

令和7年度 成果報告書

基本情報(公開)

事業名 プログラム名		戦略的イノベーション創造プログラム(SIP) 第3期 統合型ヘルスケアシステムの構築
研究開発課題名		E-2 大規模医療文書・画像の高精度解析基盤技術の開発
研究開発 担当者* 1	機関名	国立大学法人京都大学
	所属	大学院情報学研究科
	役職	特定教授
	氏名	黒橋 禎夫

実施期間*2	令和7年4月1日～令和8年3月31日
--------	--------------------

*1 委託研究開発契約書に定義

*2 年度の契約に基づき、本委託研究開発を行った期間又は中止までの期間

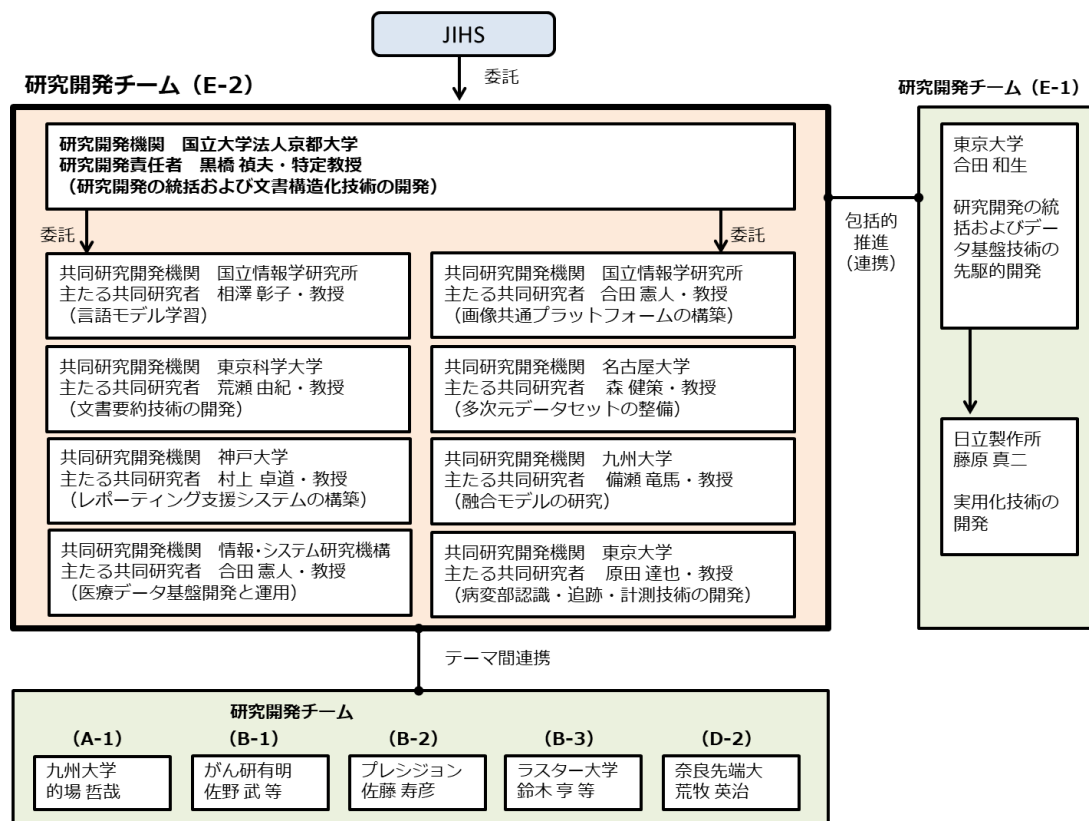
1. 研究開発テーマ概要(公開)

1.1 研究開発内容

本研究では、ヘルスケアシステムの最前線、即ち、臨床現場に於いて日々生み出される二大データ（電子カルテ等の医療文書ならびに CT/MRI 等の医療画像）をターゲットとして、高精度の自然言語処理技術と画像解析技術を確立することにより、臨床に於ける医師等の負担軽減と診療の質向上に繋げることを目指す。医療文書では、近年の発展が目覚ましい深層学習ならびに大規模言語モデル（基盤モデル）を駆使して日本語による医療 LLM を作成し、膨大な医療文書から因果関係および時系列関係を取り入れた高精度の構造化と要約を実現する自然言語処理技術を確立する。この技術により、医師による医療文書の読解・理解を飛躍的に高効率化する。医療画像では、患者を撮影した一連の画像の中から、病変部の経時的な変化を定量的に捕捉する画像処理技術の確立に挑戦する。これにより、医師による医療画像の読影効率を高める。さらに、画像解析の結果と構造化した所見文等との融合を可能とするマルチモダリティ解析技術の礎を築き、医療の大規模マルチモーダルモデル（LMM）の実現に資する。

1.2 研究開発実施体制

本研究では、研究開発責任者（黒橋）による全体統括の下、自然言語処理分野ならびに画像処理分野でトップクラスの研究開発能力を有し、更に医学分野との関係構築の経験・コミュニケーションの実績を誇る研究者（京都大学、東京科学大学、国立情報学研究所、名古屋大学、九州大学、東京大学、神戸大学、情報・システム研究機構）を結集し、研究開発に取り組む。



