[研究論文]　CNNを用いた自己進化型CADシステムの提案

安倍和弥[1]・武尾英哉[1]・永井優一[2]・縄野繁[3]

1 電気電子情報工学科
2 国立がん研究センター東病院
3 国際医療福祉大学三田病院

Proposal for self-evolving CAD system using CNN

Kazuya ABE[1], Hideya TAKEO[1], Yuuichi NAGAI[2], Shigeru NAWANO[3]

Abstract

In recent years, convolutional neural networks (CNN) have found active application in the field of computer-aided diagnosis (CAD). Typically, general-purpose, high performance detectors are designed by conducting machine learning provided with comprehensive sets of case images having different kinds of variations. The reasoning behind this is that images not included in learning are not expected to be successfully detected. In fact, a comparison of CNN trained using only relatively typical cases and CNN trained using a comprehensive set of cases showed superior performance by the latter. Such is also expected to be the case in the commercial release of CAD. This is because the cases used in development may not encompass the cases at the medical facilities that will actually operate the system.

In response to this situation, this paper proposes a self-evolving CAD system that incorporates a relearning function.

It is not possible to upgrade the detectors of previous CAD systems following their commercial release. As a remedy, we propose a mechanism for upgrading the detector, accomplished by endowing the CAD system with a misrecognition correction function that uses interpretation report information (i.e., electronic medical records), and performing relearning after certain data is applied. This function provided improved performance compared to when relearning is not performed. Moreover, multiple relearning resulted in gradual improvements, demonstrating a process by which the system evolved into a CAD system matched to the involved medical facility.

Key words: Breast Cancer, Tumor Mass, CNN, Self-evolving CAD, Relearning

1.　はじめに

近年，医療の法現場ではCTやMRIなど撮影装置の高精細化により生成される医用画像の量が増大している．撮影される画像の枚数は患者一人あたり数百枚にもおよび，それを読影する医師の負担も急増している．そのため診断の補助を行うコンピュータ画像支援診断（Computer Aided Diagnosis/Detection：CAD）の開発が非常に重要視されている．

そのCADの研究分野においても人工知能，AIを用いた研究が進められている．その中でも画像認識手法であるCNN（Convolutional Neural Network）は積極的に用いられるようになった[1)]．一般的に，機械学習には様々なバリエーションを持った症例画像を網羅的に与えて学習を行うことで，汎用的で高性能な判別器を設計する．これは，学習に含まれていない画像はうまく検出ができないと考えられるからである．これはCADのリリース時にも同様のことが考えられる．開発時に用いる症例は，実際に運用する病院での症例を網羅しているとは限らないからである．また，特に乳がん検診に用いられるマンモグラフィの場合は，各病院で様々な撮影系にて行われているため，画質が異なることがある．このような場合施設特有の画像へCADがマッチしていくことは有用と考えられる．そこで本論文では，再学習機能を組み込

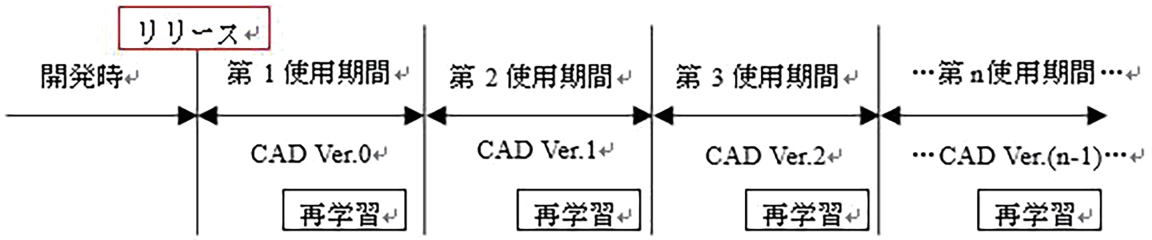


Fig.1　CAD システムの進化の流れ

んだ進化型 CAD システムを提案する．CAD システムのリリリース後に判別器の再学習が行われることはない．これは従来の CAD システムにおいては，判別器に用いる特徴量を人が作成しているためリリース後の更新は困難であったためである．しかし，CNN などの機械学習では症例画像からコンピュータが自動で特徴量を作成することが可能である．筆者らはこの観点から判別器のリリース後の更新が可能であると考える．

CNN を用いた乳がん腫瘤影の検出は Al-masni ら[2,3]や Tan ら[4]などが行っているが，これらは開発と評価に用いるデータを同一のデータベースから分ける形で使用している．そのため画質が異なる場合のあるマンモグラフィにおいては，実際のリリース時にもすべての施設の画像で同様の結果となるとは保証できない．

