



IV 腹部画像診断におけるITの技術革新と挑戦

1. 腹部領域におけるITの最新動向

1) Transformerを用いた 診断レポートの自動要約の 研究

西尾 瑞穂 神戸大学未来医工学研究開発センター

放射線科の診断レポートは画像診断の実践として重要であり、研究を補助するデータとしても活用される。近年の放射線医学の進歩により、多くの画像検査が行われ、その結果として多くの診断レポートが放射線科医によって書かれている。ほとんどの場合で診断レポートは自然言語で書かれ、非構造化データとして保存される。そのため、手動で処理する必要があり、大量の診断レポートを処理することは難しいことが多い。

自然言語処理(NLP)は、コンピュータが自然言語を処理できるようにする技術である^{1)~3)}。NLPを利用することで、電子カルテや診断レポートから構造化情報を自動的に抽出することが可能と言われる。放射線科の研究では、NLPはテキスト分類、テキスト要約、テキスト生成などに使用される⁴⁾。

Transformerやそれに基づくlanguage modelは、最近のNLPにおける大きな進展の一つとされる。language modelは大量のテキストを使用してトレーニングされたニューラルネットワークで、誤解を恐れずに言えば、畳み込みニューラルネットワークの親戚のようなものと言える。近年、Bidirectional Encoder Representations from Transformers (BERT)⁵⁾、Text-to-Text Transfer Transformer (T5)^{6), 7)}、Generative Pre-Training 1~3^{8)~10)}など、さまざまな種類のlanguage modelが開発されている。最近、しばしば話題にのぼる「ChatGPT」(<https://chat.openai.com/>)は、Generative Pre-Training 3を改善し、商用サービスとして利用可能に

したものとされる。われわれのグループでは診断レポートの処理にTransformerやlanguage modelを応用する研究を行ってきた^{11)~14)}。

一般に、診断レポートは「所見」と「impression」の2つのセクションに分かれる。もし所見セクションからimpressionセクションを自動要約することができれば、放射線科医の負担が軽減されると期待される。本研究の目的は、T5のlanguage modelを使用して、診断レポートの所見セクションをimpressionセクションへ自動要約することである¹⁵⁾。

方法

1. データセット

本研究では、日本医用画像データベース(Japan Medical Image Database: J-MID)のデータセットを利用した(<https://www.radiology.jp/j-mid/>)。J-MIDは、複数の大学病院や国立病院が協力して構築している放射線科の画像と診断レポートのデータベースである。今回は、J-MIDデータベースから日本の10の大学病院の診断レポートを抽出し、モデルの構築などに利用した。対象となったレポートは、2010年8月4日~2023年3月31日の日付のものである。所見またはimpressionセクションが欠損しているレポートは除外され、各レポートから所見とimpressionセクションのペアが集められた。J-MIDでは、実際に臨床に用いられた診断レポートが含

まれているため、各レポートのimpressionセクションが自動要約のground truthとして使用された。J-MIDから集められたレポートは、その日付に基づいて3つのセットに分割された。トレーニングセットは2010年8月4日~2022年11月30日および2022年12月10日~12月31日、検証セットは2022年12月1日~12月9日、テストセットは2023年1月1日~3月31日の日付のレポートとした。

2. Transformerおよびlanguage model

T5はエンコーダとデコーダから構成されるTransformerをベースにしたニューラルネットワークである。T5を利用する際には、事前に学習されたT5モデルをファインチューニングすることが多く、これにより文章要約、質問応答、文章分類などに活用できるとされる。この中でも、T5は特に文章要約のタスクに有用とされる。本研究では、診断レポートの文章要約にT5を利用した。T5への入力はい見セクションの文章で、出力はimpressionセクションの文章とした。T5モデルは所見セクションを要約するようにファインチューニングされた。事前に学習されたT5モデルはHugging Faceから入手可能である(<https://huggingface.co/>)。今回は、「google/mt5-base」¹⁶⁾の事前学習済みのモデルを利用した。google/mt5-baseは多言語用の事前学習済みのモデルである。