

救急医療におけるAIを用いた腹部CT画像診断支援の現状と課題に関する文献レビュー

Literature review on the current status and issues of abdominal CT imaging diagnosis support using AI in emergency medicine

江藤 芳浩^{1), 2)}, 恵谷 玲央³⁾, 品川 佳満⁴⁾, 小嶋 光明⁵⁾

1) 医療法人慈恵会西田病院

2) 修士 (医科学) 大分県立看護科学大学大学院 看護学研究科 健康科学専攻 博士後期課程

3) 博士 (保健学) 大分県立看護科学大学 看護学部 人間科学講座 環境保健学研究室

4) (大学教員) 博士 (医療情報学) 大分県立看護科学大学 看護学部 人間科学講座 健康情報科学研究室

5) (大学教員) 博士 (薬学) 大分県立看護科学大学 看護学部 人間科学講座 環境保健学研究室

Key words: Artificial Intelligence, Machine learning / Deep learning, abdominal CT, Detection of Non-traumatic lesions, acute abdomen

[Abstract]

We reviewed the literature on AI-based models used for detecting non-traumatic lesions in abdominal CT imaging to determine the current status and challenges of using AI to detect diseases of abdominal organs and acute abdominal conditions. We searched PubMed and Google Scholar and extracted 106 references. The majority of studies were aimed at detecting tumors of the liver, kidney, and colon, with detection accuracy tending to be higher for liver tumors and renal stones and lower for gastrointestinal tract disease.

Of the 15 references on diseases of acute abdomen, renal and ureteral stones and colitis accounted for 10 of them. The main challenge was that the data set was not sufficient for the detection of renal and ureteral stones. In the detection of colitis, which has relatively low detection accuracy, the method of measuring the thickness of the colon wall caused false negatives and false detections of other organs.

【要旨】

腹部CT画像診断におけるAIを用いた非外傷性病変の検出モデルに関する文献をレビューし、腹部臓器および急性腹症の疾患検出におけるAI利用の現状と課題を明らかにした。PubMedおよびGoogle scholarを用いて106件の文献を抽出した。肝臓、腎臓、大腸の腫瘍検出を目的とした研究が大半を占め、検出精度は肝腫瘍、腎結石が高く、消化管疾患が低い傾向であった。

急性腹症の疾患に関する文献15件のうち、腎・尿管結石と大腸炎が10件であった。主な課題としては、腎結石、尿管結石の検出ではデータセットが十分ではないということであった。また相対的に検出精度が低い大腸炎検出では、大腸壁の厚さを検出要件としたために偽陽性や他臓器の誤検出が増加していた。

1. 緒 言

CT画像診断は救急医療の分野において必要とされ、CT検査はその役割を果たす手段として広く認

識されている^{1), 2), 4)}。本邦のCT装置の人口当たりの保有台数は世界一である上、CT装置の目まぐるしい技術向上によって多量かつ精密なデータを短時間に収集可能となったことから、CT画像診断の需要と期待はますます高まっている。それに伴い、放射線科医師の読影件数は増加傾向であるが⁵⁾、本邦におけるCT・MRI利用件数当たりの放射線科医数は減少傾向にある⁶⁾。画像診断を専門とする放射線科医師は慢性的に不足しており、都市部と地方都市との診断格差も拡大傾向にあるなど、放射線科医師の負担増に拍車を掛けている^{7), 8)}。こうした背景から、専門外の医師による画像診断がしばしば見受けられ、診断に苦慮しているのが現状である^{9), 10)}。救急外来 (ER: emergency room) における画像診断の主な対象疾患において、急性腹症はいずれの年齢層においても高い比率を占めている^{11), 12)}。蛭名らは、ERでのCT画像診断において、急性腹症は実質臓器や管腔臓器に加え、血管系やfree air、腹水の有無まで注意深く観察しなければならないため、ER担当医師と放射線科医師の単位症例数当たりの画像診断

ETO Yoshihiro^{1), 2)}, ETANI Reo, Ph.D.³⁾, SHINAGAWA Yoshimitsu, Ph.D.⁴⁾, OJIMA Mitsuaki, Ph.D.⁵⁾

- 1) Medical corporation Jikeikai Nishida hospital
- 2) Doctoral Program, Health Sciences Graduate School, Oita University of Nursing and Health Sciences
- 3) Laboratory of Environmental Health Science, Department of Health Sciences, Oita University of Nursing and Health Sciences
- 4) Laboratory of Health Informatics and Biostatistics, Department of Health Sciences, Oita University of Nursing and Health Sciences
- 5) Laboratory of Environmental Health Science, Department of Health Sciences, Oita University of Nursing and Health Sciences

Received November 15, 2022; accepted April 13, 2023

の不一致数は頭部疾患より多く、ER担当医師にとって急性腹症の画像診断はピットホールであると報告している¹³⁾。こうした現状から、腹部CT画像診断における急性腹症などの非外傷性病変の検出において、AIシステムが放射線科医師のレベルでCT画像を解釈できれば、専門外医師の画像診断を支援することができ、放射線科医師の不在を補う手段になり得る¹⁰⁾。

現在、医療分野で利用されているAIのほとんどは機械学習 (ML: Machine Learning) に分類され、主に放射線診断学分野における画像診断支援で使用されている。近年は、ビッグデータを利用したMLに加え、自らデータ特徴量を抽出できる深層学習 (DL: Deep Learning) の普及が進み、診断支援AIアプリケーションの開発や研究が世界的規模で進められている¹⁴⁾。画像診断の特定分野においては、放射線科医師の診断能力に同等、もしくはそれ以上の成績を発揮するモデルの研究報告が見られるようになっており¹⁵⁻¹⁷⁾、画像診断支援や病状分析においてAIは有用であると考えられる。AIの臨床現場への導入には判断根拠の提示、学習データの確保と学習の最適化などの課題もあるが、導入によって患者の早期診断・治療が改善され、読影・診断業務に係る諸問題に対する解決策になる可能性がある。

