

科学研究費助成事業 研究成果報告書

令和 4 年 6 月 19 日現在

機関番号：30108

研究種目：若手研究(B)

研究期間：2017～2021

課題番号：17K15868

研究課題名(和文)統合的画像診断システム構築のための基盤整備に関する研究

研究課題名(英文)Study on infrastructure development of integrating imaging diagnosis system

研究代表者

谷川原 綾子(Yagahara, Ayako)

北海道科学大学・保健医療学部・講師

研究者番号：50711884

交付決定額(研究期間全体):(直接経費) 3,200,000円

研究成果の概要(和文):本研究では、多様な疾患の画像診断支援を実現するための統合的画像診断支援システムの構築を目的とした。このシステムは、コンピュータ支援診断(CAD)システムに含まれる病変検出機能と、病変の特徴量から疾患を推測するための画像診断オントロジーを組合わせたものである。今回の研究期間では、放射線画像レポートに必要な用語と画像解析に関連する用語を教科書や論文抄録から本システムの基盤となる正常画像解剖、疾患と画像所見、画像所見と画像解析に関する3つのモデルを構築した。

研究成果の学術的意義や社会的意義

本オントロジーは、画像データから得られる特徴量を抽出し、その出現パターンや濃度勾配等を既知の医学知見と照合し、医学的事象の判断能力の獲得を目指している。これは、人工知能技術による医療者の思考と類似した処理の実現である。本研究で構築したオントロジーは幅広い疾患に対応するCADシステムの基盤となるとともに、処理過程が明確となるため、説明可能な人工知能システムの開発にもつながると考えている。

研究成果の概要(英文): Computer-Aided Diagnosis (CAD) systems have gained attention as essential support for a clinical decision of "particular" diseases. The purpose of this study was to construct an integrating imaging diagnosis system focusing on a wide variety of diseases. This system combines the lesion detection by CAD with a diagnostic imaging ontology. Using this research period, three following ontologies were constructed: normal imaging anatomy, relations between diseases and image findings, and relations between image findings and image analysis. The integration of these three ontologies will not only be the foundation of a CAD that contributes to the detection of a wide variety of diseases but will be an explainable artificial intelligence.

研究分野：医療情報学

キーワード：自然言語処理 固有表現抽出 オントロジー 放射線検査 画像診断 コンピュータ支援診断システム

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

放射線画像におけるコンピュータ診断支援システム(Computer Aided Diagnosis; CAD)は、画像情報の解析を行い、病変の自動検出や良悪性の鑑別など医師による診断の支援として活用されている。CAD は特定の病変や疾患を対象として開発されているものが多いが、本研究のゴールは、多疾患を対象とした CAD である統合的画像診断支援システムの構築である。具体的には、CAD にて病変の特徴量を抽出し、この特徴量から疾患を推測するための画像診断オントロジーを組合せたシステムである。ここでオントロジーとは、因果関係などを論理的に記述したモデルである。画像診断機器ごとの画像所見と、疾患・臨床症状などの情報が関係づけられると、検査依頼における適切な画像検査の選定、撮影プロトコルの最適化への応用が可能となり、検査の質の向上にもつながると考えている。

2. 研究の目的

画像診断情報と、CAD システムの病変特徴量の関係をオントロジーで融合させることにより多疾患の判定が可能な統合的画像診断支援システムの基盤を構築することを目的とした。本研究期間では、正常画像解剖、疾患とその画像所見、画像所見と画像解析に関するモデルの構築を行った。

