

機械学習による胸部 X 線画像からの疾患検出

1200280 東 聖 【 知能情報学研究室 】

1 はじめに

胸部 X 線検査は健康診断で最も一般的な画像検査であるため、膨大な画像データが蓄積されており、画像診断支援の研究が盛んである [1]。文献 [1] では、複数疾患のマルチラベル分類を試みているが、ラベルが 15 種と多く、クラス毎のデータ数の差が大きいため、多クラス分類として直接機械学習を行なっても十分な精度が得られない。

そこで本研究では、疾患検出精度の向上を目的とした 2 段階検出を用いた胸部疾患の検出を行う。

2 提案手法

健常画像がデータの半分を占めるため、まず、ステップ 1 として健常画像と疾患画像の 2 クラス分類を畳み込みニューラルネット (CNN) で行う。ステップ 2 として、疾患画像のみから 14 クラスのマルチラベル分類を行う。

3 実験設定と評価指標

3.1 データセット

本研究には NIHCC (National Institutes of Health Clinical Center) のデータセット¹を用いる。このデータセットは 112120 枚の胸部 X 線画像から構成されており、1 つの画像に複数のラベル付けがされているマルチラベルである。また、このデータセットには 14 種の疾患のラベルが存在する。

3.2 推定モデル

今回の実験では VGG16 を使用した。マルチラベル分類のため識別層の活性化関数は sigmoid 関数、学習時には 5 エポック間にバリデーションの正確度が向上しない場合に学習率を減らすコールバック関数の ReduceLROnPlateau を用いる。また、病気の見逃しをできるだけ減らすため、評価指標には平均再現率を用いる。

4 実験

まず、提案手法と比較するため、全画像 112120 枚を用いた 15 クラスマルチラベル分類を行う。次に提案手法のステップ 1 である、全画像 112120 枚の画像を用いた健常画像と疾患画像の 2 クラス分類を行う。この実験では疾患がなければ健常画像、疾患があれば全て疾患画像としてラベル付を行う。そして、疾患の種類を検出するために、全画像中の疾患画像 51759 枚のみを用いて提案手法のステップ 2 である 14 クラスマルチラベル分類を行う。これら 3 つの結果を表 1 に示す。



肺浸潤 胸水 無気肺 心肥大

図 1 胸部 X 線画像と疾患の例 (NIHCC データ)

表 1 分類クラス数による再現率の違い

モデル	平均再現率
15 クラス	0.0002
2 クラス (健常, 疾患)	0.7279
14 クラス (健常なし)	0.0165
2 クラス (肺浸潤, 胸水)	0.7561
3 クラス (肺浸潤, 胸水, 無気肺)	0.5329
4 クラス (肺浸潤, 胸水, 無気肺, 結節)	0.3427

■ 15 クラスマルチラベル分類 表 1 の 15 クラスでは、平均再現率が 0.0002 となった。この結果から、単純にマルチラベル分類を行なっても疾患の検出精度が低いことが分かった。

■ 健康か疾患の 2 クラス分類 表 1 の 2 クラス (健常, 疾患) では、平均再現率が 0.7279 となり、2 クラス分類であれば分類を行うことが可能であると分かった。

■ 14 クラスマルチラベル分類 表 1 の 14 クラス (健常なし) では、平均再現率が 0.0165 となり、15 クラスマルチラベル分類と比較するとおよそ 1.6% の差が見られた。

また、クラス間のデータ数の違いより疾患検出精度に影響を与えることが考えられるため、データ数の多い疾患順に 2, 3, 4 クラスのマルチラベル分類を行なった。

表 1 の 2, 3, 4 クラスのデータより、データ数の多い順に再現率が高かった。そのため、より疾患検出精度を向上させるにはより多くのデータが必要である。

5 まとめ

本研究では、疾患検出精度の向上を目的とした 2 段階検出を用いた胸部疾患の検出を行い、二段階検出を用いない場合と比較して、およそ 1.6% の差がみられた。

参考文献

- [1] X. Wang, et al., “ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases”, IEEE CVPR, pp. 2097-2106, 2017.

¹<https://nihcc.app.box.com/v/ChestXray-NIHCC>