

令和 3 年 6 月 15 日現在

機関番号：82401

研究種目：挑戦的研究（萌芽）

研究期間：2018～2020

課題番号：18K19766

研究課題名（和文）超微細画像情報の健康インフォマティクスによる次世代健康診断法の構築

研究課題名（英文）Development of Next-Generation Health Diagnosis Method by Health Informatics Using Ultra-fine Morphological Imaging Data

研究代表者

久米 慧嗣（Kume, Satoshi）

国立研究開発法人理化学研究所・生命機能科学研究センター・客員研究員

研究者番号：30708441

交付決定額（研究期間全体）：（直接経費） 4,900,000円

研究成果の概要（和文）：本研究では、病気の潜在的状態である「未病」の状態を早期かつ超高感度検出する技術提供を目指す。血液細胞の培養細胞、あるいは正常や老齢の動物の組織を対象に、広域電顕技術やバーチャルスライドを利用したイメージング解析を実施して、それらの画像データの取得・蓄積を進めた。さらに、深層学習を用いた画像セグメンテーションなどを実施して、画像中にある生命構造を解読して、意味づけを行えるかどうかを検証した。細胞の核構造については、90%以上の精度にてその構造領域を認識できるところまで達成した。また、細胞形態の概念関係モデルにマッピングすることで、画像メタデータに基づく知識階層化のプロトタイプを構築した。

研究成果の学術的意義や社会的意義

最先端の技術開発による画像のビッグデータ化に加え、概念関係モデル（オントロジー）を用いてアノテーションした教師画像データと、深層学習等の機械学習の手法を組み合わせることで、異常部位を自動検出し、さらにその結果を人間と機械が共に可読な「知識階層」として記述する。これによって、これまで医学で見過ごされてきたような、未病時に出現する超微細病態変化「病気のタネ」を自動的に言い当てるとともに、従来の人間の解釈を自然言語で記述していた方法論を大きく刷新する。延いては、健康科学の境地から未病状態を早期かつ超高感度に検出する『健康インフォマティクス』を創生し、次世代の健康診断の概念を提唱する。

研究成果の概要（英文）：This research aims to provide a technology for the early and ultra-sensitive detection of "presymptoms", a potential state of disease. We conducted imaging analysis of cultured blood cells or normal and aged animal tissues using the wide-range electron microscopy technique and virtual slide imaging, and acquired their image data. We also conducted image segmentation using deep learning and other methods to decipher and make sense of the life structures in the image. For the nuclear structure of cells, we have achieved more than 90% accuracy in recognizing their structural regions. In addition, we built a prototype of knowledge hierarchy based on image metadata by mapping to a conceptual relational model of cell morphology.

研究分野：生物情報学

キーワード：画像ビッグデータ 電子顕微鏡 画像セグメンテーション 深層学習 未病状態 血液細胞 健康インフォマティクス 概念関係モデル

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

我々は、未病診断、あるいは健康状態の予測などの健康科学への技術提供を目指して、組織細胞内の僅かな形態学的変化を検出することが可能な、電子顕微鏡技術、特に走査型電子顕微鏡 (SEM) を用いた細胞画像診断に着目して研究を進めている。近年、SEM 装置や撮像技術の開発が進展し、広領域(例えば、数 mm X 数 mm)のナノレベル画像取得が可能となった。また、最新の情報解析技術を組み合わせることで、画像のビッグデータ解析も現実になりつつある。こうした状況の中、画像ビッグデータ解析への基盤、技術整備は遅れており、未だデータ統合などによる画像データの知識化、あるいは疾患病態の解析や診断は実現していない。そこで、我々はデータを説明するデータである「メタデータ」や、メタデータを記述する標準語彙「オントロジー」の手法を用いて、顕微鏡の画像データやその撮像条件、表現型情報などを記述する研究を進めた。その成果において、顕微鏡の画像データやその撮像条件などを記述する顕微鏡オントロジーを、「セマンティックウェブ」と呼ぶ標準化された知識記述体系で構築することに成功し、さらには、Resource Description Framework 形式を利用した世界初の微細構造画像メタデータベースの設計や開発にも成功した。