3. 研究の方法

a. 正常画像解剖オントロジーの構築

正常脳画像に焦点を当てて、画像診断に必要な解剖学や画像コントラストなどの用語を収集し、これら用語間の関係を記述した。用語収集について、放射線医学に関する英語の教科書より頭部の画像解剖に関する文章を収集した。次に、放射線読影用語集である RadLex 3.1.0 と品詞タグ付けソフトである Stanford Part-Of-Speech tagger を用いて画像解剖に関わる専門用語を抽出した。RadLex に記載されていない語は、目視にて抽出した。抽出した用語の関係性は教科書の記述から決定した。用語間の関係性を記述するためにソフトウェア Google Refine を用いて Resource Description Framework (RDF)形式にて記述され、オントロジーエディタ Protégé によってモデルの可視化を行った。なお、モデルは放射線科医師 1 名、診療放射線技師 1 名が確認した。

b. 疾患と画像所見オントロジーの構築

ここでは、脳疾患を対象としたモデル RadLex をベースに構築した。a.の項で使用した放射線医学の教科書の本文から、「病名」、「画像診断機器」および「画像所見」の項に記載されている文章を抽出し、テキスト化した。RadLex と Stanford Part-Of-Speech tagger により画像診断に関わる語を抽出した。目視にて RadLex に含まれなかった用語を抽出し、これらの語は、RadLex を参考に構築したオリジナルのオントロジーモデルに組み込んだ。このモデルは、最上位に“KRM New Class”を、そのサブクラスに“Disease_name(疾患名)”、“Modality(画像診断機器)”、“Radiologic_sign(画像所見)”を設定し、これら 3つのサブクラスの下位に RadLex に含まれなかった用語が記述された。なお、これらの階層構造は“SubClassOf”によって表現された。RadLex では、疾患名を含んだ“clinical findings”と画像所見の関係は“may_be_caused_by”、画像所見と画像診断機器の関係は“related_modality”と記述されているため、KRM New Class と RadLex に含まれる用語間の記述においてもこれら 2つの関係を採用した。モデルは放射線科医師 1 名、診療放射線技師 1 名が確認した。

c. 画像所見と画像解析オントロジーの構築

まず、PubMed にて、“computed aided diagnosis”と“computed aided detection”をキーワードとして 2018 年までに出版された CAD 関連の研究論文の抄録を収集した。次に、放射線医学と画像解析に関する専門用語を抽出するために、RadLex3.1 と画像処理の用語であるコンピュータビジョンと画像処理の辞書である Dictionary of Computer Vision and Image Processing (DCVIP) の 2つの用語を辞書として使用した。最後に、画像所見と画像解析に関連する用語をリンクさせるため“isAnalyzedBy”という関係を新たに定義し、オントロジーを構築した。

4. 研究成果

a. 正常画像解剖オントロジーの構築

教科書の文章から 1237 の用語が抽出され、そのうち 424 (34.3%) の用語が RadLex に収載されている用語と一致した。RadLex は、“anatomical entity” や “imaging observation” などの 11 のカテゴリが存在するが、それぞれのカテゴリごとの抽出率を図 1 に示す。RadLex に含まれていない用語は、手動的にこれらのカテゴリに分類された。用語間の関係性は “has_Radiologic_Sign” や “has_nerve” など、40 を超える関係が抽出された。構築されたモデルの一部を図 2 に示す。

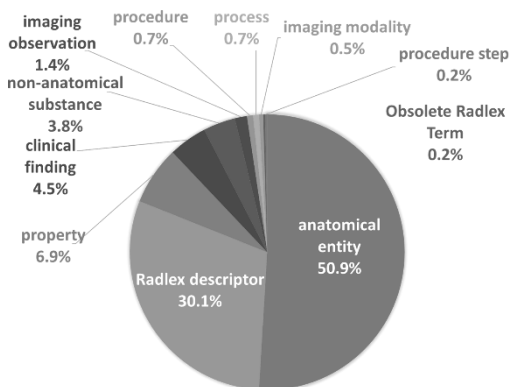


図 1 抽出用語の割合



図 2 正常画像解剖情報に関する知識表現モデルの例

b. 疾患と画像所見オントロジーの構築

教科書から 1271 文抽出され、143 の疾患名と 1128 の画像所見、11 の画像診断機器の用語が同定され、その中で RadLex と完全一致した語はそれぞれ 33 (23.1%) と 65 (5.8%) であった。RadLex と一致しなかった語は KRM new class の下に追加された。関係性の記述について、新たに、1121 の “may_be_caused_by”、392 の “related_modality” が記述された。図 3 に構築されたオントロジーの一部を示す。

