

ポスター

ポスター9

医療データ解析・評価

2017年11月22日(水) 13:30 ～ 14:30 L会場（ポスター会場1）（12F ホワイエ）

[3-L-3-PP9-2] X線画像を用いた Neural Networkの学習手法に関する基礎的研究

小笠原 克彦¹, 辻 真太郎¹, 山田 晋太郎², 張 洪健², 樋口 謙介³, 阿部 司⁴, 佐藤 友昭⁵ (1.北海道大学大学院保健科学研究院, 2.北海道大学大学院保健科学院, 3.日本電気株式会社 社会公共ビジネスユニット 医療ソリューション事業部, 4.NECソリューションイノベータ株式会社 パブリック事業本部 第一医療ソリューション事業部, 5.NECソリューションイノベータ株式会社 プラットフォーム事業本部 第二PFソフトウェア事業部)

背景・目的

医用画像は自然画像と比較して収集可能な枚数が少なく、データ拡張などの手法によって枚数を増加させる必要がある。しかしながら、Neural Networkの学習モデルを構築する際、最初から大量の医用画像を用いて学習させることは、データの拡張の手間もあり容易な方法とは言い難い。そこで本研究では医用画像における Neural Networkの学習の効果に影響を与える要因を明らかにすることを目的とし、X線画像の識別時の Neural Networkの層数、医用画像の部位、画像サイズ、画像の回転による影響の検討を行なった。

対象・方法

本研究で用いた X線画像は、副鼻腔画像(ウォーターズ法)、頭部正面撮影(PA方向)、膝正面画像、肘正面画像、胸部正面画像(PA方向)の5種類(それぞれ以下、W、H、K、E、Cとする)とした。各部位の画像は500枚とし、学習用及び検証用の画像をそれぞれ450枚、50枚とした。画像識別及び学習モデルの構築には NEC社製画像学習ソフトウェア Rapidを使用した。まず初めに事前実験として、画像の学習回数を表すエポック数(1～30)の決定を行なった。次いで、本実験として X線画像(5部位、計2,250枚)を、Neural Networkの層数を4、19、34、64としてそれぞれ学習モデルを構築し、残りの250枚を検証用として正答率を求めた。さらに画像のサイズによる学習への影響を検討するために、画像のサイズを元画像(1,020×1,275)の10%にした条件で同様に正答率を求めた。最後に、学習のロバストネス性を検証するために、画像を90度回転させた条件で正答率を求めた。

結果・考察

本研究結果から、層数4、19、34、64における部位 W、H、K、E、Cの画像の識別率は画像のサイズによらず全て100%となった。画像の向き90度に対する識別率は、画像サイズが100%で W(84%-90%)、H(0%-10%)、K(2%-18%)、E(0%-16%)、C(36%-100%)、10%で W(84%-88%)、H(0%-14%)、K(2%-18%)、E(0%-18%)、C(34%-100%)となり Wは層数に関わらず画像の回転に対して高い識別率を維持し、H、K、Eについては層数によらず低い識別率となり、Cに関しては総数が増加するにつれ識別率が低下した。

X 線画像を用いた Neural Network の学習手法に関する基礎的研究

辻真太郎^{*1}、小笠原克彦^{*1}、山田晋太郎^{*2}、張洪健^{*2}、遠藤晃^{*3}、樋口謙介^{*4}、阿部司^{*5}、佐藤友昭^{*6}

*1.北海道大学大学院保健科学研究院、*2.北海道大学大学院保健科学院、*3.北海道大学病院 医療情報企画部、
*4.日本電気株式会社 社会公共ビジネスユニット 医療ソリューション事業部、
*5.NECソリューションイノベータ株式会社 パブリック事業本部 第一医療ソリューション事業部、
*6.NECソリューションイノベータ株式会社 プラットフォーム事業本部 第二PFソフトウェア事業部。

A Basic Study of Learning Methods of Neural Network Using X-ray Images

Shintaro Tsuji¹, Katsuhiko Ogasawara¹, Shintaro Yamada², Kenko Chyo²,
Akira Endoh³, Kensuke Higuchi⁴, Tsukasa Abe⁵, Tomoaki Satou⁶.

*1 Faculty of Health Sciences, Hokkaido University, *2 Graduate School of Health Sciences, Hokkaido University,

*3 Affiliation for the third author, Department of Medical Informatics, Hokkaido University Hospital,

*4 NEC Corporation. Public Solutions Business Unit, Medical Solution Division,

*5 NEC Solution Innovators, Ltd. Public Solutions Business Unit, 1st Medical Solution Division,

*6 NEC Solution Innovators, Ltd. Platform Solutions Business Unit, 2nd Platform Software Division.