上記の研究により、顕微鏡オントロジーを利用して画像データにメタデータを付与し、他の生命科学データなどと結びつけけることで、画像データを機械可読で解釈でき、形態学を含め系統的に管理する基盤作りができつつある。しかし、日々大量のデータが産生される生命科学分野においては、メタデータの付与を人間の手作業で行うには限界があり、画像データに人間が理解できる語彙を付与する新しい方法論が必要である。そこで、我々は、画像データからデータの特徴等のプロパティ情報を自動的に抽出し、それらを階層化して解釈するような人工知能のプラットフォームの構築を進めている。このような画像ビッグデータを適正に管理する次世代ビッグデータ解析の基盤を実現できれば、病健康科学の境地から未病状態を早期かつ超高感度に検出する『健康インフォマティクス』を創生し、次世代の健康診断システム開発へ繋がるのではないかと本研究テーマを着想した。

2. 研究の目的

未病状態とは、加齢、あるいはストレスや疲労が負荷されることによって、健康な状態から病気に移行するまでの潜在的な状態のことを指す。未病状態は「病気」そのものとは定義されないが、慢性的な未病状態が続けば、生活の質 (QOL) が著しく低下するばかりでなく、糖尿病、高血圧、脳卒中、うつ病など、さまざまな疾患を引き起す一因となる。このようなストレスや疲労が蓄積された未病状態で起きる僅かな病態変化を早期に検出する技術が確立されれば、軽減対策に資する食品素材やサプリメントを早期に摂取するなど、先制医療の観点から様々な改善策を講じることができる。その結果、疾患の予防やその発症遅延など、医療の負担や経済的コストは劇的に低減され、社会的にも大きな貢献と期待される。

我々は、未病診断・健康状態の予測など、健康科学分野への技術提供を目指して、組織細胞内の超微細な形態学的変化を高感度に検出可能な SEM を用いた細胞画像診断に着目した。特に、SEM で得られる画像データは生命現象の基礎をなす核構造やミトコンドリアなどの細胞内小器官の微細形態について極めて詳細な生体情報を提供する点で、ごく僅かな変動でも、非常に重要な健康指標となりうる。従来の撮像技術では限定された組織領域しか撮像することができず、組織網羅的な撮像は困難であったが、広域切片 SEM 法の登場により、超広域のナノレベル撮像が高速に可能となった。また、SEM で得られる画像ビッグデータの利活用に向けて、画像データの機械可読化処理の実現は極めて重要かつ火急な課題である。

本研究では、モデル動物やヒト検体を対象に、病気の潜在的状態である「未病」の状態を早期かつ超高感度に検出することを目指して、同広域 SEM 法を用いて、組織や血液細胞の超微細な異常状態を網羅的に検出し、計算機が自動的に生物学的解釈を行う技術を開発する。これによって、これまで見過ごされてきたような、未病時に出現する超微細病態変化「病気のタネ」を自動的に探索する。

3. 研究の方法

(1) 未病時の組織や血液細胞における SEM 画像データの蓄積、およびメタデータの付与

疲労・ストレス負荷を施した未病モデル動物、および正常動物 (マウスあるいはラット) を対象に、末梢血液細胞をはじめ、脳、腎臓や肝臓などの臓器を対象に、未病検出を実現する技術開発を行う。具体的には、1) 動物の血液細胞や臓器を取得する。2) マイクロスライサー等を用いて薄切した組織試料を作成し、前処理を施し、ウラニル・鉛などを使用して脂質染色法の 1 つである電子染色を行い、SEM 組織標本を作製する。3) SEM を用いて、生体試料表面を細い電子線で走査し、試料から出る信号を検出することで、試料由来の画像データを取得して、データを蓄積する。

次に、機械学習技術を適用する準備として、撮像した画像から種百枚程度を選び出し、我々が

構築している顕微鏡オントロジーを利用して、各画像データにメタデータを付与する作業を行う。これを、機械学習用の教師データとして整備する。

(2) 機械学習法を用いた画像メタデータの自動生成技術に関する検証実験

メタデータが付与された画像データを用いて、深層学習 (DL) 法によって画像データごとにメタデータの自動生成を試みる。まずは、数十枚程度の画像データの処理から始め、画像データに対して、生物学的解釈を含むメタデータとの関係性を学習させ、画像データの分類体系を作成する。最終的には、数千〜数万枚の画像処理に向けて、準備を進める。

(3) 機械学習と画像メタデータに基づく知識語彙の階層構造構築に関する検証実験

DL 技術によって得られたメタデータと画像データとの統合を検討する。例えば、画像セグメンテーションや画像分類を実施することで、画像データを定量解析してメタデータと紐づけることで、「異常」等の上位の集合概念や階層構造を再構築し、オントロジーとの整合性を確認する。このような検証実験によって、実際の未病者あるいは健康脆弱者から取得された血液細胞の試験を見据えた上で、未病時の変調などを系統的に人間の知識の枠組みに再現できるようなプロトタイプシステムを構築する。

