

令和 5 年 10 月 25 日現在

機関番号：32610

研究種目：研究活動スタート支援

研究期間：2021～2022

課題番号：21K21320

研究課題名（和文）深層学習を利用した子宮頸がん細胞自動識別手法の確立

研究課題名（英文）Establishing a method for Automatic identification of Cervical Cancer Cells by a Deep Neural Network

研究代表者

坪下 幸寛 (Tsuboshita, Yukihiro)

杏林大学・その他部局等・教授

研究者番号：50910634

交付決定額（研究期間全体）：（直接経費） 2,300,000円

研究成果の概要（和文）：本研究では専門技師の識別性能を超える自動がん細胞識別の実現を目指し、深層学習を子宮頸部の細胞診標本に適用し実験を行った。まず、細胞診標本をデジタルカメラで撮影し、異型細胞の有無と程度を付与した大規模な画像データセットの作成を行った。データセットの画像数は約2,619枚、患者数は550人となった。次に、深層学習による予測実験を行った。実験には代表的な深層学習モデルである、ResNet、MobileNet、DenseNetを用いた。最高性能のモデルは、DenseNetであり、正確度0.90、感度0.87、特異度0.94となった。これは、専門技師の識別性能（感度0.7、特異度0.9）を上回った。

研究成果の学術的意義や社会的意義

細胞診画像のデータベースの作成を行い、最終的に患者数550人、画像数2,619枚の医療分野としては大規模なデータベースを作成した。これらの画像には、細胞診の診断結果のみではなく、HPV検査結果、組織診の検査結果も付与されている点で、これまでの細胞診とは一線を画している。さらに、深層学習を用いた分類実験を行い、一番性能が良かったDenseNet121を用いた実験では、正確度0.895、感度0.870、特異度0.943という結果を得た。これは、専門技師の認識性能を上回る結果となり、今後の実用化に向けて強固な橋頭保を得ることができた。

研究成果の概要（英文）：In the present study, we aimed to realize automatic cancer cell identification that surpasses the identification performance of specialist technicians, and therefore we performed experiments by applying a deep learning to cytological specimens of the uterine cervix. First, cytological specimens were photographed with a digital camera, and a large-scale data set with the presence and degree of atypical cells was created. The number of images in the database was about 2,619, and the number of patients was 550. Next, using this dataset, we conducted a prediction experiment using deep learning. ResNet, MobileNet, and DenseNet were used for the experiment. The best-performing model was DenseNet with the accuracy of 0.90, the sensitivity of 0.87, and the specificity of 0.94, surpassing the performance of morphological examinations performed by specialist technicians (sensitivity of 0.7, specificity of 0.9).

研究分野：機械学習

キーワード：深層学習 機械学習 自動診断 子宮頸がん検診 細胞診

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

ヒト乳頭腫ウイルス (human papillomavirus: HPV) ワクチンの接種推奨が中止されている日本において、子宮頸がん患者の増加を抑える唯一の手段は、一次検診で上皮内病変を早期発見することである。一次検診はこれまで形態学的検査法の細胞診が担ってきたが、その感度はせいぜい70%と指摘されており、その感度の低さが憂慮されて高感度 (>90%) の遺伝子検査による HPV 検出法に変更する旨の指針案が提示された。しかし、HPV 検出法は細胞診に比較して高価で特異度が低いため、一次検診法として採用には至っていない。

一方、深層学習 (Deep Neural Network) は従来のシステムをはるかに凌駕する性能で、画像から対象物を検出・認識することができ、年々深層学習を使った新たな研究成果が多数報告されている。深層学習を積極的に活用することで、形態学的検査法の感度の飛躍的な向上が期待される。

### 2. 研究の目的

本研究の目的は、子宮頸がん検診で得られた細胞診標本に深層学習を適用し、従来まで問題とされてきた専門技師による細胞診の低感度問題の解決を目指すことである。

これまで、子宮頸部細胞診標本の深層学習は、専門技師の形態学的認識に基づいた既知の異型細胞の検出を行う深層学習に留まっており、ヒトの細胞識別能力を超える成果は得られていない。それに対して本研究では、既知の異型細胞だけを標的とせず、病変の確定診断 (組織診) 結果および HPV 検出結果に基づいて分類された細胞診標本上の細胞群に対しても深層学習を実施することで、専門家の細胞識別能力を凌駕する実用的な性能を有する手法の確立を目指す。

