乳がん病変検出のための深層学習を用いた計算機支援画像診断システム

鈴木 真太郎*・張 暁 勇**・本 間 経 康***
市 地 慶*・高 根 侑 美****・柳 垣 聡***
川 住 祐 介***・石 橋 忠 司***・吉 澤 誠*****

A Deep Learning-based Computer-aided Diagnosis System for Mammographic Lesion Detection

Shintaro SUZUKI*, Xiaoyong ZHANG**, Noriyasu HOMMA***,
Kei ICHIJI*, Yumi TAKANE****, Satoru YANAGAKI***,
Yusuke KAWASUMI***, Tadashi ISHIBASHI*** and Makoto YOSHIZAWA*****

In recent years, deep convolutional neural network (DCNN) has widely been applied for image recognition, and shown a remarkable performance in various natural image-related applications. However, for medical image-related application such as computer-aided diagnosis (CAD), due to the limitation of training data and the modality difference between the natural and medical images, training the DCNN for medical image recognition is still a research topic. In this paper, we propose a DCNN-based method for lesion detection in mammograms. The proposed method consists of the following two steps. Given a mammogram, lesion candidates are firstly detected from the mammogram based on their intensity characteristics. Secondly, a transfer learning-based method is applied for training an existing DCNN to classify the lesion candidates into lesions or normal tissues. The proposed method is tested on a public mammogram database. Compared with several previous studies, our proposed method achieved a higher true positive rate and a lower false positive in lesion detection.

Key Words: deep convolutional neural network (DCNN), transfer learning, computer-aided diagnosis (CAD), mammogram

1. はじめに

近年、乳がん罹患率の急激な上昇¹⁾と早期発見の重要性が周知されてきたことを背景に、乳房 X 線撮影（マンモグラフィ）を用いた画像検診受診者数が増加傾向にある。しかしこれに伴い読影医の負担も増加し、疲労に起因した病変の見落

としなどの発生が懸念されている。この対策として期待されているのが、コンピュータを用いて乳房 X 線画像上の病変候補を自動検出し、第 2 の意見として医師に提示することで診断を補助する、計算機支援診断（computer-aided diagnosis: CAD）システムの開発である²⁾。

従来の CAD システムの多く^{3)~6)}は、医師の読影論理に倣った画像解剖学的な病変の特徴を手動で設計し、抽出された特徴量をもとに病変か正常かを識別する。しかし、医師の読影論理は専門知識のほか経験にもとづいて定性的に表現される場合が多く、病変特徴を定量的に表現することはきわめて難しいため、CAD システムの病変検出性能向上は容易ではなかった。特に病変のみを正確に表わす特徴の定義は難しく、病変を見逃さないように検出すると、同時に似たような特徴をもつ正常組織も多く誤検出してしまう問題があった。

このような特徴設計の困難さを解決するものとして近年注目を集めているのが、深層学習の一種である deep convolutional neural network (DCNN) である。DCNN は、特徴および識別規則を学習的に自動獲得することができ、自然画像認識の分野においては従来法を上回る高い性能を示すことが知られている⁷⁾。このような DCNN に医師の複雑な読影論理を学習させることができれば、従来手動設計がきわめて困

* 東北大学大学院工学研究科 仙台市青葉区荒巻字青葉 6-6

** 仙台高等専門学校 仙台市青葉区愛子中央 4-16-1

*** 東北大学大学院医学系研究科 仙台市青葉区星陵町 2-1

**** 東北大学病院 仙台市青葉区星陵町 1-1

***** 東北大学サイバーサイエンスセンター

仙台市青葉区荒巻字青葉 6-3

* Graduate School of Engineering, Tohoku University, 6-6 Aramaki Aza Aoba, Aoba-ku, Sendai

** National Institute of Technology, Sendai College, 4-16-1 Ayashichuo, Aoba-ku, Sendai

*** Tohoku University Graduate School of Medicine, 2-1 Seiryomachi, Aoba-ku, Sendai

**** Tohoku University Hospital, 1-1 Seiryomachi, Aoba-ku, Sendai

***** Cyberscience Center, Tohoku University, 6-3 Aramaki Aza Aoba, Aoba-ku, Sendai

(Received November 27, 2017)

(Revised March 11, 2018)

Fig. 1 Overview of candidate ROI detection

難であった病変特徴を自動獲得することができ、高い病変検出性能を実現できると考えられる。このため、医用画像へのDCNNの適用は近年大きな研究トピックとなっている。たとえば、乳房X線画像のカテゴリ分類にDCNNを適用した報告もある⁸⁾。この研究では、約72万枚の乳房X線画像を用いた学習により、カテゴリ分類能力を獲得させている。しかしカテゴリ分類では、CADシステムとして一般的な「検出」とは異なり、画像上での異常の位置を提示することができない。また、分類対象が比較的容易な陰性や良性などの所見に限定されており、診断上重要な悪性病変の存在が疑われる所見を除外しているなど、CADシステムとしては改善の余地が多い。