腹部CT画像診断においても非外傷性病変や、特にピットホールとされる急性腹症の検出を支援するAIが導入されれば、放射線科医師や専門外医師の読影業務を支援することで、病変の見落とし防止や読影精度の向上、読影時間の削減などの業務効率化も期待でき、その有用性は高いと推察される。特に、AIによる急性腹症のCT画像診断支援は、腹部臓器や病変の特徴から対象物を識別することが容易ではなく、かつ短時間に結果を出力する必要があるため積極的に研究対象にされていないと推察され、これまで腹部CT画像診断におけるAIを使用した急性腹症の病変検出の精度、成績や臨床応用の可能性などに関連する研究の現状を明らかにした報告はない。しかし、医師の負担軽減が一層求められる現在において、その手段の一つとなり得るAI活用の現状を明らかにしておくことは今後の研究に有用である。

2. 目的

本研究では、腹部CT画像診断におけるAIを用いた非外傷性病変の検出に関する文献レビューを行い、特に急性腹症に該当する疾患検出におけるAI利用の現状と課題を明らかにすることを目的とした。

3. 方法

3-1 文献収集方法と検索期間

文献のデータベース検索システムはPubMedおよびGoogle scholarを用いた。検索における適合基準は、腹部CT画像診断におけるAIを利用した非外傷性病変の検出支援に関する文献とした。今回の検索では、腹部領域における非外傷性病変のAIによる検出支援について、網羅的に現在の状況をレビューすること、また一部用いられたデータセット数が不明のものがあることから、データセットの件数による条件は設けないこととした。また除外基準は、①CT検査以外のモダリティ (病理学診断・内視鏡検査・超音波検査・MRI・生化学検査) を主な研究対象とした文献②腹部以外の領域 (肺・心臓・頭部) を対象とした文献③骨・筋・血管病変を対象としたもの——とした。以上により、AIを利用した腹部CT画像診断における非外傷性病変の検出支援に関する文献を抽出できるようにした。なお、検索期間はDLが最初に開発・提案され、医療分野に利用されるようになった2010年以降¹⁸⁾とし、2010年1月1日から2022年9月30日までとした。

適合基準および除外基準によるPubMedの検索項目をTable 1に示す。その結果、対象臓器ごとの検索数合計262件の文献を検索した。次いで検索された全ての論文について論文タイトルを確認し、AIによる各臓器のCT画像診断支援を対象とした論文のみ164件を抽出した。さらに採用された各文献の引用文献リストの論文を確認して、各基準を満たす文献を21件追加した。その後、著者が計185件の文献の要旨・抄録を読んでから、適合基準・除外基準を満たした文献と、要旨・抄録では判然としない文献については本文を読

Table 1 Search items in PubMed

Compliant items		"Computed tomography"
	AND	"Machine learning" OR "Computer aided"
AND		liver OR gallbladder OR pancreas OR kidney OR stomach OR "small intestine" OR colon OR appendix OR uterine OR ovary OR bladder
	NOT	pathology OR endoscope OR "blood data" OR us (ultrasound) OR mri (magnetic resonance imaging) OR "blood vessel" OR muscle OR bone OR lung OR heart OR head

Match and exclusion criteria items in PubMed and Google Scholar search formulas

み、病変検出を目的とする71件の文献を選別した。

Google Scholarでは、PubMed検索と同様に行い、検索された265件の文献全ての論文タイトルを確認し、PubMedと重複しない46件を抽出した。さらに著者が全ての要旨・抄録と必要に応じて本文を読み、適合基準・除外基準を満たした病変検出を目的とする35件の文献を選別した。

以上の結果、106件（肝臓疾患：52件、胆嚢病変：2件、膵臓病変：4件、腎臓病変：19件、胃病変：1件、小腸病変：2件、大腸病変：20件、虫垂病変：1件、子宮病変：2件、卵巣病変：1件、膀胱病変：2件）を文献レビューの対象として採用し整理した。さらにこの対象文献から、急性腹症疾患¹⁹⁾に該当する胆石症（胆嚢炎）、小腸疾患、腎・尿管結石、大腸炎、急性虫垂炎の5疾患、15件の文献^{10)、20)・33)}を抽出し、これらを整理した。

3-2 分析方法

1) 腹部病変の検出に関する文献の整理

抽出された文献について、各種臓器と器官、検出対象とする疾患名、文献数、モデルの性能評価について整理し、腹部臓器などの病変検出に対するAI利用の現状について概要を整理した。

2) 急性腹症疾患の検出に関する文献の整理

腹部病変の検出に関する文献のうち、急性腹症に該当する5疾患の検出に関する文献について、疾患名・研究目的、疾患データの収集に用いた患者数、使用データ数とラベル付け（トレーニング・テストデータの選定）の方法、性能評価、課題について文献ごとに整理した。

4. 結果

4-1 腹部病変の検出に関する研究の概要

文献レビューの対象としたAIを利用した腹部CT画像診断支援に関する文献検索の結果、対象とした腹部臓器や消化管が全て研究対象とされていた。病変の検出を目的とした106件の文献について臓器ごとに対象疾患、文献数、性能評価としてテストデータセットの精度、感度、特異度、AUC（Area Under the Curve）および病変のセグメンテーション評価のIoU（Intersection over Union）、ダイス係数（Dice coefficient）について整理しTable 2に示す。

腹部CT検査におけるAIによる病変検出に関する106件の文献のうち、腫瘍の検出に関する文献が86件

で最も多く、次いで結石の検出が7件、炎症疾患の検出が5件であった。

AIによるCT画像診断支援に関する文献のうち、13疾患（14項目）の文献で精度による評価が示され、感度が示されていたのが10疾患、特異度が9疾患、AUCが9疾患であった。その他、臓器のセグメンテーションと組み合わせて性能評価が行われていた肝腫瘍、腎腫瘍、小腸疾患に関する文献については、少数であるがIoUとダイス係数が示されていた。