A convolutional neural network (CNN) in medical images achieved high accuracy of a detecting lesion and diagnosis compare to a computer aided diagnosis (CAD) but difference between accuracies of each sites is not verified. This study aims to clear a difference of accuracy about sites of X-ray images stored in a picture archiving and communication system (PACS) toward optimization of training CNN in medical images. We defined ten classifications: 1)Waters (W), 2)Paranasal sinus (Ps), 3)Right elbow (RtE), 4)Left elbow (LtE), 5)Right knee (RtK), 6)Left knee (LtK), 7)Chest of twenties male (C20M), 8)Chest of twenties female (C20F), 9)Chest of sixties male(C60M), and 10)Chest of sixties female(C60F). And 15,000 images were extracted from augmentation of 3,936 X-ray images in the PACS. Finally, 12,000 and 3,000 images were used for training and tests of CNN developed by NEC Ltd. Our CNN achieved the accuracies of 100%: W, Ps, RtK, LtK and RtE. In contrast, the accuracies of LtE, C20M, C20F, C60M, and C60F were 99.3%, 96.3%, 98.0%, 98.0% and 99.0%. The results of this study indicated that the classifications of age or sex in chest X-ray images were more difficult than that of the other sites.

Keywords: Deep learning, Convolutional Neural Network(CNN), Picture archiving and communication system(PACS), X-ray images, Classification.

1. 諸論

近年、人工知能の一つである深層学習(Deep learning)の技術革新が著しく、特に画像を中心とした畳み込みニューラルネットワーク(Convolutional Neural Network、以下 CNN)が普及しており、医用画像においても病変抽出や画像診断などへの応用が試みられている¹⁻⁶⁾。

しかし、風景や人工物のなどの画像認識とは異なり、医用画像には画像毎に撮影の部位や方向などの指示があるが、分類結果に与える影響を明らかにした研究はない。また、オーダ情報と紐づけられた医用画像を利用するには、大量の医用画像を蓄積している Picture archiving and communication system(以下、PACS)を用いる必要がある。

従って、臨床における CNN の普及には、PACS の医用画像を利用した分類を行い、その分類結果を CNN の学習方法にフィードバックすることが必要である。

2. 本研究の目的

本研究では、医用画像における CNN の学習方法の最適化を目指して、撮影情報が紐づけられた PACS の医用画像を対象に、医用画像の撮影部位に関する CNN の分類精度の違いを明らかにすることを目的とした。

3 方法

3.1 対象

CNN の画像分類に利用した医用画像は、北海道大学病院および北海道大学大学院保健科学研究院の倫理委員会

の承認を得た平成 23 年 1 月から平成 28 年 12 月までの 5 年間に北海道大学病院で撮影された X 線画像とし、PACS から手動的に抽出を行った(表 1)。画像の選定の際には、同一患者の画像が含まれない様に撮影時のオーダを確認した。また、CNN に分類させる画像は表 1 に示す 10 分類(頭部 2 種類、肘 2 種類、膝 2 種類、胸部 4 種類)とした。次に、画像拡張を行うため ImageJ を用いて各画像について $\pm 10^\circ$ 、 $\pm 20^\circ$ 、 $\pm 30^\circ$ の回転を行い元画像数の 6 倍の画像を得た。さらに、これらの画像の中からランダムに画像を抽出して CNN の学習用画像 1,200 枚、テスト用画像 300 枚を決定した。

3.2 用いた CNN と分類結果の評価指標

画像の分類には NEC 社製の CNN である RAPID を使用した。CNN の学習と分類を行う際の画像サイズは、縦方向を 256pixel として、全ての部位の撮影画像について縦横比を固定した。CNN の学習におけるパラメータの設定は、識別率が最大となる様に調整を行い、学習画像の学習回数を 30 回、ネットワーク規模を 19 とした。次に各分類における識別率(式 1)と確信度(式 2)、確信度の標準偏差を算出した。確信度は、CNN が行った X 線画像の 10 分類結果において各画像に付される値である。例えば、胸部、頭部、四肢の 3 分類の場合は、CNN から得た分類結果が胸部の場合、その確信度は胸部 97%、頭部 1%、四肢 2%などが考えられ、確信度の合計は 100%となる。

$$\text{識別率(\%)} = \frac{\text{CNNで正しく分類した画像枚数}}{\text{CNNで分類を行った枚数}} \times 100 \dots \text{式 1}$$

$$C_1 + C_2 + C_3 + \dots + C_i = 100(\%) \dots \text{式 2}$$

i: 分類の数, C_1 : 分類 1 の確信度, C_2 : 分類 2 の確信度, C_3 : 分類 3 の確信度, C_i : 分類 i の確信度.