自己進化型 CAD システムは，評価を行った症例を追加して再学習を行うことにより，学習回数が上がるにつれて使用する施設の症例に最適な判別器に進化することが可能である．再学習の流れを Fig.1 に示す．開発時の CAD を一定期間使用し（第 1 使用期間），その後に集まった症例を追加して再学習(Ver.1)．Ver.1 を一定期間使用し（第 2 使用期間），集まった症例を追加して再学習(Ver.2)．このように一定期間ごとに再学習を行い，より使用施設に最適な CAD へと進化させる．

一方，CNN の特性として，データの質による性能の低下も考えられる．再学習の際に誤りデータなどが混入した場合の性能低下も考えられる．そのため，追加症例の質による性能変化の検討も併せて行う．

本論文のポイントを以下に示す．

① 再学習時の追加症例の質の検討
② 自己進化型 CAD システムの検討

本論文では，2 章で学習データセットの質的検討，3 章で自己進化型 CAD システムの検討・評価の結果を示す．4 章にて考察，5 章にて本研究のまとめについて述べる．

2. CNN 開発における学習データセットの検討

2.1 乳がん腫瘤影判別用 CNN

乳がん腫瘤影判別の CNN は Al-masni らの手法[2]や藤田らの書籍[5]を参考に yolo[6]を用いて作成した．OS は ubuntu16.04，cuda8.0 を用いて GPU を利用した学習を行っている．

2.2 誤認識訂正機能

現在の医療現場において読影レポート情報（電子カルテ）はほぼ全ての現場にて使用されている．電子カルテには医師の所見を書き込むシェーマ機能があり，乳がん腫瘤影においても実際の乳房の形状を基にスケッチを行える（Fig.2）．そのスケッチ情報を読み込み，医師のマーク座標から正解座標を算出する．CAD の検出した腫瘤影の座標と正解座標を比較して誤検出領域の除去と検出ミスの領域の拾い直しを行う．このような電子カルテの情報を利用した学習型 CAD を作成することにより，誤認識訂正を行っての再学習が可能となる．

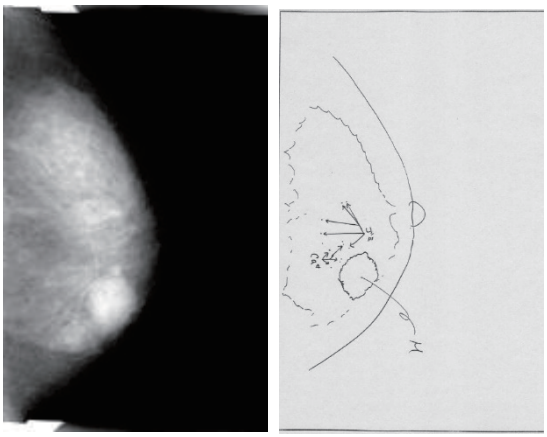


Fig.2　シェーマ図

2.3 症例の質による CNN の性能変化

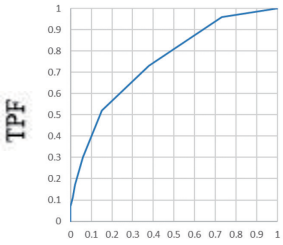
自己進化型 CAD の開発にあたって学習データセットの質の検討を行った．自己進化型学習を行う場合，CAD にて評価した画像を一定量貯めた後，学習データとして追加して再学習を行う．この追加する症例の質による CNN の性能変化を検討した．

開発時の CNN はがん症例 40 例で学習，第 1 使用期間後の評価としてがん症例 40 例，非がん症例 40 例を使用した．その後，評価に使用した画像を，FP 率が 10%の時の検出結果を基に誤認識訂正を行う場合と行わない場合での 2 通りで画像を追加し再学習を行い，質の違いによる性能の変化を調べた．誤認識訂正を行った場合の追加数は，がん症例 40 例，行わなかった場合は，がん症例 15 例と非がん症例 4 例が追加されている．

