

RESEARCH ARTICLE

A Novel Cascade Classifier for Automatic Microcalcification Detection

Seung Yeon Shin¹, Soochahn Lee^{2*}, Il Dong Yun^{3*}, Ho Yub Jung³, Yong Seok Heo⁴, Sun Mi Kim⁵, Kyoung Mu Lee¹

1 Department of Electrical and Computer Engineering, ASRI, Seoul National University, Seoul, Republic of Korea, **2** Department of Electronic Engineering, Soonchunhyang University, Asan, Republic of Korea, **3** Division of Computer and Electronic Systems Engineering, Hankuk University of Foreign Studies, Yongin, Republic of Korea, **4** Department of Electrical and Computer Engineering, Ajou University, Suwon, Republic of Korea, **5** Department of Radiology, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

* sclsch@sch.ac.kr (SL); yun@hufs.ac.kr (IDY)



OPEN ACCESS

Citation: Shin SY, Lee S, Yun ID, Jung HY, Heo YS, Kim SM, et al. (2015) A Novel Cascade Classifier for Automatic Microcalcification Detection. PLoS ONE 10(12): e0143725. doi:10.1371/journal.pone.0143725

Editor: Zhaohong Deng, Jiangnan University, CHINA

Received: April 28, 2015

Accepted: November 8, 2015

Published: December 2, 2015

Copyright: © 2015 Shin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Individual μ C annotations of the mini-MIAS DB are available at <http://dx.doi.org/10.5061/dryad.jm6k3>. Mammograms and annotations of SNUBH-MDB- μ Cc are available upon request due to ethical restrictions. Interested researchers may submit requests to Sun Mi Kim (kimsmlms@daum.net). The authors do not own individual μ C annotations of the MIAS DB. This data is third party data, owned by A. Oliver (aoliver@atc.udg.edu).

Funding: SL has been funded by the Soonchunhyang university (<http://home.sch.ac.kr/english/index.jsp>) research fund. IDY was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF, <http://www.nrf.or.kr>)

Abstract

In this paper, we present a novel cascaded classification framework for automatic detection of individual and clusters of microcalcifications (μ C). Our framework comprises three classification stages: i) a random forest (RF) classifier for simple features capturing the second order local structure of individual μ Cs, where non- μ C pixels in the target mammogram are efficiently eliminated; ii) a more complex discriminative restricted Boltzmann machine (DRBM) classifier for μ C candidates determined in the RF stage, which automatically learns the detailed morphology of μ C appearances for improved discriminative power; and iii) a detector to detect clusters of μ Cs from the individual μ C detection results, using two different criteria. From the two-stage RF-DRBM classifier, we are able to distinguish μ Cs using explicitly computed features, as well as learn implicit features that are able to further discriminate between confusing cases. Experimental evaluation is conducted on the original Mammographic Image Analysis Society (MIAS) and mini-MIAS databases, as well as our own Seoul National University Bundang Hospital digital mammographic database. It is shown that the proposed method outperforms comparable methods in terms of receiver operating characteristic (ROC) and precision-recall curves for detection of individual μ Cs and free-response receiver operating characteristic (FROC) curve for detection of clustered μ Cs.

Introduction

Breast cancer is the most common cancer in women worldwide and the second most common cancer overall [1]. In the US, nearly 1 in 36 (2.8%) women die of breast cancer related illnesses, and about 1 in 8 (12%) will develop invasive breast cancer during their lifetime [2]. The majority of cases of breast cancer are associated with clusters of microcalcifications (μ C), which are tiny calcium deposits that appear as little white spots on a mammogram [3]. Thus, identifying clusters of μ Cs is crucial for the detection of breast cancer. Thus, a significant amount of research on computer aided detection (CAD) of μ Cs in mammograms has been conducted.

www.nrf.re.kr/nrf_eng_cms/), funded by the Ministry of Education, Science and Technology (2013R1A1A2A10004550, 2015R1A5A7036384). HYJ was supported by Hankuk University of Foreign Studies Research Fund of 2015. SMK was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korean government (MEST, <http://english.msip.go.kr/english/main/main.do>) (grant code: 2012R1A1A3008621). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Fig 1 shows a digital mammogram containing a cluster of μ Cs. The detection of individual μ Cs is generally difficult due to their diversified morphologies, small size and surrounding tissues. Moreover, many false positives are likely to occur due to dense breast tissue, cysts, or noise with similar local appearance with μ Cs.

Previous works can be classified into two categories, *image processing based methods* and *learning based methods*. Many early works were based on image processing techniques, including image enhancement [4, 5], filtering [6–8], morphological operations [9], and coding [10–

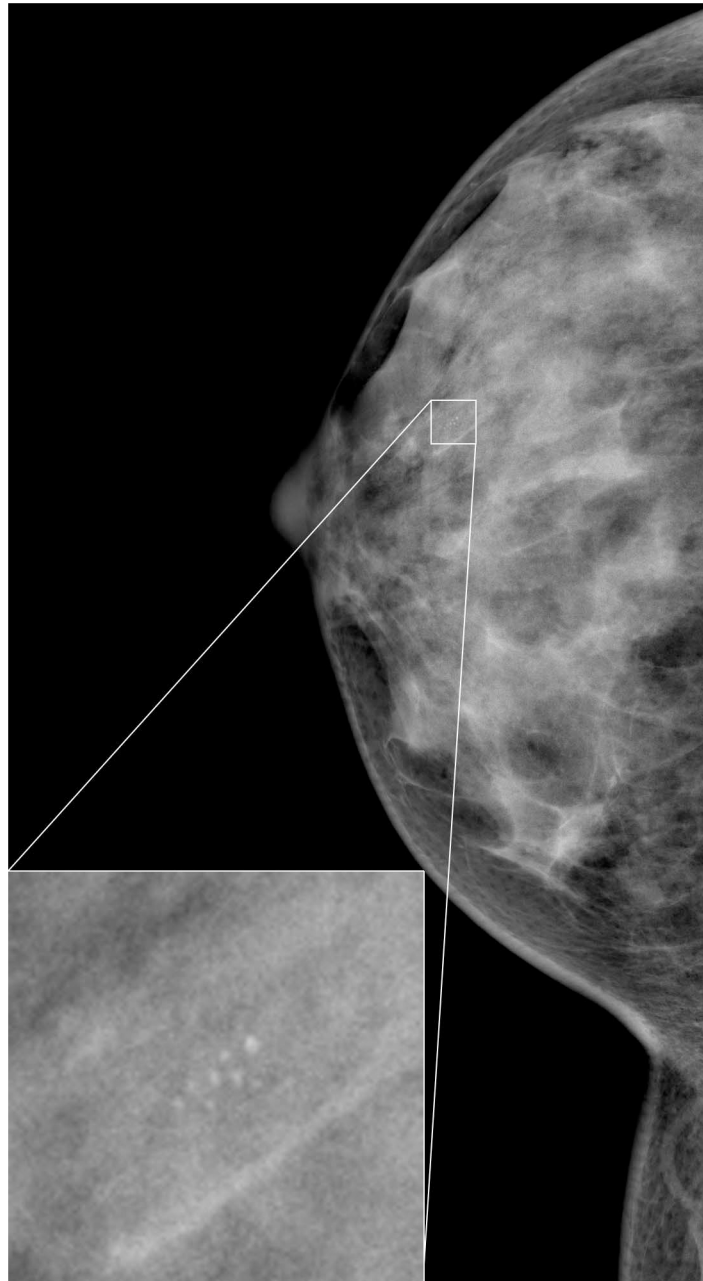


Fig 1. An example mammogram and the enlarged view of a contained microcalcification cluster. Intensity windowing is applied for improved visibility.

doi:10.1371/journal.pone.0143725.g001

[12]. For the most part, these methods are designed by experts with specific targets regarding the characteristic appearance that it aims to distinguish. Thus, while these techniques are successful for the targeted types of μ Cs, they may be limited for μ Cs with subtle local appearance variations, since it is extremely difficult to manually examine and consider all possible cases for a large amount of real data.

More recently, several learning based methods have been developed based on developments in discriminative local feature descriptors and machine learning methods for classification. During the past decade, methods using different machine learning methods such as boosting [13], support vector machines (SVM) [14] and relevance vector machine (RVM) [15] have been proposed. In these methods, selecting the appropriate feature is a critical issue for improving detection quality. The appropriate feature must effectively represent the discerning characteristics between the μ Cs and normal tissue. Usually, the feature selection procedure needs expert knowledge and requires time-consuming hand-tuning. For example, in the method by Oliver et al. [13], features extracted from a bank of filters are selectively used in a boosted classifier framework.

Previous works can also be classified by the specific configuration of μ Cs in which to detect. Malignant tissues often occur as closely clustered μ Cs within a limited region. Therefore, detection of clusters of μ Cs is of higher clinical importance than the detection of individual μ Cs.

Several works thus aim to directly detect μ C clusters. An example is the work by Papadopoulos et al. [16, 17], where a μ C cluster is detected by classifying features characterizing either an individual, or a group of, μ Cs in a particular region of interest (ROI).

Other works aim to detect individual μ Cs first, and then detect μ C clusters. The assumption here is that, compared to clusters of μ Cs, the appearance of individual μ Cs is more consistent and thus can be better modelled. And, based on robust detections of individual μ Cs, the μ C clusters can accurately be detected as well. In the work by El-Naqa et al. [14], a method comprising a pixel-wise SVM classifier trained to detect individual μ Cs and a successive enhancement learning (SEL) algorithm to enhance accuracy is proposed. In the work by Wei et al. [15], a linear classifier and a non-linear RVM classifier are combined in a two-stage network to enhance accuracy and efficiency. In both works, the evaluation of individual μ C detection is conducted not on actual clinical cases where only an extremely small number of pixels are μ Cs but on comparatively well-balanced test samples. In the work by Bria et al. [18], a ranking based cascade μ C detector based on classifiers for Haar-like features is followed by a classifier for μ C clusters. Their evaluations on custom datasets show competitive performance compared to high-end commercial systems.

In this paper, we present a new cascaded classification framework for automatic detection of μ Cs and μ C clusters. Our framework comprises three classification stages, where the first and second stages enable accurate and robust detection of individual μ Cs, while the final stage aggregates the individual μ Cs to detect μ C clusters. More specifically, a random forest (RF) classifier [19, 20] learned using simple features capturing the second order local structure of μ Cs is applied to efficiently eliminate non- μ C pixels in the target mammogram. μ C candidates determined in the first stage are then further classified using a more complex discriminative restricted Boltzmann machine (DRBM) classifier [21] to obtain individual μ C detections. By dividing the detection process into two separate and complementary classification stages, we are able to utilize implicit characteristics learned from the representation learning of the DRBM as well as explicit characteristics of μ C appearance in the first stage. Finally, we detect μ C clusters by aggregating the detected individual μ Cs with two different rules.