4-2 各腹部臓器などにおける病変検出に関する文献について

各腹部臓器などにおいてレビュー対象となった疾患において、最も多かった腫瘍86件のうち肝腫瘍、腎細胞癌、大腸ポリープが76件と大半を占めていた。また結石については腎・尿管結石が7件のうち6件、炎症性疾患については5件のうち4件が大腸炎であった。

各臓器などにおける疾患とその文献数、およびモデルの信頼性や妥当性を示す性能評価については以下の通りであった。

1) 肝臓

肝臓病変の検出を目的とした文献を52件抽出した。そのうち肝腫瘍の検出に関する文献が47件、脂肪肝検出に関する文献が3件、肝炎または肝硬変検出に関する文献が2件であり、肝腫瘍の検出に関する研究が非常に多い結果であった。肝腫瘍の19文献のうち14文献が精度90%以上であり、おおむね高精度の病変検出モデルを構築していた。また脂肪肝検出については1文献が複数の条件に基づく検出を行っており、感度にバラツキがあるものの、特異度は0.942~0.957と高値であった。肝硬変などに関する文献ではAUC0.85~0.871であった。

2) 胆嚢

胆嚢病変の検出を目的とした文献は2件であった。それぞれ黄色肉芽腫胆嚢炎と胆石症の検出を目的としたものであった。また胆石症の検出精度については結石の性状によって精度に差が認められ、粒状結石で0.927と高値であった。

3) 膵臓

膵臓病変の検出を目的とした文献は、膵腫瘍の検出に関するものが4件で、うち精度評価は2件で0.90と0.97、その他、AUC0.75とダイス係数0.71がそれぞれ1件ずつであった。

4) 腎臓

腎臓病変の検出を目的とした文献は19件検索され

Table 2 Literature on AI-based diagnostic support for abdominal CT imaging

organs	disease	number of articles	accuracy	performance			
				sensitivity/ specificity	AUC	IoU	Dice coefficient
liver	Tumor	47	0.82 ~ 0.998	0.818 ~ 0.966 /0.909 ~ 1.0		0.954	0.625 ~ 0.963
	Fatty liver	3		0.295 ~ 0.914 /0.942 ~ 0.957			
	Cirrhosis	2			0.85 ~ 0.871		
gall bladder	cholecystitis	1	0.837		0.817		
	Gallstone	1	granular 0.927 sludge 0.803				
pancreas	Tumor	4	0.902 ~ 0.972	0.945 ~ 0.971 /0.977 ~ 0.975	0.75 ~ 0.945		0.71
kidney	Renal cell carcinoma	13	0.739 ~ 0.998	0.64 ~ 0.98 /0.83 ~ 0.93	0.803 ~ 0.96	0.938	0.63 ~ 0.82
	Renal stone	5	0.63 ~ 0.997	0.86 ~ 0.959 /0.91 ~ 0.98	0.95		
	Ureter stone	1			0.997		
stomach	cancer	1	0.769				
small intestine	Small Intestinal obstruction	1	0.73	0.72 /0.80	0.73		
	Small intestine lesion	1				0.86	
colon	Polyps / cancer (Colonography)	16		0.62 ~ 1.00 /0.85	0.85 ~ 0.91		
	Colitis	4	0.70	0.727 ~ 0.937 /0.733 ~ 0.95	0.986		
appendix	Appendicitis	1	0.915	0.902 /0.92			
bladder	Bladder cancer	2	0.892	0.849			
Uterus	Uterine cancer	2	0.8 ※				
Ovary	Ovarian cancer	1	0.95 ~ 0.97				

※ 1 文献は PET (Positron Emission Tomography) との精度比較で平均 6.06% 向上

Of the 106 articles on AI-based lesion detection in abdominal CT, 86 were related to the detection of tumors, followed by 7 for stone detection and 5 for inflammatory diseases. Regarding performance evaluation, 13 diseases (14 items) were evaluated by accuracy, 10 diseases by sensitivity, 9 diseases by specificity, 9 diseases by AUC.

た. そのうち腎結石が5件, 尿管結石が1件, 腎細胞癌が13件であった. 性能評価については腎結石が精度0.63~0.997, 感度0.86~0.959, 特異度0.91~0.98, AUC0.95, 尿管結石がAUC0.997, 腎細胞癌が精度0.739~0.998, 感度0.64~0.98, 特異度0.83~0.93, AUC0.803~0.96であった.

5) 小腸

小腸疾患の検出を目的とした文献は2件で, 小腸閉塞と小腸病変(小腸リンパ腫, 炎症性腸疾患, 小腸非リンパ腫)の検出を目的としたものであった. 性能評価では小腸閉塞が精度0.73, 感度0.72, 特異度0.80, 小腸病変検出がIoU0.86であった.

6) 大腸

大腸病変の検出を目的とした文献は20件であった. 大腸炎検出が4件で, うち3件は同一著者であった. 性能評価は精度0.7, 感度0.727~0.937, 特異度0.733~0.95であった. その他16件は全てコロノグラフィーによる大腸癌やポリープの検出で, 感度0.62~1.00, 特異度0.85であった.

7) 虫垂

虫垂病変の検出を目的とした文献は, 急性虫垂炎の

検出に関するものが1件で, 精度0.915, 感度0.902, 特異度0.92であった.

8) 膀胱

膀胱病変の検出を目的とした文献は, 膀胱癌の検出が2件で, それぞれ精度0.892と感度0.85であった.

9) 子宮

子宮病変の検出を目的とした文献は, 子宮頸癌の検出が2件で, 精度評価されたもので0.80であった.

10) 卵巣

卵巣病変の検出を目的とした文献は, 卵巣癌の検出のものが1件で, 精度0.95~0.97であった.