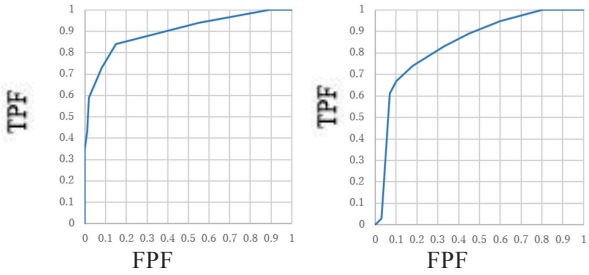
結果の ROC グラフを Fig.3 に，Az(ROC グラフの下面

積）を Table1 に示す．ROC は学習用とは別に用意したテストデータ，がん症例 40 例，非がん症例 40 例を使用して評価している．

Table1 及び Fig.3(a)(b)より初期学習の CNN と誤認識訂正を行った CNN を比較すると判別性能が著しく向上しているのが見てとれる．一方，Fig.3(a)(c)より誤認識訂正を行わなかった場合においても性能は上昇したが，訂正を行った場合ほどの向上は見られなかった．また，再学習の回数が増加するにつれて非がん症例の個数も増加していくため，回数が増えるにつれ性能が低下する傾向が見られた．よって進化型 CAD には誤認識を訂正する機能が求められる．



(a) 開発時の CNN



(b) 誤認識訂正あり　(c) 誤認識訂正なし

Fig.3　追加症例の質の評価

Table1　誤認識訂正の有無による評価の比較

	Az	
	誤認識訂正を行う場合	誤認識訂正を行わない場合
開発時	0.75	
再学習 1 回目	0.90	0.84

3.　自己進化型 CAD システムの性能評価

2 章の結果より，自己進化型 CAD は誤認識症例の訂正を行うことで性能の上昇が確認された．

本章では 2 章と同様に初期の CNN を作成，以降段階的に症例データを追加し CNN の再学習による性能向上についての検討を行った．

評価にあたり，用意した学習データを腫瘤影の大小と，腫瘤内と辺縁での濃淡の差を基に次の A～D の 4 等分に分割した．

・A：サイズ 5cm 以上，濃淡差が 80 以上
・B：サイズ 5cm 以上，濃淡差が 80 未満
・C：サイズ 5cm 未満，濃淡差が 50 以上
・D：サイズ 5cm 未満，濃淡差が 50 未満

開発時の学習を A のみ 40 例で学習した CNN と ABCD を各 10 例ずつの 40 例で学習した CNN を作成した．これは，症例の偏りを使用施設による症例画像の質の違いに見立てて検討を行うためである．本来であれば施設ごとに症例画像を用意し判別器の製法変化を検討する予定であったが，プライバシー保護の問題などから複数施設の画像を必要数用意できなかったためこの方法を採用した．以降の再学習は ABCD 各 10 例の 40 例を追加する形で，計 3 回追加を行い性能の評価を行った．

各々の結果の ROC グラフを Fig.4，Fig.5 に，Az(ROC グラフの下面積)を Table2 に示す．ROC は学習用とは別に用意したテストデータ，がん症例 100 例，非がん症例 100 例を使用して評価している．

Table2 及び Fig.4,Fig.5 より A のみでの開発時,ABCD での開発時の両 CNN ともに性能向上が見られる．再学習 3 回では A のみからの学習でも Az が 0.92 以上，ABCD 網羅からでは Az が 0.96 以上と高性能の CNN に進化していることが確認できる．A のみで開発した CNN では著しく検出性能が向上しており，再学習によって実際に使用する施設の症例により適応した CNN に進化していると示唆できる．また，ABCD を網羅して開発した CNN においても上り幅こそ緩やかではあるか向上が見られる．これらの結果より，施設による症例画像の質の偏りに影響されることなく，再学習によって検出性能の向上が見込めることが確認できる．

Table2　再学習による最適化

	Az	
	A のみで学習	ABCD を網羅
開発時	0.75	0.92
再学習 1 回目	0.90	0.93
再学習 2 回目	0.89	0.97
再学習 3 回目	0.92	0.96

4.　考察

2 章の Table1 より，CNN に追加する症例の質の重要性が確認できる．また，自己進化型 CAD では再学習の際に追加する症例の質により性能が低下する可能性が確認された．非がん症例が少数でも加わってしまうと Az が大きく低下する．これに対して誤認識訂正機能を導入し，誤認識を訂正した場合では検出性能の向上が見られた．この誤認識訂正機能は，ほぼすべての病院で導入されている電子カルテの情報を組み込むことにより可能であると考えられる．