The key advantages of the proposed framework are as follows: 1) Improved accuracy of individual μ C detection based on the DRBM which automatically learns the detailed morphology of μ C appearances. 2) Improved efficiency of individual μ C detection from the fast RF

classification. Not only does this stage improve efficiency, it enhances the second stage by focusing the discriminative power of the DRBM exclusively for difficult cases with subtle differences in appearance. 3) Improved accuracy of μC cluster detection based on more accurate and robust individual μC detection. We note that this work is an expansion of our previous work [22], which only described the DRBM classification stage.

The validity of the proposed method is evaluated and compared with relevant methods, on digitized mammograms from the original Mammographic Image Analysis Society (MIAS) database [23] and the mini-MIAS database [24], a processed version of the original, as well as digital mammograms obtained from Seoul National University Bundang Hospital (SNUBH).

Methods

Training The Individual Microcalcification Detector

Cascaded Classification Framework. On a typical mammogram, only an extremely small fraction of pixels are actually μC s, as shown in Fig 1. Also, only a small number of pixels that are not μC s have similar local appearance to μC s. The appearances of most other background pixels are significantly different. Given the high resolution of recent mammography systems, it is imperative that unlikely candidates are excluded as efficiently as possible. We thus propose a two-stage cascaded classification network to speed up the whole procedure. Fig 2 shows the schematic diagram of our framework. This speed-up is achieved in the first stage by learning a highly efficient RF classifier. Even with the limited computational burden, the RF classifier has enough discriminative power to exclude most non- μC pixels. The input dataset to the second stage is then small enough to be evaluated by the complex DRBM classifier. By combining the two stages, not only are we able to achieve higher efficiency, but also higher accuracy, since two classifiers cooperate to distinguish false positives.

Stage-1: Random Forest Classifier with Hessian Eigenvalue Features. The use of eigenvalues of the Hessian matrix as a measure of local appearance characteristics was proposed by Frangi et al. in [25]. Their focus was on tubular structures for vessel enhancement. Here, we apply a similar concept for the blob-like structure of μC s. We briefly introduce the concepts suggested in [25] and explain how we utilize it to detect μC s.

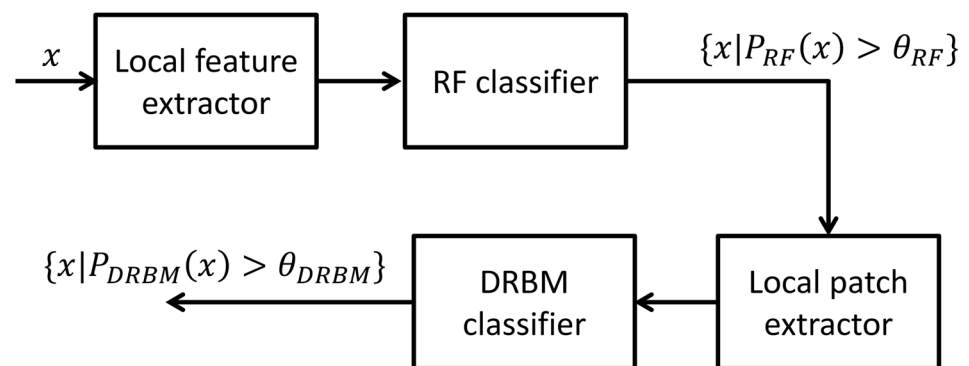


Fig 2. Schematic diagram of the proposed cascaded classification framework. Extracted local features from x , which are all pixels in a region of interest (ROI), are classified by the first RF classifier to produce μC probabilities $P_{RF}(x)$ and μC candidates are detected by thresholding these probabilities with a certain value θ_{RF} . The μC candidates are classified again by the second more sophisticated DRBM classifier. The input of the DRBM classifier is a local image patch and final detection results are again acquired by thresholding the probabilities $P_{DRBM}(x)$ with a certain value θ_{DRBM} .

doi:10.1371/journal.pone.0143725.g002

Given an image I , the Hessian matrix of the image computed in the neighborhood of a pixel

$$p \text{ is defined as } \mathbf{H}_{p,\sigma} = \begin{bmatrix} \frac{\partial^2 I(p, \sigma)}{\partial x^2} & \frac{\partial^2 I(p, \sigma)}{\partial x \partial y} \\ \frac{\partial^2 I(p, \sigma)}{\partial x \partial y} & \frac{\partial^2 I(p, \sigma)}{\partial y^2} \end{bmatrix}. \text{ The image differentiation is calculated as a}$$

convolution with Derivatives of Gaussians (DoG) and the standard deviation of the Gaussian, σ , should be tuned to correspond to the size of the μC to detect. By eigen-decomposition of the Hessian $\mathbf{H}_{p, \sigma} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$, we can compute the eigenvalues $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2)$, ($|\lambda_1| < |\lambda_2|$) and their corresponding eigenvectors $\mathbf{Q} = (\mathbf{q}_1, \mathbf{q}_2)$. The eigenvalues λ_1 and λ_2 represent the image gradients in the principal directions represented by the eigenvectors \mathbf{q}_1 and \mathbf{q}_2 , and thus represent local appearance characteristics, as summarized in Table 1. We define λ_1^i and λ_2^i together with their ratio $R^i = \frac{|\lambda_1^i|}{|\lambda_2^i|}$ and the Frobenius norm of the Hessian $\|\mathbf{H}\|_F^i = \sqrt{\sum_{j \leq 2} \lambda_j^2}$ as the feature

vector components $\mathbf{v}^i = \{\lambda_1^i, \lambda_2^i, \|\mathbf{H}\|_F^i, R^i\}$ of the single i_{th} scale σ_i . The reason we use the eigenvalues together with their ratio and norm is to maximize the discriminative power of the random forest. By including the norm, which represents the image gradient magnitude, and the ratio, which represents whether the local appearance is more blob-like or tubular, we can increase discriminative power and possibly reduce the depth of the decision trees in the random forest. To enable detection of μCs of multiple sizes, we construct the total feature vector by concatenating components from multiple scales as $\mathbf{v} = \{\mathbf{v}^i | i = 1, \dots, n_\sigma\}$, where n_σ is the number of examined scales.

Random forest (RF) [19, 20] is an ensemble classifier comprising multiple decision trees. Each tree is learned using a randomly sampled subset S of the full training data set T . This subset S is inserted into the root node and recursively partitioned at each node by an optimized decision into disjoint subsets, which correspond to the left and right child nodes. The optimized decisions at each node is defined by the parameters $\theta = (\varphi, \tau)$, where φ is a particular dimension index among \mathbf{v} and τ is a threshold of a binary test. Depending on the possible values of φ and τ , there can be a very large number of decision candidates. The optimal decision is defined as that which maximizes the expected information gain, defined as:

$$\theta^{opt} = \arg \max_{\theta \in \Theta_j} I(S_j, \theta), \tag{1}$$

where

$$I(S_j, \theta) = H(S_j) - \sum_{i \in \{L,R\}} \frac{|S_j^i|}{|S_j|} H(S_j^i). \tag{2}$$

S_j denotes the subset of data samples which has reached the node j . Θ_j is the set of decision candidates for node j , which is often a randomly sampled subset of the all possible decisions. L

Table 1. Patterns of the value of the eigenvalues λ_k (H = high, L = low, N = noisy, no clear tendency and usually small, +/- indicate the sign of the eigenvalue). ($|\lambda_1| < |\lambda_2|$).

| λ_1 | λ_2 | pattern |
|-------------|-------------|------------------------------|
| N | N | noisy |
| L | H- | tubular structure (bright) |
| L | H+ | tubular structure (dark) |
| H- | H- | blob-like structure (bright) |
| H+ | H+ | blob-like structure (dark) |

doi:10.1371/journal.pone.0143725.t001

and R mean left and right child nodes, respectively. $H(S_j) = -\sum_{c \in C} p(c) \log(p(c))$, where c is a particular class among the set of classes C , is the Shannon entropy, which represents the information gain for classification tasks. Partitioning is repeated until the subset at a particular node has homogeneous labels or its size is smaller than some predefined minimum, or if the node depth is larger than the predefined maximum. The distribution of labels of the subset at each leaf node defines the posterior class probabilities $p(c|\mathbf{v})$ for a given feature vector \mathbf{v} that reaches that node. The recursive splits refine the class distributions in their child nodes so that input cases are better distinguished. The forest output is computed as the average of individual trees:

$$p(c|\mathbf{v}) = \frac{1}{n} \sum_i^n p_i(c|\mathbf{v}), \tag{3}$$

where \mathbf{v} denotes the feature vector of the test instance and n denotes the number of trees in our forest. We note that, in our case we have binary classes, μC or non- μC . We thus obtain classification by thresholding the posterior μC probability. Although it should theoretically be 0.5, we can control this threshold to control the ratio of true μC s to false positives in the results.

Stage-2: Discriminative Restricted Boltzmann Machine Classifier. The detected candidates from the stage-1 are further classified by the DRBM classifier. Here, the raw image patch centered on the candidate pixel is used as the feature for the DRBM classifier, rather than the Hessian eigenvalue feature of the previous stage. We briefly explain the concept and model of DRBM below. We refer the reader to [21] for further details.

RBM [26] is a generative model based on a layer of hidden variables to capture a distribution over visible variables. Since RBM can effectively model common characteristics as well as discriminative features over the training data, they have been used to train features for other learning algorithms such as neural networks and SVMs. Thus, RBM is well suited for our problem of modeling the morphological ambiguity of μC s. Given a training set $D = \{(\mathbf{x}_i, y_i), i = 1, 2, \dots, |D|\}$ comprising the input image patches \mathbf{x}_i and their binary class labels $y_i \in \{0, 1\}$, we minimize the

negative log-likelihood $L(D) = -\sum_{i=1}^{|D|} \log p(y_i, \mathbf{x}_i)$ to train a generative model. The joint distribution between the hidden variables \mathbf{h} , visible variables \mathbf{x} , and class label y is modeled as:

$$p(y, \mathbf{x}, \mathbf{h}) \propto e^{-E(y, \mathbf{x}, \mathbf{h})}, \tag{4}$$

where

$$E(y, \mathbf{x}, \mathbf{h}) = -\mathbf{h}^T \mathbf{W} \mathbf{x} - \mathbf{b}^T \mathbf{x} - \mathbf{c}^T \mathbf{h} - \mathbf{d}^T \vec{y} - \mathbf{h}^T \mathbf{U} \vec{y}. \tag{5}$$

Here, \vec{y} denotes a vectorized version of the class label y . Specifically, given C possible class labels for y , \vec{y} is a C dimensional vector with all zero elements except the y_{th} element, which is set to 1. The matrices \mathbf{W} and \mathbf{U} represent the weights representing the relation between \mathbf{x} and \mathbf{h} , and \mathbf{x} and \vec{y} , while \mathbf{b} , \mathbf{c} , and \mathbf{d} are the vectors representing the biases for \mathbf{x} , \mathbf{h} , and \vec{y} , respectively. The objective of training is to optimize all the weight matrices and vectors according to $L(D)$. The graphical model of the RBM is illustrated in Fig 3.