本研究では、DCNNを用いた乳房X線画像上の病変検出システムを提案する。しかし、DCNNに病変検出を学習させるために必要な大量の医用画像を収集することは、文献8)の場合以上に困難である。この原因として、つぎの2点が挙げられる。第一に、病変特徴の獲得には大量の病変例を用いて学習を行なう必要があるが、正常例に対して病変を含む症例の割合は非常に小さく、数十万件規模の症例画像の収集は容易ではない。第二に、病変検出の学習および適切な評価のためには、画像に加えて画像上における正解病変領域の情報が必要である。このような領域情報の作成は、十分な知識と経験をもつ医師の協力に頼るほかなく、大量に作成することは難しい。

この問題を解決するため、本研究ではDCNNの学習に転移学習⁹⁾を適用し、限られた数の画像から病変特徴の効率的な獲得を試みた。また、一般に乳房X線画像は縦横数千画素程度の大きなデータであるのに対し、DCNNの入力画像は計算量などの制約により縦横数百画素程度の大きさが一般的で

ある。画像の大幅な縮小は微細な病変の見落としに繋がるため、DCNNには原画像の一部を切り出したregion of interest (ROI) 画像を入力し、ROIに病変が含まれるか否かを識別させる。病変の位置および大きさは多様であるから、原画像全体に渡りさまざまな大きさのROIを走査し、基本的にはそれら膨大な数のROIすべてをDCNNにより識別する必要がある。しかし、このような網羅的な手法は計算量の観点からきわめて非効率的である。

このため本研究では、はじめに乳房X線画像上では病変が周囲に比べて高輝度、すなわち高コントラストを有することなど、従来法で蓄積された特徴の知見や手法をもとにしたシンプルな前処理により、病変サイズや位置情報を保持したまま大まかな病変候補の自動選定を行なう。このような従来知見もとづく特徴は多くの病変に共通するため、真陽性（病変を正しく検出する）数は多い一方、従来法の問題で述べたように病変だけを正確に表現することはできないため、偽陽性候補も多く選定される。そこで、これらをDCNNにより正しく識別するような2段階検出法を構成し、最終的な偽陽性数を低減する。前述の網羅的な手法と異なり、提案手法のDCNNは病変、もしくは従来知見の意味で病変にある程度類似したROI画像のみを対象として集中的に学習を行なう。このため、少ない訓練データでも病変とそれに類似した正常組織のわずかな特徴の違いを効率的に獲得することができると考えられる。また、一般にDCNNによる識別論理はブラックボックスとなってしまうが、医用システムとしては病変識別の論理が説明できることが望ましい。本研究では、DCNNが病変の識別に有用な特徴を獲得し、適切な論理にもとづいて病変の識別を行なっていることを確認するため、DCNNの識別論理の分析を行なった。

Fig. 2 GMM-based breast segmentation and target region extraction

2. 提案手法

2.1 病変候補検出

前処理としての病変候補検出の目的は、DCNN では違いが吸収されてしまう病変の位置や大きさを検出・保持したまま、大まかな病変候補の選定を行なうことである。したがって、多くの病変に共通するシンプルな特徴にもとづいてできる限り多くの病変を検出することを優先し、偽陽性（誤検出）を許容するものとする。

提案する病変候補検出法は、Fig. 1 に示すように、検出対象領域の抽出（target region extraction）、病変候補領域の抽出（candidate region extraction）、病変候補領域（ROI）画像の作成（candidate ROI extraction）の3つのステップからなる。

2.1.1 検出対象領域の抽出

乳房 X 線画像では、Fig. 1 左端の原画像のように、しばしば乳房領域外に撮影時の諸情報を記録したラベルなどが存在する。これらは以降の検出対象領域抽出の妨げとなるため、事前に乳房領域のみを抽出しておく必要がある。本研究では、原画像を大津の閾値法¹⁰⁾を用いて二値化し、最も面積の大きい領域を乳房領域として抽出することで、Fig. 2 (a) に示すような乳房領域のみの画像を得る。

乳房は解剖学的に乳腺と脂肪の主に2つの組織からなり、乳房 X 線画像の輝度値は主にこの2つの透過的重ね合わせによって構成される。すなわち輝度値は X 線透過方向における両者の厚さで決定するといえるが、撮影は乳房を均一な厚みに押しつぶした状態で行なわれるため、全域に渡って厚さが一定に保たれていると仮定できる。この場合、輝度値は両者の割合で決まり、乳腺の割合が大きいほど高輝度になるといえる。ここで、輝度にもとづいて以下のように画像解剖学的

な乳腺密度分布（割合）ごとに4つの領域に分割する。

Class 1: 乳房と背景領域の境界（marginal area）

Class 2: 脂肪組織（fat tissue）

Class 3: 脂肪組織と乳腺組織の混合組織
（fat and fibroglandular tissue）

Class 4: 高密度乳腺組織（dense tissue）

すなわち class 1 が最も低輝度で、class 2, 3 の順に高輝度となり、class 4 が最も輝度が高い。

Fig. 2 (a) に示す画像のヒストグラムを Fig. 3 に示す。一般に乳房 X 線画像のヒストグラムは3~4つ程度の極大値をもち、おのおのの両側はなだらかに減少する正規分布的な形状となっており、おのおのの分布が各 class に相当するものと推測される。そこで各 class から求めた確率密度関数が正規分布であると仮定し、その確率密度関数をガウス混合モデル（gaussian mixture model: GMM）を用いて4つのガウス分布の混合により近似することで、画像の乳房領域を上記4つの class に分割する^{11), 12)}。