各臓器などの計18疾患を対象とした文献のうち, 精度が示された13疾患(14項目)の全ての文献の精度を疾患ごとにプロットしたものをFig.1に示す. 腫瘍および結石症(胆泥を除く)の文献はいずれも85%以上の高い精度を示していた. 一方, 消化管疾患については虫垂炎の文献を除いて全て70~77%であり, 相対的に低い傾向であった.

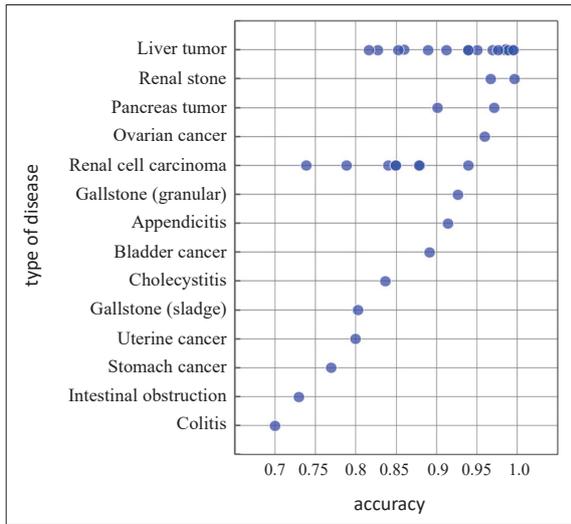


Fig.1 Accuracy comparison of literature (by disease)

Accuracy comparison of literature on 13 diseases that showed accuracy out of a total of 18 diseases of each organ. The literature on tumors and stone diseases showed a high accuracy of more than 85%, but all gastrointestinal diseases were 70~77% except the literature on appendicitis.

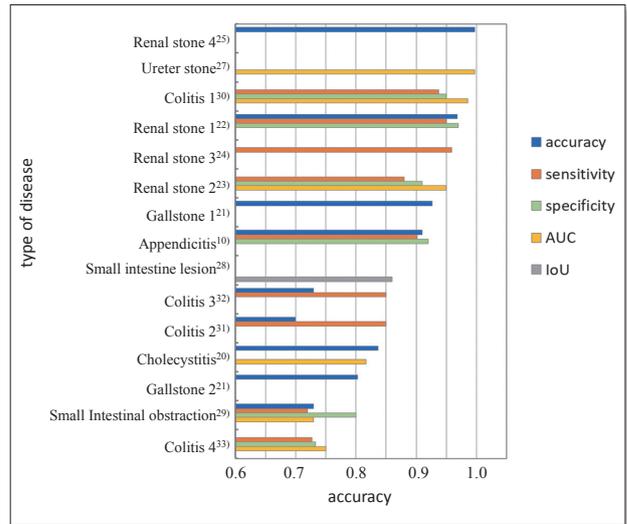


Fig.2 Performance of each literature on the detection of acute abdominal illness

The performance on renal and ureteral stones and granular cholelithiasis was generally higher than 90%, and colitis, cholecystitis, biliary mud, small bowel lesions, and small intestinal obstruction tended to be low.

4-3 急性腹症疾患の検出に関する文献について

急性腹症に該当する5疾患の検出を目的とした15件について、疾患名、文献番号、文献タイトル、目的、疾患データを抽出した患者数、使用データ数、データラベル付けの方法、モデル性能、課題について個別に整理し、Table 3にまとめた。また性能評価において精度が示された胆嚢炎、胆石、腎結石、小腸閉塞、大腸炎、虫垂炎の9件、感度が示された腎結石、小腸閉塞、大腸炎、虫垂炎の9件、特異度が示された腎結石、小腸閉塞、大腸炎、虫垂炎の6件、AUC評価が示された胆嚢炎、腎結、尿管結石、小腸閉塞、大腸炎の6件について、Table 3の各疾患における評価方法と性能値をFig.2にまとめた。いずれの評価方法においても腎・尿管結石と粒状胆石症の性能がおおむね90%以上と高く、大腸炎、胆嚢炎、胆泥、小腸病変、小腸閉塞が低い傾向であった。消化管疾患の文献の中で、大腸炎の1文献 (colitis1) が感度、特異度、AUCが95%を超えており、虫垂炎は精度、感度および特異度が90%を超えていた。

トレーニングやテストに使用されたデータについては、対象患者数やデータ数と疾患の間に明らかな関連性は認められなかった。トレーニングデータ数については、腎結石の1文献が90画像で最も少なく、交差検証も行っていない。またテストデータについては、大腸炎の2文献がそれぞれ40画像と44画像、小腸閉塞が

44画像、尿管結石が88画像と少なく、いずれも交差検証は行っていない。また全ての文献でデータのラベル付けは放射線科医師または専門医が行っていた。

各文献における課題については、胆嚢炎検出では、炎症とがんの合併症があった場合は判別ができないと報告されていた。腎・尿管結石の検出では、6文献中4文献においてデータセットが十分でないことや、動脈硬化による血管の石灰化と誤認することを共通した課題として挙げていた。大腸炎については、解剖学的特徴から大腸の誤検出が生じやすいこと、またCTのスキャンスライス厚 (解像度) に影響を受けること、大腸炎による壁肥厚が十分でない場合は偽陰性になることが報告されていた。虫垂炎については、手動で作成した虫垂領域のデータを使用しており、臨床使用では虫垂領域のローカリゼーションと分類アルゴリズムの開発が必要であると報告されていた。

5. 考察

5-1 腹部CT画像診断におけるAI活用に関する研究の現状

AIを利用した腹部CT画像診断支援に関する文献は231件検索され、病変検出を目的とした106件以外の125件の文献の研究目的は、臓器のセグメンテーションや臓器体積の予測、疾患の病期や重症度の推定、他