DRBM is a variant of the RBM made suitable for classification. For the DRBM, the conditional probability $p(y|\mathbf{x})$ is optimized instead of the joint probability $p(y, \mathbf{x})$ as:

$$L(D) = -\sum_{i=1}^{|D|} \log p(y_i|\mathbf{x}_i). \tag{6}$$

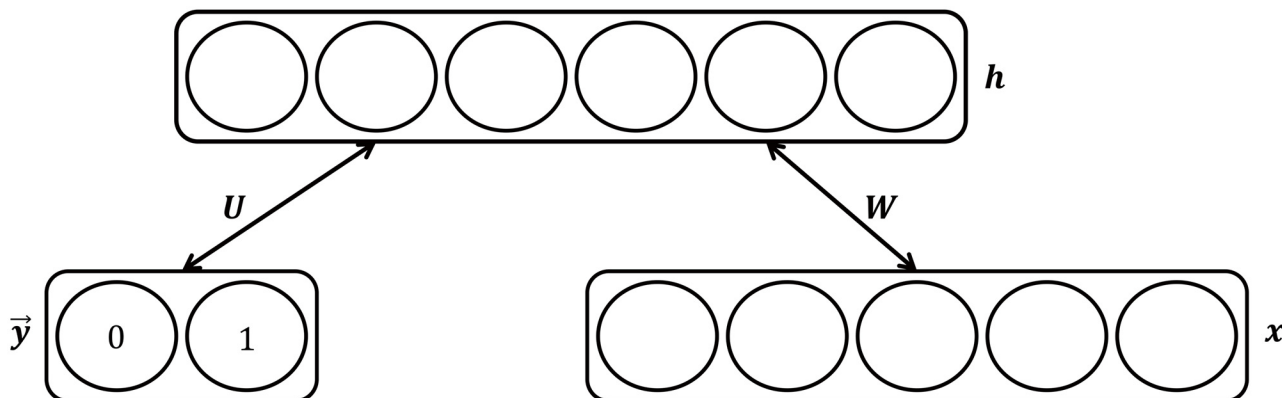


Fig 3. Restricted Boltzmann Machine. In training, the optimal values for weights W , U , which define the relations between visible variables x and hidden variables h , h and the class label \vec{y} , respectively, and b , c , d , which are biases in x , h , and \vec{y} , respectively, are computed. When new test data is inserted, the model computes the class label \vec{y} by using the learned parameters. We note that this figure was previously presented in [21] and is reprinted here for the description of RBM.

doi:10.1371/journal.pone.0143725.g003

This is then used to infer the label for an input image patch. Since the DRBM is used to model features from detection candidates obtained from stage-1, the particular features that distinguish its false positives from actual μ Cs are learned. The μ C probability is obtained from the DRBM, from which we can control the final detection results by adjusting the probability threshold.

Detecting Microcalcification Clusters

From the individual μ Cs detected from the cascaded classifier, we detect clusters of μ Cs using two different rules. The first rule, proposed by Kallergi et al. [27] and previously applied to μ C cluster detection by Wei et al. [15], defines a group of detected individual μ Cs as a true positive cluster if: 1) the distances between each pair of the μ Cs are less than d_{c1} , and 2) three or more true μ Cs are located within some localized region A_{c1} . If none of the μ Cs are true positives, then the cluster is assigned as a false positive. The second rule, proposed by Oliver et al. [13], is based on the cumulative local probability p_C from individual μ C detection probabilities p_{ind} which can be written as:

$$p_C(x, y) = \int_{\Omega} p_{ind}(x', y') dx' dy', \tag{7}$$

where Ω denotes again some localized region A_{c2} for which the p_{ind} s are accumulated. Thresholding p_C with a certain value yields cluster detections. If the center of the detected cluster is inside an annotated cluster, we regard it as a true positive. In this paper, we evaluate the results based on both rules.

Microcalcification Detection for a New Image

Given a new image, the individual μ Cs and their clusters are detected by the following process:

Breast Region Segmentation. As can be seen in Fig 1, a large portion of mammograms represent empty space. Since we are interested only in the breast region, we must first segment this region. Here, we apply a method similar to the one by Papadopoulos et al. [16], which comprises the following steps on an input image: (1) adaptive thresholding; (2) morphological dilation to smooth resulting binary labels obtained from (1); and (3) connected component analysis

to detect the largest region. For the adaptive thresholding, we construct an intensity histogram and determine the threshold based on the maximum count bin b_{max} . Since pixels in empty space have very similar intensities and occupy nearly half the image, they invariably correspond to b_{max} . A fixed number of bins, to the right of b_{max} are examined and the threshold is determined as the intensity value that delineates the bins with the maximum declivity of the bin count. The inclusion of adaptive thresholding compared to the method of [16] helps to improve robustness since each subject has different tissue, and thus different intensity distributions.

Classification of Individual μ Cs. We first apply a learned RF classifier to all pixels in the segmented breast region in a sliding window manner. We compute Hessian matrices for each pixel, and then calculate its eigenvalues. The feature vector \mathbf{v} is constructed from the eigenvalues for each pixel, which is then given as the input for classification by the RF.

Due to the sliding window approach, multiple positive pixel classifications can constitute a single μ C. Thus, the RF classification of the whole image results in a label map with multiple connected components. Since we want to obtain single coordinates corresponding to the center of each μ C, we perform non-maximum suppression, by determining the point that has the maximum normalized cross correlation value with a two-dimensional Gaussian kernel for each connected component.

We further classify the extracted points from the non-maximum suppression with a learned DRBM classifier. A fixed-sized raw image patch centered on the μ C candidate coordinate is given as the input vector for the classifier. Final detections are determined by thresholding the inferred probabilities obtained by the classifier.

Classification of Clusters of μ Cs. To detect μ C clusters by the rule of Kallergi et al. [27] and Wei et al. [15], for each individual μ C detection, we determine other μ C detections that are within d_{c1} . If there are more than three of these μ Cs, we determine these μ Cs as a cluster. In case of overlapping clusters, we select the one with the largest number of individual μ C detections.

To detect clusters by the rule of Oliver et al. [13], for each pixel, we compute the cumulative local probability p_C of a window of fixed size, centered at that pixel, and perform thresholding to determine local μ C cluster windows. In case pixels with μ C cluster windows form a connected component, the windows form a combined μ C cluster where the center and the bounding box are that of the connected component.

Results

Mammogram Datasets

The proposed method is evaluated on three different datasets. The details of each dataset are given below.

SNUBH-MDB. The Seoul National University Bundang Hospital Digital Mammographic database (SNUBH-MDB) comprises 319 digital mammograms, all with spatial resolution of 0.1 mm/pixel, pixel resolution of 1914×2294 pixels, and 12 bit depth, from 175 clinical cases all obtained from the SNUBH using a GE Senographe 2000D digital mammography system.

Ethics approval for the study was obtained from the SNUBH for this retrospective study, and the requirement for informed consent was waived. All patient records and information were anonymized and de-identified prior to analysis.

All cases were examined by radiologists and biopsies were performed on subjects with inconclusive results. Through these clinical procedures, all mammograms were labeled as malignant, benign, or normal. For each mammogram, we achieved the bounding boxes of existing clusters of μ Cs from manual annotations by radiologists. For 49 randomly selected mammograms, each from different clinical cases, the pixel coordinates of individual μ Cs were additionally provided by the same radiologists. When training the individual μ C detector,

positive data were sampled from coordinates of individual μC , while negative data were sampled randomly from breast regions excluding the bounding box of μC clusters. For brevity, we refer to these 49 mammograms as SNUBH-MDB- μCi and the remaining 270 mammograms, from 126 clinical cases, as SNUBH-MDB- μCc from this point on. The SNUBH-MDB- μCi is composed of 20 benign (from 20 clinical cases) and 29 malignant (from 29 clinical cases) mammograms. The SNUBH-MDB- μCc is composed of 3 types of data, specifically, 106 benign (from 54 clinical cases), 62 malignant (from 32 clinical cases), and 102 normal (from 40 clinical cases) mammograms. While biopsies were not performed on normal cases, they can be further divided into cases with and without μCs . μCs in normal cases have unclear boundaries and are hard to detect. Overall, the SNUBH-MDB- μCc contains more challenging and wide range of patients than the SNUBH-MDB- μCi .

MIAS DB. The Mammographic Image Analysis Society (MIAS) database [23] comprises 207 normal mammograms, 90 mammograms containing abnormalities including masses, spiculated masses, and architectural distortions, but without μCs , and 25 mammograms containing a total of 28 clusters of μCs . The total 322 images are of spatial resolution 0.05 mm/pixel and optical density linear in the range 0 – 3.2 and quantised to 8 bits. The size of the mammograms were either small (1600×4320 pixels), medium (2048×4320 pixels), large (2600×4320 pixels), or extra-large (5200×4320 pixels). For images containing μCs , the center coordinate and the radius of the circle enclosing the μC cluster are included in the original database. To train our individual μC detector, we used the annotations of the individual μCs used in [13], provided by its authors.

Mini-MIAS DB. The mini-MIAS database [24] is a reduced, clipped and padded version of the original, where the spatial resolution is reduced to 0.2 mm/pixel, and the sizes of all images are fixed to 1024×1024 pixels. While it includes the same 322 images, μC cluster annotations are provided for only 23 mammograms. We note that due to the differences in image configuration, we obtained annotations of individual μCs (as pixel coordinates) from radiologists, independently from that we had obtained for the original MIAS DB.

Access to DBs. Mammograms and all annotations of the SNUBH-MDB- μCi , mammograms of the MIAS DB, and individual μC annotations of the mini-MIAS DB are available at <http://dx.doi.org/10.5061/dryad.jm6k3>. Mammograms and annotations of SNUBH-MDB- μCc are available upon request for researchers who meet the criteria for access to confidential data. Requests should be made to Sun Mi Kim (kimsmlms@daum.net), see <http://cv.snu.ac.kr/research/cascade-mc-detector15/index.html> for more information. Mammograms and μC cluster annotations of the mini-MIAS DB are available at <http://peipa.essex.ac.uk/info/mias.html>. Individual μC annotations of the MIAS DB are available at <http://eia.udg.edu/~aoliver/mammoWeb/lesMic.html>.

Experimental Settings

Training Sets. A positive and negative set of feature vectors of individual μCs are required for training the individual μCs classifier. Since we have obtained expert annotations for the SNUBH-MDB, MIAS DB, and the mini-MIAS DB, it is straightforward to construct the positive sets, which comprises the features at the annotated μC coordinates. This same set is applied for both the stage-1 RF and stage-2 DRBM classifiers.