1.3 研究推進のスケジュール

項目	令和5年度	令和6年度	令和7年度	令和8年度	令和9年度
(サブテーマ1) 大規模医療文書の高精度 解析基盤技術の開発	1 a 		医療ドメインへの適応	医療言語モデルの学習・高度化	
	1 b 		構造化システムの設計・ 試作・有効性評価	構造化システム試用・フィードバック実証実験	
	1 c 		要約システムの設計・ 試作・有効性評価	要約システム試用・フィードバック実証実験	
	要約学習データ構築	補正予算によるLLM 構築事業との連携			
(サブテーマ2) 大規模医療画像の高精度 解析基盤技術の開発	2 a 		経時的画像データの整備 自動アノテーション付与	多次元医療データの整備	
	2 b 		モデルの精度向上 特徴量自動取得・計測	病変部の自動認識・追跡・計測	
	2 c 		融合モデルの探索	融合モデルの精度評価	
	データ格納方式 検証	補正予算によるLLM 構築事業との連携			
(サブテーマ3) LLM/LMMを用いた包括的 な画像診断のレポート 支援システムの構築		補正予算による取組み	放射線レポート自動作成 支援LLM	LMMによるJ-MIDからのデータセットの自動 作成に関する研究	
(サブテーマ4) 医療 LLM/LMM 利活用を 促進する医療データ基盤開 発と運用		補正予算による取組み	医療データ基盤の管理運用・医療データ基盤への学習データの追加、更新 モデル・学習データの 来歴管理設計と実装	医療データ基盤上でのサービス提供とフィードバック	

2. 本年度の成果・進捗の概要(公開)

本研究には二つの柱があり、第一に京都大学、東京科学大学、国立情報学研究所のチームによる大規模構言語モデル(基盤モデル)を駆使した、膨大な医療文書から因果関係および時系列関係を取り入れた高精度の構造化と要約を実現する自然言語処理技術の確立を目指した研究開発(サブテーマ 1)、第二に国立情報学研究所、名古屋大学、九州大学、東京大学の画像解析チームが取り組む、一連の医療画像の中から、病変部の経時的な変化を定量的に捕捉する画像処理技術の研究開発(サブテーマ 2)である。

具体的なサブテーマでの取り組みは以下のとおりである。

- (サブテーマ 1a) 大規模医療文書に基づく医療言語モデル構築技術の開発
- (サブテーマ 1b) 言語モデルと知識の融合利用による大規模医療文書構造化技術の開発
- (サブテーマ 1c) 言語モデルと知識の融合利用による大規模医療文書要約技術の開発
- (サブテーマ 2a) 医療画像に基づく経時変化解析のための学習データセットの整備技術の開発
- (サブテーマ 2b) 病変部を自動で認識・追跡・計測するシステムの開発
- (サブテーマ 2c) 画像と他のモダリティを組み合わせた融合モデルによる知識発見基盤の構築
- (サブテーマ 3) LLM/LMM を用いた包括的な画像診断のレポート支援システムの構築

(サブテーマ 4) 医療 LLM/LMM 利活用を促進する医療データ基盤開発と運用

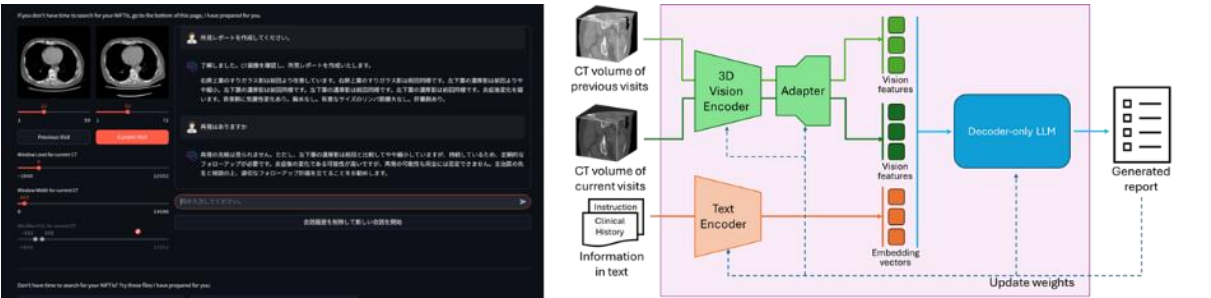
(成果・進捗の概要)