ガウス混合モデルは、入力データの分布をガウス分布の線形加重和から得られる混合分布で表現するモデルである。適切な数のガウス分布を用い、線形結合する重みの係数と各分布の平均、分散を適切に求めることで連続分布を任意の確度で近似できる。まず混合確率密度関数 $G(\mathbf{x})$ を以下のように定義する。

$$G(\mathbf{x}) = \sum_{i=1}^K w_i g_i(\mathbf{x} | \mu_i, \sigma_i^2) \quad (1)$$

ここで、 $\mathbf{x} = (x_1, x_2, \dots, x_N)$ は入力データ、 N はデータ数である。 K は混合数、 $w_i > 0$ は混合率で $\sum_{i=1}^K w_i = 1$ 、 μ_i は class i の平均値、 σ_i^2 は分散を表わし、 $g_i(x_j) > 0$ はそれぞれの確率密度関数である。各確率密度関数はガウス関数 $g_i(x_j)$

Fig. 3 Histogram approximation by the GMM

$$g_i(x_j) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(x_j - \mu_i)^2}{2\sigma_i^2}\right) \quad (2)$$

で表わされる．この混合分布が入力データの分布に近づくようパラメータの推定を行なう．パラメータ推定には expectation maximization (EM) アルゴリズム¹³⁾を用いた．なお初期値としては，乳腺分布の4つの class を表わすため $K = 4$ ，最大画素値を M とし， $i = 1, 2, 3, 4$ に対して以下のように設定した．

$$w_i = \frac{1}{K} \quad (3)$$

$$\mu_i = \frac{iM}{K+1} \quad (4)$$

$$\sigma_i^2 = \frac{1}{KN} \sum_j (x_j - \bar{\mu})^2 \quad (5)$$

$$\bar{\mu} = \frac{1}{N} \sum_j x_j \quad (6)$$

Fig. 3 上の4本の実線および破線は，それぞれ4つの class に対応するガウス分布と，それらの混合により近似された確率密度分布を表わす．画像の各画素値 x_j を，4つのガウス分布のうち確率密度が最も大きいものに対応する class $A(x_j) \in \{1, 2, 3, 4\}$ に分類する．

$$A(x_j) = \arg \max_i \left(\frac{w_i g_i(x_j)}{G(x_j)} \right) \quad (7)$$

画像上の乳房領域を分割した結果を Fig. 2 (b) に示す．前述の画像解剖学的な乳腺分布ごとの4つの領域に，おおよそ分割されていることがわかる．ここで，乳がん病変は一般に乳腺組織に発生することから，病変検出の対象 (Target) を Fig. 3 に示すように乳腺が含まれる class 3 以上の高輝度領域に限定し，Fig. 2 (c) のような検出対象領域を抽出する．これにより誤検出の削減と計算コストの低減が期待される．

2.1.2 病変候補領域の抽出

ここでは，Fig. 1 左端の原画像より Fig. 4 (d) のような病変候補領域を示す二値画像を生成する．

はじめに，Fig. 1 左端の原画像より，脂肪組織などの非病変の背景組織を形態学的フィルタの一種であるトップ・ハット変換¹⁴⁾を用いて除去した画像 Fig. 4 (a) を作成する．トップ・ハット変換は，原画像に対してオープニング画像¹⁴⁾を差し引く処理であり，オープニングに用いた構造要素よりも小さな要素を原画像上から抽出する．乳房 X 線画像上において，一般に病変は局所的に高い輝度値をもつものに対し，脂肪組織などの背景組織は低周波成分が支配的である．このため構造要素を適切に設計することにより，トップ・ハット変換によって背景組織の輝度成分のみを減じることができると期待される．

このような手法は乳房 X 線画像上の多くの病変に対して有効と考えられるが，構造要素の設計には特定の検出対象の想定が必要である．以下では代表的な乳がん所見の1つである腫瘍陰影の検出を例に説明する．

乳房 X 線画像上において，腫瘍は類円形の高輝度領域として観察される．また，腫瘍陰影は一般に直径 5 cm 未満であることから，本実験では構造要素に直径 5 cm の円を用いた．Fig. 1 左端の原画像と背景除去後の画像 Fig. 4 (a) を比べると，背景組織の低周波成分が除去され，局所的に高コントラストな領域がより明瞭に強調されていることが確認できる．

つぎに，画像 Fig. 4 (a) に対し，2.1.1 項で得た検出対象領域 Fig. 2 (c) との論理積を取ることで Fig. 4 (b) のように病変検出の対象となる領域のみを抽出する．

続いて，一般に病変は局所的に高い輝度値をもつことに着目し，Fig. 4 (b) で強調された領域をもとに閾値処理により二値化された Fig. 4 (c) のような病変候補領域を抽出する．閾値決定の手法については後述する．

つぎに，高輝度領域の二値化画像 Fig. 4 (c) に対しオープニングとクロージング処理により高輝度領域の整形を行ない，ここでの病変候補領域として Fig. 4 (d) のような二値化画像を得る．オープニングおよびクロージング処理は数値形態学的な平滑処理であり¹⁴⁾，対象の大きさ (面積) をあまり変化させずに平滑化を行なうことができる．乳房 X 線画像上の病変領域は，乳腺組織などとの重なりにより複雑な輝度分布をもつことが多く，二値化の閾値によってはしばしば複数の高輝度領域の集まりとして抽出される．このため，高輝度領域内の小さな穴状の欠損や領域間の隙間を埋める働きのあるクロージング処理により，複数の分裂した高輝度領域を1つの病変領域に整形する．また，二値化画像 Fig. 4 (c) 上には乳房 X 線画像上のノイズや正常乳腺に起因する微小な高輝度領域が存在する．これらの明らかに非病変の領域を除去するため，前述のクロージング処理後の画像に対し，微小な高輝度領域を除去する働きをもつオープニング処理を行なう．これらの処理により，Fig. 4 (d) に示すように整形された病変候補領域のみを抽出することができる．