To construct the negative set of the first stage RF classifier, we randomly sample coordinates from the whole breast region except the bounding boxes of μC cluster, for the SNUBH-MDB and the mini-MIAS DB. For the MIAS DB, the obtained expert annotations also include coordinates of negative samples. Thus we use these coordinates to construct the negative set in this case.

For the negative set of the stage-2 DRBM classifier, we sample from the false positives of the stage-1 RF classifier. This is the same for all datasets. We set the size of the negative set to be the same as that of the stage-1 RF classifier.

Parameters. Parameter values are empirically set for each different dataset. The set of scales in stage-1 RF are fixed as $\sigma = \{1\}$, $\sigma = \{1, 2, 3, 4, 5\}$, $\sigma = \{0.5, 1, 1.5, 2, 2.5\}$, for the SNUBH-MDB, MIAS DB, and the mini-MIAS DB, respectively. The RF in the stage-1 RF, comprises 100 trees, each with maximum depth 30. For each tree, the training is stopped when the maximum depth is reached or when no node can be split without resulting in a child with corresponding dataset size smaller than the 1% of the training set size. The probability threshold of the RF classifier, which can be adjusted to optimize the stage-2 DRBM classification performance, is also empirically determined as 0.3, 0.5, and 0.5 for the SNUBH-MDB, MIAS DB, and the mini-MIAS DB, respectively. We note that, the threshold value 0.3 for the SNUBH-MDB is smaller than 0.5, which means that the μC probability $p_{\mu C}^{RF}$ can become smaller than the non- μC probability $p_{\mu C}^{RF} = 1 - p_{\mu C}^{RF}$. This is because the emphasis in this stage is on minimising false negatives, even if some non- μC s incorrectly pass, since they can be excluded in the next stage.

For the stage-2 DRBM classifier, which is trained by contrastive divergence [28], we have to tune many hyper-parameters including the size of the visible layer, the size of the hidden layer, the learning rate, and the number of iterations over the training set [21]. The optimal values for the hyper-parameters were determined by a grid-like search on a validation set. The values of visible layer size (input patch size), hidden layer size, learning rate, and the number of iterations, used for evaluation were, 225 (15×15), 50, 0.02, 100 for the SNUBH-MDB, 841 (29×29), 200, 0.02, 30 for the MIAS DB, and 81 (9×9), 200, 0.04, 30 for the mini-MIAS DB, respectively. When training, we use “mini-batches” for faster convergence. For more details regarding this process, we refer the reader to [21] and [29].

For μC cluster detection, we set as $d_{c1} = 0.4cm$ and $A_{c1} = 1cm^2$ for the first rule of Wei et al. [15] and $A_{c2} = 1cm^2$ for the second rule of Oliver et al. [13], all of which were the values used in the respective original works. In terms of the mammograms used in our experimental evaluation, the area of $1cm^2$ corresponds to 100×100 , 200×200 , and 50×50 pixels for the SNUBH-MDB, MIAS DB, and mini-MIAS DB, respectively.

Quantitative Evaluation

We evaluate the accuracy of 1) the individual μC detection step and 2) the μC cluster detection framework. Performance of our method is compared to the methods of Wei et al. [15] and Oliver et al. [13]. For both methods, we use our own implementations. We have carefully followed the descriptions in the papers during implementation, and adapt the parameters to each database for the best results.

Evaluation of individual μC detection framework. Since annotations of individual μC s must be given and the images containing μC s are limited, especially for the MIAS DB and mini-MIAS DB, we only evaluate individual μC detection accuracy on the SNUBH-MDB- μC i, by 10-fold cross validation. In cross validation, the whole dataset is divided into 10 different groups, where one group is designated for the test, and the remaining groups are used to train the classifier. This process is repeated 10 times so that testing is performed at least once for each group.

Quantitative performance is measured by constructing the receiver operating characteristic (ROC) curve as well as the precision-recall curve. The ROC curve is a plot of sensitivity, i.e., the ratio of μC s actually detected among all μC s present in the images, against (1-specificity), i.e., the ratio of false positive detections among all non- μC pixels. Precision-recall curve is a plot of precision, i.e., the ratio of true μC s among all positive detections, against recall, which is

identical to sensitivity. Different values of sensitivity/specificity and precision/recall are obtained at different operating points. Here, the operating point is controlled by a single parameter, namely, the threshold for the probability inferred by the stage-2 DRBM classifier.

Fig 4 shows the ROC and precision-recall curves for the methods in [13, 15], the stage-1 RF classifier, and the proposed cascaded classifier. For both ROC and precision-recall, a higher curve represents better performance. We can easily see that our first stage RF classifier has slightly higher accuracy than the methods of [13] and [15], while the proposed cascaded classifier clearly performs better than the previous methods. In terms of the ROC curve, with the same sensitivity value of 0.9 as the methods of [13, 15], the stage-1 RF classifier, and the proposed cascaded classifier, the (1-specificity) values are 0.00146, 0.00027, 0.00018, and 0.00007, respectively. This means that when the number of true positives is 869 among a total of 966 μ Cs, the number of false positives are 75036, 13902, 9025, and 3819, respectively. In terms of the precision-recall curve, precisions are 0.01, 0.06, 0.09, and 0.19 for a same recall value of 0.9 for the methods of [13, 15], the stage-1 RF classifier, and the proposed cascaded classifier, respectively. This means that to correctly detect 869 true positives among all target μ Cs, the total numbers of detections obtained are 63414, 14435, 9719, and 4665. This also means that the proposed cascaded classifier only gives 6% and 28% of false positives compared to [13] and [15], respectively, when detecting a similar rate of μ Cs.

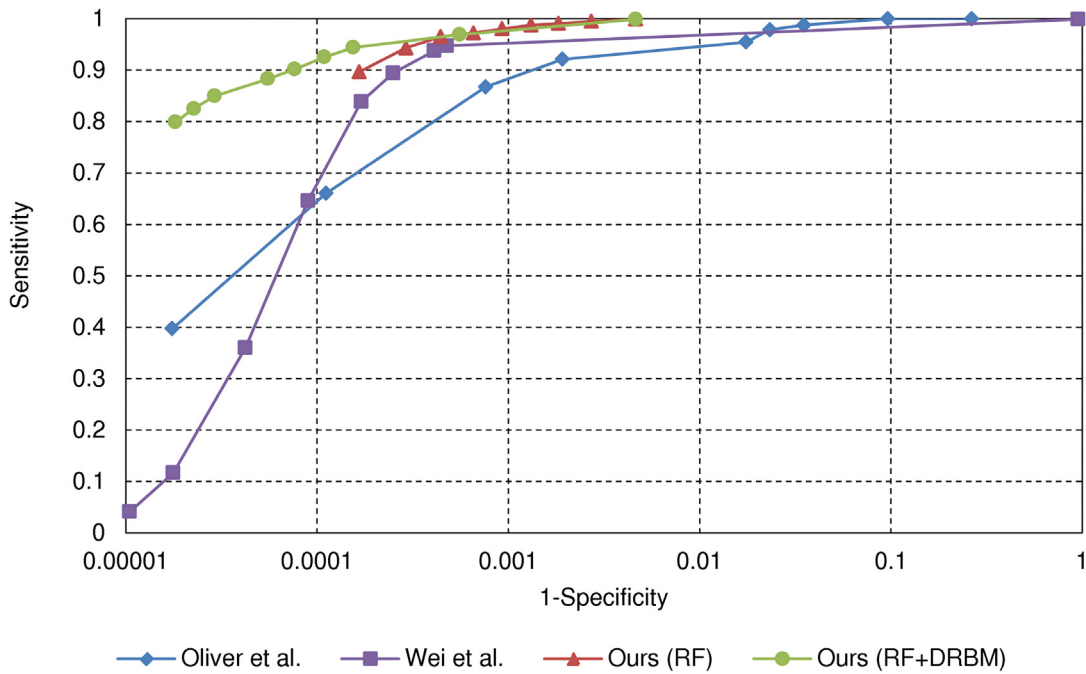
Evaluation of μ C cluster detection framework. The detection accuracy is evaluated by comparing the free-response receiver operating characteristic (FROC) curves [30] of the different methods. The FROC curve plots the sensitivity on the cluster detection, i.e., the ratio of μ C clusters actually detected among all μ C clusters present in the images, against the average number of false positive clusters per image. A higher FROC curve represents better performance.

We evaluate the detection accuracy based on the first rule, from [15, 27], on the SNUBH-MDB- μ C_i, by 10-fold cross validation. For this rule, the operating point is controlled by thresholding the probability for individual μ C detection. Using the second rule from [13], the evaluations are performed on the SNUBH-MDB- μ C_i, SNUBH-MDB- μ C_c, MIAS DB, and mini-MIAS DB. Again, 10-fold cross validations are performed for SNUBH-MDB- μ C_i, MIAS DB, and mini-MIAS DB. We note that, for the MIAS DB and mini-MIAS DB, images with and without μ Cs are each divided into 10 groups of uniform size, so that testing is performed on images without μ Cs as well. For the SNUBH-MDB- μ C_c, training is performed on the SNUBH-MDB- μ C_i, thus constituting a more realistic experimental setup. For the second rule, the operating point is controlled by the threshold for the cumulative local probability p_C .

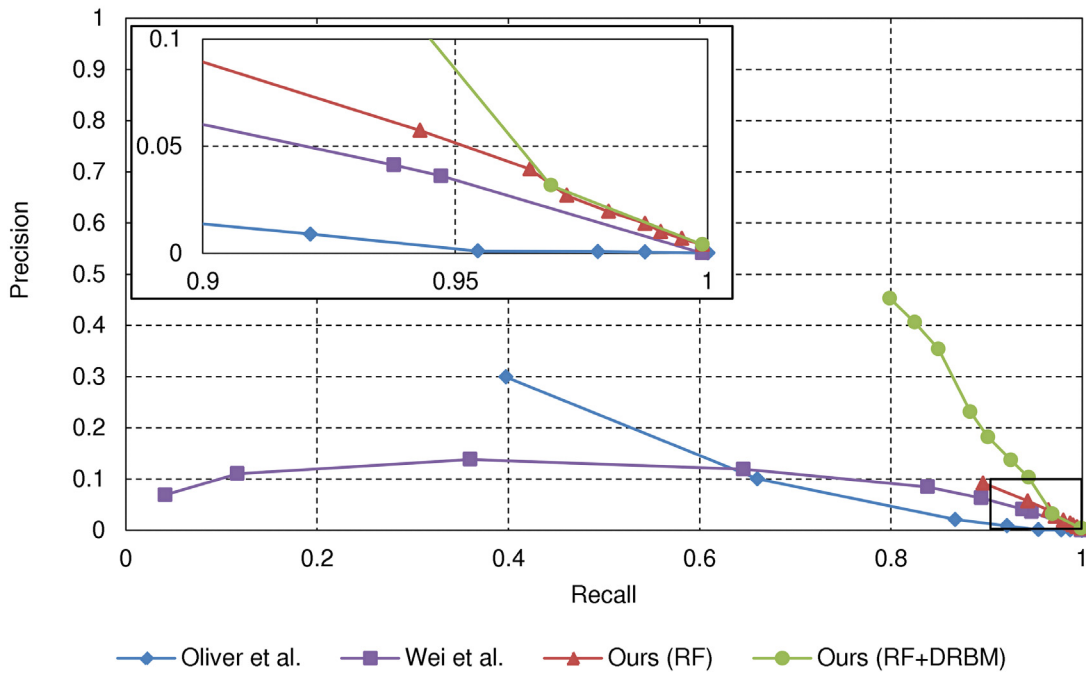
Fig 5 shows the FROC curves for the SNUBH-MDB- μ C_i. In Fig 5(a), the cluster detection results based on the individual detection results of the stage-1 RF classifier and the cascade classifier, using the same rule as in [13], are compared with the method of [13]. The AUC values of each method are 0.82, 0.92, and 0.79. In Fig 5(b), the results using the same rule as in [15] are compared with the method of [15]. The AUC values of each method are 0.82, 0.91, and 0.84. Overall, the cluster detection results based on the stage-1 RF classifier are comparable to the methods of [13] and [15]. On the other hand, the results based on the cascade classifier clearly outperform the previous methods. The fraction of true positives is 3%–20% higher than the method of [13], and 3%–25% higher than the method of [15], for a similar number of false positives.

