i=1}^m \delta(y_i = j) \cdot (c_j = x_i)}{1 + \sum_{i=1}^m \delta(y_i = j)} \quad (4)$$

であり, α は Center Loss における学習率を表す. 本稿では $\alpha = 0.001$ とした.

表 1 弱学習器の構造

Layer	Input size	Output size
Fully connected	1000	1000
Fully connected	1000	2
Softmax Layer	2	2

Triplet Loss は式(5)で定義される損失関数である.

$$L_{Triplet Loss} = \|(d_+, 1 - d_-)\|_2^2 \quad (5)$$

ただし,

$$d_+ = \frac{e^{\|Net(x) - Net(x^+)\|_2}}{e^{\|Net(x) - Net(x^+)\|_2} + e^{\|Net(x) - Net(x^-)\|_2}} \quad (6)$$

$$d_- = \frac{e^{\|Net(x) - Net(x^-)\|_2}}{e^{\|Net(x) - Net(x^+)\|_2} + e^{\|Net(x) - Net(x^-)\|_2}} \quad (7)$$

であり, x^+ , x^- はそれぞれ x と同じラベルを持つ特徴量, 異なるラベルを持つ特徴量を表す.

識別用の損失関数として Softmax Loss を用いる. Softmax Loss は多クラス分類に用いられる手法で最後の層を Softmax Layer とし, 交差エントロピー誤差を用いたものである. 交差エントロピー誤差は式(8)で定義される.

$$L_{Cross Entropy} = - \sum_{k=1}^K t_k \log y_k \quad (8)$$

ただし K はクラス数, t_k は正解ラベルを表す. また, Softmax 関数は式(9)で定義される.

$$y_k = \frac{\exp(z_k)}{\sum_{k'=1}^K \exp(z_{k'})} \quad (9)$$

2.3 アンサンブル学習

アンサンブル学習は複数の弱学習器で構成される. それぞれの弱学習器が異なる特徴量を学習することでより高い精度を発揮することが知られている.

本論文で用いたバギング法は, それぞれの弱学習器を元のデータセットから再度サンプリングした新しいデータセットをそれぞれの弱学習器を学習するために用い, 学習を行う. その後, テスト用のデータを用いることにより予測を行うことができる. そのため, リサンプリングを行う必要があり, リサンプリングの結果, 同じような特徴量を学習してしまう恐れがある. そのため, リサンプリングのサンプル生成器の評価が必要となる.

本稿ではリサンプリングを行わず, それぞれのネットワークの学習するデータセットは同じものを用い, 異なる特徴量の学習を行うことを可能にする. そのため, それぞれのネットワークの距離学習の損失関数を式(10)として定義する.

$$loss_i = \alpha \times L_{Triplet Loss} + (1 - \alpha) \times L_{Center Loss} \quad (10)$$

ただし,

$$\alpha = \frac{i - 1}{n - 1} \quad (11)$$

とする. ここで n は弱学習器の数, i は弱学習器の i ($1 \leq$

$i \leq n$) 番目であることを表す。

本論文では弱学習器の数を3 ($n = 3$)とし、3つのネットワークの出力結果をもとに多数決により最終的な出力を決定する。

2.4 最適化

本論文では最適化として確率的勾配降下法(Stochastic Gradient Descent, SGD)を用いる。なお、学習率0.001、モメンタム0.9とする。