1. サブテーマ 1a: 大規模で高性能な医療言語モデルは、医療現場における業務支援 AI の高機能化やアプリケーションの開発の効率化の要となる。ここで、医療言語モデルの構築に必要な文書コーパスやモデルのパラメタ値を最適化するためには、言語モデルの性能評価手法を確立する必要がある。しかし、日本語の医療言語モデルについては、医師国家試験5年分を構造化した IgakuQA しか評価指標が存在せず、十分な評価が行えなかった。そこで過年度において、バランスがとれた大規模な日本語医療言語モデル向け評価用データセットを構築するためのフレームワークの開発に取り組み、成果として、言語モデル間の差異を吸収する手法を提案して、複数言語(日英)やサイズの異なるモデルに対しても横断的に比較評価が行えることを示した。また、「統合型ヘルスケアシステムの構築における生成 AI の活用」のテーマ1「安全性・信頼性を持つオープンな医療 LLM の開発・社会実装」との連携により、実際に大規模なベンチマークセットを構築して公開、さらに開発したモデルの性能指標として活用した。さらに 2025 年度においては、日本語医療言語モデルのための学習コーパス設計に関する理論および実証的な分析・評価、臨床タスクに適合したモデルを構築するための、入力長の拡大および指示チューニングデータセットの構築の構築に取り組んだ。これらの成果として、実効パラメタ数が同クラスの海外モデルと互角の性能を達成してプレビュー版モデルとして公開した。一方、これまでの開発リソースはウェブ上の公開情報が中心であり、実際の臨床現場におけるユースケースとの乖離が課題として意識された。そこで、診療 DX ソリューション開発に既に大きな実績のある B-2 と連携することで、実臨床に即した医療言語モデルへの高度化を実現できることが期待された。以上の経緯および 2025 年度実施のステージゲートを踏まえ、本サブテーマの成果の一部を継承・発展させる形で 2026 年度よりテーマ B-6 が新設された。
2. サブテーマ 1b: J-CaseMap 症例報告の構造化要約は医師の症例理解を効率化することから、診療業務における医師の負担を軽減する上で非常に重要な実用的価値を持つ。2024 年度までに追加事前学習によるドメイン適応技術を用いた医療特化 8x13B MoE ベースの構造化モデルは医師による評価(100 件の症例報告構造化の平均点)において 91.8 点の高スコアを獲得した。2025 年度は、構造化の対象をより複雑な構造の電子カルテの退院サマリデータへと拡張し、退院サマリに含まれる知見や看護師向けのケア上の注意点等を補完する知識補完要約についての取り組みを進めた。この過程で、最新のモデルは Few-shot で追加学習を行わずとも、患者の病状から看護師向けの注意点を提案する能力があることを示した。現在、医師らの視点を取り入れることで知識補完要約の更なる改善を目指しつつ、実際の医療現場における LLM 活用の可能性を模索している。
3. サブテーマ 1b において、2024 年度は、様々な医療タスクについて J-CaseMap 構造化タスクを組み合わせることによる影響について調査を行い、J-CaseMap 構造化タスクが他の医療タスクに資する医療知識の獲得に寄与することを確認した。2025 年度は、約 20,000 症例の構

造化済データに基づく医療ナレッジベースを構築した。この医療ナレッジベースは、症例データ全体における関係三つ組の出現頻度に基づくエッジ重みを持ち、病態、所見、薬剤等の医療エンティティ間の関係の妥当性や希少性を評価することができる。現在、診断などの他の医療タスクについてナレッジベースを活用する手法についての検討を進めている。

4. サブテーマ 1c: 大規模医療文書要約技術の開発では、自治医科大から提供を受けた退院サマリデータを用い、退院サマリに含まれるショートサマリを自動生成するモデルの構築に取り組んだ。SIP 医療言語モデル、Qwen3 によるファインチューニング、Few-shot プロンプトによる要約手法を開発した。さらにこれら大規模言語モデルを評価器として用いる LLM-as-a-Judge 手法の開発も行った。退院サマリが満たすべき観点(忠実性、網羅性、関連性、完結性、安全性)を盛り込んだルーブリックを設計した。本 LLM-as-a-Judge 評価の結果、SIP モデル、Qwen3 ともに人手の要約より顕著に高い品質と判断され、これは医師による評価と一致する。負荷の高い退院サマリ作成の大規模言語モデルによる補助が十分可能であることが示された。
5. サブテーマ 1c では、退院サマリ生成に活用することを目指し、医療分野の科学論文から医療知識を自動抽出する手法の開発に取り組んだ。投与した薬剤が有効に作用する条件(患者の性別や身体特性、他の疾患等)を約千件の科学論文の概要にアノテーションしたデータセットを構築した。医学分野、生命科学分野の修士課程学生に依頼し、ターゲットの疾患に対し薬剤が有効に作用する条件を特定、タグ付けを行った。さらに本データセットを用い、薬剤が作用するための条件を自動抽出する手法を大規模言語モデルを用いて開発した。提案手法は既存のアプローチを顕著に上回る抽出性能を達成した。本成果は Findings of Annual Meeting of the Association for Computational Linguistics (Findings-ACL2026) に採択されている。
6. サブテーマ 2a では、国立情報学研究所の運営する医療画像ビッグデータクラウド基盤を活用し、画像の撮影日時に基づいた経時的学習データの整備に取り組んだ。その結果、72 万を超える検査、症例数にして 17 万人余り(症例あたりの平均検査回数 4.2)の経時的 CT 画像を学習用データセット候補として収集した。その中で画像と所見文が揃っていて、かつ、同一患者に異なる日の検査がある症例に絞り込み約 3.9 万人(データセット候補の約 23%)を整備した。
7. サブテーマ 2a、2b: 2024 年度までに大規模医療画像の経時的変化解析ならびに病変解析等に必要とされる、画像認識のためのセグメンテーション基盤モデルの構築、および多臓器セグメンテーション向けの画像認識基盤モデルの構築を行った。国立情報学研究所医療ビッグデータ研究センターで収集された約 7800 万枚のCT断面画像(26 万例の3次元CT画像)を用いて、自己教師学習の手法によって3次元CT画像処理向けの基盤モデルを構築した。このモデルを用いた多臓器セグメンテーションモデルの構築を行うとともに、リンパ節抽出モデルの構築を行った。7800 万枚のCT画像で事前学習を行ったモデルを用いたリンパ節抽出手法を用いた結果、少数の学習データを用いるのみで非常に精度の高いリンパ節抽出が行えることを確認した。さらに、経時撮影された2つのCT画像を入力すると所見文を出力するLMMの初期バージョンの構築を