(4) 未病者スクリーニング、およびヒト血球細胞の超微細構造解析による検証試験

これら知見に基づいて、ヒト血液細胞を対象にした検証試験を行う。当研究所の診療所でスクリーニングされた未病者あるいは健康脆弱者から取得された血液検体を用いて、SEM を用いた広域超微細構造解析を試みる。モデル動物で得られた画像データの学習モデルを基礎として、実際のヒト血液細胞の形態情報を新たに融合させることで、ヒトに最適な未病検出技術を構築する。また、アンケートや体組成、自律神経機能、脳機能などの健康情報とも複合解析を実施することで、細胞形態情報から様々な関係性を見出す。

4. 研究成果

血液由来の培養細胞株、あるいは正常や老齢のモデル動物の組織を対象に、広域切片 SEM 法やバーチャルスライド技術を利用したイメージング解析を実施して、それらの画像データの取得・蓄積を進めた。これらの画像データには、我々が構築した顕微鏡オントロジーを利用して、メタデータ付加等を実施した。このような大規模な画像データを簡便に扱うために、誰でも直感的に使用できるツールが必要であったので、Web 経由で使える画像ビューアを試作して、画像データの可視化を進めた。このようなシステムによってデータ共有をしたことで、様々な専門家やスタッフとも知識共有ができ、画像データへのアノテーションなどがスムーズに行うことができた。

次に、これら組織や細胞株について、機械学習・深層学習 (DL) を用いた画像セグメンテーション用の教師データセットの準備を行い、画像中にある生命構造を解読して、意味づけを行えるか否かを検証した。同データセットを用いて、画像セグメンテーションの解析モデルを構築して、画像のピクセル毎でメタデータ判別を行い、メタデータごとの注目画像領域 (ROI) を得た。この予測精度については、DL モデルを検討することで最終的に 95%以上の正解率に達した。特に、細胞画像に対する結果では、細胞体や核、ミトコンドリアなどの ROI ごとにメタデータを自動付与するとともに、各 ROI の形状情報とも紐付けることができた。また、知識語彙の階層構造構築に関する検証実験としては、共同編集型の知識基盤である Wikidata を利用して、SPARQL クエリ言語によって「細胞」に関するデータ構造を取得して、その階層関係をネットワーク構造として明らかにした。Wikidata 以外にも、疾患関連のオントロジーでのデータ構造の階層関係を調べて、組織形態の知識構造構築に利用する準備を進めた。

当初予定していた二次元 (2D) 画像データの DL 解析に加えて、三次元 (3D) 電顕画像データについても、DL 解析モデルを構築して検証試験を行った。具体的には、マウス神経系の 3D 画像データセットを用いて、画像のピクセル毎でのメタデータ判別を試みた。特に、神経細胞体の膜構造を認識するようにモデル構築、および解析を実施したところ、細胞膜構造の立体像が 95%以上の精度で得られた。一方、このような画像メタデータ付与の試みでは、その解析モデルの再現性を担保する上で、モデル学習に必要な、バイオ画像データの共有利用も重要なミッションである。そこで、画像データ共有の一環として、教師ラベルを含むバイオ画像データセットを、生命科学系データの解析ツールを開発するコミュニティである Bioconductor にて公開 (パッケージ名「BioImageDbs」) を行った。

ヒト検体を用いた検証試験では、神戸リサーチコンプレックス事業と共同して、ヒト血液検体や各種健康計測データの取得を実施した。さらに、この健康計測データでスクリーニングされた未病者あるいは健康脆弱者から取得された血液検体を用いて、広域切片 SEM 法を用いた微細構造解析を試みた。その結果、未病者から得られた血液細胞は、健常者のそれと比較して細胞内に形態学的な変化が見られた。また、アンケートや体組成、自律神経機能などの健康情報とも複合解析を進めた。また、ヒト血液検体では、バーチャルスライド技術によるイメージングも進めて、光学顕微鏡レベルでの形態像の取得や評価も実施した。さらには研究進展で、超微細構造解析のスループットの向上が不可欠であることも分かり、多検体の電顕前処理技術を開発するために、前処理装置のプロトタイプを開発して、その動作検証・条件検討等を実施することで、前処理装置の実用化を進めた。