Fig 6 shows the FROC curves for the SNUBH-MDB- μ C_c, using the rule in [13]. The rule in [15] is inapplicable since there are no annotations of individual μ Cs for this dataset. Here, we can see that although the stage-1 RF classifier results are much worse, the cascade classifier results are much better than that of [13], and slightly better than [15], with the AUC value 0.63 for the proposed method and 0.60 for [15], respectively.

We note that the substantial performance decrease of all methods for the SNUBH-MDB- μ C_c compared to that for the SNUBH-MDB- μ C_i has two main causes. First, the



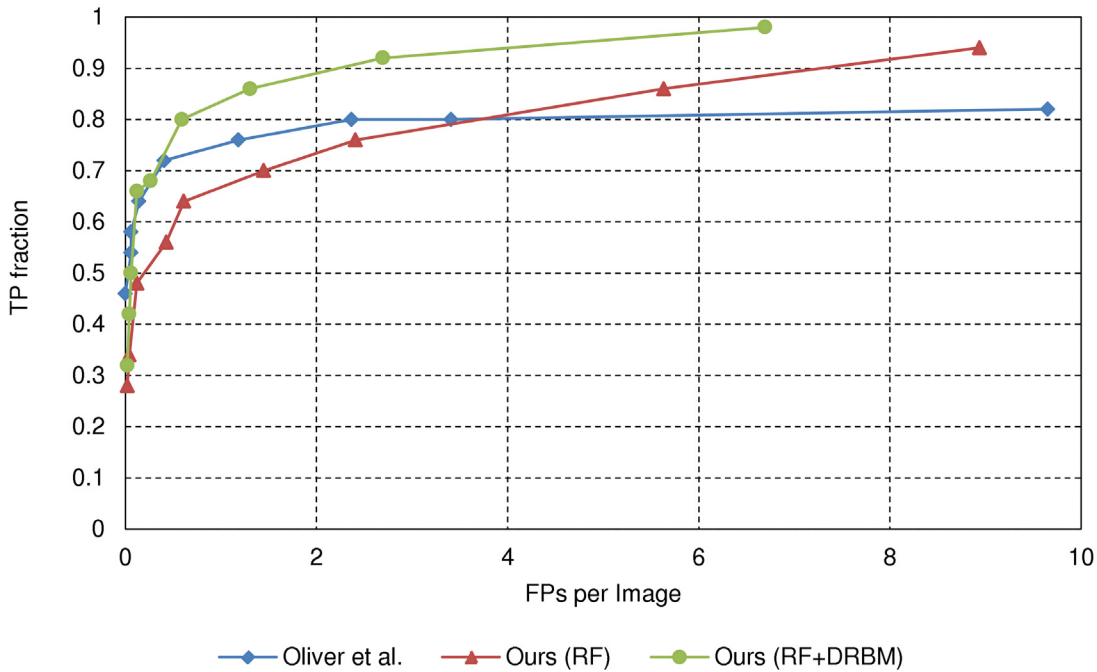
(a)



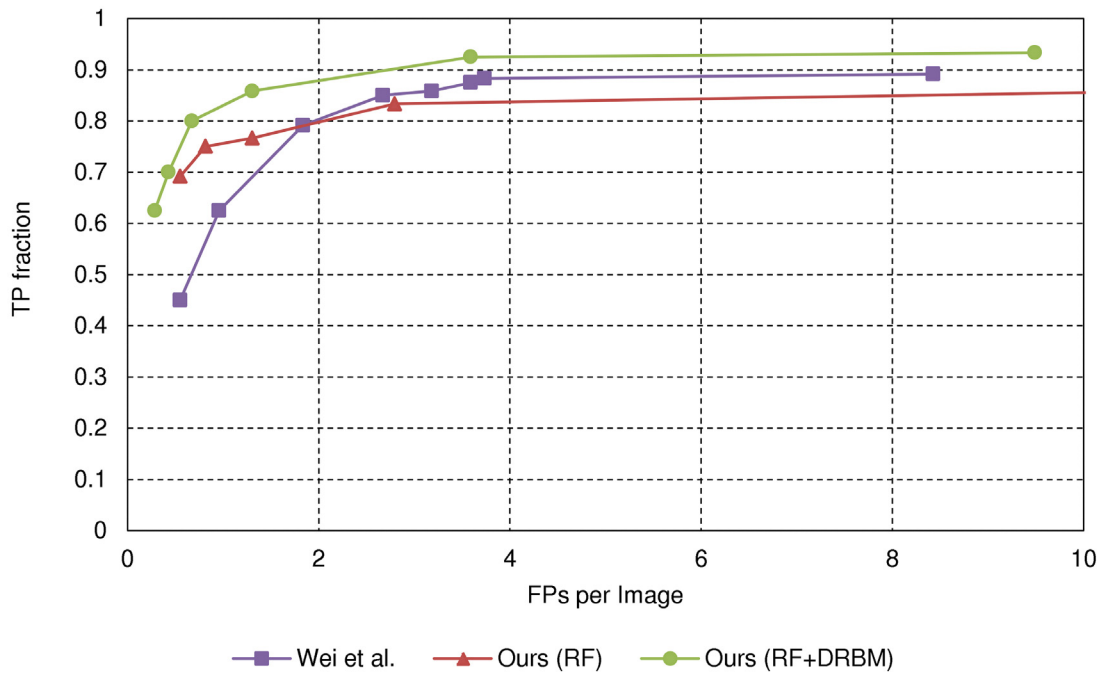
(b)

Fig 4. Quantitative results on the individual μ C detection in the SNUBH-MDB- μ Ci. (a) ROC curve. The x-axis (FP rate) is represented on a log-scale. (b) Precision-Recall curve.

doi:10.1371/journal.pone.0143725.g004



(a)



(b)

Fig 5. FROC curves obtained by 10-fold cross-validation on the SNUBH-MDB- μ CI. Two representative rules for the cluster detection are utilized here. (a) FROC curves by the rules used in [13]. (b) FROC curves by the ruled used in [15].

doi:10.1371/journal.pone.0143725.g005

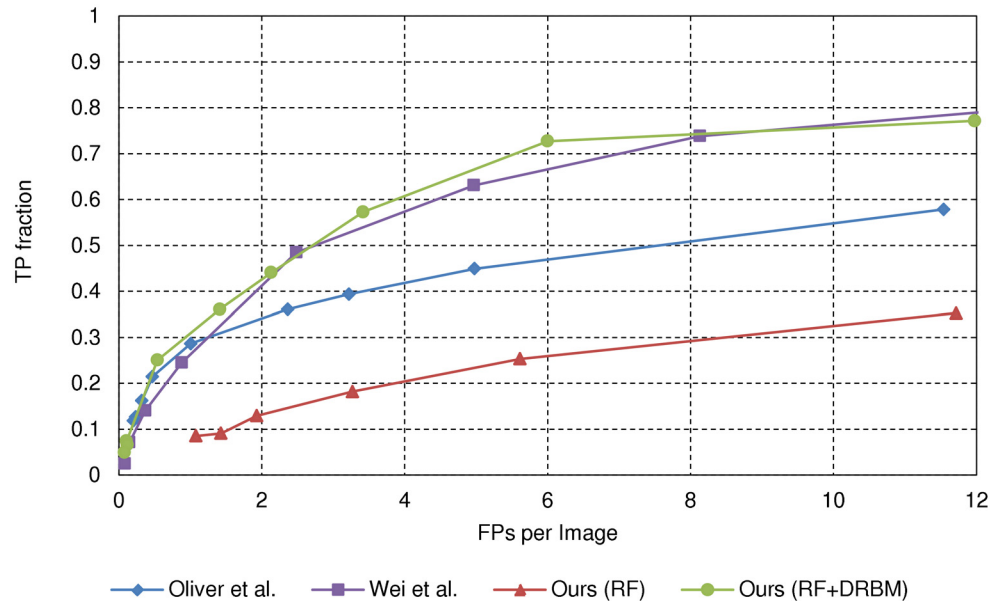


Fig 6. FROC curves for the SNUBH-MDB- μ Cc. The SNUBH-MDB- μ Cc can be evaluated only by the rule in [13] since we have only cluster annotations for the SNUBH-MDB- μ Cc.

doi:10.1371/journal.pone.0143725.g006

SNUBH-MDB- μ Cc also contains normal cases, which are not included in the SNUBH-MDB- μ Ci. Since these cases were not included in the training, the trained classifier may be insufficient to handle these cases. Second, while the methods were trained with the SNUBH-MDB- μ Ci, which only contains craniocaudal (CC) view images, the SNUBH-MDB- μ Cc also contains mediolateral oblique (MLO) view images. Since positive and negative training samples are only collected from the CC view images, the trained classifier may be inadequate for local appearances exclusive to MLO view images.

Fig 7 shows the FROC curves for the MIAS DB. It shows that the proposed cascade classifier and the method of [13] show comparable accuracy. These methods clearly outperform the stage-1 RF classifier and the method of [15]. To analyse the significance of the results, we compute 95% confidence intervals, following the parametric procedure of [31]. At sensitivity 0.8, the false positives are in the range of [1.04, 2.88] by the proposed cascade method and [0.96, 1.73] for the method of [13]. At sensitivity 0.9, the ranges are [1.4, 3.19] and [3.23, 5.52], respectively. We can see that the proposed method outperforms the method of [13] for higher sensitivity values. The proposed cascade classifier is able to return all true μ C clusters, if more than 3.3 false positives per image are allowed on average. Overall, the proposed method performs slightly better, in terms of area under the curve (AUC) value, with 0.84, compared to [13], with 0.81. We note that, for this dataset, the results of the method [13], provided by the authors, are from their original implementation.