生成AIテーマ2と協働して行った。この成果について 2025 年 4 月 24 日に名古屋大学・国立情報学研究所から共同プレスリリースを行った。2025 年度は、CT 画像枚数を約 2 億枚に拡張して画像基盤モデルの開発を進めた。経時 3D CT 画像所見文生成モデルでは所見文生成において入力 CT 画像に加えて臓器領域認識結果も合わせて入力する新たなモデル SIGMA を構築し、ベンチマークでは唯一比較可能な CT2RepLong と比較してすべてのスコアで上回ることを確認した。さらに、所見文の誤り検出に向けた取り組みに着手した。



図：経時 3D CT 画像所見文生成モデルとモデル構造

	BLEU-1	BLEU-2	BLEU-3	BLEU-4	METEOR	ROUGE-L
CT2RepLong	0.298	0.216	0.174	0.144	0.276	0.235
SIGMA	0.331	0.248	0.199	0.162	0.312	0.253

図：経時 3D CT 画像所見文生成モデルベンチマーク

- サブテーマ 2b では、2024 年度の SIP 生成 AI テーマ 2 の続きとして、日本語汎用視覚言語モデルおよび医用視覚言語モデルの構築を行った。昨年度開発した日本語汎用視覚言語モデル「Asagi」の性能向上のため、Mixture of Expert (MoE)モデルでの開発と、日本語キャプションと画像のペアから構成される学習データセットの拡張を行った。学習データを 2,000 万件から 8,000 万件程度まで大幅に増加のうえ学習させたことで、さらなる性能向上を確認した。一方、開発した Asagi の知見を活かし、医用視覚言語モデルの開発を行っている。ここでは、PubMed Central Open Access データを利用した。しかし、これらは英語のキャプションと画像のペアであるため、これらを翻訳し、1200 万ペアのデータセットを構築した。このデータで学習した結果、推論モデルにおいて GPT-5 に次ぐ精度を達成した。
- サブテーマ 2c において、九州大学病院循環器内科と連携し、CT 画像および臨床データを用いた冠動脈石灰化切削デバイスの使用可否予測モデルを開発した。石灰化領域の情報を集約するクラストークンを臨床データに基づいて動的に生成する Hypernetwork を導入することで、AUC を 0.68 から 0.71 に改善し、本成果は IEEE ISBI に採択された。さらに血管の曲率情報を特徴量として導入することで、AUC 0.764 を達成した。また、生成モデルを用いた血管中心線抽出手法 (VesselFusion) を開発し、公開 CT データセットにおいて既存手法 VesselFormer を上回る精度と、より自然な血管構造の推定を確認した。本成果についても IEEE ISBI に論文が採択されている(計 2 本)。現在は、CT のみでは限界がある課題に対応す

るため、術中 IVUS 画像との統合を目指し、最適輸送と生成モデルを組み合わせたクロスモダリティ推定手法の検討を開始している。循環器領域との連携を軸に、マルチモーダル AI および基盤モデルの方法論的発展も視野に研究を推進している。

10. サブテーマ3では、生成 AI を活用した画像診断の支援技術の高度化において、複数の具体的成果が得られた。第一に、画像診断レポート内の誤字・脱字検出に関する研究を開始し、zero-shot で LLM を用いた二値判別で ROC-AUC 0.85 を達成した。第二に、胸部 X 線画像における結節検出において、CNN の判別結果を LMM(RadVLM)の入力に統合する手法を開発し、単独モデルを上回る正診率を示し、マルチモーダル統合の有効性を確認した。第三に、画像診断の 53 冊の教科書と画像診断以外の領域の 40 冊を合わせた 93 冊の日本語の教科書を用いた LLM の継続学習を実施し、日本医学放射線学会専門医試験で 43.8%から 48.6%へと正答率を向上させ、画像診断特化型 LLM の実現可能性を示した。SIP の研究成果の商品化にむけ、複数社と調整を進めている。