最後に，二値化のための閾値決定の流れを Fig. 5 に示す．ここでは，前処理の目的に沿って病変領域の確実な抽出を優先し非病変領域の誤抽出はある程度許容するが，極端に多く

Fig. 4 Mass candidate region extraction

Fig. 5 The flowchart of threshold decision

の誤抽出領域は計算量の観点からも好ましくない。また、一般に乳房 X 線画像 1 枚あたりの腫瘍病変は 1, 2 個程度であり、3 箇所を超えることはまれである。そこで、誤抽出も考慮に入れつつ 1 枚あたりの病変候補領域数の上限目安を 5 とし、これを基準として閾値を決定する。すなわち、画像 I に対し閾値 $t = \{0, 1, \dots, 255\}$ で閾値処理を行なった際の整形後の候補領域数 $f(I, t)$ が 5 以上の場合はその閾値 t を採用し、5 より少ない場合は領域数が最大となるような閾値 t を採用する。

$$t^*(I) = \begin{cases} t & (\text{if } f(I, t) \geq 5) \\ \arg \max_t \{f(I, t)\} & (\text{otherwise}) \end{cases} \quad (8)$$

このような条件を満足する $t^*(I)$ は複数存在するので、それらの最大値を最終的な閾値 $T(I)$ とする。

$$T(I) = \max_t t^*(I) \quad (9)$$

2.1.3 病変候補 ROI 画像の作成

続いて、抽出された各病変候補領域の位置と大きさに従い、DCNN に入力する病変候補 ROI 画像を 2.1.2 項で述べた

Fig. 4 (a) のような背景組織除去後の乳房 X 線画像上より切り出す。本研究で用いる DCNN の入力サイズは 227×227 画素の正方形画像であり、ROI 画像のサイズをこれに一致させる必要がある。また、病変候補が ROI からはみ出してしまうと識別が非常に困難となることから、ROI 画像の一辺の長さ L_{ROI} は病変候補の水平、垂直方向の長さ L_x, L_y に対し余裕をもたせて、 $L_{ROI} = 1.5 \times \max(L_x, L_y)$ の正方形とした。ROI の位置は、病変領域の水平、垂直方向それぞれの中央を中心とするが、ROI が原画像上からはみ出す場合には、ROI の大きさは変えず位置を ROI が原画像に収まるまで内側に移動する。DCNN は入力画像上での対象の位置変動に対して非常に高いロバスト性を有することが知られており、候補領域が ROI 内に収まっていれば識別性能への影響は小さいと考えられるためである。各 ROI は切り出し後、前述の DCNN の入力サイズに合わせてリサイズされる。補間法にはバイキュービック補間を用いた。これにより Fig. 1 右端に示すように、中心付近に病変候補の高輝度領域を含む ROI 画像を得る。

2.2 DCNN による病変候補の識別

2.2.1 DCNN の基本構造

本研究で病変識別器として用いた DCNN は、Krizhevsky et al. による AlexNet⁷⁾ と同様の構造をもつものである。Fig. 6 に示すように、この DCNN は特徴抽出を担う 5 層の畳込み層 (conv1~conv5 層)、次元と計算量の削減および特徴の位置変動に対するロバスト性の獲得を担う 3 層のプーリング層 (pool1, pool2, pool5 層)、およびクラス識別を行なう 3 層の全結合層 (fc6~fc8 層) からなり、深層ニューラルネットワークとしては比較的簡単な構造である。DCNN への入力として 3 チャンネル (R, G, B) の画素値を与え、出力層は入力画像が ImageNet¹⁵⁾ における 1,000 種類のクラスそれぞれに属する確率を出力する。なお、AlexNet について詳細は参考文献 7) を参照されたい。

Fig. 6 Architecture of the *AlexNet* consisting of 5 convolution layers (conv1~conv5), 2 normalization layers (norm1, norm2), 3 pooling layers (pool1, pool2, and pool5), and 3 fully-connected layers (fc6~fc8)

Fig. 7 The transfer learning for mass detection in mammograms. (a) The DCNN is pre-trained by using a large-scale image dataset for 1,000 classes. (b) We modify the last layer into two neurons, corresponding mass and normal classes respectively, and subsequently retrain the DCNN using a small-scale mammogram dataset

2.2.2 DCNN の転移学習

転移学習とは、あるタスクで学習した「知識」（対象の識別に有用な特徴や識別規則）を別のタスクに転用する手法である⁹⁾。DCNNでは、まず浅い層でエッジや色といった局所的で単純な特徴が抽出され、より深い層に伝わる過程でそれらが組み合わせられることで大域的で複雑な特徴を構成する。このため、深い層では特定のタスクに固有な特徴が表われるのに対し、浅い層の特徴は多くの画像認識タスクに対して汎用性の高いものとなる。目的とするタスクに対して十分な量の学習データを用意できない場合、このような単純で汎用性の高い特徴を別のタスクで獲得し、これを目的タスクに転用、さらに目的タスクの学習データを用いて微調整を行なうことで、限られた量の学習データから特徴や識別規則を効率的に獲得できると期待される。本研究では、まず大量に収集可能な自然画像でDCNNを学習し、獲得した知識を病変候補の識別に転用する。