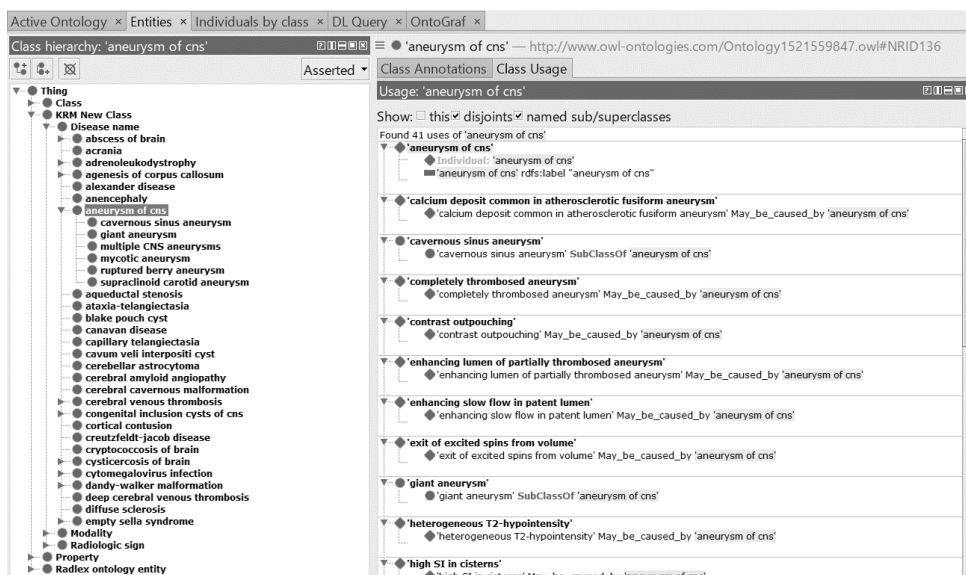


図 3 画像所見情報に関するオントロジーの例

c. 画像所見と画像解析オントロジーの構築

抄録に含まれる用語の中で、2つの用語集に収載されている用語と標記が一致した語は1861であり、Radlexは1134(61%)、DCVIPは672(36%)、両方に掲載されている語は55(3%)であった。誤分割により抽出漏れが生じたのは23語(2.3%)のみであった。

“isAnalyzedBy”によって267パターン of 画像解析と画像所見の組み合わせが記述された。例として、画像所見“nodule”には“edge”、“intensity”、“histogram”などの画像特徴量に関する語を含め117語が紐づけられた。

モデルの精緻化に向けて、次の2点について更なる検討が必要であった：①用語の抽出エラーの対応、②同義語の取り扱い。①について、テキストから用語を抽出する際に、単語の過分割、誤分割による用語抽出エラーが存在した。先行研究にて、専門用語の単語数を考慮した文字列マッチング手法が提案されている。この手法を参考に、現在、辞書用語を正確に抽出することを目指した専門用語抽出器の開発を進めている。②について、本手法では、同一文字列の用語を同一概念として認識し、同義語は別概念として扱われてしまう。これを改善するため、分散表現等を利用した同義語の自動検出ツールの開発も併せて進めている。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

| | |
|--|---------------------------|
| 1. 著者名 Yagahara Ayako, Sato Tetta | 4. 巻 76 |
| 2. 論文標題 Evaluation of the Automatic Full Form Retrieval Method from Abbreviation Using Word2vec for Terminology Expansion | 5. 発行年 2020年 |
| 3. 雑誌名 Japanese Journal of Radiological Technology | 6. 最初と最後の頁 1118 ~ 1124 |
| 掲載論文のDOI（デジタルオブジェクト識別子） 10.6009/jjrt.2020_JSRT_76.11.1118 | 査読の有無 有 |
| オープンアクセス オープンアクセスとしている（また、その予定である） | 国際共著 - |

