HOKKAIDO UNIVERSITY

Title	胃X線画像を用いた機械学習に基づく胃がん発症リスクの自動分類に関する研究［論文内容及び審査の要旨］
Author(s)	石原, 賢太
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第13086号
Issue Date	2018-03-22
Doc URL	https://hdl.handle.net/2115/70550
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/4.0/
Type	doctoral thesis
File Information	Kenta_Ishihara_abstract.pdf, 論文内容の要旨

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称　博士（情報科学）　氏名　石原 賢太

学 位 論 文 題 名

胃 X 線画像を用いた機械学習に基づく胃がん発症リスクの自動分類に関する研究
（Machine Learning-based Automatic Classification of Gastric Cancer Risk from X-ray Images）

本論文は, 機械学習に基づいて胃 X 線画像から胃がん発症リスクを自動で分類する研究を行い, その成果をまとめたものである.

我が国のがん死亡者数の第 3 位は, 胃がんである. その主な発症リスクは,Helicobacter pylori(以降,H. pylori) の感染に伴う胃炎により高まることが, 医学的研究から判明している. さらに,H. pylori 除菌治療により胃がん発症リスクが減少することも判明している. これを受けて我が国では, 胃がんの死亡率の減少を目的に, リスク診断を多くの対象者に対して行うための方針が打ち出された. 現状のリスク診断では, 医師は血液から取得したリスクの層別化に有効な測定値に着目しながら, 胃 X 線検査および胃内視鏡検査といった画像検査を行う. 胃 X 線検査においては, 規定の複数体位より撮影された胃 X 線画像に対して, ひだの形や粘膜面の模様の変化として現れる症状に注目しながら, 複数枚の画像を総合的に診断する. さらに,2 名以上の医師による読影が原則化されているため, 医師にかかる負担は大きい.

医師の負担軽減のためには, 機械学習を用いて画像を解析し, 医師の診断を支援することが有効である. 既に, マンモグラフィを用いた乳がん検診においては, 機械学習に基づく解析手法を搭載した診断支援システムが普及し, 一定の有効性が確認されている. 一方, 胃 X 線画像を用いた胃がん発症リスクの診断支援システムは, 実用化には至っていないものの, 胃のひだの症状を捉える画像特徴や特徴選択に基づくリスク検出手法が提案されている. しかしながら, 従来手法では, 解析対象が単一体位の画像のみに留まっており, 医師が診断時に着目する他の体位の画像や血液の測定値といったデータを十分に考慮した手法は実現されていない. そのため上記データから, 機械学習を活用して胃がん発症リスク分類に重要な情報を抽出することで, 更なる精度向上が期待される.

そこで本論文では, 単一体位の胃 X 線画像のみを分析対象とした手法の精度限界を打破するために, まずは,1) 複数体位の胃 X 線画像の総合的な判断に基づく胃がん発症リスクの検出手法を提案する. 続いて,2) 複数体位の胃 X 線画像に加えて, 血液の測定値を効果的に活用して胃がん発症リスクの段階を分類する手法を提案する. さらに,3) 深層学習に基づき画像特徴の弁別性を高め, 胃がん発症リスク検出の高精度化を実現する手法を提案する. 以下で各々の手法の詳細について説明する.

まず 1) では, 複数体位の画像を総合的に判断するために, 各体位に現れている症状に加え, 各体位のリスク検出に対する重要度を考慮する. これは, 体位間で詳細に描出される症状が異なるため, 医師は各体位に合った症状の評価を行い, 複数体位の画像に対する総合的な診断を行うためである. 具体的に 1) では, 各体位の画像に現れている症状を捉えるために, 胃 X 線画像から算出した複数種の画像特徴を Multiple Kernel Learning に基づき, 各特徴の重要度を考慮しながら統合し, 体位毎のリスク検出器を構築する. また, 複数体位の総合的判断を行うために, 各体位のリスク検出に対する重要度として検出器の精度を考慮しながら複数の検出結果を統合する. 特徴毎の重要度および, 各体位のリスク検出に対する重要度を考慮することにより, 各体位に合った症状を重視することに加え, 体

位間の関係を総合的に評価した胃がん発症リスク検出が可能となる.

次に 2) では, 高精度なリスク分類のために, 血液の測定値を活用して, 画像特徴からリスクの変化を捉えた特徴を抽出する. これは, リスクの増加に伴い, 画像中の症状および血液の測定値に変化が現れることが判明したためである. 具体的に 2) では, 画像特徴と血液の測定値間の相関関係を, カーネル正準相関分析により解析することで, 胃がん発症リスクの変化に関連が高い画像特徴を抽出する. この特徴抽出を撮影体位毎に行い, リスク分類器を構築する. さらに, 複数の分類結果を体位毎の精度を考慮しながら統合することで, 胃がん発症リスクを分類する. これらの処理を「発症リスクの有無の分類時」および「発症リスクの度合いの分類時」に対して階層的に適用することで, 高精度な胃がん発症リスクの段階分類が可能となる.

最後に 3) では, 各体位における発症リスク分類の精度向上を目的に, 深層学習の一種である Deep Convolutional Neural Network を活用して弁別性の高い特徴を抽出する. 通常, 深層学習には大量の学習データが必要とされるため, データの取得が容易ではない医用画像においては, 直接, 深層学習を適用することは困難である. そこで, 医用画像解析においては, 学習データ量を増やすために, 原画像を複数の小さな画像 (以降, パッチ) に分割する方法が広く採用されている. ただし, 胃がん発症リスクの診断時には, 発症リスクのみを記録するため, 医師の診断に寄与した画像中の領域情報は存在しない. そこで 3) では, 胃がん発症リスクとパッチ間の関連度を定義し, それに基づき発症リスクと関連の高いパッチを選択する. さらに, 選択したパッチから症状全体を捉えた特徴を抽出する. これにより, 画像に対する領域情報を学習データとして使用せずに, 発症リスクに関連する領域のみからの弁別性の高い特徴抽出が可能となり, 胃がん発症リスク検出の高精度化が実現される.

以下に本論文の構成を示す. 第 1 章では, 本論文の研究背景および目的を述べる. 第 2 章では, 本研究の関連研究として機械学習を用いた医用画像の解析手法を紹介し, 本論文で解決すべき問題を明らかにする. 第 3 章では, 複数体位の胃 X 線画像を用いた胃がん発症リスクの検出手法を提案する. 各体位の画像に現れている症状の把握に加え, 各体位のリスク検出に対する重要度の考慮によって, 胃がん発症リスクを高精度に検出可能であることを実験により確認する. 第 4 章では, 血液の測定値と画像特徴間の関連性を活用した胃がん発症リスクの段階の分類手法を提案する. 両者の関連性を活用することで, 胃がん発症リスクの有無のみでなく, リスクの度合いまで高精度に分類可能であることを実験により確認する. 第 5 章では, 深層学習を用いた領域抽出に基づいて胃がん発症リスク検出を高精度化する手法を提案する. 実験では, まずリスクに関連する領域が抽出可能であるかを確認する. 次に, 抽出した領域のみから得られる特徴によって, 検出精度が向上することを確認する. 第 6 章では, 本研究の成果を要約し, 論文全体のまとめとする.

以上を要約すると, 本論文は, 複数枚の胃 X 線画像や血液の測定値より, 機械学習を活用して医師の診断基準に基づいた重要な情報を抽出する手法を提案し, 胃がん発症リスク分類を実現している. また, 深層学習の効果的な導入による高精度化を実現している. さらに, 本手法を実際の検査において取得されたデータに適用する実験を行い, その有効性を示している.