学習の具体的な手順は以下のとおりである。まずFig. 7(a)に示すように約120万枚の自然画像からなる *ImageNet*¹⁵⁾ データセットを用いてDCNNを学習する。この操作を事前学習と呼ぶ。つぎに、DCNNの出力 (fc8) 層を病変識別のタスクにおける分類クラス (病変/正常) に対応する2つのニューロンからなる新たな全結合層で置換する。続いて、Fig. 7(b)に示すように病変候補ROI画像を用いて病変候補の識別を学習する。この操作を *fine-tuning* という。

すなわち、事前学習により獲得した conv1~fc7 層のパラ

メータ (fc8 層はランダムに初期化) を初期値として病変識別の学習 (*fine-tuning*) を行なう。この手順により、自然画像認識で獲得した汎用性の高い単純な特徴を転用し、これを改変して病変識別に適した特徴を構成することが可能となり、限られた数の医用画像から病変特徴を効率的に獲得できると期待される。

3. 実験結果

3.1 実験データセット

DCNNの学習 (*fine-tuning*) および病変検出性能評価は、乳房X線画像のデータベースである *Digital Database for Screening Mammography* (DDSM)¹⁶⁾ の画像を用いて行なった。原画像のサイズは画像ごとに異なるがおおむね 5000 × 3000 画素前後、解像度は1画素あたり 50 μm である。原画像の濃度分解能は12 bit であるが、これを8 bit にスケールしたものをを用いた。乳がん罹患者487名の頭尾方向撮影 (craniocaudal: CC) 画像のうち、DCNNの学習 (*fine-tuning*) に用いるものとして417名、検出性能評価に用いるものとして70名の画像に分け、それぞれの画像に対し2.1節の手法で病変候補ROI画像を作成し、学習ならびに評価を行なった。ROI画像数の内訳をTable 1に示す。

3.2 実験条件

DCNNの事前学習では *ImageNet* の自然画像約120万枚を用いて *AlexNet* を学習するため、膨大な計算資源と計算時間が必要となる。このため本研究では、事前学習を行なう代わ

Table 1 Experimental data (number of ROIs)

	Training data	Test data	Total
Mass	509	91	600
Normal	1,624	258	1,882
Total	2,133	349	2,482

りに *Caffe*¹⁷⁾により提供されている学習済みの *AlexNet* モデルを用いた。

fine-tuning は, MATLAB 上で動作する深層学習ライブラリである *MatConvNet*¹⁸⁾を用いて実装し, 学習を行なった。学習回数は 50 epochs に設定し, 学習係数は 0.001 より 0.0001 まで, epoch ごとに対数的に等間隔に減衰させるものとした。また, バッチサイズは 256, 重み減衰の係数は 0.0005, モメンタム係数は 0.9 とし, fc6 層ならびに fc7 層には係数 0.5 のドロップアウトを適用した。

3.3 病変検出性能の評価

提案システムによる病変検出結果の例を **Fig. 8** に示す。各画像上の矩形は病変として検出された ROI の位置および大きさを表わし, 真陽性検出を TP, 偽陽性検出を FP と表記している。また, 医師の指定した真の病変領域を破線で示している。Fig. 8(a) の例では, 提案システムにより 2 つの ROI が病変として検出されたことがわかる。うち 1 つの ROI が真の病変領域と一致しており, 検出すべき病変を適切に検出できたことを示し, もう一方の ROI は偽陽性検出である。Fig. 8(b) は比較的大きな腫瘤を有する乳房 X 線画像の例であるが, 大きな 2 つの ROI に跨る形で病変が適切に検出されている。この例では偽陽性検出はない。Fig. 8(c) は, 小さな腫瘤をもつ例である。病変の大きさに一致した小さな ROI により, 病変が適切に検出されていることが確認できる。また, この例でも偽陽性検出はない。

病変検出性能の定量的な評価指標として, CAD システムの性能評価で一般的に用いられる free-response receiver operating characteristic (FROC) 曲線¹⁹⁾を用いた。一般に, 病変と正常の閾値を変更することにより, 病変を正しく病変として検出する割合である真陽性率 (true positive rate: TPR) は変動するが, この真陽性率と, 正常組織を病変と誤検出する偽陽性数 (false positives per image: FPI) は互いにトレードオフの関係にある。**Fig. 9** に示すように, FROC 曲線は横軸に乳房 X 線画像 1 枚あたりの偽陽性検出数, 縦軸に真陽性率をとる。したがって, FROC 曲線が左上を通るほど病変検出性能が高いことを表わし, 理想的には図中縦軸 0~100% に一致する直線となる。Fig. 9 には, 比較のため提案システムのほか, 手動での特徴設計による従来システム 3 手法^{12), 20), 21)}の FROC 曲線をあわせて示した。提案システムは, 真陽性率が 80% のときの偽陽性数が他手法の約 1/4 となり, 実用的な真陽性率を保ったまま偽陽性数を大幅に低減できることが示唆された。