### 3. 研究の方法

本研究は、次に示す二つの方法により研究を進めた。

#### 子宮頸がん検診による細胞診標本の大規模学習データセットの作成

既に組織診結果、HPV 検出結果に基づいて分類された細胞診標本の細胞群のデジタル画像は 2,000 検体分が準備済みであり、これをベースとして学習データセットの作成を進めた。ヒトが形態学的認識に基づいて判別した既知の異型細胞だけでなく、組織診、HPV 検査結果に基づいた分類が付与されている大規模データセットの作成はこれまで前例はなく、学術的価値が大きいものと考えられる。

#### 深層学習モデルを用いた評価実験

畳み込みニューラルネットワークを用いた代表的な深層学習モデルである ResNet<sup>[3]</sup>、DenseNet<sup>[4]</sup>、MobileNetV3<sup>[5]</sup>を用いて評価実験を行った。性能向上の施策として、大規模画像データベース ImageNet の事前学習済みモデルを用いた Fine-tuning を行った。

### 4. 研究成果

#### 4.1. 細胞診標本データセットの作成

この研究で用いる標本検体は、細胞診検査の結果が既に専門の技師によって得られたサンプルである。現在杏林大学保健学部には、標準化された検体が 550 枚準備済みであり、これを深層学習用に順次データベース化した。

表 1 組織診検査の分類種別

分類	意味
正常	-
CIN1	軽度異形成
CIN2	中等度異形成

CIN3	高度異形成・上皮内がん
------	-------------

標本検体を専用デジタルカメラ OLYMPUS DP20 を装着した顕微鏡 OLYMPUS BX20 を用いて写真を撮ることで、標本検体画像のデジタル化を行った。使用した顕微鏡の対物レンズの倍率は 10 倍、開口絞りレバーを 0.2 に設定し、写真撮影時には、写真用の光に設定したうえで 1 つの標本検体につき、それぞれ撮る場所を変えて約 2~5 枚撮影した。

今回の実験のデータセットは、異型細胞部分をピクセルごとにセグメントする等のアノテーションは行わず、画像単位で分類の付与を行った。これは、検出やセグメンテーションのための位置を指定するアノテーションには手間がかかるという点と、注視点を可視化する手法の発達により、分類の学習をするだけで、ある程度異型細胞の検出が行うことができるものと考えている点からである。また、細胞標本画像内で、全ての細胞が正常であった場合は正常、1 つでも異型細胞があった場合は異常と分類し、その中で最も進行度の高い分類を採用する。深層学習に判断させる際には、組織診検査は確定診断であるので、細胞診検査の結果ではなく、組織診検査の結果を使用した。組織診の分類は、表 1 を参照のこと。その際、組織診を行わなかった検体を正常、細胞診検査上では CIN2 と CIN3 の判断が極めて難しいことから、CIN2 と CIN3 をまとめて CIN2 以上として、正常、CIN1、CIN2 以上の 3 つに分類した。現在は、患者数 550 人分の 2,619 枚の作業が完了しており、このうち正常な検体画像は 1,504 枚、CIN1 は 360 枚、CIN2 以上は 755 枚である。

今回の実験では、細胞診検査の結果の分類だけでなく、組織診検査の結果の分類、HPV 検査の結果もデータセットに付与している。これにより、細胞診で得られたデータから、組織診の結果を推定するというような実験も可能である。表 2 は今回の実験のデータセットと他の存在する細胞診データセット<sup>[1][2]</sup>との比較である。我々のデータセットの方が、画像枚数は上回っており、さらに、HPV 検査結果の分類、組織検査結果の分類も付与されている。

表 2 他の細胞診データセットとの比較

	画像枚数	細胞診分類	HPV 分類	組織診分類
提案データセット	2,619	○	○	○
SIPAKMED <sup>[1]</sup>	966	○	×	×
Herlev <sup>[2]</sup>	917	○	×	×

#### 4.2. 深層学習を用いた評価

ここでは、現在構築が完了したところまでのデータセットを用いた深層学習の評価の結果を示す。結果の分類は、組織診の分類(正常/CIN1/CIN2 以上)を用いた。

##### 4.2.1. 実験設定

深層学習モデルは、代表的な畳み込みニューラルネットワークモデルである ResNet<sup>[3]</sup>、DenseNet<sup>[4]</sup>、MobileNetV3<sup>[5]</sup>を採用した。さらにそれぞれの深層学習モデルにおいて、転移学習の一つである ImageNet<sup>[6]</sup> Pretraining Model を用いた Fine-tuning<sup>[8]</sup> を行っている。また、CIN2 以上という分類は、以降 CIN2>という表記を用いて表す。分類ごとのサンプル数を表 3 に示す。