3 章の Table2 より，自己進化型 CAD は再学習を行うことにより性能が向上することが確認できた．また，開

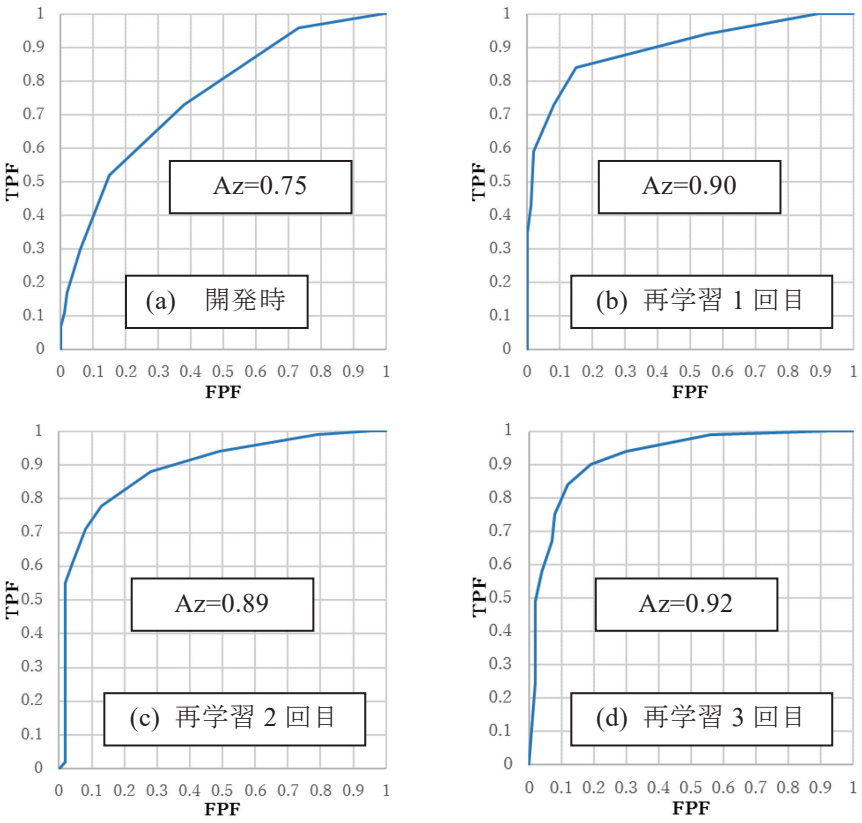


Fig. 4 A のみで開発した CNN の進化

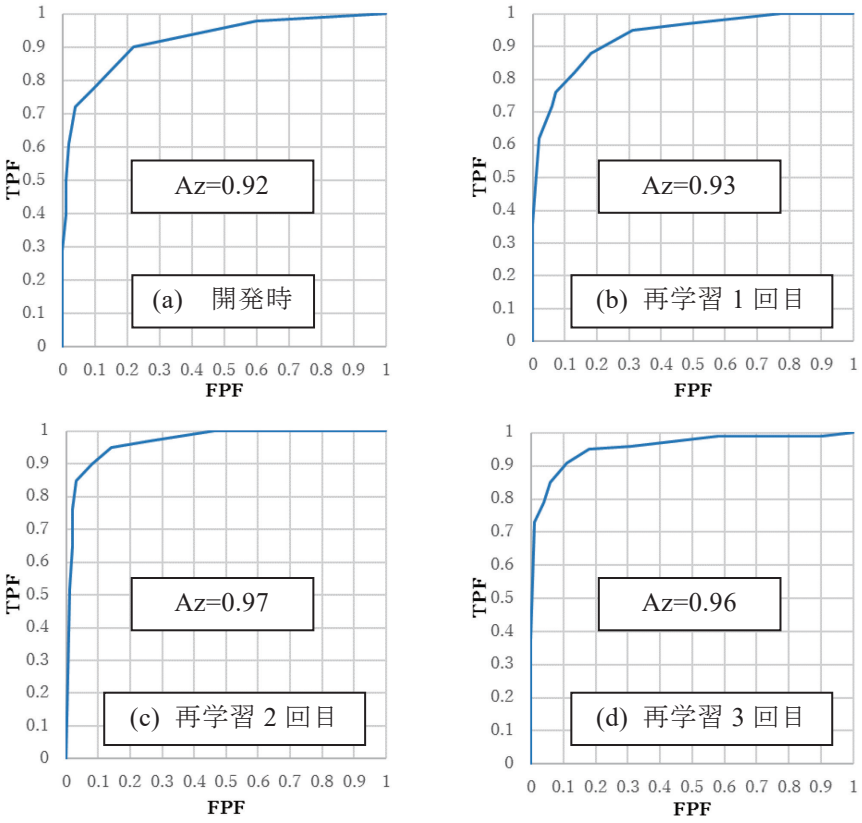


Fig. 5 ABCD 網羅で開発した CNN の進化

発機関の持つ症例データと実運用を行う施設での症例に差があった場合においても，再学習ごとに性能が向上し，より使用する施設に最適な CNN へと進化することが確認できた．ABCD を網羅した CNN では途中で性能が飽和しているように見えるが，低性能での飽和ではないため再学習における CAD の性能保障になると考えられる．しかし，本結果は症例の偏りを施設の症例の違いに見立てたものであり，実際に施設ごとに収集を行った場合においてのものと異なる可能性はある．これについては引き続き検討を行う必要がある．