3.4 考察

通常, DCNN が獲得した特徴および識別規則はブラックボックスとなってしまうが, 医用システムとしては識別の根拠が説明可能であることが望ましい。このため本研究では, class model visualization (CMV)^{22), 23)}の手法を用い, 提案法により DCNN が腫瘤と正常組織の識別に必要な特徴を適切に獲得できていることを確認した。

CMV は, DCNN の特定のニューロンの出力を最大化するような入力画像 (モデル画像) を生成する手法である。モデル画像 I は, つぎの目的関数を最大化するよう入力画像を繰り返し修正することにより生成される。

$$I = \arg \max_I \{S_c(I) - \lambda_{l2} \|I\|_2^2 - \lambda_{tv} \|\nabla I\|_2^2\} \quad (10)$$

はじめの項 $S_c(I)$ は, 入力画像 I を与えた際の注目クラス c に対応するクラススコアである。 $\lambda_{l2} \|I\|_2^2$ および $\lambda_{tv} \|\nabla I\|_2^2$ の項は画像 I に制約を加える正則化項である。次式で定義される $\|I\|_2$ は L-2 ノルムと呼ばれ, 画素値が過大なることを抑制する。

$$\|I\|_2 = \sum_{i,j} \sqrt{x_{ij}^2} \quad (11)$$

x_{ij} は入力画像 I の位置 (i, j) における画素値を表わす。 $\|\nabla I\|_2$ は total variation (TV) と呼ばれ, つぎの式で表わされるように隣り合う画素値の差の絶対値の総和である。

$$\|\nabla I\|_2 = \sum_{i,j} \sqrt{(x_{i,j+1} - x_{ij})^2 + (x_{i+1,j} - x_{ij})^2} \quad (12)$$

この項を最小化することにより, エッジを残しつつ不要な輝度変化を抑制し, 画像を平滑化する効果が得られる。なお, λ_{l2} および λ_{tv} は各項の効果を調節する係数である。(10) 式の目的関数を最小化することにより, 注目クラスのスコアを最大化しつつ不要な陰影が抑えられた典型的なモデル画像を得る。

この手法を出力層の各ニューロンに適用することで, 各クラススコアを最大化するような入力画像 (これを「クラスモデル」と呼ぶ) を得る。このクラスモデルにより, DCNN がどのような特徴にもとづき各クラスを識別しているかを推察することができる。

本研究ではこの手法を用いて腫瘤および正常組織のクラスモデルを生成し, DCNN が獲得した特徴の妥当性を確認した。**Fig. 10(a)** に CMV により生成された腫瘤のクラスモデルを示す。なお, 各モデル画像は陰影の視認性向上のためグレースケール化したのちガンマ補正 ($\gamma = 0.75$) を加えたものである。腫瘤のモデル画像からは, 乳房 X 線画像上の腫瘤陰影に類似した類円形の高輝度領域が複数確認できるほか, 腫瘤本体部分より放射状に伸びる線状の構造が確認できる。これは悪性の腫瘤にしばしば見られる「スピキュラ」に類似した特徴と考えられる。**Fig. 10(b)** に正常クラスモデルを示す。類円形の高輝度領域や放射状の陰影のような明確な

Fig. 8 Examples of detection result. The rectangles are true positive (TP) detections and false positive (FP) detections, and the ground truths are delineated by the dashed curves. The true negative (TN) detections are not displayed

Fig. 9 FROC curves of the proposed system and conventional systems. The better the performance, the closer the curve is to the upper left corner of the graph, where a higher true positive rate with a less false positives per image

構造は確認できず、線状もしくは帯状の淡い陰影が確認できる。これらは、正常な乳腺組織やそれらの重ね合わせによる陰影に類似した特徴と考えられる。これらの結果は、DCNNが腫瘍と正常組織の識別に有用な特徴を適切に獲得できたことを示唆するものといえる。

続いて、転移学習 (fine-tuning) 前後での DCNN の変化を比

較することを目的とし、fine-tuning 前後の DCNN の conv4, conv5, fc7 層のニューロンについて、それぞれモデル画像を生成した。各層は非常に多数のニューロンからなり、そのすべてを解析することは現実的でないため、DCNN に Fig. 10 (a) の腫瘍モデル画像を入力した際、各層で出力が最大のニューロンを選択し、その出力が最大となるモデル画像を生成した。結果を Fig. 11 に示す。DCNN は、浅い (入力側に近い) 層で局所的で単純な特徴を抽出し、それらがより深い (出力側に近い) 層に伝わるにつれ大域的で複雑な特徴を扱えるようになることが知られている。Fig. 11 から、浅い conv4 層は単純な特徴に、深い fc7 層は複雑な特徴に強く反応することが読み取れる。また fine-tuning 前後の変化に注目すると、いずれの層でも大まかな特徴を維持しつつ若干の変化が加わっていることが確認できる。特に変化が大きいのは出力層直前の fc7 層である。Fig. 11 (e) に示す fine-tuning 前の画像では鳥類の頭部に類似した陰影が確認できるのに対し、Fig. 11 (f) の fine-tuning 後では特徴が明らかに変化していることが見て取れる。これらの変化は、fine-tuning において、事前学習で獲得した特徴をある程度維持しつつ腫瘍の識別に有用な特徴につくりかえられたことを示唆するものであり、定性的な視覚評価ではあるが転移学習が意図どおりに働いていることを示す結果といえる。これらの特徴変化の定量評価は今後の課題である。

4. おわりに

本研究では、乳房 X 線画像上の病変検出を行なう CAD システムとして、まず多くの偽陽性検出を許容したシンプルな前処理で病変候補を抽出し、これらを DCNN により正しく

Fig. 11 The model images of conv4, conv5, and fc7 layers (gamma correction ($\gamma = 0.75$) was applied).

