

2020年11月20日

報道関係者各位

国立大学法人 奈良先端科学技術大学院大学

画像診断結果の自動作成を支援する医師の読影レポートを集積して公開 AIを導入した医療診断の研究を加速

【概要】

奈良先端科学技術大学院大学（学長：横矢直和）先端科学技術研究科 ソーシャル・コンピューティング研究室の荒牧英治教授らの研究グループは、株式会社ワイズ・リーディングと合同で、CTなどの画像診断の結果を記載した文書を自動作成するための研究を加速するため、公開読影レポートデータセットを開発・公開いたしました。

近年、医療AIを用いて、単純X線（レントゲン）、CT、MRIなどの医用画像から読影レポートを自動生成する研究が盛んになっています。しかし、現存する評価方法で生成物に点数付けをすると、医学的な妥当性とは必ずしも一致しません。これは、肝心の目標である「よい読影レポートとは何か？」の定義が難しいということでもあり、分野の発展の大きな障害となっています。

私達は、その原因の一つとして、同じ診断内容であっても文章としてさまざまな書き方ができること（表現の多様性）に着目しました。医療現場では、一つの画像に対して一つの読影レポートしか作成されません。そのため、ただ単に医療機関に保存されている読影レポートを集めただけでは、同じ診断内容に対して人間がどのような書き方をするのかを十分に知ることはできません。

そこで私達は、同一のCT画像に対して複数名の読影医に読影レポートを作成していただき、集積して文章表現の適切な理解に役立てるという新しい方法で読影レポートデータセット「Japanese Medical natural language processing Standard datasets: Radiography Reports (J-MedStd: RR)」を作成し、一部をサンプルしてWebサイト (<https://sociocom.naist.jp/j-medstd/rr/>) にて公開することにいたしました。

このデータセットは将来、医療AIが単純X線（レントゲン）、CT、MRIなどに対してよりよい自動診断を下すためのお手本となる学習データとして広く使用されることを目的としています。

さらに今後、サンプル数の拡張を続け、本データを大規模化する予定です。本データが、専門性の高い医療文書の品質を判定したり、医用画像から自動診断を行ったりする医療AIの一助となり、ひいては医療分野における人間とAIの協働をますます加速させることが期待されます。

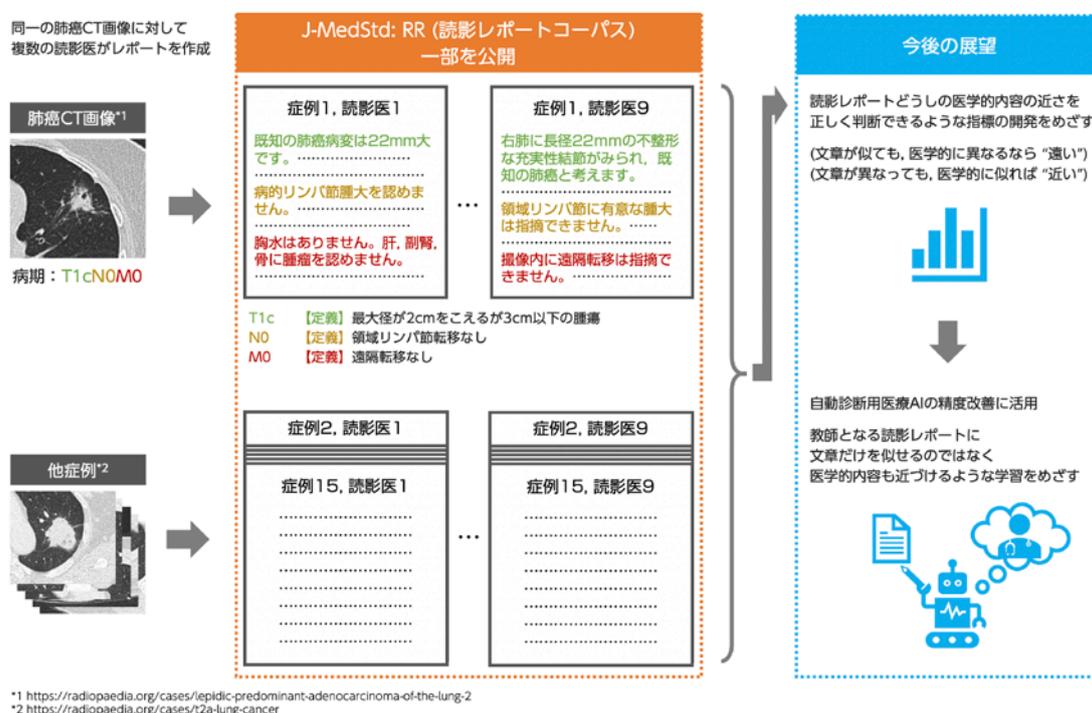
つきましては、関係資料を配付いたします。

【ご連絡事項】

- (1)本件につきましては、奈良先端科学技術大学院大学から奈良県文化教育記者クラブをメインとし、学研都市記者クラブ、大阪科学・大学記者クラブに、同時にご連絡しております。
- (2)取材希望がございましたら、恐れ入りますが下記までご連絡願います。
- (3)プレスリリースに関する問い合わせ先

奈良先端科学技術大学院大学 企画総務課 渉外企画係
TEL : 0743-72-5026 E-mail : s-kikaku@ad.naist.jp

J-MedStd: RR (読影レポートデータセット)



(図) J-MedStd: RR の作成方法の概要、および今後の展望

【背景と目的】

医療 AI を用いて、単純 X 線（レントゲン）、CT、MRI などの医用画像から読影レポートを自動生成する研究がさかんになっていますが、肝心の読影レポートの品質を点数づけして評価する方法は確立されていません。その原因は、同じ医学的内容が医師によって表記に微妙な違いがあったり、逆に異なる医学的内容なのに文章としては似ていたりすることにあります。

このため私達は、同一の CT 画像に対して複数名の読影医に読影レポートを作成していただくという新しい方法で読影レポートデータセット「Japanese Medical natural language processing Standard datasets: Radiography Reports (J-MedStd: RR)」を作成しています。

今回は、CT 画像から肺がんの進行度を国際対がん連合 (UICC) が提案する TNM 分類によって判定するという一つの要素に絞り、医学的内容に応じてどのような表現方法が存在するのかを収集しました。具体的には、放射線医学関連情報・教育サイト「Radiopaedia (<http://radiopaedia.org>)」で無料公開されている肺がん CT 画像を 15 個選び、それぞれに対して 9 名の読影医に読影レポートを作成していただき、計 135 文書を収集しました。

その一部をサンプルとして Web サイト (<https://sociocom.naist.jp/j-medstd/rr/>) にて公開しています。

【今後の展開】

本データを通じて、医療 AI が医学的内容の正しさを考慮しながら読影レポートの自動作成ができるようになったり、医療 AI が医学的内容を読み取ったうえで読影レポートの品質を評価したりできるようになることが期待されます。

【用語解説】

TNM分類：癌の進行度（病期、ステージ）を決定するために広く用いられている方法。癌のサイズと周囲臓器への浸潤度（**T** 因子）、リンパ節転移（**N** 因子）、遠隔転移（**M** 因子）の3因子を評価し、これらの因子を組み合わせることによって癌の進行度を決定する。たとえば **T** 因子が **T1c**、**N** 因子が **N0**、**M** 因子が **M0** となる癌の進行度を **TNM** 分類で表現すると「**T1cN0M0**」となる。

【本プレスリリースに関するお問い合わせ先】

奈良先端科学技術大学院大学 企画総務課 渉外企画係

TEL : 0743-72-5026 E-mail : s-kikaku@ad.naist.jp