表 1 CNN が分類した 10 種類の X 線画像

分類した画像	PACS から抽出した画像 (枚)
Waters 法	592
副鼻腔正面	403
肘関節正面(左)	225
肘関節正面(右)	320
膝関節正面(左)	299
膝関節正面(右)	300
胸部正面 20 代女	500
胸部正面 20 代男	500
胸部正面 60 代女	500
胸部正面 60 代男	297
合計	3,936

4 結果

各 X 線画像における CNN の識別率と確信度を示す(表 2)。

表 2 各 X 線画像における CNN の識別率と確信度

X 線画像	識別率(%)	確信度の平均値±標準偏差
Waters 法	100	1.00±0.001
副鼻腔正面	100	1.00±0.0008
肘関節正面(左)	99.3	0.999±0.003
肘関節正面(右)	100	0.998±0.02
膝関節正面(左)	100	1.00±0.0005
膝関節正面(右)	100	0.999±0.003
胸部正面 20 代女	98.0	0.997±0.02
胸部正面 20 代男	96.3	0.989±0.05
胸部正面 60 代女	99.0	0.990±0.05
胸部正面 60 代男	98.0	0.992±0.04

5 考察

本研究で用いた CNN は、頭部、四肢、胸部を含む 10 種類の X 線画像の分類に対して識別率 96%以上、確信度の平均値 99%以上で分類を行うことが確認された。CNN の分類から得られた確信度とその標準偏差から、CNN の学習が適切に行われていることが確認できる。

一方、肘関節正面(左)の識別率が他の識別率より低下した原因は、テスト画像の中に介助者の手が入った小児の画像等が含まれていたことが影響していると考えられる。また、胸部画像における年齢・性別に関する分類では、20 代の識別率(男女平均 97.2%)が 60 代の識別率(男女平均 98.5%)より低く、性別の分類については男性の識別率(20 代と 60 代の平均 97.15%)が、女性の識別率(20 代と 60 代の平均 98.5%)より低い結果となった。この原因は、20 代の男性の胸部画像

の分類における識別率が低いことが考えられる。現在調査中であるが、識別できなかった画像について分類の根拠となった部分を明らかにし、その要因を明らかにする必要があると考える。さらに胸部画像の分類については、類似した先行研究に骨年齢の評価⁷⁾があり、0 歳から 18 歳までの手部の X 線画像を年齢・性別毎に分けてそれぞれ症例を 72~608 症例を準備し、約 5,800 症例を学習画像として 90%以上の症例を評価可能であったとの報告がある。本研究胸部画像では 20 代・60 代の 2 つの年代を対象として各 1,200 枚(合計 4,800 枚)を学習画像とした結果、96%以上の識別率を得たことから、学習の枚数と得られた識別率・確信度から考慮して妥当な分類結果を得たと考えられる。一方、「類似した撮影部位」や「左右の部位」の分類の識別率・確信度が限りなく 100%であることから、「胸部画像の年齢・性別の分類」より難易度が低いと推測される。また、本研究では検査頻度の多い胸部の側面像や腹部領域の画像の分析を行っていないため今後同様の検証を行う必要がある。加えて今後は、先行研究で行われているような股関節変形性関節症の診断⁸⁾などの画像診断の領域における難易度の高い分類に対する検証を行う。

6 結論

本研究では、PACS に蓄積されている X 線画像を対象として CNN で分類を行った。本研究により高い識別率を CNN に要求する場合は、学習・テスト画像に含まれる変則的な画像の選定に注意を要することが明らかとなった。また、本研究で対象とした X 線画像においては、「類似した部位の分類」、「左右の部位の分類」と比べて「同部位における性別や年齢の分類」の識別率が低くなる傾向が明らかとなった。

7 参考文献

- 1) Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. IEEE Trans Med Imaging. 2016;35(5):1207-1216. doi:10.1109/TMI.2016.2535865.
- 2) Jain A, Levy D. DeepMammo Breast Mass Classification using Deep Convolutional Neural Networks. Stanford cs231n. 2016.
- 3) Pereira S, Pinto A, Alves V, Silva CA. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. IEEE Trans Med Imaging. 2016;35(5):12401251. doi:10.1109/TMI.2016.2538465.
- 4) Lee S, Choi M, Choi H, Park MS, Yoon S. FingerNet: Deep learning-based robust finger joint detection from radiographs. In: 2015 IEEE Biomedical Circuits and Systems Conference (BioCAS). IEEE; 2015:1-4. doi:10.1109/BioCAS.2015.7348440.
- 5) Beura S, Majhi B, Dash R. Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer. Neurocomputing. 2015;154:1-14. doi:10.1016/j.neucom.2014.12.032.
- 6) Sun W, Zheng B, Lure F, et al. Prediction of near-term risk of developing breast cancer using computerized features from bilateral mammograms. Comput Med Imaging Graph. 2014;38(5):348-357. doi:10.1016/j.compmedimag.2014.03.001.
- 7) Lee H, Tajmir S, Lee J, et al. Fully Automated Deep Learning System for Bone Age Assessment. Journal of Digital Imaging, 2017;30(4):427-441. doi:10.1007/s10278-017-9955-8
- 8) Xue Y, Zhang R, Deng Y, Chen K, Jiang T. A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis. PloS one, 2017;12(6):e0178992. doi:10.1371/journal.pone.0178992. eCollection 2017.