5. 主な発表論文等

〔雑誌論文〕 計4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 久米 慧嗣、村川 泰裕	4. 巻 55
2. 論文標題 SEMによる切片の広域撮像技術と生体組織の網羅的形態解析への展望	5. 発行年 2020年
3. 雑誌名 顕微鏡	6. 最初と最後の頁 13~17
掲載論文のDOI（デジタルオブジェクト識別子） 10.11410/kenbikyo.55.1_13	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 久米慧嗣、田村勝、小林紀郎、村川泰裕	4. 巻 37
2. 論文標題 電子線・X線イメージング技術が織りなす画像ビッグデータ時代の病理学	5. 発行年 2019年
3. 雑誌名 病理と臨床	6. 最初と最後の頁 617-623
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 小口 綾貴子、村川 泰裕、柳田 素子、久米 慧嗣	4. 巻 13
2. 論文標題 広域電子顕微鏡解析からみる腎病理の未来図	5. 発行年 2021年
3. 雑誌名 月刊『腎臓内科』	6. 最初と最後の頁 502-509
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Kume S, Nishida K	4. 巻 NON
2. 論文標題 BioImageDbs: Bio- and biomedical imaging dataset for machine learning and deep learning (for ExperimentHub).	5. 発行年 2021年
3. 雑誌名 Bioconductor package version 3.13	6. 最初と最後の頁 NON
掲載論文のDOI（デジタルオブジェクト識別子） 10.18129/B9.bioc.BioImageDbs	査読の有無 無
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16件（うち招待講演 7件 / うち国際学会 5件）

1. 発表者名 久米 慧嗣
2. 発表標題 生化学者・生物物理学者からみた電顕解析のおもしろさ
3. 学会等名 第10回 植物電顕若手ワークショップ（理化学研究所 横浜キャンパス 交流棟ホール）（招待講演）
4. 発表年 2019年

1. 発表者名 Yasuhiro Murakawa, Norio Kobayashi, Satoshi Kume
2. 発表標題 Deep analysis of images from ultra-microstructural microscopy using machine learning
3. 学会等名 The 6th RIKEN-KI-SciLifeLab Symposium: Biomedical Data for Artificial Intelligence (RIKEN Yokohama Campus)（招待講演）
4. 発表年 2019年

1. 発表者名 Satoshi Kume
2. 発表標題 Comprehensive analysis of biological tissues using scanning electron microscopy and machine learning
3. 学会等名 RIKEN-CAS Young Scientists Seminar (Shenzhen Institutes of Advanced Technology (SIAT), Chinese academy of sciences (CAS), Shenzhen, China)（招待講演）
4. 発表年 2019年

1. 発表者名 Satoshi Kume, Yosky Kataoka
2. 発表標題 Comprehensive analysis of ultra microstructural microscopy images using machine learning
3. 学会等名 A3 Foresight Program Meeting (Seoul National University College of Medicine, Seoul, Korea)（招待講演）
4. 発表年 2019年

1. 発表者名 Satoshi Kume, Yosky Kataoka
2. 発表標題 Comprehensive Morphology Analysis Using The Wide-range Electron Microscopy Imaging and Machine Learning
3. 学会等名 理化学研究所-広島大学 合同シンポジウム「イメージングから理論」(広島大学)
4. 発表年 2019年

1. 発表者名 久米 慧嗣, 片岡 洋祐
2. 発表標題 広域超微細構造イメージング法とAI技術を用いた網羅的生体組織形態解析
3. 学会等名 第51回日本臨床分子形態学会総会・学術集会 シンポジウム2「AI を用いた形態学的研究の新知見」(日本顕微鏡学会とのジョイントシンポジウム)(久留米シティプラザ)(招待講演)
4. 発表年 2019年

1. 発表者名 Satoshi Kume
2. 発表標題 Semantic Segmentation for Image Metadata Creation in RIKEN Imaging Metadatabase
3. 学会等名 BioHackathon 2019 Symposium
4. 発表年 2019年

1. 発表者名 久米慧嗣, 村川泰裕, 田村勝, 小林紀郎
2. 発表標題 走査型電子顕微鏡の広域撮像技術を用いた生体組織の網羅的解析
3. 学会等名 日本顕微鏡学会 第75回学術講演会「S-3 走査電顕を用いた樹脂包埋切片観察法の基礎と生物試料への応用」(名古屋国際会議場)(招待講演)
4. 発表年 2019年