11. サブテーマ4では、2024 年度の SIP 生成 AI テーマ 4 の続きとして、医療データ基盤の運用ならびに各種学習データの蓄積、更新を進めた。医療データ基盤の運用においては、医療データ基盤上のハードウェア資源をセキュアな状態で医療機関へ提供する「医療用仮想プライベートクラウド環境」を高度化し、当該環境から情報・システム研究機構内の他拠点に設置された GPU サーバを利用可能とするための環境構築を行った。また、参画する特定の研究者およびシステム管理者以外が、対象となる機微な医療データにアクセスできないよう、SINET L2VPN を活用したセキュアなネットワーク接続方式およびユーザ管理手法を自治医科大学と共同で確立した。これにより、従来は「医療用仮想プライベートクラウド環境」上で管理するデータを、各研究拠点内に GPU 環境を構築して学習・推論する必要があったが、同一環境上でデータの蓄積・抽出から学習・推論まで、他の利用者から隔離された状態で安全に実行することが可能となった。本環境の整備により、研究対象及び規模の拡大が期待できる。

学習データの蓄積、更新においては、大規模 Web クローリングを継続し、学習用日本語文書 17.5 億件(HTML および PDF)と、学習用画像データ 9.5 億件を LLM/LMM 開発向けに新たに提供した。画像データについては、LMM 開発における前処理負荷を低減するため、情報量の乏しい画像を除去すると共に、画像が貼り込まれた HTML データと対にしたデータセットとして整備した。

3. 成果物の公表(公開)

3.1 論文など(原著論文、学位論文、プロシーディングス、総説、解説、速報など)

論文数(総数)	(内国際誌)	(内国内誌)
30	25	5

京都大学

1. Chengzhi Zhong, Fei Cheng, Qianying Liu, Yugo Murawaki, Chenhui Chu, Sadao

- Kurohashi. Language Lives in Sparse Dimensions: Toward Interpretable and Efficient Multilingual Control for Large Language Models. In Proceedings of 19th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2026) 7958-7970 2026 年 3 月.
2. Causal Tree Extraction from Medical Case Reports: A Novel Task for Experts-like Text Comprehension, Sakiko Yahata, Zhen Wan, Fei Cheng, Sadao Kurohashi, Hisahiko Sato, and Ryoza Nagai, In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, pages 25849–25867, Suzhou, China. 2025/11/4-2025/11/9, 国外, ポスター
 3. Chengzhi Zhong, Qianying Liu, Fei Cheng, Junfeng Jiang, Zhen Wan, Chenhui Chu, Yugo Murawaki, Sadao Kurohashi. What Language Do Non-English-Centric Large Language Models Think in?. In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025): Findings Volume 26333-26346 2025 年 7 月

国立情報学研究所

4. Mohit Gupta, Akiko Aizawa, Rajiv Ratn Shah: “Med-CoDE: Medical Critique based Disagreement Evaluation Framework,” Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 4: Student Research Workshop), 112–119 (2025.4).
5. Kon Woo Kim, Rezarta Islamaj, Jin-Dong Kim, Florian Boudin, Akiko Aizawa: “Repurposing Annotation Guidelines to Instruct LLM Annotators: A Case Study,” Natural Language Processing and Information Systems — 30th International Conference on Applications of Natural Language to Information Systems (NLDB 2025), 140–151 (2025.7).
6. Kazuma Kobayashi, Zhen Wan, Fei Cheng, Yuma Tsuta, Xin Zhao, Junfeng Jiang, Jiahao Huang, Zhiyi Huang, Yusuke Oda, Rio Yokota, et al.: “Leveraging High-Resource English Corpora for Cross-lingual Domain Adaptation in Low-Resource Japanese Medicine via Continued Pre-training,” Findings of the Association for Computational Linguistics: EMNLP 2025, 11469–11488 (2025.11).
7. Jiahao Huang, Fei Cheng, Junfeng Jiang, Kazuma Kobayashi, Akiko Aizawa: “Language Bias in Multilingual RAG: A Case Study in the Japanese Medical Domain,” Proceedings of the 2025 9th International Conference on Natural Language Processing and Information Retrieval (NLPIR 2025) (2025.12).
8. 小林 和馬, 相澤 彰子. 医療分野における大規模言語モデルの可能性.循環器ジャーナル 74 巻 1 号 (2026 年 1 月発行)(査読なし)

東京科学大学

1. R. Miyano and Y. Arase. Adaptive LoRA Merge with Parameter Pruning for Low-Resource Generation. in Findings of the Association for Computational Linguistics: ACL 2025, pp. 19353–19366 (July 2025).
2. G. Luo and Y. Arase. MedSummRAG: Domain-Specific Retrieval for Medical Summarization. in Proceedings of the Workshop on Biomedical Language Processing (BioNLP), pp. 27–33 (Aug. 2025).
3. G. Luo, N. Nishida, Y. Matsumoto, and Y. Arase. Applicability Condition Extraction for Therapeutic Drug-Disease Relations. In Proceedings of Findings of Annual Meeting of the Association for Computational Linguistics (Findings-ACL2026). (July 2026, to appear).

