

所見文書の要約を用いた深層学習による矯正歯科治療の自動診断

大塚 琢生[†] 梶原 智之[†] 谷川 千尋[‡] 清水 優仁[‡] 長原 一^{*} 二宮 崇[†]

[†]愛媛大学大学院理工学研究科 [‡]大阪大学歯学部附属病院 ^{*}大阪大学データビリティフロンティア機構

1 はじめに

歯科医師は、豊富な知識と経験に基づき診断を実現している。経験の浅い歯科医師による問題の見落としを防ぎ、専門医の負担を軽減するために、本研究では自然言語処理の技術を用いて矯正歯科治療の自動診断に取り組む。

近年、様々な自然言語処理タスクにおいて深層学習モデルが高い性能を達成している。しかし、矯正歯科治療における自動診断の先行研究[1-2]は、Bag of Words (BoW) の特徴抽出および Support Vector Machine (SVM) の分類器という、従来の機械学習のアプローチを採用している。これは、入力となる矯正歯科治療の所見文書が以下の3つの特徴を持つために、深層学習モデルの単純な適用では扱いが難しいからである。

- ① 非常に長い文章 (平均 1,886 語)
- ② 不統一なスタイル (箇条書きなど)
- ③ 専門用語の頻繁な使用

特に課題①は、長距離の依存関係を捉えることの難しさ[3]やメモリ量の制約のために、矯正歯科治療における自動診断タスクへの深層学習モデルの適用を困難にしている。

本研究では、入力として所見文書の要約を用いることにより、上記の課題①および②を解決し、矯正歯科治療における自動診断の性能を改善する。先行研究[1-2]と同様に、大阪大学歯学部附属病院に所蔵されている970件の所見文書を用いて評価実験を行った結果、畳み込みニューラルネットワーク (CNN) が既存手法を上回る性能を達成した。また、所見文書の要約を用いることで、再帰ニューラルネットワーク (RNN) および自己注意ネットワーク (SAN) を含む全てのモデルの性能が改善できることを明らかにした。

Deep Learning-Based Automated Orthodontic Diagnosis from Findings Summaries

Takumi Otsuka[†] (otsuka@ai.cs.chime-u.ac.jp)

Tomoyuki Kajiwara[†] (kajiwara@cs.chime-u.ac.jp)

Chihiro Tanikawa[‡] (ctanika@dent.osaka-u.ac.jp)

Yuuji Shimizu[‡] (yjshimizu@dent.osaka-u.ac.jp)

Hajime Nagahara^{*} (nagahara@ids.osaka-u.ac.jp)

Takashi Ninomiya[†] (ninomiya@cs.chime-u.ac.jp)

[†] Graduate School of Science and Engineering, Ehime University

[‡] Graduate School of Dentistry, Osaka University

^{*} Institute for Datability Science, Osaka University

表1: 入力テキストの単語数

	所見文書	要約
平均単語数	1,886	179
最大単語数	6,379	467
最小単語数	312	61

2 提案手法

近年の自然言語処理では、深層学習に基づく手法が広く活躍している。特に、本研究で扱うテキスト分類のタスクでは、2010年代の前半からCNNおよびRNNのニューラルネットワークが使用され、2010年代後半からはBERT[4]などの自己注意ネットワークに基づくモデルが用いられている。本研究では、これらの深層学習モデルを矯正歯科治療の自動診断に初めて適用する。

しかし、先述のとおり、所見文書は自然言語処理で通常扱うテキストと比べて非常に長い文章で構成されているため、単純に深層学習モデルを適用するのでは十分な性能が得られない。例えば、BERTへの入力は最大512単語に制限されているが、表1に示すとおり、本研究で扱う所見文書は平均1,886単語で最大6,379単語という非常に長いテキストである。また、所見文書は視覚的なわかりやすさのために、各項目が文として成立しない箇条書きであったり、空白文字を用いたインデントを使用したりと、様々なスタイルが混在した文章となっている点も、深層学習モデルの性能低下の要因と考えられる。

そこで我々は、所見文書の要約を用いた矯正歯科治療の自動診断を提案する。本研究で扱う要約は、平均9文程度の長さの文章で構成されており、表1に示すとおり、最大のものでもBERTに全文を入力可能な長さである。また、箇条書きやインデントを使用せず統一したスタイルで記述している点も深層学習モデルにとっての利点であると考えられる。なお、要約は大阪大学歯学部附属病院の専門医が人手で記述した。

3 実験設定

データセットには、先行研究[1-2]でも用いられた970件の矯正歯科治療に関する日本語の所見文書を用いた。ただし、本研究では新たに全ての所見文書に対して要約を付与した。各所見文書には、患者の症状を表す複数のラベルが割り