(a) The mass class model (b) The normal class model

Fig. 10 The class model images generated by the class model visualization (gray-scaled and gamma correction ($\gamma = 0.75$) was applied)

識別する 2 段階検出法を提案した。第 1 段階の前処理において病変候補をある程度絞ることで効率的かつ効果的な訓練データを提示可能とし、手動ではきわめて困難である特徴設計を行なうことなしに病変検出性能の大幅な向上を実現すると同時に、病変の位置ならびに大きさの提示が可能な CAD システムを実現した。近年、本研究のように DCNN を乳房 X 線画像上の病変検出に応用した例がいくつか報告されている^{24), 25)}。評価に用いている画像データが異なるため厳密な比較は困難であるが、いずれも従来より良好な検出性能を示しており、DCNN の乳がん病変検出に対する有効性を示唆するものである。

今後データ数を増やして信頼性を向上させることにより、提案システムを用いて医師の診断効率および診断精度の向上が実現され、乳がんの早期発見・早期治療による生活の質の向上ならびに医師の負担軽減による医療費削減への効果が期待される。

謝辞 本研究は JSPS 科研費 17K17582, 26540112 ならびに東北大学スマート・エイジング学際重点研究センターの助成を受けた。

参 考 文 献

- 1) A. Matsuda, T. Matsuda, A. Shibata, K. Katanoda, T. Sobue, H. Nishimoto and The Japan Cancer Surveillance Research Group: Cancer Incidence and Incidence Rates in Japan in 2008: A Study of 25 Population-based Cancer Registries for the Monitoring of Cancer Incidence in Japan (MCIJ) Project, *Japanese Journal of Clinical Oncology*, **44-4**, 388/396 (2013)
- 2) J. Tang, R.M. Rangayyan, J. Xu, I.E. Naqa and Y. Yang: Computer-Aided Detection and Diagnosis of Breast Cancer With Mammography: Recent Advances, *IEEE Trans. Inf. Technol. Biomed.*, **13-2**, 236/251 (2009)
- 3) H.P. Chan, D. Wei, M.A. Helvie, B. Sahiner, D.D. Adler, M.M. Goodsitt and N. Petrick: Computer-aided classification of mammographic masses and normal tissue: Linear discriminant analysis in texture feature space, *Phys. Med. Biol.*, **40-5**, 857/876 (1995)
- 4) B. Sahiner, H.P. Chan, N. Petrick, D. Wei, M.A. Helvie, D.D. Adler and M.M. Goodsitt: Classification of mass and normal breast tissue: A convolutional neural network classifier with spatial domain and texture images, *IEEE Trans. Med. Imag.*, **15-5**, 598/610 (1996)
- 5) N.R. Mudigonda, R.M. Rangayyan and J.E.L. Desautels: Gradient and texture analysis for the classification of mammographic masses, *IEEE Trans. Med. Imag.*, **19-10**, 1032/1043 (2000)
- 6) J. Wei, B. Sahiner, L. Hadjiiski, H. Chan, N. Petrick, M. Helvie, M. Roubidoux, J. Ge and C. Zhou: Computer aided detection of breast masses on full field digital mammograms, *Med. Phys.*, **32-9**, 2827/2837 (2005)
- 7) A. Krizhevski, I. Sutskever and G.E. Hinton: ImageNet Classification with Deep Convolutional Neural Networks, *Proc. NIPS*, 1097/1105 (2012)

