

肺炎画像による転移学習を用いた胸部 X 線画像じん肺分類モデル

檜垣 明良 【 知能情報学研究室 】

1 はじめに

胸部 X 線画像を用いたじん肺分類に関する研究は近年活発に進められており、深層学習を用いることで高い識別精度を達成した例も報告されている [1]。しかし、医療画像は一般に収集コストが高く、学習に用いる画像枚数が限られる場合が多い。加えて撮影機関や撮影条件などドメイン間の差異も相まって、学習時とは異なるデータに対して十分な汎化性能が得られないという問題も指摘されている。

本研究では、これらの課題に対する手法として、じん肺と画像的特徴が類似する肺炎画像を用いた転移学習に着目した。類似疾患を事前に学習させることで、少数のじん肺画像においても安定した分類性能が得られる可能性があると考え、その有効性と限界について検討する。

2 実験

2.1 データセット

事前学習用データセットとして、中国・広州市の広州婦人小児医療センターにおいて、1~5歳の小児患者を対象に日常診療の一環として撮影された胸部 X 線画像を用いる。本データセットは肺炎画像と健常者画像を同数含み、学習用各 1,100 枚、検証用各 200 枚、テスト用各 234 枚の計 3,068 枚で構成される。

転移学習先のじん肺・健常者データセットには、高知大学医学部 (KM)、National Institutes of Health Clinical Center (NIHCC)、および National Institute of Occupational Safety and Health (NIOSH) で収集された胸部 X 線画像を用いる。内訳は、KM がじん肺 91 枚、NIHCC が健常者 94 枚、NIOSH が健常者 23 枚・じん肺 28 枚であり、各画像には NF またはじん肺のラベルが付与されている。

2.2 実験内容

本研究では、VGG16 を基盤とした深層学習モデルを用いて、じん肺・健常者の二値分類を行った。ImageNet 事前学習モデルと、肺炎・健常者識別モデルを事前学習として用いた転移学習モデル (図 1) を比較し、分類層を新たに構築した上で、一部の畳み込み層を段階的に解凍して学習を行った。

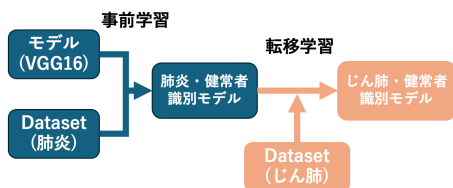


図 1 本研究における転移学習の概念図

表 1 じん肺分類モデルの感度 (学習対象層別比較)

事前学習	FC	+B5	+B4	+B3	+B2	+B1
ImageNet	0.78	0.93	0.95	0.93	0.90	0.92
肺炎	0.92	0.98	0.97	0.82	0.87	0.88

注: FC は全結合層, +Bk は Blockk までを学習対象層とした条件を表す。

3 結果・考察

まず、NIOSH のみで構成されたデータセットを用いた実験結果を表 1 に示す。既存の ImageNet 事前学習モデルと比較して、肺炎事前学習モデルでは感度 (Recall) の向上が確認された。一方で、学習対象層を拡張した場合、両モデル間の性能差は縮小し、一部の条件では ImageNet 事前学習モデルの方が高い性能を示した。

これらの結果から、肺炎画像による事前学習は、肺野領域の形状や濃淡分布といった胸部 X 線画像に共通する構造をあらかじめ獲得できるため、分類層と高次特徴層を学習対象とした少数データ環境下において有効であると考えられる。一方で、学習対象層を広げすぎることによって、転移元である肺炎に由来する情報と識別対象であるじん肺の情報が適切に分離されず、結果として識別性能が低下した可能性が示唆された。

次に、撮影機関の異なるデータを用いて汎化性能の検証を行った。KM 由来のじん肺画像および NIHCC 由来の健常者画像をテストデータとして用いた条件では、高い識別性能が確認された。一方で、NIHCC および KM 由来の画像を学習データとし、NIOSH 由来画像をテストデータとした条件では、肺炎事前学習モデルは ImageNet 事前学習モデルと比較して、分類精度および ROC-AUC において大きな改善を示さなかった。

4 まとめ

本研究では、胸部 X 線画像を用いたじん肺・健常者分類において、類似疾患である肺炎画像を用いた転移学習の有効性について検討した。その結果、NIOSH のみで構成された小規模データセットにおいては、肺炎事前学習モデルにより感度の向上が確認された。一方で、他ドメイン画像を含む条件下では、転移学習のみでは撮影機関や撮影条件の違いに起因する差を十分に吸収できないことが示された。

参考文献

- [1] 東, “胸部 X 線画像からの機械学習を用いた塵肺検出に関する研究”, 高知工科大学修士論文, 2021.