Table 3 Literature on CT image diagnosis support by AI for acute abdomen

disease	reference	theme	purpose	number of patients	data set (images)	labeling	accuracy	sensitivity/specificity	AUC	IoU	task	
cholecystitis	20	Machine Learning-Based Radiological Features and Diagnostic Predictive Model of Xanthogranulomatous Cholecystitis ; Qiao-Mei Zhou. et al (2022)	黄色肉芽腫性胆嚢炎と胆嚢がんのCT画像特性から鑑別診断モデルを確立する	153	training 153	radiologist	0.837		0.817		黄色肉芽腫性胆嚢炎と胆嚢炎が合併している場合は区別できない	
Gallstone 1	21	A novel YOLOv3-arch model for identifying cholelithiasis and classifying gallstones on CT images ; Shanchen Pang. et al (2019)	胆石症を検出し胆石を分類する	1389	training 4000 inspection 986 test 1000	radiologist	granular 0.927					胆泥結石は粒状胆石より検出精度が低下する
Gallstone 2							sludge 0.803					
Renal stone 1	22	Deep learning model for automated kidney stone detection using coronal CT images ; KadirYildirim. et al (2021)	深層学習技術を用いて、CT画像を使用した腎結石の自動検出	433	training 1163 inspection 290 test 346	radiologist urologist	0.968	0.95 / 0.97			多様なデータセットを使用する必要がある	
Renal stone 2	23	A deep learning system for automated kidney stone detection and volumetric segmentation on noncontrast CT scans ; Daniel C. et al (2022)	高精度再構成による高ノイズCT画像における腎結石の検出とサイズの定量化	91	training 90 inspection 6185 test 90	training&test : radiologist inspection : external facility		0.88 / 0.91	0.95		腎洞の小さなアテローム性動脈硬化症のプラークは偽陽性の原因となる	
Renal stone 3	24	Automatic Detection and Scoring of Kidney Stones on Noncontrast CT Images Using S.T.O.N.E. Nephrolithometry: Combined Deep Learning and Thresholding Methods ; Yingpu Cui. et al (2021)	腎結石検出およびスコアリングのための深層学習ベースのモデル開発	131	segmentation 334 classification 564 test 234	radiologist			0.959		<ul style="list-style-type: none"> 多くの施設から多様なデータを利用する必要がある 術後や異常のある腎はデータから除外している 動脈硬化による血管の石灰化が誤検出される 結石の組成が検出に影響する可能性がある 	
Renal stone 4	25	Exemplar Darknet19 feature generation technique for automated kidney stone detection with coronal CT images ; Mehmet Baygin. et al (2022)	新たな転移学習ベースの画像分類により、CT画像から腎結石を検出する方法 (ExDark19) の提案	165	training 1453 test 346	open data	0.997				<ul style="list-style-type: none"> 多様なデータセットを使用する必要がある 腎結石の位置を示すことはできない 	
Renal stone 5	26	Deep learning model-assisted detection of kidney stones on computed tomography ; Caglayan A. et al (2022)	腎結石の検出におけるディープラーニングモデルの有用性	455	training 953 test 200	radiologist	0.63 ~ 0.93				<ul style="list-style-type: none"> 結石の組成が考慮されていない 治療の成功を予測するAIアルゴリズムの効果を検証していない 	
Ureter stones	27	Computer aided detection of ureteral stones in thin slice computed tomography volumes using Convolutional Neural Networks ; MartinLångkvist. et al (2018)	CT画像から尿管結石を識別するためのアルゴリズム開発	465	training 349 test 88	radiologist			0.997		多くの施設の多様なデータを利用する必要がある。自然排斥されるような小結石や遠位尿管に達することがない大きな結石はデータから除外している	
Small intestine lesion	28	Segmentation and Identification of Small Intestinal Lesions Based on U-Net with Batch Normalization and Virtual Sample ; Xufeng Liu. et al (2021)	炎症性腸疾患などの小腸病変のU-Netに基づく小腸セグメンテーションと病変の識別	143	training 143 test 143	experienced doctor				0.86	U-Netの構造をコーディングやデコーディングすることによるデータ情報の一部損失を軽減することが必要	
Small Intestinal obstruction	29	Machine Learning Based Prediction Model for Closed-Loop Small Bowel Obstruction Using Computed Tomography and Clinical Findings ; Goyal, Riya MD. et al (2022)	CT画像と臨床所見を統合した小腸閉塞の予測モデルの開発	223	training 179 test 44	radiologist	0.73	0.72 / 0.80	0.73		ほとんどのデータが手術所見に基づいた小腸閉塞症例の陽性症例で陰性、偽陰性症例が制限されていた	
Colitis 1	30	Detection and diagnosis of colitis on computed tomography using deep convolutional neural networks ; Jiamin Liu. et al (2017)	腹部CTスキャンによる大腸炎診断支援のためのSVM分類器の提案	80	training 3003 test 1001	radiologist		0.937 / 0.95	0.986		<ul style="list-style-type: none"> 結腸の形態的特徴から炎症領域を検出することは困難 結腸以外の臓器の誤検出対策が必要 スキャンスライス厚が検出に影響する 	
Colitis 2	31	Colitis Detection on Abdominal CT Scans by Rich Feature Hierarchies ; Jiamin Liu. et al (2016)	畳み込みニューラルネットワークを利用してCT画像から大腸炎を検出	26	training 220 test 40	radiologist	0.7		0.85		記述なし	
Colitis 3	32	Colitis Detection on Computed Tomography using Regional Convolutional Neural Network ; Jiamin Liu. et al (2017)	畳み込みニューラルネットワークを利用してCT画像から大腸炎を検出	56	training 404 test 44	radiologist	0.73		0.85		便を伴う結腸、腎臓および肝臓は大腸炎と類似しているため誤検出される可能性がある	
Colitis 4	33	Computer-Aided Detection of Colitis on Computed Tomography Using a Visual Codebook ; Zhuoshi Wei. et al (2013)	大腸炎検出のためのコンピュータ診断支援の提案	22	training 1000 test 1000	radiologist		0.727 / 0.733	0.75		アコーディオン兆候が明確でない場合や病変領域が小さい場合に偽陰性が生じる	
Appendicitis	10	Convolutional-neural-networkbased diagnosis of appendicitis via CT scans in patients with acute abdominal pain presenting in the emergency department ; Jin Joo Park. et al (2020)	虫垂炎の診断	215	training 584 test 83 external inspection 100	radiologist	0.91	0.902 / 0.92			<ul style="list-style-type: none"> 虫垂切除後や虫垂腫瘍のデータを除外している 放射線科医が手動抽出した虫垂領域のデータを使用したため、臨床利用では領域ローカリゼーションと分類アルゴリズムの開発が必要 	