ABCD を網羅した CNN の再学習 3 回目時の FROC を Fig. 6 に示す．これまでに示した ROC グラフは CAD システムの判別処理（腫瘤影候補検出後の FP 除去処理）の性能についてのものである．一方，CAD システム全体の性能を検討するために，候補検出処理を含めた検出性能を FROC で示した．Fig. 6 では，過去に筆者らによって開発された候補検出処理 7)を含めた CAD システム全体を FROC として示したものである．なおこの手法では，乳がん腫瘤影の候補領域を，適応リングフィルタを用いて検出を行っている．FP 数が 0.5[個/image]で約 90%の TP 率となった．現在の一般的な乳がん CAD が同 FP 数で約 85%程度であるため，本 CAD は良好な検出結果であったいえる．

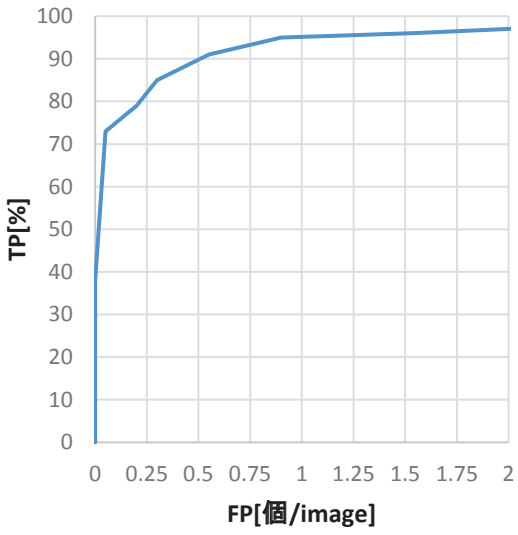


Fig. 6　システム全体の FROC 性能

5.　まとめ

本論では，学習に用いるデータデットの質の検討，およびそれを用いた再学習による自己進化型 CAD システムについての提案と検討を行った．その結果，CNN はやはり矛盾が起こるデータを少数でも加えてしまうと性能が著しく低下することが確認できた．それに対し電子カルテ情報を利用した誤認識訂正機能を組み込むことにより，再学習における矛盾データの混入を防ぐ手法を提起した．

さらに，開発機関と実運用を行う施設で症例データに差があった場合においても，複数回再学習を行うことにより段階的な精度向上が見られることが確認できた．また，意図的に網羅した症例での CNN においても，上り幅は緩やかではあったものの再学習による性能の向上は確認できた．

本提案システムは，国内薬事（薬機法）や FDA 認可時には問題が発生する懸念がある．認可の審査時の性能が再学習を行った際に保証されない可能性があるためである．本論での実験においては開発時性能と比較しての再学習による検出性能の低下は見られなかったが，可能性としては否めないものである．

参考文献

[1] Robertson S, Hartman J, Robertson S et al: Digital image analysis in breast pathology-from image processing techniques to artificial intelligence. Translational Research Vol.194, pp.19-35, (2018)

[2] Al-masni M. A., Al-antari M. A., Park J et al: Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. Computer Methods and Programs in Biomedicine Vol.157, pp.85-94, (2018)

[3] Al-masni M. A., Al-antari M. A., Park J et al: Detection and classification of the breast abnormalities in digital mammograms via regional Convolutional Neural Network. IEEE Conf Proc. No.EMBC. pp.1230-1233, (2017)

[4] Tan Y. J., Sim K. S., Ting F. F: Breast cancer detection using convolutional neural networks for mammogram imaging system. IEEE Conf Proc. No.ICORAS. pp.1-5, (2017)

[5]藤田一弥，高原歩：実装ディープラーニング．オーム社，東京， pp150-167，(2016)

[6] Joseph R, Santosh D, Ross G et al: You Only Look Once: Unified, Real-Time Object Detection. Computer Vision and Pattern Recognition, arXiv:1506.02640, (2015)

[7] 武尾英哉，志村一男，今村貴志，他：乳房 CR 画像を用いた異常陰影検出システムの開発と性能評価．Med Imag Tech 22(4)：201-214，(2004)