表 3 分類ごとのサンプル数

分類	正常	CIN1	CIN2 >
サンプル数	1,504	360	755

提供された画像サイズは、1600x1200dpi であったが、メモリの制約から、実際に学習、テストを行う際には 512x512dpi に縮小した。

学習率は 0.01 を初期値として AdamW<sup>[7]</sup> で制御するとともに指数減衰させた。また、損失関数は分類課題で標準的に利用される soft-max cross entropy loss<sup>[8]</sup> 関数を使用した。今回は、サンプル数が少ないこともあり、validation set は定義していない。したがって、Early stopping<sup>[8]</sup> で学習回数を制御するなどのことは行っておらず、一律 300 回のエポック数だけ学習させたモデルの性能を評価した。

## 4.2.2. 性能評価

実験結果を表4に示す。ResNetの正確度の平均は0.81、DenseNetでは0.84、MobilenetV3では0.88となった。特にDenseNet121は正確度0.895、感度の平均0.870、特異度の平均0.943、MobilenetV3 smallは正確度0.868、感度の平均0.820、特異度の平均0.924、MobilenetV3 largeは正確度0.901、感度の平均0.856、特異度の平均0.939となり、専門技師の行う形態学検査の性能(感度0.7、特異度0.9)を上回る結果となった。現在の学習セットのデータ数は約2,619枚程度の規模であり、今後学習データ数の増加につれて、さらなる性能向上が見込まれる。

表4 実験結果

	正確度	感度			特異度		
		negative	CIN1	CIN2	negative	CIN1	CIN2
ResNet18	0.826	0.890	0.591	0.789	0.805	0.948	0.934
ResNet34	0.810	0.916	0.502	0.744	0.758	0.952	0.936
ResNet50	0.808	0.917	0.504	0.746	0.746	0.948	0.945
ResNet101	0.812	0.932	0.493	0.733	0.745	0.952	0.946
DenseNet121	0.895	0.926	0.808	0.878	0.925	0.966	0.939
DenseNet169	0.814	0.901	0.551	0.769	0.777	0.946	0.935
DenseNet201	0.817	0.908	0.573	0.756	0.785	0.937	0.938
MobileNetV3 small	0.868	0.934	0.711	0.815	0.845	0.975	0.938
MobileNetV3 large	0.901	0.955	0.740	0.873	0.889	0.985	0.943

## 4.3. 研究成果のまとめと今後の課題

本研究では、最初に細胞診画像のデータベースの作成を行い、最終的に患者数550人、画像数2,619枚になった。これらの画像には、細胞診の診断結果のみではなく、HPV検査結果、組織診の検査結果も付与されている。次に、深層学習を用いた分類実験を行い、一番性能が良かったDenseNet121を用いた実験では、正確度0.895、感度0.870、特異度0.943という結果となった。

今回の実験で、性能面では満足のいく結果が得られたので、今後はこの方法を用いてどのように細胞診における専門技師のサポートを行ってゆくかを具体的に検討してゆきたい。

## 文 献

- [1] Plissiti M. E., et al., "Sipakmed: a new dataset for feature and image based classification of normal and pathological cervical cells in Pap smear images," In 25th IEEE Int. Conf. on Image Processing (ICIP), 3144–3148, 2018.
- [2] Jantzen, J., Dounias, G., Analysis of Pap-smear Image Data. In Proc. of the Nature-Inspired, 2006.
- [3] He, K., Zhang, X., Ren, S., Sun, J., "Deep residual learning for image recognition," CVPR, pp. 770, 2016.
- [4] G.Huang, Z.Liu, L.van der Maaten, K.Q.Weinberger. Densely Connected Convolutional Networks. IEEE Conference on Pattern Recognition and Computer Vision (CVPR), 2016.
- [5] Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H., Searching for MobileNetV3. In arXiv [cs.CV]: 1905.02244, 2019.
- [6] Beyer, L., Hénaff, O., Kolesnikov, A., Zhai, X., Oord, V., "Are we done with ImageNet?," arXiv: 2006.07159, 2020.
- [7] Loshchilov, I., and Frank H., "Decoupled Weight Decay Regularization." arXiv [cs.LG]: 1711.05101, 2017.
- [8] Goodfellow, I., Bengio, Y., Aaron Courville, A., Deep Learning, The MIT Press, 2016.

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 坪下幸寛, 大河戸光章
2. 発表標題 深層学習を用いた子宮頸がん細胞の自動分類
3. 学会等名 電子情報通信学会パターン認識・メディア理解研究会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

AI用いて新しい子宮頸がん検診法 確立めざす <a href="https://www.kyorin-u.ac.jp/univ/research/health/566/">https://www.kyorin-u.ac.jp/univ/research/health/566/</a>
---

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
研究協力者	大河戸 光章  (Okodo Mitsuaki)  (10276206)	杏林大学・保健学部・准教授    (32610)	

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