名古屋大学

1. 森 健策. 視覚言語モデル(VLM)とその循環器分野への応用, 環器ジャーナル.Vol.74, No.1, pp.135-141, (2026/1/1)
2. 森 健策, 小田 昌宏. 眼底画像から健康関連バイオマーカーを推測する AI. あたらしい眼科 Vol.43 No.01(2026年01月号)
3. Qi Chuanyu, Masahiro Oda, Kensaku Mori, Quality Evaluation Framework for Automatically Generated Answers in Japanese Medical VQA Dataset Construction. 電子情報通信学会技術研究報告(MI), vol. 125, no. 395, MI2025-90, pp. 112-115
4. Yuta Kawamura, Tomohiro Kikuchi, Masahiro Oda, Cheng Wang, Yuichiro Hayashi, Holger R. Roth, Dong Yang, Kensaku Mori, Automatic estimation of mediastinal lymph node map in CT images Proceedings. Volume 13926, Medical Imaging 2026: Computer-Aided Diagnosis; 139261Q (2026)
5. Cong Khang Nguyen, Cheng Wang, Zongxuan Shi, Yue Heng, Luyang Zhang, Chuanyu Qi, Masahiro Oda, Kensaku Mori. Joint training to improve report generation capability from longitudinal CT volumes with a vision-language model, Proceedings Volume 13926, Medical Imaging 2026: Computer-Aided Diagnosis; 139261O (2026)
6. Yue Heng, Cheng Wang, Yuichiro Hayashi, Masahiro Oda, Kensaku Mori An investigation of segmentation-pretrained and image-text pretrained models for medical VQA Proceedings Volume 13926, Medical Imaging 2026: Computer-Aided Diagnosis; 1392606 (2026)

九州大学

1. Takumi Okuo, Shinnosuke Matsuo, Shota Harada, Kiyohito Tanaka, and Ryoma Bise, Weakly-Supervised Domain Adaptation with Proportion-Constrained

- Pseudo-Labeling, International Joint Conference on Neural Networks (IJCNN), 2025
2. Kaito Shiku, Shinnosuke Matsuo, Daiki Suehiro, and Ryoma Bise, Learning from Majority Label: A Novel Problem in Multi-class Multiple-Instance Learning, Pattern Recognition, 112425, 2025.
 3. Yuki Shigeyasu, Shota Harada, Akihiko Yoshizawa, Kazuhiro Terada, Naoki Nakazima, Mariyo Kurata, Hiroyuki Abe, Tetsuo Ushiku, Ryoma Bise, Domain Generalization of Pathological Image Segmentation by Patch-Level and WSI-Level Contrastive Learning, International Conference on Machine Vision Applications (MVA), 2025
 4. Kaito Shiku, Kazuya Nishimura, Shinnosuke Matsuo, Yasuhiro Kojima, and Ryoma Bise, Auxiliary Gene Learning: Spatial Gene Expression Estimation by Auxiliary Gene Selection, The 40th Annual AAAI Conference on Artificial Intelligence (AAAI), 2026.
 5. Soichi Mita, Shumpei Takezaki, and Ryoma Bise, VesselFusion: Diffusion Models for Vessel Centerline Extraction from 3D CT Images, IEEE International Symposium on Biomedical Imaging (ISBI), 2026.
 6. Kaito Shiku, Ichika Seo, Tetsuya Matoba, Rissei Hino, Yasuhiro Nakano, and Ryoma Bise, Hypernetwork-Based Adaptive Aggregation for Multimodal Multiple-Instance Learning in Predicting Coronary Calcium Debulking, IEEE International Symposium on Biomedical Imaging (ISBI), 2026.

神戸大学

1. Umeno A, Nishio M, Matsuo H, et al. Identification of high-priority radiology reports with unexpected findings using fine-tuned large language models. Eur Radiol. 2026;36(5):3766-3774. doi:10.1007/s00330-025-12252-2 査読有り
2. Matsuo H, Nishio M, Fujimoto K, et al. Artificial Intelligence for Chest Radiography: An Overview of Techniques, Challenges, and Future Directions. npj Health Systems. Accepted, in press. 査読有り
3. Nishio M. Clinical Usefulness of Multimodal Generative Artificial Intelligence on Chest Radiography. AJR Am J Roentgenol. 2025;225(4):e2533500. doi:10.2214/AJR.25.33500 査読無し
4. Matsumoto R, Matsuo H, Sugimoto M, Matsunaga T, Nishio M, Kono AK, Yamasaki G, Takahashi M, Kondo T, Ueno Y, et al. Deep Learning-Based Detection of Intracranial Hemorrhages in Postmortem Computed Tomography: Comparative Study of 15 Transfer-Learned Models. Applied Sciences. 2025; 15(19):10513. <https://doi.org/10.3390/app151910513> 査読