〔学会発表〕 計11件（うち招待講演 0件/うち国際学会 5件）

| |
|--|
| 1. 発表者名 Sato T, Yagahara A, Tanikawa T. |
| 2. 発表標題 Determining the Optimal Model for Automatic Conversion from Abbreviation to Full form using Word2vec. |
| 3. 学会等名 第76回日本放射線技術学会総会学術大会 |
| 4. 発表年 2020年 |

| |
|--|
| 1. 発表者名 Yagahara A, Yokohama N, Sato T. |
| 2. 発表標題 Comparison of the Accuracy of Automatic Synonym Detection among Distributed Representation Models in Radiological Technology Domain |
| 3. 学会等名 第76回日本放射線技術学会総会学術大会 |
| 4. 発表年 2020年 |

| |
|--|
| 1. 発表者名 Yagahara A, Uesugi M, Ando D, Tha KK, EndoA A, Fujita K |
| 2. 発表標題 Automatic Prediction of Optimal MRI Protocols Using Encoder-Decoder Model |
| 3. 学会等名 RSNA2019（国際学会） |
| 4. 発表年 2019年 |

| |
|---|
| 1. 発表者名 谷川原綾子, 上杉正人, 安渡大輔, タキンキン, 遠藤晃, 藤田勝久 |
| 2. 発表標題 Encoder-Decoder Model による最適 MRI プロトコル予測システムの開発 |
| 3. 学会等名 第39回医療情報学連合大会 |
| 4. 発表年 2019年 |

| |
|--|
| 1. 発表者名 佐藤哲太, 川合美帆, 永田龍, 谷川原 綾子 |
| 2. 発表標題 英語抄録を用いた放射線技術用語の同義語自動抽出-人工知能技術 Word2vec を用いた検討- |
| 3. 学会等名 日本放射線技術学会第75回北海道部会秋季大会 |
| 4. 発表年 2019年 |

| |
|---|
| 1. 発表者名 佐藤哲太, 谷川原綾子 |
| 2. 発表標題 word2vecによる同義語自動抽出 - 放射線関連医学用語を対象として - |
| 3. 学会等名 第19回日本医療情報学会北海道支部会学術大会 |
| 4. 発表年 2019年 |

| |
|--|
| 1. 発表者名 Yagahara A, Uesugi M, Tha KK, Ando D, EndoA A, Fujita K |
| 2. 発表標題 What is the Best Word Segmentation Method for the Encoder-Decoder Model to Predict Optimal MRI Protocols? |
| 3. 学会等名 ECR2020 (国際学会) |
| 4. 発表年 2020年 |

| |
|---|
| 1. 発表者名 谷川原綾子, ミハウブタシンスキ, 辻真太郎, 上杉正人 |
| 2. 発表標題 チャンキング及び体言判別を用いた専門用語の自動抽出手法-放射線技術学関連の医学用語抽出への応用- |
| 3. 学会等名 第23回日本医療情報学会春季学術大会 |
| 4. 発表年 2019年 |

| |
|--|
| 1. 発表者名 Yagahara A, Tha KK, Jiang G |
| 2. 発表標題 Construction of a RadLex-based knowledge representation model for brain disorders using text mining |
| 3. 学会等名 ECR2019 (国際学会) |
| 4. 発表年 2019年 |

| |
|---|
| 1. 発表者名 Yagahara A, Tha KK, Jiang G |
| 2. 発表標題 Construction of a knowledge representation model focused on normal radiologic anatomy and interpretation |
| 3. 学会等名 SNR2018 (国際学会) |
| 4. 発表年 2018年 |

| |
|---|
| 1. 発表者名 Yagahara A, Tsuji S, Fukuda A, Nishimoto N, Jiang G, Ogasawara K |
| 2. 発表標題 Extraction of image processing and diagnosis terms for computer-aided diagnosis ontology construction -Morphological analysis using RadLex and image processing terminology- |
| 3. 学会等名 ECR2018 (国際学会) |
| 4. 発表年 2018年 |

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

| | 氏名 (ローマ字氏名) (研究者番号) | 所属研究機関・部局・職 (機関番号) | 備考 |
|--|---------------------------|-----------------------|----|
|--|---------------------------|-----------------------|----|

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

| 共同研究相手国 | 相手方研究機関 | | | |
|---------------|-------------|--|--|--|
| United States | Mayo Clinic | | | |