Fifteen articles aimed at detecting acute abdomen found no clear association between the number of patients or data used for training or testing and disease. In all articles, data were labeled by radiologists or specialists. In addition, one of the most common problems with each article was that the dataset was insufficient.

疾患との鑑別、良悪性の鑑別、がん転移や再発予測など多岐にわたっていた。また病変検出を目的とした106件から急性腹症15件を除いた91件のうち、86件を腫瘍・ポリープ検出が占めていた。

腫瘍やポリープの検出は虫垂以外の臓器に共通しており、特に肝腫瘍の検出を目的とする文献数が47件で他に比べて非常に多く、これは10mm径以上の肝腫瘍の診断において造影CT検査が推奨されていることや³⁴⁾、肝臓が比較的大きく、肝実質のCT値が均一で腫瘍を判別しやすいことがAIによる検出に適していることが考えられた。肝腫瘍の次に多いのは腎・尿管結石の6件で、結石のCT値が周囲軟部組織と大きく異なるためAIによる判別に適していると考えられた。

大腸炎の文献は4件であったが、このうち3件が同一著者であり、研究者としては2人であった。また同一著者の3件に使用された説明図は同一のものであったが、データが同一のものであるかは不明であった。大腸は盲腸から直腸まで非常に長く走行が複雑で、かつ個人差があり、特に横行結腸やS状結腸は可動性に富み、経時的に形態が変化する。また上行結腸・下行結腸は後腹膜に固定され可動性が低いとはいえ、便やガスなどの内容物により容易に形態が変化する。こうした消化管の特徴はAIによる分類と病変検出を難しくしている^{30,33)}。そのため全ての文献でヨード造影剤を使用したCT画像を用いていたが、ヨード造影剤の投与についてはアレルギー反応や腎機能などに注意が必要である。また経口造影剤については閉塞や穿孔の有無を確認し慎重に投与する必要がある。

婦人科疾患や胆嚢疾患の文献数が少ないのは、婦人科疾患の画像検査では、CT検査はがんの病期を評価するための主要な方法であるが、正確な診断を行うためにはヨード造影剤を使用し、目的によっては経口造影剤を投与する必要がある³⁵⁾、消化管検査と同様に慎重な投与が必要である。またCT検査は放射線被ばくの問題や、MRIと比較して軟部組織のコントラストが低いという欠点があることから³⁵⁾、一般的にMRIやUSが優先されるためであると考えられた。また急性胆嚢炎の診断については、超音波検査が低侵襲性、普及度、簡便性、経済性などから第一選択的検査法とされており、CT検査は臨床所見や他検査によって確定診断が困難な場合や局所合併症が疑われる場合において、造影ダイナミックCTによって施行することが推奨されているため³⁶⁾、CTが疾患検出の方法として優先されないためであることが考えられた。

5-2 急性腹症疾患のCT画像診断におけるAI研究の現状

今回、急性腹症に該当する疾患として検索されたのは5疾患であったが、5疾患以外に緊急手術を要する腹部大動脈破裂、肝がん破裂、異所性妊娠、重症急性胆管炎、内蔵動脈瘤破裂をはじめ、胃・十二指腸潰瘍、膵炎、腎梗塞、腎盂腎炎、卵巣捻転、尿路感染など¹⁹⁾があり、いずれも検索されなかった。これらの疾患は、緊急手術適用のため時間的制約があるもの、画像識別での特徴に乏しいもの、救急医療の中では重要性が比較的高くないものなど、さまざまな理由から研究対象になっていないと推察された。

現状における急性腹症疾患のCT画像診断におけるAI研究について、Table 3に整理した結果を基に「放射線科医師によるラベル付け」「疾患と性能」「同一疾患の文献における性能差」「性能向上のための課題」の観点から考察した。

1) 放射線科医師によるラベル付けについて

救急医療において急性腹症の対象疾患である胆石症(胆嚢炎)、急性膵炎、小腸閉塞、尿管結石、大腸炎、急性虫垂炎の診断は、CT画像から速やかに異常を検出することが求められ、特に放射線科医師不在時に正確な診断を行うためにAIへの期待は高い。トレーニングやテストデータとする画像のラベル付けは、ほぼ全ての文献において放射線科医師が行っており、AIの診断精度の目標を放射線科医師の診断精度相当としていた。本邦においても、厚生労働省の「人工知能(AI)を用いた診断、治療等の支援を行うプログラムの利用と医師法第17条の規定との関係について」(平成30年12月19日 医政医発1219第1号)に示されている通り、AIを利用した診断・治療支援を行うプログラムを使用して診療を行う場合、診断、治療等を行う主体は医師であり、医師はその最終的な判断の責任を負うとされており、その上で画像診断における画像データのラベル付けは、画像診断を専門とする放射線科医師によって行われることが適切である。

2) 疾患と性能について

腎・尿管結石の文献は、いずれも非常に高い精度、感度、特異度、AUCを示しており、臨床への活用が十分に期待される。また胆石も粒状胆石については精度0.93と高い精度を示しており、結石症は臓器分類が高精度であれば、結石と周辺組織との大きなCT値差が検出に有用と考えられ、高い精度のモデルを構築でき