有り

5. Sofue K, Ueno Y, Yabe S, et al. Advantages of deep learning reconstruction algorithm in ultra-high-resolution CT for the diagnosis of pancreatic cystic neoplasm. *Jpn J Radiol.* 2025;43(10):1652-1662. doi:10.1007/s11604-025-01804-7 査読有り
6. 塚本 一真, 西尾 瑞穂, 黒崎 涼, 松尾 秀俊, 野上 宗伸, 今岡 いずみ, 大村 和元, 日下 亜起子, 村上 卓道, 村垣 善浩. 深層学習モデルを用いた脳出血分類 AI に対する有用性の検討. 日本診療放射線技師会誌 72(9) 1014-1014. 査読無し
7. Sofue K, Ueshima E, Ueno Y, Yamaguchi T, Hori M, Murakami T. Improved Image Quality of Virtual Monochromatic Images with Deep Learning Image Reconstruction Algorithm on Dual-Energy CT in Patients with Pancreatic Ductal Adenocarcinoma. *J Imaging Inform Med.* 2026;39(1):815-826. doi:10.1007/s10278-025-01514-6 査読有り

3.2 学会発表など((国内・国際)学会口頭発表・ポスター発表、招待講演で成果を公表したもの)

京都大学

1. 症例報告テキストの因果木抽出のための自動評価手法の検討. 八幡 早紀子 (京大), Fei Cheng (京大), 黒橋 禎夫 (京大), 佐藤 寿彦 (プレジジョン), 永井 良三 (自治医大). 言語処理学会 第 32 回年次大会 (NLP2026)

国立情報学研究所

1. 相澤彰子. 臨床応用を見据えた 基盤型日本語医療 LLM の構築. 第 45 回日本画像医学会「シンポジウム AI」(2026 年 2 月 17 日)
2. 小林 和馬, 相澤 彰子. 医療分野における大規模言語モデルの可能性. 循環器ジャーナル 74(1) 20-26 (2026 年 1 月)
3. 相澤彰子. 日本語医療ドメインに適応したオープンな LLM の開発. 第 45 回医療情報学連合大会 (2025 年 11 月 15 日)
4. 相澤彰子. 日本語大規模言語モデルの医療ドメインへの適応と課題. 第 79 回日本臨床眼科学会 (2025 年 10 月 9 日)
5. 相澤彰子. 大規模医療言語モデルと信頼される AI. 第 64 回日本生体医工学会大会(2025 年 6 月 6 日)
6. 小林 和馬. 医療における人工知能技術の活用と課題. 土木学会 第 8 回応用数理・AI セミナー (2025 年 7 月 14 日)
7. 小林 和馬. Human-in-the-Loop 型アプローチによる医療 AI の研究開発. [理研 AIP 成果報告イベント] 生成 AI とその医療応用(2025 年 7 月 18 日)
8. 小林 和馬. SIP プロジェクトにおける日本語医療 LLM の研究開発. 第 2 回 内閣府 BRIDGE 事業 成果報告会(2025 年 8 月 20 日)

9. 小林 和馬. Human-in-the-Loop 型アプローチで創る医療 AI. 第 53 回 日本磁気共鳴医学会大会(2025 年 8 月 30 日)
10. Kazuma Kobayashi. Governance, Integration, and Innovation: Building Trustworthy Medical AI in Japanese Healthcare. The 28th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2025)(2025 年 9 月 26 日)
11. 小林 和馬. 医療分野における人工知能の進歩と課題. 第 53 回日本放射線技術学会秋季学術大会(2025 年 10 月 19 日)
12. 小林 和馬. 信頼される日本語医療 LLM を目指して: 透明性のある構築プロセスの重要性. 第 1 回 J-STAGE セミナー(JST-STM ジョイントセミナー)(2025 年 11 月 4 日)
13. 小林 和馬. 安全性・信頼性を持つ医療生成 AI の社会実装に向けた技術的・社会的課題の抽出とガバナンス構築への展望. 第 45 回医療情報学連合大会(2025 年 11 月 15 日)
14. Kazuma Kobayashi. Developing Medical AI via a Human-in-the-Loop Approach. 4th WPI-PRIME International Symposium(2026 年 2 月 6 日)
15. 小林 和馬. 生成 AI は医療現場でどこまで活用されているのか!? PMDA シンポジウム「生成 AI の医療活用の最前線」(2026 年 3 月 6 日)
16. 小林 和馬. 医師と共創する AI. 北海道大学 CLAP&PRISM-HU 令和 7 年度 成果発表会 (2026 年 3 月 19 日)

名古屋大学

1. RLHF-based Q&A Pairs generation from Finding Texts of Abdominal CT images for Fine-Tuning of Medical Large Multimodal Models, Chuanyu Qi, Cheng Wang, Masahiro Oda, Kensaku Mori, IFMIA2025, P-100, IF250083, かがわ国際会議場, (2025/3/21) 国内, 口頭
2. Uncertainty-aware semi-supervised learning for enhanced multi-organ segmentation in CT volumes, Guotai Dong, Masahiro Oda, Yuichiro Hayashi, Kensaku Mori, Kazunari Misawa, SPIE Medical Imaging 2025, Paper 13408-28, Town and County Resort, San Diego, USA, 2025/2/16-2025/2/20,(2025/02/19) 国外, 口頭.