In comparison, we observe that the FROC curves in Fig 8 are considerably lower than those from the MIAS DB in Fig 7. As previously mentioned, due to the reduced resolution, the μ Cs in the mini-MIAS DB have considerably smaller pixel area compared to the original MIAS DB. This diminishes spatial cues required to distinguish μ Cs from noise or other tissue with similar appearance. This may result in limited accuracy for low resolution images. Thus, the proposed method requires high resolution images for optimal performance. Although the proposed cascade classifier and the method of [15] show comparable accuracy, the proposed method

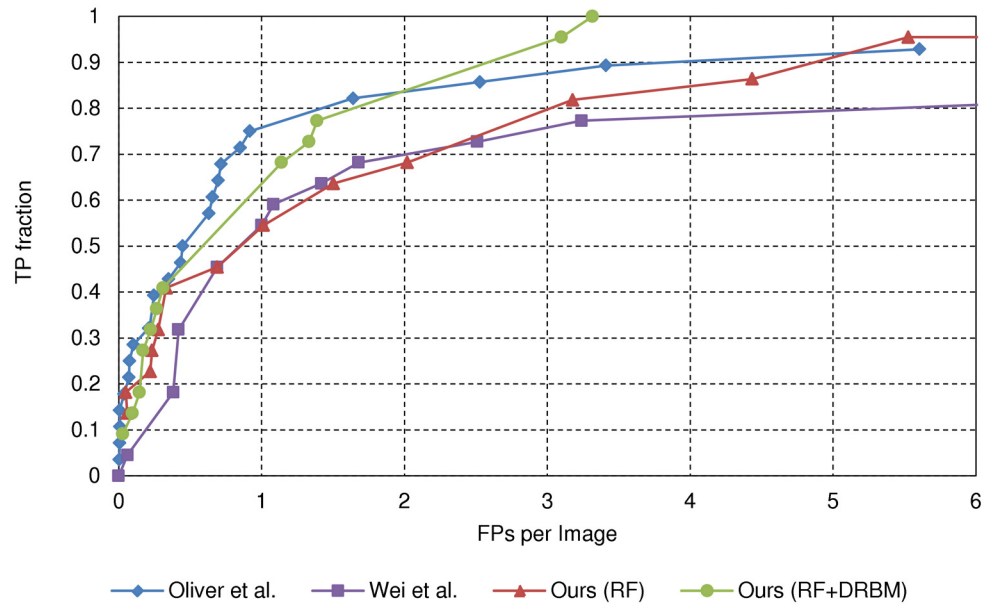


Fig 7. FROC curves for the MIAS database. All 322 images including mammograms with μ Cs and without μ Cs are evaluated by the rule in [13].

doi:10.1371/journal.pone.0143725.g007

performs better at operating points with less than four false positives per image. Overall, the proposed method, with AUC 0.62, performs better than the method of [15], with AUC 0.57.

Qualitative Evaluation

Evaluation of individual μ C detection framework. Fig 9 shows individual μ C detection results for three example cases from the SNUBH-MDB- μ Ci in each row. The whole and

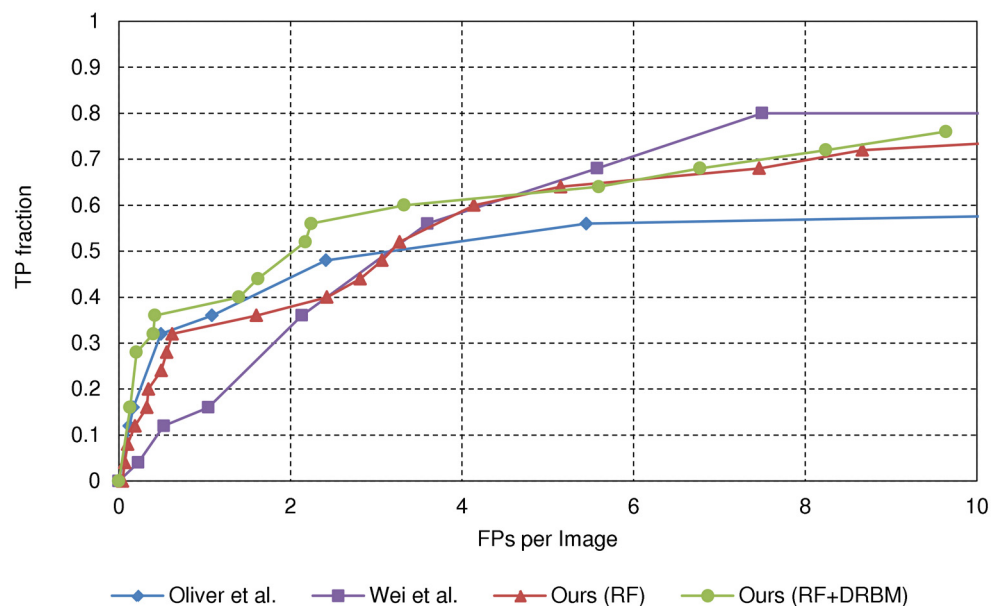


Fig 8. FROC curves for the mini-MIAS database. All 322 images including mammograms with μ Cs and without μ Cs are evaluated by the rule in [13].

doi:10.1371/journal.pone.0143725.g008

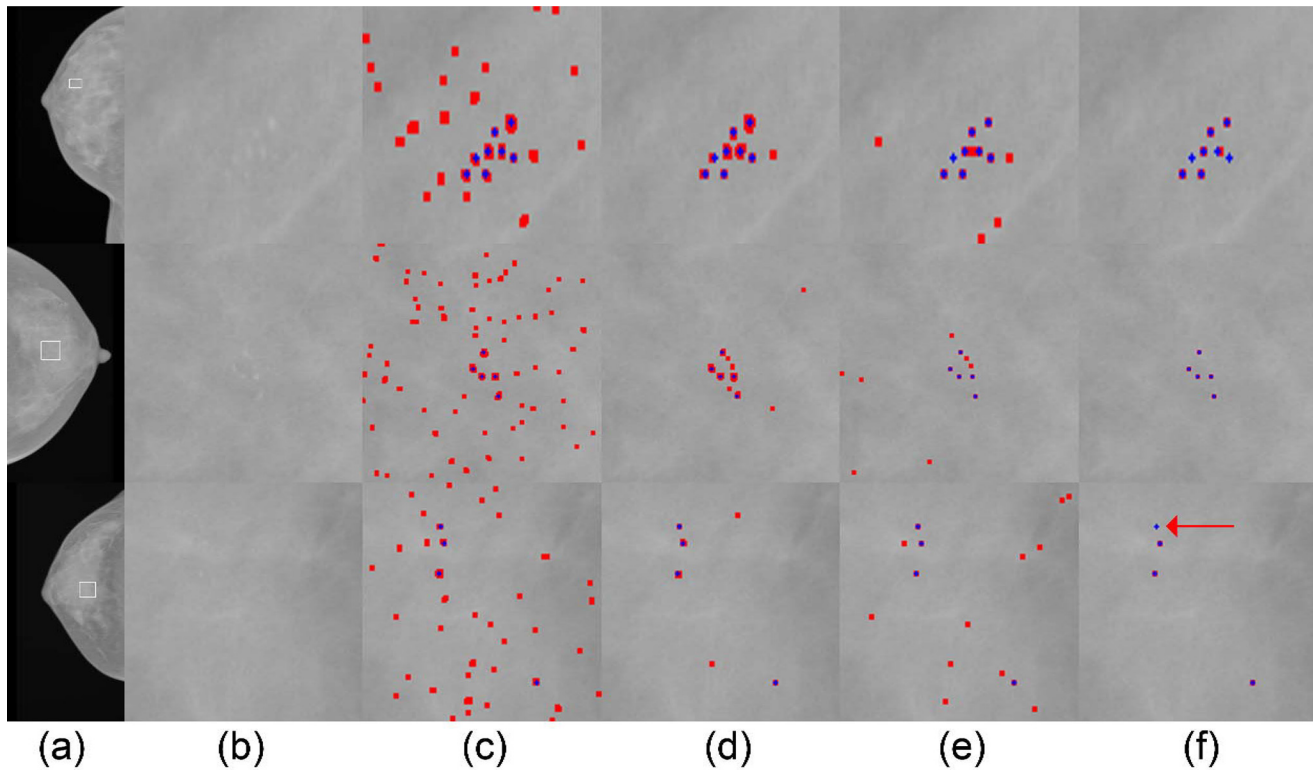


Fig 9. Qualitative results on the individual μ C detection in the SNUBH-MDB- μ Ci. Each row represents a different case and each column shows (a) Test images, (b) Expanded views of ROIs (bounded by white boxes in (a)) containing annotated μ Cs and corresponding detection results by the methods of (c) Oliver et al. [13], (d) Wei et al. [15], (e) Ours (RF), (f) Ours (RF+DRBM). Blue crosses are annotated individual μ Cs and red pixels in the detection results are detected points by the methods. Best viewed in color.

doi:10.1371/journal.pone.0143725.g009

zoomed views of the images are shown in Fig 9(a) and 9(b), respectively. We can see that the μ Cs are difficult to localize due to its unclear boundaries. Compared to the results of the stage-1 RF classifier, shown in Fig 9(e), the numbers of false positives are greatly reduced in the cascade classifier, shown in Fig 9(f). This shows the effectiveness of the higher level DRBM classifier. For these particular cases, the method of [13] detects more false positives compared to the method of [15], but both methods detect higher number false positives compared to the proposed cascade classifier. However, there is still room for improvement. In the case shown in the bottom row, the cascade classifier fails to detect a true μ C, highlighted by an arrow, which results in a μ C cluster detection failure.

Evaluation of μ C cluster detection framework. Fig 10 shows μ C cluster detection results for the same cases as in Fig 9. Here, the cluster detection rule of [13] is used for the method of [13], while the rule of [15] is used for the method of [15]. For the proposed cascade classifier, results using both rules are shown. For the top and middle row cases, all methods correctly detect the μ C cluster, but the proposed method with the rule of [13] detects the smallest number of false positives. For the bottom row case, all methods detect the correct μ C cluster, except the proposed, using the rule of [15]. The missed detection is caused by the aforementioned missed individual μ C detection. Again, comparing the number of false positives and bounding box sizes, the proposed method with the rule of [13] shows the best results.