ると考えられた。一方、大腸炎検出の研究では1文献を除き、精度、感度共0.7~0.85と相対的に低かった。大腸炎の文献はいずれも大腸壁の厚さによって大腸炎を識別しているため³⁰⁻³³⁾、大腸の形態的特徴から炎症領域を検出することが他臓器に比べて難しく、周辺臓器との分類不良による誤検出も生じやすい^{30,32)}。また壁肥厚が軽度の場合は偽陰性になりやすいことや³³⁾、CTのスキャンスライス厚(解像度)が判別に影響することがモデル構築での課題とされていた³⁰⁾。こうした複数の要因が、他臓器疾患の文献に比べて高い性能を実現することが難しい要因である。さらに正常と異常を判別する大腸壁の厚さをどのように定義するのか、また壁の厚さを認識するために使用されていた経口造影剤やヨード造影剤は、患者の身体的負担やアレルギーなどの副作用を考慮しなければならず、検索された文献では臨床において造影剤を使用できない場合は大腸炎を認識することができない。そうしたことから、大腸壁の厚さ以外の検出方法の検討が必要であると考えられた。虫垂はサイズが小さく形態も複雑であり、CT画像から検出することは難易度が高い器官であるが、今回、検索した文献では放射線科医師が手動で虫垂領域を抽出したデータを使用していたため、90%以上の精度、感度、特異度が示されていた。

3) 同一疾患の文献における性能差について

大腸炎の同じ著者の3文献において、colitis 1の文献に対してcolitis 2, colitis 3の文献の感度が低かった。これは、colitis 1の文献が16層VGG (Visual geometry group) ネットを使用したFaster R-CNN (Faster region based convolutional neural networks) モデルであるのに対して、colitis 2, colitis 3はAlexNetを使用したR-CNNモデルで、さらに患者数、データ数共にcolitis 1より少ないことが要因であると考えられ、同じ疾患の検出であっても、アルゴリズムの選択とトレーニング・テストデータ数がモデル性能に影響すると考えられた。

4) 性能向上のための課題について

性能向上の課題として、腎・尿管結石の4文献^{22,24-26)}が他施設のデータを用いることを挙げており、より多く多様なデータでトレーニングすることによって、さらなる性能向上が望めることを示唆していた。大腸炎の文献は腎・尿管結石の文献よりもデータ数が少ないにもかかわらず、データ数やその多様性について課題として記載されていなかったが、大腸の解剖学的特徴

から腎臓以上に多様なデータでトレーニングする必要があると考えられた。

6. 結語

腹部CT画像診断におけるAIを利用した非外傷性病変の検出に関する文献レビューを行った結果、対象とした全ての腹部臓器、消化管の病変検出についてAIによるCT画像診断支援の研究が行われており、病変検出を目的とするものは、肝腫瘍、腎細胞癌、大腸ポリープが大半を占めていた。モデルの性能については腫瘍や結石症の文献において高い性能を示していたが、消化管病変の検出性能は相対的に低い傾向であった。

急性腹症の病変検出性能の向上のためには、適切なアルゴリズムの選択と多様なデータをより多く用いることが必要である。今回、急性腹症に該当する疾患検出モデルについては、胆嚢炎、腎・尿管結石、小腸疾患、大腸炎、虫垂炎以外は検索されなかったが、他の急性腹症疾患についても今後の研究が望まれる。消化管の炎症性疾患などについては相対的に検出精度が低く、最も検出精度が低かった大腸炎の検出においては、周囲臓器の誤検出や肥厚した大腸壁の偽陰性を避けるため、腸管壁の厚さを検出要件としないモデル構築の研究に期待したい。

利益相反

筆頭著者および共著者全員に開示すべき利益相反はない。

表の説明

Table 1	PubMedの検索項目 PubMedおよびGoogle Scholarの検索式における適合基準と除外基準の項目
Table 2	AIによる腹部CT画像診断支援に関する文献 腹部CT検査におけるAIによる病変検出に関する106件の文献のうち、腫瘍の検出に関する文献が86件で最も多く、次に結石の検出が7件、炎症疾患が5件であった。また性能評価については13疾患(14項目)の文献で精度による評価が示され、感度評価が10疾患、特異度が9疾患、AUCが9疾患であった。
Table 3	急性腹症に対するAIによるCT画像診断支援に関する文献 急性腹症の検出を目的とした15件の文献において、トレーニングやテストに使用された対象患者数やデータ数と疾患の間に明らかな関連性は認められなかった。全ての文献でデータのラベル付けは放射線科医師または専門医が行っていた。また各文献の課題として多かったのはデータセットが不十分であることであった。

図の説明

- Fig.1 文献の精度比較 (疾患別)
各臓器などの計18疾患のうち、精度が示された13疾患の文献の精度比較。腫瘍および結石症の文献は85%以上の高い精度を示していたが、消化管疾患は虫垂炎の文献を除いて全て70~77%であった。
- Fig.2 急性腹症の検出に関する各文献の性能
腎・尿管結石と粒状胆石症の性能がおおむね90%以上と高く、大腸炎、胆嚢炎、胆泥、小腸病変、小腸閉塞が低い傾向だった。