九州大学

1. 冠動脈 CT 画像を用いた石灰化の切削の必要性を推定する Multimodal Model, 瀬尾惟周, 志久開人, 仲野泰啓, 的場哲哉, 備瀬竜馬, 画像の認識・理解シンポジウム(MIRU), 2025/8/1, 国内, ポスター.
2. 拡散モデルによる 3 次元 CT 画像からの血管中心線推定, 見田壮一, 竹崎隼平, 備瀬竜馬, 電子情報通信学会技術研究報告(PRMU2025-36, 2026 年 1 月研究会), 2026/1/30, 国内, 口頭 (2025 年度 PRMU 研究奨励賞).

東京大学

1. Coronary Artery Segmentation with Partial Annotations in Coronary CT Angiography Images, Yusuke Kurose, Haruto Chiba, Lin Gu, Junichi Iho, Youji Tokunaga, Makoto Horie, Yusaku Hayashi, Keisuke Nishizawa, Yasushi Koyama, Tatsuya Harada, IEEE International Symposium on Biomedical Imaging (ISBI), 2025.
2. 医療用大規模日本語視覚言語モデルの構築, 安道健一郎, 黒瀬優介, 上原康平, 菊池智博, 牧元久樹, 小寺聡, 小林和馬, 村尾晃平, 吉田浩, 田村孝之, 合田憲人, 喜連川優, 原田達也, 言語処理学会 第 32 回年次大会(NLP2026), 2026.
3. 医療用大規模日本語視覚言語モデルの構築, 安道健一郎, 黒瀬優介, 上原康平, 菊池智博, 牧元久樹, 小寺聡, 小林和馬, 村尾晃平, 吉田浩, 田村孝之, 合田憲人, 喜連川優, 原田達也, 医用画像情報学会, 2025.
4. 推論過程に能動的質問を組み込んだ日本語 V&L データセットの構築, 上原康平, 黒瀬優介, 安道健一郎, Jiali Chen, Fan Gao, 金澤 爽太郎, 坂本 拓彌, 竹田 悠哉, Boming Yang, Xinjie Zhao, 村尾晃平, 吉田浩, 田村孝之, 合田憲人, 喜連川優, 原田達也, 画像の認識・理解シンポジウム(MIRU2025), 2025.

神戸大学

1. AI-Powered MRI Analysis for Predicting Muscle Invasion in Bladder Cancer. M. Nishio. ACAR2026. 2026/03/19 国際, 口頭
2. Cost-efficient localisation of open large language models via textbook-based continuous pretraining: evaluation on medical licensing and radiology board exams. H. Matsuo, M. Nishio, T. Matsunaga, T. Murakami. ECR2026 2026/03/04 国際, ポスター
3. Development and Evaluation of a Deep Learning-Assisted Diagnostic Support System for Intracranial Hemorrhage. K. Tsukamoto, M. Nishio, R. Kurosaki, I. Imaoka, K. Ohmura, A. Kusaka, Y. Muragaki, T. Murakami. ECR2026 2026/03/04 国際, ポスター
4. Comparison of LLaVA 1.5 and 1.6 for Radiology Report Generation: The Impact of LoRA Fine-Tuning on Medical Image Analysis. H. Matsuo, T. Matsunaga, M. Nishio, T. Murakami CARS2025 2025/06/19 国際, ポスター
5. Evaluation of the Effectiveness of Large Language Models in Automatic Structuring of Chest X-ray Radiology Reports. T. Matsunaga, M. Nishio, H. Matsuo, T. Murakami CARS2025 2025/06/19 国際, ポスター

3.3 診療ガイドライン、省令、基準、日本薬局方、添付文書改訂、国の技術文書(通知)等への反映

該当なし

3.4 研修プログラム、カリキュラム、シラバス、教材、e-learning 等の公表

該当なし

3.5「国民との科学・技術対話」に対する取り組み

該当なし

3.6 その他

【プレスリリース】

- ・世界初、CT 画像を AI が比較・解析し日本語の所見文を生成 ～放射線科医の経過観察支援に期待～，名古屋大学および国立情報学研究所，2025 年 4 月 24 日
- ・オープンな医療用マルチモーダルモデルを開発—142 億パラメータを持つ日本語に特化した医療用視覚言語モデル—，東京大学および理化学研究所，2026 年 3 月 6 日

【モデル公開等】

- ・2025 年 4 月：日本語医療 LLM(SIP-jmed-llm-2)のオープンソースライセンス・モデルおよびコーパス公開(NII)
- ・2025 年 11 月：最新版の日本語医療 LLM(SIP-jmed-llm-3)の公開(NII)
- ・2026 年 3 月：医療用視覚言語モデルの公開(東京大学)

以上