1. 発表者名 久米慧嗣, 村川泰裕, 陳桂花, 小口綾貴子, 田村勝, 小林紀郎, 榎屋啓志, 須賀三雄, 前田光代, 片岡洋祐
2. 発表標題 広域電子顕微鏡技術による生体組織のビッグデータ解析
3. 学会等名 第108回 日本病理学会総会「シンポジウム4 組織を見る新技術」(東京国際フォーラム)(招待講演)
4. 発表年 2019年

1. 発表者名 Satoshi Kume, Norio Kobayashi, Yasuhiro Murakawa
2. 発表標題 Ultra-fine Imaging Cytodiagnosis
3. 学会等名 A3 Foresight Program Meeting 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Kobayashi, N., Kume, S., Josh Moore, and Jason R. Swedlow
2. 発表標題 OME Ontology: A Novel Data and Tool Integration Methodology for Multi-Modal Imaging in the Life Sciences
3. 学会等名 SWAT4HCLS 2018: 11th International Conference on Semantic Web Applications and Tools for Health Care and Life Sciences (国際学会)
4. 発表年 2018年

1. 発表者名 村川泰裕, 小林紀郎, 久米慧嗣
2. 発表標題 臨床における医療イノベーションの必要性
3. 学会等名 画像ビッグデータが切り拓く健康・医学の新時代
4. 発表年 2018年

1. 発表者名 M. Suga, H. Nisioka, M. Nakamura, K. Suzuki, K. Konishi, Y. Suzuki, T. Nonaka, S. Kume, M. Maeda, Y. Kataoka, and K. Ohta
2. 発表標題 Quantitative analysis of cell nuclei in mouse Cerebellar cortex using Array Tomography and Deep neural network
3. 学会等名 2018 ASCB EMBO Meeting (国際学会)
4. 発表年 2018年

1. 発表者名 M. Suga, H. Nisioka, M. Nakamura, K. Suzuki, K. Konishi, T. Nonaka, S. Kume, M. Maeda, Y. Kataoka, and K. Ohta
2. 発表標題 Quantitative analysis of morphological feature of cell nuclei in Cerebellum cortex using Array Tomography and Deep Learning
3. 学会等名 19th International Microscopy Congress 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 M. Suga, H. Nisioka, M. Nakamura, K. Suzuki, K. Konishi, T. Nonaka, S. Kume, M. Maeda, Y. Kataoka, K. Ohta
2. 発表標題 Three Dimensional Structure Analysis of Cell Nuclei in Mice Cerebellar Cortex using Array Tomography
3. 学会等名 Microscopy & Microanalysis 2018 Meeting (国際学会)
4. 発表年 2018年

1. 発表者名 須賀三雄, 小西功記, 前田光代, 野中崇雄, 武居俊輔, 久米慧嗣, 西岡秀夫, 片岡洋祐, 太田啓介
2. 発表標題 Array tomographyとdeep learningを用いた細胞核の形態的特徴解析
3. 学会等名 日本顕微鏡学会 第77回学術講演会
4. 発表年 2021年

〔図書〕 計1件

1. 著者名 久米慧嗣、小林紀郎、片岡洋祐	4. 発行年 2019年
2. 出版社 技術情報協会	5. 総ページ数 598
3. 書名 ストレス・疲労のセンシングとその評価技術（第5章 バイオマーカーの検出と疲労、ストレス検知への応用 第3節 疲労・ストレス・未病時における分子代謝・超微細構造に基づくバイオマーカー探索研究の取り組み）	

〔産業財産権〕

〔その他〕

<p>理研微細構造画像メタデータベース（理研MIM） http://clst.multimodal.riken.jp/RIKENImageDBjpn/</p> <p>Bio- and biomedical imaging dataset for machine learning and deep learning (for ExperimentHub) https://bioconductor.org/packages/devel/data/experiment/html/BiolImageDbs.html</p>

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
研究分担者	小林 紀郎 (Kobayashi Norio) (20415160)	国立研究開発法人理化学研究所・情報システム本部・ユニットリーダー (82401)	
研究分担者	村川 泰裕 (Murakawa Yasuhiro) (50765469)	国立研究開発法人理化学研究所・生命医科学研究センター・チームリーダー (82401)	
研究分担者	水野 敬 (MIZUNO KEI) (60464616)	大阪市立大学・大学院医学研究科・特任准教授 (24402)	

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
---------	---------