参考文献

- 1) 穂山雄次:救命救急医療におけるマルチスライスCTの役割. 放射線撮影分科会誌, 39, 8-10, 2002.
- 2) Cheng Fang, et al.: Radiological anatomy of the abdomen. SURGERY, 39(2), 61-74, 2021.
- 3) 鈴木卓也, 他:急性腹症の患者の診断におけるCT検査の有用性に関する研究. 日腹部救急医学会誌, 30, 875-81, 2010.
- 4) David J. Winkel, et al.: Evaluation of an AI-Based Detection Software for Acute Findings in Abdominal Computed Tomography Scans. Investigative Radiology, 54, 55-59, 2019.
- 5) Nakajima Y, et al.: Radiologist supply and workload: international comparison. Radiation medicine, 26(8), 455-465, 2008.
- 6) Matsumoto M, et al.: Geographic distribution of radiologists and utilization of teleradiology in Japan: a longitudinal analysis based on National Census Data. PLoS One, 10(9), 2015.
- 7) 川本誠一:画像診断とAI (人工知能). 大阪急総医誌, 41 (1), 3-19, 2019.
- 8) 藤田広志:AI画像診断の全体像と将来の展望-医師を助ける"第三の目"-. 情報処理, 62 (2), 1-8, 2021.
- 9) 高本聖也, 他:診療放射線技師による時間外救急CTの緊急所見報告の効果. 日本臨床救急医学会雑誌, 22 (5), 703-708, 2019.
- 10) Jin Joo Park, et al.: Convolutional-neural-network-based diagnosis of appendicitis via CT scans in patients with acute abdominal pain presenting in the emergency department. Scientific Reports, 10, 9556, 2020.
- 11) 恩田秀賢, 他:当院高度救命救急センターにおける高齢者医療の現状. 日医大医会誌, 9 (2), 129-134, 2013.
- 12) 加藤博之:高齢者に多い救急疾患. 日本老年医学会雑誌, 48 (4), 312-316, 2011.
- 13) 蛭名正智, 他:ER型救命センターでのCT画像診断における診断精度の現状と質向上の試み. 日臨救医誌, 18, 1-4, 2015.
- 14) 中田典生:画像診断における人工知能活用の現状と展望. 日腎会誌, 59 (7), 1064-1070, 2017.
- 15) Adam Yala, et al.: A Deep Learning Model to Triage Screening Mammograms: A Simulation Study. Radiology, 293(1): 38-46, 2019.
- 16) Zhichao Feng, et al.: Machine learning based on clinical characteristics and chest CT quantitative measurements for prediction of adverse clinical outcomes in hospitalized patients with COVID-19. European Radiology, published online, 31(10), 7925-7935, 2021.
- 17) Becker AS, et al.: Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. Invest Radiol, 52, 434-440, 2017.
- 18) 山根友絵, 他:日本における医療分野での人工知能 (AI) 活用に関する文献検討. 豊橋創造大学紀要, 25, 61-70, 2021.
- 19) 急性腹症診療ガイドライン出版委員会:急性腹症診療ガイドライン2015. 医学書院, 20, 2015.
- 20) Zhou QM, et al.: Machine Learning-Based Radiological Features and Diagnostic Predictive Model of Xanthogranulomatous Cholecystitis. Front Oncol, 24(12), 792077, 2022.
- 21) Pang S, et al.: A novel YOLOv3-arch model for identifying cholelithiasis and classifying gallstones on CT images. PloS One, Jun 18, 14(6), 2019.
- 22) Yildirim K, et al.: Deep learning model for automated kidney stone detection using coronal CT images. Comput Biol Med, Aug, 135, 2021.
- 23) Elton DC, et al.: A deep learning system for automated kidney stone detection and volumetric segmentation on noncontrast CT scans. Med Phys, 49(4), 2545-2554, 2022.
- 24) Cui Y, et al.: Automatic Detection and Scoring of Kidney Stones on Noncontrast CT Images Using S.T.O.N.E. Nephrolithometry: Combined Deep Learning and Thresholding Methods. Mol Imaging Biol, Jun; 23(3), 436-445, 2021.
- 25) Baygin M, et al.: Exemplar Darknet19 feature generation technique for automated kidney stone detection with coronal CT images. Artif Intell Med, May, 127, 2022.
- 26) Caglayan A, et al.: Deep learning model-assisted detection of kidney stones on computed tomography. Int Braz J Urol, Sep-Oct, 48(5): 830-839, 2022.
- 27) Långkvist M, et al.: Computer aided detection of ureteral stones in thin slice computed tomography volumes using Convolutional Neural Networks. Comput Biol Med, Jun 1, 97, 153-160, 2018.
- 28) Liu X, et al.: Segmentation and Identification of Small Intestinal Lesions Based on U-Net with Batch Normalization and Virtual Sample. In: 2021 6th International Conference on Image, Vision and Computing (ICIVC). IEEE, 134-139, 2021.
- 29) Goyal R, et al.: Machine Learning Based Prediction Model for Closed-Loop Small Bowel Obstruction Using Computed Tomography and Clinical Findings. J Comput Assist Tomogr, Mar-Apr 01, 46(2), 169-174, 2022.
- 30) Liu J, et al.: Detection and diagnosis of colitis on computed tomography using deep convolutional neural networks. Med Phys, Sep, 44(9), 4630-4642, 2017.
- 31) Liu J, et al.: Colitis detection on abdominal CT scans by rich feature hierarchies. In: Medical Imaging 2016: Computer-Aided Diagnosis, SPIE, 9785, 423-429, 2016.
- 32) Liu J, et al.: Colitis detection on computed tomography using regional convolutional neural networks. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), IEEE, 863-866, 2016.
- 33) Wei Z, et al.: Computer-aided detection of colitis on computed tomography using a visual codebook. In: 2013 IEEE 10th International Symposium on Biomedical Imaging, IEEE, 141-144, 2013.
- 34) Azeez Osho, et al.: Role of imaging in management of hepatocellular carcinoma: surveillance, diagnosis, and treatment response. Hepatoma Res, 6,55, 2020.
- 35) Elzbieta Luczynska, et al.: Diagnostic imaging in gynecology. Ginekologia Polska, 93(1), 63-69, 2022.
- 36) 急性胆管炎・胆嚢炎診療ガイドライン改訂出版委員会:急性胆管炎・胆嚢炎診療ガイドライン2018. 99, 医学図書出版株式会社, 2018.