Iranian Journal of Medical Physics
Vol. 13, No. 4, December 2016, 261-268
Received: August 15, 2016; Accepted: December 03, 2016

Original Article

A Hierarchical Classification Method for Breast Tumor Detection

Mojtaba Mohammadpoor[1], Afshin Shoeibi [2*], Hoda zare[3], Hasan Shojaee[4]

Abstract

Introduction
Breast cancer is the second cause of mortality among women. Early detection of it can enhance the chance of survival. Screening systems such as mammography cannot perfectly differentiate between patients and healthy individuals. Computer-aided diagnosis can help physicians make a more accurate diagnosis.
Materials and Methods
Regarding the importance of separating normal and abnormal cases in screening systems, a hierarchical classification system is defined in this paper. The proposed system is including two Adaptive Boosting (AdaBoost) classifiers, the first classifier separates the candidate images into two groups of normal and abnormal. The second classifier is applied on the abnormal group of the previous stage and divides them into benign and malignant categories. The proposed algorithm is evaluated by applying it on publicly available Mammographic Image Analysis Society (MIAS) dataset. 288 images of the database are used, including 208 normal and 80 abnormal images. 47 images of the abnormal images showed benign lesion and 33 of them had malignant lesion.
Results
Applying the proposed algorithm on MIAS database indicates its advantage compared to previous methods. A major improvement occurred in the first classification stage. Specificity, sensitivity, and accuracy of the first classifier are obtained as 100%, 95.83%, and 97.91%, respectively. These values are calculated as 75% in the second stage
Conclusion
A hierarchical classification method for breast cancer detection is developed in this paper. Regarding the importance of separating normal and abnormal cases in screening systems, the first classifier is devoted to separate normal and tumorous cases. Experimental results on available database shown that the performance of this step is adequately high (100% specificity). The second layer is designed to detect tumor type. The accuracy in the second layer is obtained 75%.

Keywords: Mammography, Breast Cancer, Classification, Computer Aided Diagnosis

1- Electrical & Computer Dept., University of Gonabad, Gonabad, Iran
2-Medical Physics Dept., Gonabad University of Medical Sciences, Gonabad, Iran
*Corresponding author: Tel:+985157223028; Fax:+985157223027; E-mail: afshin.shoeibi@gmail.com

3- Medical Physics Research Center, Mashhad University of Medical Sciences, Mashhad, Iran
4- Basic Sciences Dept., Gonabad University of Medical Sciences, Gonabad, Iran

1. Introduction

Cancer incorporates a group of diseases characterized by uncontrolled growth and spread of abnormal cells. If the spread is not managed properly, it can result in death. Early detection can help reduce the rate of mortality due to cancer and make the treatment less extensive and more likely to be successful [1].

Cancer is one of the most common and fatal diseases worldwide. Breast cancer is the most prevalent cancer among women (34%) and it is the second cause of cancer death in women, with a total of 11,090 new invasive cases and 1,180 deaths expected to occur in 2016. During 2003-2012, breast cancer death rates decreased by 1.65% per year. Overall, breast cancer mortality rates declined by 36% from 1989 to 2012 due to advancements in early diagnosis and treatment.

Breast cancer can be detected by different kinds of screening tests such as mammography, ultrasonography, computed tomography scan, and magnetic resonance imaging. Mammography is the preferred screening examination for breast cancer, however, it is less accurate in women with dense breast tissue. Moreover, similarities between early signs of breast cancer and normal structures in mammography images make diagnosis of breast cancer difficult. Therefore, mammography alone cannot prove that a suspicious area is tumorous and whether the tumor is malignant or benign; thus, the tissue has to be removed using breast biopsy techniques [1].

Devising methods in automatic classification of suspicious areas in mammography is essential. Computer-aided diagnosis (CAD) techniques were presented in recent years to help physicians, reduce false positive rate (FPR), and perform diagnosis action faster, more accurately, and more easily [2-9, 13, 14]. There are sets of automatic or semi-automatic tools to help radiologists with detection and classification of breast abnormalities [2]. Generally, a CAD system incorporates pre-processing, feature extraction, and classification stages [3].

Breast Imaging-Reporting and Data System (BI-RADS), established by the American College of Radiology, is a scheme for putting the findings from mammogram screening (for breast cancer diagnosis) into a small number of well-defined categories [4]. The BI-RADS assessment categories are

(0) incomplete, (1) negative, (2) benign findings, (3) probably benign, (4) suspicious abnormality, (5) highly suspicious of malignancy, (6) known biopsy with proven malignancy.

There are several computer-aided diagnosis efforts for detection of breast cancer, especially in its early stages. Kwon et al. proposed a time-domain microwave imaging system for early detection of breast cancer. In order to improve image quality, they made use of gaussian bandpass filter (BPF) [5].