- 8) K.J. Geras, S. Wolfson, Y. Shen, S.G. Kim, L. Moy and K. Cho: High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Network, arXiv preprint, arXiv:1703.07047 (2017)
- 9) S.J. Pan and Q. Yang: A Survey on Transfer Learning, *IEEE Transaction on Knowledge and Data Engineering*, **22**-10, 1345/1359 (2010)
- 10) N.A. Otsu: A threshold Selection Method from Gray-Level Histograms, *IEEE Trans. Sys. Man. Cyber.*, **9**, 62/66 (1979)
- 11) R.J. Ferrari, R.M. Rangayyan, R.A. Borges and A.F. Frilere: Segmentation of the Fibro-glandular Disc in Mammograms Using Gaussian Mixture Modeling, *Medical and Biological Engineering and Computing*, **42**, 378/387 (2004)
- 12) 小形, 本間, 石橋, 張, 大橋, 長谷川, 川住, 阿部, 杉田, 吉澤: 構造情報に基づく乳房 X 線画像上の腫瘍陰影検出法, 計測自動制御学会東北支部第 285 回研究集会予稿集, 285-2 (2013)
- 13) A.P. Dempster, N.M. Laird and D.B. Rubin: Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society, Series B (Methodological)*, **39**-1, 1/38 (1977)
- 14) 小畑秀文: モルフォロジー, コロナ社 (1996)
- 15) J. Deng, W. Dong, R. Socher, L.J. Li, K. Li and L. Fei-Fei: ImageNet: A Large-Scale Hierarchical Image Database, *CVPR09*, 248/255 (2009)
- 16) M. Heath, K. Bowyer, D. Kopans, R. Moore and W.P. Kegelmeyer: The Digital Database for Screening Mammography, *Proceedings of the Fifth International Workshop on Digital Mammography*, 212/218 (2001)
- 17) Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama and T. Darrel: Caffe: Convolutional Architecture for Fast Feature Embedding, arXiv preprint, arXiv:1408.5093 (2014)
- 18) A. Vedaldi and K. Lenc: MatConvNet: Convolutional Neural Networks for MATLAB, *Proc. of the ACM Int. Conf. on Multimedia*, 689/692 (2015)
- 19) P.C. Bunch, J.F. Hamilton, G.K. Sanderson and A.H. Simmons: A Free Response Approach to the Measurement and Characterization of Radiographic Observer Performance, *SPIE*, **127**, 124/135 (2000)
- 20) R. Campanini, D. Donquavanni, E. Iampieri, M. Masotti, G. Palermo, A. Riccardi and M. Roffilli: A Novel Featureless Approach to Mass Detection in Digital Mammograms based on Support Vector Machines, *Phys. Med. Biol.*, **49**-6, 961/975 (2004)
- 21) N.H. Eltonsy, G.D. Tourassi and A.S. Elmaghraby: A Concentric Morphology Model for the Detection of Masses in Mammography, *IEEE Trans. Med. Imag.*, **26**-6, 880/889 (2004)
- 22) A. Mahendran and A. Vedaldi: Understanding Deep Image Representations by Inverting Them, *CVPR*, 5188/5196 (2015)
- 23) K. Simonyan, A. Vedaldi and A. Zisserman: Deep Inside Convolutional Networks: Visualizing Image Classification Models and Saliency Maps, *ICLR*, arXiv: 1312.6034 (2014)
- 24) T. Kooi, G. Litjens, B. van Ginneken, A. Gubern-Mérida, C.I. Sánchez, R. Mann, A. den Heeten and N. Karssemeijer: Large scale deep learning for computer aided detection of mammographic lesions, *Medical Image Analysis*, **35**, 303/312 (2017)
- 25) D. Ribli, A. Horváth, Z. Unger, P. Pollner and I. Csabai: Detecting and classifying lesions in mammograms with Deep Learning, arXiv preprint, arXiv:1707.08401v3 (2017)

[著者紹介]

鈴木 真太郎

2016 年東北大学工学部情報知能システム総合学科卒業。同年東北大学大学院工学研究科電気エネルギーシステム専攻入学, 現在に至る。

張 曉 勇

2011 年東北大学大学院工学研究科博士後期課程修了。2011 年同大サイバーサイエンスセンター, 医学研究科博士研究員, 2015 年同大工学研究科助教。2017 年仙台高等専門学校総合工学科助教, 現在に至る。画像処理と医療応用の研究に従事。IEICE 会員, IEEE Senior 会員。

本間 経 康 (正会員)

1995 年東北大学大学院工学研究科博士後期課程修了。2000 年 Saskatchewan 大学客員教授, 2003 年東北大学医学部助教授などを経て, 2013 年同大学院医学系研究科教授, 現在に至る。人工神経回路網, 複雑系, 計算知能の医療応用などの研究に従事。IEEE, AAPM などの会員。

市 地 慶 (正会員)

2014 年東北大学大学院工学研究科博士後期課程修了。同年東北大学電気通信研究所研究教育支援者, 2015 年 JSPS 特別研究員 PD。2017 年東北大学大学院工学研究科電気エネルギーシステム専攻助教, 現在に至る。医用システムのための時系列信号処理の研究に従事。AAPM 会員, IEEE 会員。

高 根 侑 美

2009 年東北大学医学部保健学科卒業。2015 年東北大学大学院医学系研究科博士後期課程修了。同年九州大学大学院医学研究院保健学部助教。2016 年東北大学病院入職, 現在に至る。医用画像を用いた画像解析および画質評価の研究に従事 (保健学博士)。

柳 垣 聡

2014 年東北大学医学部医学科卒業。2017 年東北大学大学院医学系研究科放射線診断学分野入学。現在に至る。放射線画像診断の臨床, 研究に従事。

川 住 祐 介

2002 年東北大学医学部卒業。2008 年東北大学大学院医学系研究科博士過程修了。2009 年東北大学大学院医学系研究科保健学専攻助教。2016 年同専攻講師。現在に至る。放射線画像診断の臨床、教育、研究に従事（放射線診断専門医、医学博士）。

石橋忠司

1977 年東北大学医学部卒。84 年医学博士，日本医学放射線学会専門医取得。2005 年東北大学医学部保健学科教授，現在に至る。画像診断学（腹部，乳房領域），インターベンショナルラジオロジーの研究に従事。

吉澤 誠（正会員）

1978 年東北大学大学院工学研究科電気及通信工学専攻博士後期課程修了。工学博士。東北大学工学部助手、助教授、豊橋技術科学大学工学部助教授などを経て、2000 年サイバーサイエンスセンター教授。生体計測制御・遠隔医療に関する研究などに従事。

|||||