3. 実験と結果

本稿では、少数のデータを用いた歯根吸収の自動分類のための深層学習による画像診断支援法を提案した。比較実験として次のようなモデルを用いた。

モデル1: アンサンブル学習なし, 距離学習なし

モデル2: アンサンブル学習あり, 距離学習なし

モデル3: アンサンブル学習なし, 距離学習の損失関数として Center Loss

モデル4: アンサンブル学習あり, 距離学習の損失関数としてすべてに Center Loss を用いたモデル

モデル5: アンサンブル学習なし, 距離学習の損失関数として Triplet Loss を用いたモデル

モデル6: アンサンブル学習あり, 距離学習の損失関数としてすべてに Triplet Loss を用いたモデル

モデル7: アンサンブル学習なし, 距離学習の損失関数として Center Loss + Triplet Loss としたモデル

モデル8: アンサンブル学習あり, 距離学習の損失関数として Center Loss + Triplet Loss としたモデル

モデル9: 提案手法

実験は一つの画像をテストデータとし、それ以外を学習データとする leave-one 検証を実施した。評価は精度、歯根吸収ありの画像に対し歯根吸収ありと正しく正答する確率である真陽性率 TPR(True Positive Rate)と歯根吸収なしの画像に対し歯根吸収ありと誤答する確率である偽陽性率 FPR(False Positive Rate)を用いた。

実験結果は表2に示す。

表2 実験結果

モデル	精度	真陽性率	偽陽性率
1	0.722	0.130	0.139
2	0.777	0.072	0.058
3	0.416	0.536	0.112
4	0.410	0.420	0.592
5	0.807	0.029	0.010
6	0.807	0.014	0.007
7	0.667	0.101	0.201
8	0.672	0.101	0.194
9	0.697	0.091	0.166

本実験より Center Loss を距離学習の損失関数として用いたモデルでは真陽性率が他のモデルに比べ高い結果となった。そのため、少数のデータに対し対応できていることが分かる。一方、それ以外のモデルでは真陽性率が低い結果となった。原因としてデータセットのデータ数に偏りが因であると考えられる。本論文では正常画像(歯根吸収なし)の画像枚数に比べ、異常画像(歯根吸収あり)の画像枚数が多いため、偽陽性率が低く、真陽性率が低い結果となった。また、真陽性率が高い結果となったモデルは距離学習の損失関数として Center Loss を用いたモデルであることも分かった。そのため、Center Loss を用いることにより歯根吸収ありと歯根吸収なしの画像を分離することができることも確認できた。

距離学習の損失関数が同じアンサンブル学習なしとアンサンブル学習ありのモデルを比較することにより、アンサンブル学習を用いた結果、真陽性率が下がり、偽陽性率が下がる結果となった。アンサンブル学習を行うことにより歯根吸収なしと識別することが増加することが分かる。弱学習器に与えるデータセットを元のデータセットと同じものを用いているため、歯根吸収なしのデータが豊富に含まれているためだと考えられる。

4. おわりに

本論文ではアンサンブル学習と距離学習を併用した深層学習により、パノラマX線画像からの歯根吸収の有無を判定する画像解析法を提案し、実データによる性能評価を行った結果、精度0.697、真陽性率0.091、偽陽性率0.166を得た。今後は、さらなる精度向上を図るため、アンサンブル学習の弱学習機のネットワーク構造の改良を検討する必要がある。

参考文献

- [1] Alberto Consolaro, Laurindo Zanco Furquim, "Extreme root resorption associated with induced tooth movement: A protocol for clinical management", Dental Press J Orthod, No. 19, Vol. 5, pp. 19-26, 2014.
- [2] Martin D. Levin, George Jong, "The Use of CBCT in the Diagnosis and Management of Root Resorption", 3D Imaging in Endodontics, pp. 131-143, 2016.
- [3] Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv preprint arXiv:1409.1556v6, 2015(access 2022/10/10).
- [4] Jia Deng, et al., "ImageNet: A large-scale hierarchical image database", IEEE Conference Computer Vision Pattern Recognition, pp.24-255, 2009
- [5] Bastian Leibe, et al., "A discriminative feature learning approach for deep face recognition", Computer Vision-

European Conference on Computer Vision 2016, Vol. 9911,
pp.499-515, 2016.

- [6] Elad Hoffer et al., “Deep Metric Learning using Triplet
Network”, arXiv preprint arXiv:1412.6622v4, 2018 (access
2022/10/10).