Tan et al. investigated the association between the bilateral mammographic image feature changes over time and the risk trend for early cancer detection in women after negative screening. They used different features, such as Weber Local Descriptor (WLD), Gabor Directional Similarity (GDS) features, Run Length Statistics (RLS), and Gray Level Co-occurrence Matrix (GLCM) [6]. Rouhi et al. proposed two methods for classification of benign and malignant breast tumors in mammogram images. They used various techniques for segmentation and artificial neural network (ANN) for classification. They performed the proposed techniques on Mammographic Image Analysis Society (MIAS) [12] and DDSM databases and obtained sensitivity, specificity, and accuracy rates of 92.70%, 90.54%, and 90.16%, respectively, by applying their technique on MIAS database [2].

Albarqouni et al. suggested applying deep learning to generate a ground-truth from non-expert annotations in the biomedical context [7].

Singh et al. presented a method for diagnosis of malignant/non-malignant breast tissues by using Polar Complex Exponential Transform (PCET) moments as texture descriptors. They extracted their algorithm's input region-of-interest (ROI) manually and introduced Adaptive Differential Evolution Wavelet Neural Network (ADEWNN). By applying their algorithm on mammographic images from MIAS database, they obtained accuracy, sensitivity, and

specificity of 97.965%, 98.196%, and 97.194%, respectively [8].

2. Materials and Methods

Recently, Pak et al. presented state-of-the-art breast cancer detection and classification. In their study (in the pre-processing step), they used Non-Subsampled Contourlet Transform, Super Resolution, and high-pass filtering for image quality improvement, prior to image's region-of-interest detection [9]. In the feature extraction step, they used several well-selected features such as area, compactness, and fractal as regional descriptor features (which are central moment, eccentricity and spread as boundary descriptor features, and average gray-level as density descriptor feature). Then, they calculated skewness of each feature and concatenated all of them to build the final feature vector. They divided mammographic images into three general categories, namely, normal (including BI-RADS 1), benign (including BI-RADS 2 & 3), and malignant (including BIRADS 4 & 5). They used AdaBoost for classification. Figure 1 shows the architecture of their method.

In this paper, a method for improving Pak-Rashidi-Alikhassi [9] algorithm is proposed. In this way, firstly two cancerous groups were considered as a unique group called abnormal and a classifier separated normal and abnormal categories. In the second stage, another classifier distinguished benign and malign groups from each other.

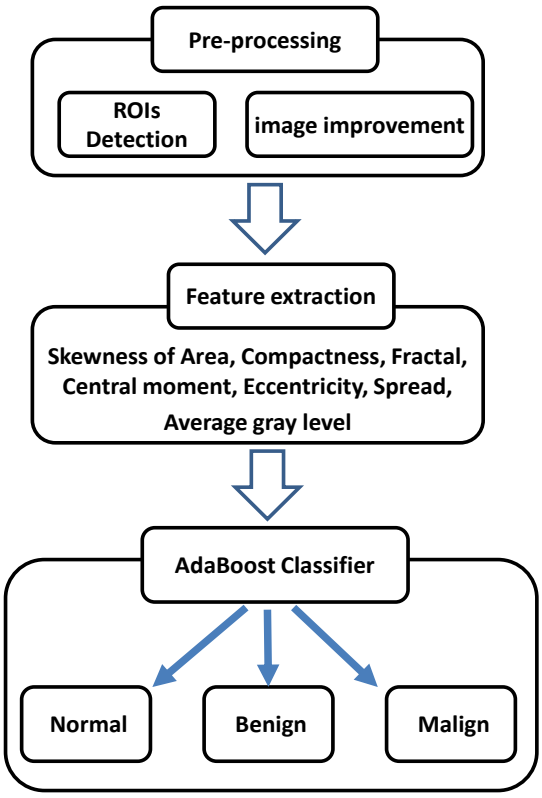

Figure 1. The architecture of Pak-Rashidi-Alikhassi method [9]

2.1. The Boosting Algorithm Adaboost

Boosting is an approach based on the idea of creating a stronger classifier by linear combination of several weak classifiers. A weak AdaBoost was the first practical boosting algorithm and remains one of the most widely used [15, 16]. Pseudocode for AdaBoost is presented in Figure 2. The algorithm get the input train data (x_1, y_1), ..., (x_m, y_m), where x_i is an observed feature and y_i are their two class labels. For simplicity, y_i is considered as +1 or -1. A weak classifier is a learner who could predict at least better than random selection. The weak learner is called in a series of rounds $t= 1, \dots, T$ to predict its weak h_t, $X \in \{-1, 1\}$. The algorithm maintains its distribution or set of weights over the training set D_t.