Fig 11 shows μ C cluster detection results for three example cases from the MIAS DB in each row. Comparing the bounding boxes and number of false positives, we can see that the proposed cascaded classifier gives the best results for the top and bottom row cases. For the middle

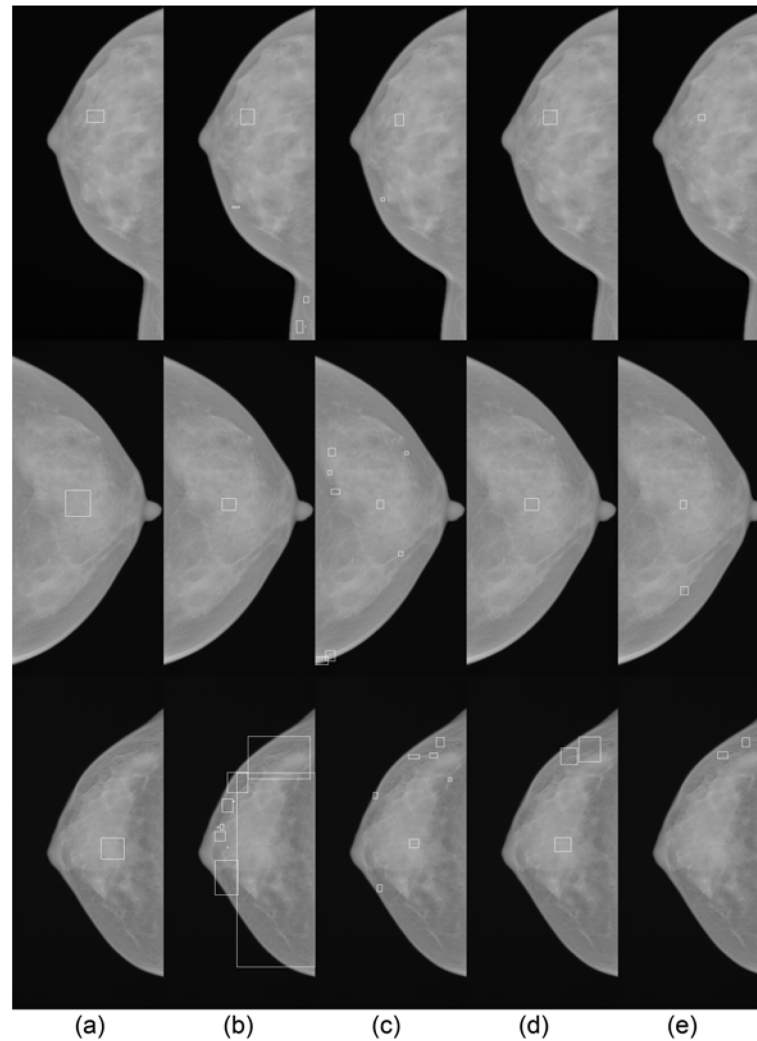


Fig 10. Qualitative results on the μ C cluster detection in the SNUBH-MDB- μ Ci. Each row represents a different case which corresponds to the case in the same row in Fig 9 and each column shows (a) annotated clusters and cluster detection results by the methods of (b) Oliver et al. [13], (c) Wei et al. [15], and (d) Ours (RF+DRBM) with the criterion of Oliver et al. [13], and (e) Ours (RF+DRBM) with the criterion of Wei et al. [15].

doi:10.1371/journal.pone.0143725.g010

row case, all three methods are comparable. While quantitative results of the method [13] were from the implementation original authors, these results are from our own implementation.

Also, Fig 12 shows μ C cluster detection results for the same example cases, but from using the images of the mini-MIAS DB, shown in Fig 11. We can see that the proposed cascaded classifier gives the best results for all three cases. The performance degradation compared to the results of the MIAS DB can also be observed.

Computational Complexity

The computational complexity at test time of the proposed stage-1 RF classifier is linearly proportional to the number of pixels as each pixel is classified independently. Most of the computational burden is in computing the hessian matrix, where Gaussian convolution is performed for numerical stability, as in [25]. This convolution depends on the number of scales and the σ

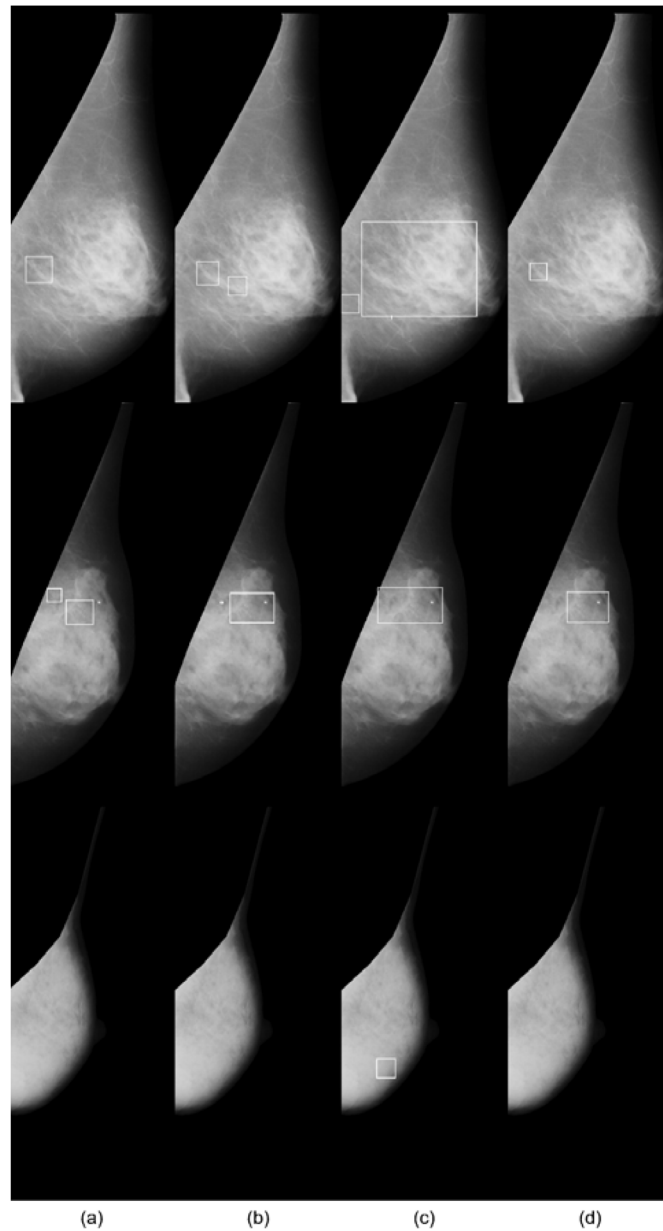


Fig 11. Qualitative results on the μ C cluster detection in the MIAS database. Each row represents a different case and each column shows (a) annotated clusters and cluster detection results by the methods of (b) Oliver et al. [13], (c) Wei et al. [15], and (d) Ours (RF+DRBM). The cluster detections of all methods are conducted by the rule in [13]. For the bottom row case, there are no μ C clusters.

doi:10.1371/journal.pone.0143725.g011

values. Once the hessian is computed, the eigenvalues can be computed by two equations. Also, the prediction using the random forest requires a small number of boolean operations for each pixel, which is the maximum tree depth, 30, multiplied by the number of trees, 100.

The stage-2 DRBM classifier generally requires more computation for a single input than the previous stage, though it depends on the input patch size, the number of hidden nodes, etc. However, since the number of predictions is only a fraction (less than 1%, on average) of the total number of pixels, this stage requires significantly less computation than the previous stage.

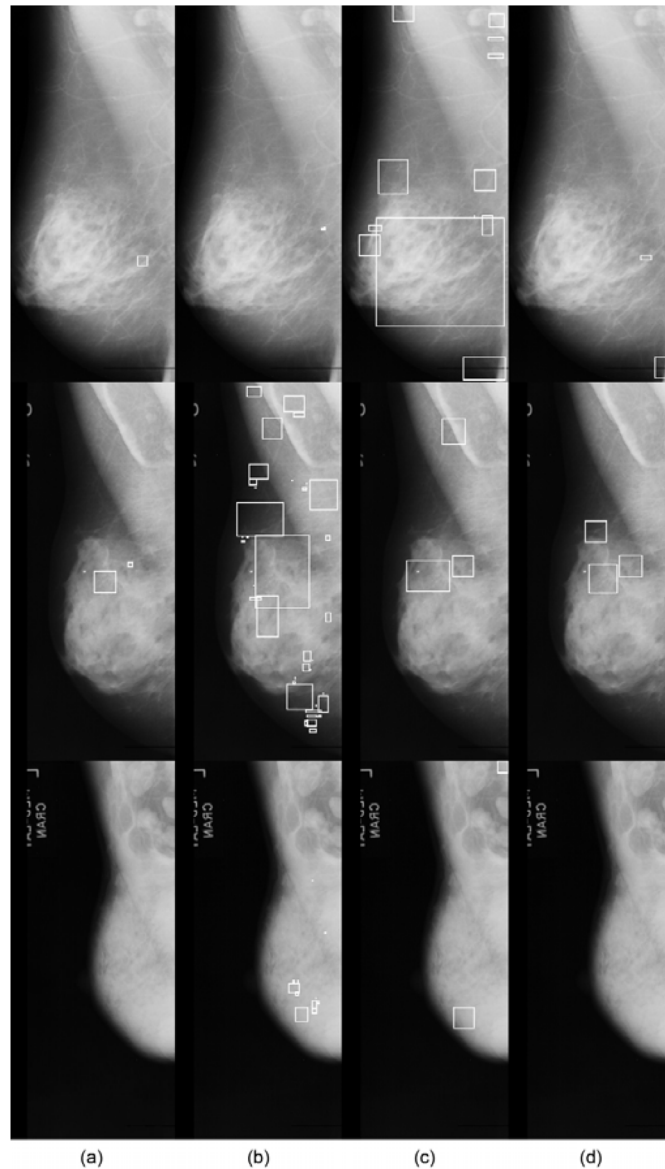


Fig 12. Qualitative results on the μ C cluster detection in the mini-MIAS database. Each row represents a different case which is a processed version of the case in the same row in Fig 11 and each column shows (a) annotated clusters and cluster detection results by the methods of (b) Oliver et al. [13], (c) Wei et al. [15], and (d) Ours (RF+DRBM). The cluster detections of all methods are conducted by the rule in [13] same as in the MIAS database.

doi:10.1371/journal.pone.0143725.g012

While stage-1 RF and stage-2 DRBM classifiers are both quite efficient at test time, training does require much more computation. However, this can be performed offline, and does not affect the inference at test time.