表2: 深層学習モデルのハイパーパラメータ

モデル	埋込層	隠れ層	層数	Dropout
CNN	256	256	1	0.2
RNN	256	256	1	0.2
SAN	256	512	2	0.1
SAN (BERT)	768	768	12	0.1

当てられている。症状のラベルは322種類あり、各所見文書は平均12個のラベルを持つ。本研究では、このようにひとつの入力テキストが複数のラベルを持つマルチラベル文書分類のタスクとして自動診断を行う。評価は、適合率および再現率を求め、その調和平均(F値)を計算する。

深層学習モデルとして、CNN・RNN・SANの代表的な3種類のニューラルネットワークを用いる。ただし、RNNにはゲート付きRNNの一種であるLSTMを使用し、SANにはフルスクラッチで訓練するSANに加えて、単語穴埋めの事前訓練を行ったBERT¹も用いる。各モデルは表2に示すハイパーパラメータを採用し、PyTorch²を用いて実装した。

比較手法として、先行研究[1-2]のBoW+SVMを用いる。比較手法では、Binary Relevance[5]によってマルチラベル分類を擬似的なシングルラベル分類タスクに変換して解く。SVMはscikit-learn³を用いて実装し、RBFカーネルを採用した。

データセットの前処理として、要約前の所見文書に含まれる改行文字は空白文字に置換し、全角文字は半角文字へ正規化した。単語分割にはSudachi⁴のA単位を用いたが、BERTのみ専用のサブワード分割を適用した。

4 実験結果

実験結果を表3に示す。所見文書をそのまま入力する場合、深層学習モデルの中ではCNNのみが比較手法を上回り、RNNおよびSANは著しく低い性能を示した。特に、事前訓練済みのBERTよりもフルスクラッチで訓練したSANの方が高い性能を示すことから、矯正歯科治療の所見文書に含まれるテキストが、一般的な日本語テキストと大きく異なる特性を持つことがわかる。

短い文長と統一したスタイルの要約を入力する場合、元の所見文書を入力するのと比較して、全てのモデルの性能改善を確認できた。特に、事前訓練済みのBERTが最も大きく性能を改善でき、フルスクラッチで訓練したSANの性能を上回った。依然として先行研究[1-2]のBoW+SVMモデルは強力なものの、要約を入力とするCNNモデル

表3: 実験結果 (F値)

モデル	所見文書	要約
BoW+SVM [1-2]	0.41	0.46
CNN	0.44	0.48
RNN	0.20	0.31
SAN	0.29	0.38
SAN (BERT)	0.27	0.43

が本実験の中で最高性能を達成した。そのため、本研究で提案する深層学習モデルと要約テキストの組み合わせが、矯正歯科治療の自動診断のために有効であることを確認できた。

5 おわりに

本研究では、自然言語処理による矯正歯科治療の自動診断において、深層学習モデルの入力として所見文書の要約を用いることで診断性能を改善した。実験の結果、本タスクにおいてはCNNモデルが最も強力であること、入力として要約を用いることで全ての自然言語処理モデルの性能を改善できること、特に事前訓練済みモデルにおける性能改善が著しいことなどが明らかになった。また、要約を入力とするCNNモデルが最高性能を達成した。

今後の課題として、要約の自動生成によって専門医の診断コストを更に軽減したい。また、顔画像やX線画像などの画像データを併用し、マルチモーダル自動診断による性能改善を目指す。

参考文献

- [1] 西原大貴, 梶原智之, 谷川千尋, 清水優仁, 長原一. 矯正歯科治療における所見文書からの自動診断に向けて. 情報処理学会第83回全国大会, pp.591-592, 2021.
- [2] Yuujin Shimizu, Chihiro Tanikawa, Tomoyuki Kajiwara, Hajime Nagahara, Takashi Yamashiro. The Validation of Orthodontic Artificial Intelligence Systems That Perform Orthodontic Diagnoses and Treatment Planning. *European Journal of Orthodontics*, Vol.44, No.4, pp.436-444, 2022.
- [3] Shuai Li, Wanqing Li, Chris Cook, Ce Zhu, Yanbo Gao. Independently Recurrent Neural Network (IndRNN): Building A Longer and Deeper RNN. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.5457-5466, 2018.
- [4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp.4171-4186, 2019.
- [5] Grigorios Tsoumakas, Ioannis Katakis. Multi-Label Classification: An Overview. *International Journal of Data Warehousing and Mining*, Vol.2007, pp.1-13, 2007.

¹ <https://huggingface.co/cl-tohoku/bert-base-japanese-v2>

² <https://pytorch.org/>

³ <https://scikit-learn.org/>

⁴ <https://github.com/WorksApplications/SudachiPy>