Table 2 shows the average computation time, in seconds, for detection of individual μ Cs from images of the SNUBH-MDB, for the methods of [13, 15] and the proposed cascade method. All the methods are implemented in C++ and tested on a 3.3 GHz CPU with 16GB RAM and no GPU support. The processing time for μ C cluster detection is excluded since it is identical for all methods. We note that our proposed method is 55 \times faster than the method of

Table 2. The computation times of our and baseline algorithms. The Test time is the average time for detecting μ Cs in one image. The time is represented in seconds.

| | Oliver et al. | Wei et al. | Ours (RF) | Ours (RF+DRBM) |
|---------------|---------------|------------|-----------|----------------|
| Test time (s) | 2020.86 | 35.92 | 35.05 | 36.03 |

doi:10.1371/journal.pone.0143725.t002

[13] and is comparable to that of [15] while outperforming both methods in terms of accuracy. Within the proposed method, the stage-2 DRBM classifier constitutes less than 3% of the total time.

Discussion

We have shown that the proposed cascade classifier outperform previous methods [13, 15], in terms of accuracy, for the SNUBH-MDB and the MIAS DB. However, this is less evident for the mini-MIAS DB, which has a much lower spatial resolution, and thus, gives lower detection accuracy for all three methods. Based on these results, we observe that not only does our method perform better, it has a higher comparative advantage, for mammograms with higher spatial resolution. The performance degradation of the proposed method at lower resolutions is due to the instabilities of Hessian eigenvalues along with the difficulties in capturing subtle differences of structures with smaller pixel dimensions. Thus, mammograms with high spatial resolution should be achieved to maximize the performance of the proposed method.

For future works, the stage-1 RF classifier can be significantly accelerated by utilizing a GPU, as described in [32]. Also, we hope to utilize different methods [33–35] for training the classifier. The proposed method can also be extended for tomosynthesis.

Acknowledgments

We thank Arnau Oliver for providing the set of expert annotations of individual μ Cs together with the mammograms of the original MIAS DB, and for providing results of their method for comparison.

Author Contributions

Conceived and designed the experiments: SS SL IDY. Performed the experiments: SS. Analyzed the data: SS SL IDY SMK. Contributed reagents/materials/analysis tools: SS SL IDY HYJ YSH SMK KML. Wrote the paper: SS SL IDY HYJ YSH.

References

1. World Cancer Research Fund International. Breast cancer statistics; 2013. Available from: <http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/breast-cancer-statistics>.
2. American Cancer Society. What are the key statistics about breast cancer?; 2015. Available from: <http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-key-statistics>.
3. Picca DA, de Paredes ES. Calcifications in the breast: A radiologic perspective. *Applied Radiology*. 2003; 32(9):29–37.
4. Papadopoulos A, Fotiadis DI, Costaridou L. Improvement of microcalcification cluster detection in mammography utilizing image enhancement techniques. *Comput Biol Med*. 2008 Oct; 38(10):1045–1055. Available from: <http://dx.doi.org/10.1016/j.combiomed.2008.07.006>. doi: 10.1016/j.combiomed.2008.07.006 PMID: 18774128
5. Salvado J, Roque B. Detection of calcifications in digital mammograms using wavelet analysis and contrast enhancement. In: *Proceedings of IEEE International Workshop on Intelligent Signal Processing*; 2005. p. 200–205.
6. Oporto-Díaz S, Hernández-Cisneros R, Terashima-Marín H. Detection of microcalcification clusters in mammograms using a difference of optimized gaussian filters. In: *Proceedings of the Second*

International Conference on Image Analysis and Recognition; 2005. p. 998–1005. Available from: http://dx.doi.org/10.1007/11559573_121.

7. Quintanilla-Dominguez J, Cortina-Januchs MG, Ojeda-maga na B, Jevtić A, Vega-Corona A, Andina D. Microcalcification detection applying artificial neural networks and mathematical morphology in digital mammograms. In: Proceedings of World Automation Congress; 2010. p. 1–6.
8. Zhang X, Homma N, Goto S, Kawasumi Y, Ishibashi T, Abe M, et al. A hybrid image filtering method for computer-aided detection of microcalcification clusters in mammograms. *Journal of Medical Engineering*. 2013; 2013. doi: [10.1155/2013/615254](https://doi.org/10.1155/2013/615254)
9. Diaz-Huerta CC, Felipe-Riveron EM, Montaña-Zetina LM. Quantitative analysis of morphological techniques for automatic classification of micro-calcifications in digitized mammograms. *Expert Systems with Applications*. 2014; 41(16):7361–7369. doi: [10.1016/j.eswa.2014.05.051](https://doi.org/10.1016/j.eswa.2014.05.051)
10. Eddaoudi F, Regragui F. Microcalcifications detection in mammographic images using texture coding. *Applied Mathematical Sciences*. 2011; 5(8):381–393.
11. Sankar D, Thomas T. A new fast fractal modeling approach for the detection of microcalcifications in mammograms. *Journal of Digital Imaging*. 2010; 23(5):538–546. Available from: <http://dx.doi.org/10.1007/s10278-009-9224-6>. doi: [10.1007/s10278-009-9224-6](https://doi.org/10.1007/s10278-009-9224-6) PMID: [19618243](https://pubmed.ncbi.nlm.nih.gov/19618243/)
12. Huang J, Li J, Liu T. A new fast fractal coding method for the detection of microcalcifications in mammograms. In: Proceedings of International Conference on Multimedia Technology; 2011. p. 4768–4771.
13. Oliver A, Torrent A, Lladó X, Tortajada M, Tortajada L, Sentís M, et al. Automatic microcalcification and cluster detection for digital and digitised mammograms. *Knowledge-Based Systems*. 2012; 28:68–75. doi: [10.1016/j.knsys.2011.11.021](https://doi.org/10.1016/j.knsys.2011.11.021)
14. El-Naqa I, Yang Y, Wernick MN, Galatsanos NP, Nishikawa RM. A support vector machine approach for detection of microcalcifications. *IEEE Transactions on Medical Imaging*. 2002; 21(12):1552–1563. doi: [10.1109/TMI.2002.806569](https://doi.org/10.1109/TMI.2002.806569) PMID: [12588039](https://pubmed.ncbi.nlm.nih.gov/12588039/)
15. Wei L, Yang Y, Nishikawa RM, Wernick MN, Edwards A. Relevance vector machine for automatic detection of clustered microcalcifications. *IEEE Transactions on Medical Imaging*. 2005; 24(10):1278–1285. doi: [10.1109/TMI.2005.855435](https://doi.org/10.1109/TMI.2005.855435) PMID: [16229415](https://pubmed.ncbi.nlm.nih.gov/16229415/)
16. Papadopoulos A, Fotiadis DI, Likas A. An automatic microcalcification detection system based on a hybrid neural network classifier. *Artificial Intelligence in Medicine*. 2002; 25(2):149–167. doi: [10.1016/S0933-3657\(02\)00013-1](https://doi.org/10.1016/S0933-3657(02)00013-1) PMID: [12031604](https://pubmed.ncbi.nlm.nih.gov/12031604/)
17. Papadopoulos A, Fotiadis DI, Likas A. Characterization of clustered microcalcifications in digitized mammograms using neural networks and support vector machines. *Artificial Intelligence in Medicine*. 2005; 34(2):141–150. doi: [10.1016/j.artmed.2004.10.001](https://doi.org/10.1016/j.artmed.2004.10.001) PMID: [15894178](https://pubmed.ncbi.nlm.nih.gov/15894178/)
18. Bria A, Karssemeijer N, Tortorella F. Learning from unbalanced data: a cascade-based approach for detecting clustered microcalcifications. *Medical Image Analysis*. 2014; 18(2):241–252. doi: [10.1016/j.media.2013.10.014](https://doi.org/10.1016/j.media.2013.10.014) PMID: [24292553](https://pubmed.ncbi.nlm.nih.gov/24292553/)
19. Breiman L. Random forests. *Machine learning*. 2001; 45(1):5–32. doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)
20. Criminisi A, Shotton J. *Decision forests for computer vision and medical image analysis*. Springer; 2013.
21. Larochelle H, Bengio Y. Classification using discriminative restricted boltzmann machines. In: Proceedings of the 25th International Conference on Machine Learning; 2008. p. 536–543.
22. Shin S, Lee S, Yun ID. Classification based micro-calcification detection using discriminative restricted boltzmann machine in digitized mammograms. In: Proceedings of SPIE Medical Imaging; 2014. p. 90351L–90351L–6.
23. Suckling J, Parker J, Dance D, Astley S, Hutt I, Boggis C, et al. The mammographic image analysis society digital mammogram database. In: *Exerpta Medica. International Congress Series*. vol. 1069; 1994. p. 375–378.
24. Pilot European Image Processing Archive. Available from: <http://peipa.essex.ac.uk/info/mias.html>.
25. Frangi AF, Niessen WJ, Vincken KL, Viergever MA. Multiscale vessel enhancement filtering. In: Proceedings of Medical Image Computing and Computer Assisted Intervention; 1998. p. 130–137.
26. Smolensky P. Information processing in dynamical systems: Foundations of harmony theory. In: Rumelhart DE, McClelland JL, PDP Research Group C, editors. *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA, USA: MIT Press; 1986. p. 194–281.
27. Kallergi M, Carney GM, Gaviria J. Evaluating the performance of detection algorithms in digital mammography. *Medical Physics*. 1999; 26(2):267–275. doi: [10.1118/1.598514](https://doi.org/10.1118/1.598514) PMID: [10076985](https://pubmed.ncbi.nlm.nih.gov/10076985/)
28. Hinton GE. Training products of experts by minimizing contrastive divergence. *Neural Computation*. 2002; 14(8):1771–1800. doi: [10.1162/089976602760128018](https://doi.org/10.1162/089976602760128018) PMID: [12180402](https://pubmed.ncbi.nlm.nih.gov/12180402/)

29. Hinton GE. A practical guide to training restricted boltzmann machines. In: *Neural Networks: Tricks of the Trade*. Springer; 2012. p. 599–619.
30. Bunch P, Hamilton J, Sanderson G, Simmons A. Free response approach to measurement and characterization of radiographic observer performance. *American Journal of Roentgenology*. 1978; 130(2):382–382.
31. Bornefalk H, Hermansson AB. On the comparison of FROC curves in mammography CAD systems. *Medical Physics*. 2005; 32(2):412–417. doi: [10.1118/1.1844433](https://doi.org/10.1118/1.1844433) PMID: [15789587](https://pubmed.ncbi.nlm.nih.gov/15789587/)
32. Sharp T. Implementing decision trees and forests on a GPU. In: *Proceedings of European Conference on Computer Vision*; 2008. p. 595–608.
33. Lee H, Grosse R, Ranganath R, Ng AY. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In: *Proceedings of International Conference on Machine Learning*; 2009. p. 609–616.
34. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: *Proceedings of Advances in Neural Information Processing Systems*; 2012. p. 1106–1114.
35. Jiang Y, Chung FL, Ishibuchi H, Deng Z, Wang S. Multitask TSK fuzzy system modeling by mining inter-task common hidden structure. *IEEE Transactions on Cybernetics*. 2015; 45(3):548–561. PMID: [24988602](https://pubmed.ncbi.nlm.nih.gov/24988602/)