Given: $(x_1, y_1), \dots, (x_m, y_m); x_i \in X, y_i \in \{-1, 1\}$ (1)

Initialize weights $D_1(i)= 1/m$.

For $t= 1, \dots, T$:

1. Train weak learner using distribution D_t
2. Get weak hypothesis h_t, $X \in \{-1, 1\}$ with error ε_t, where

$$\varepsilon_t = \Pr\{h_t(x_i) \neq y_i\} \quad (2)$$

3. Choose $\alpha_t = 0.5 \ln(\frac{1-\varepsilon_t}{\varepsilon_t})$ (3)

4. Update

$$D_{t+1}(i) = \left[D_t(i) \exp\left(\alpha_t y_i h_t(xi)\right)\right] / Z_t \quad (4)$$

where Z_t is a normalization factor (chosen so that D_{t+1} will be a distribution).

5. Output the final hypothesis:

$$H(x) = sign(\sum_{t=1}^{T} \alpha_t h_t(x)) \quad (5)$$

Figure 2. The boosting algorithm AdaBoost

2.2. The Proposed Method

Due to interclass similarities between malignant and benign classes, the classifiers have difficulty in detection, and cannot predict the true class properly. This phenomenon decreases the CAD system's accuracy. Furthermore, the most important detection demand is increasing the systems specificity, because misclassification of a patient as healthy may lead to patient death.

On the other hand, classifiers are often developed to distinguish between just two classes of objects, and generalization from two-class classification to multi-class classification is not straightforward. Although several efforts have been made in this regard [10, 11], most of classifiers like AdaBoost face some restrictions in multi-class classification, which should be eliminated [11]

In order to overcome this problem, a two-layer classification tree is proposed in this part. For this purpose, the overall data is divided into two main categories, namely, normal and abnormal. The normal category includes BI-RADS 1, while the abnormal category comprises of all patients, meaning all candidates having benign or malignant tumors, which are BI-RADS 2 to 5. The main advantage of the proposed method over Pak-Rashidi-Alikhassi method is its hierarchical sequence. Figure 3 presents the architecture of the proposed method, where the common parts with the Pak-Rashidi-Alikhassi method are shown in dashed boxes. For improving the detection accuracy, eigenvalue and eigenvectors of the feature matrix is used instead of original data. By analyzing different classifiers, AdaBoost is found to have the best performance, as shown also in the study by Pak, Rashidi, and Alikhassi [9].

Some of algorithms have tested over MIAS database. In order to evaluate the proposed method, this database is used in this paper also. Some sample figures are shown in Figure 4.

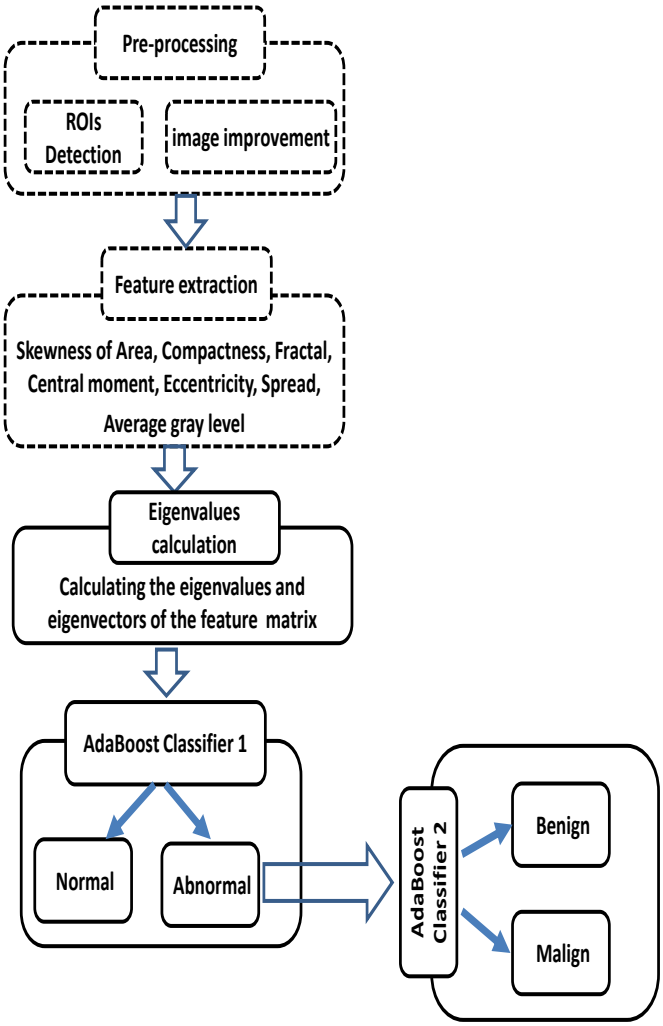

Figure 3. The proposed method architecture

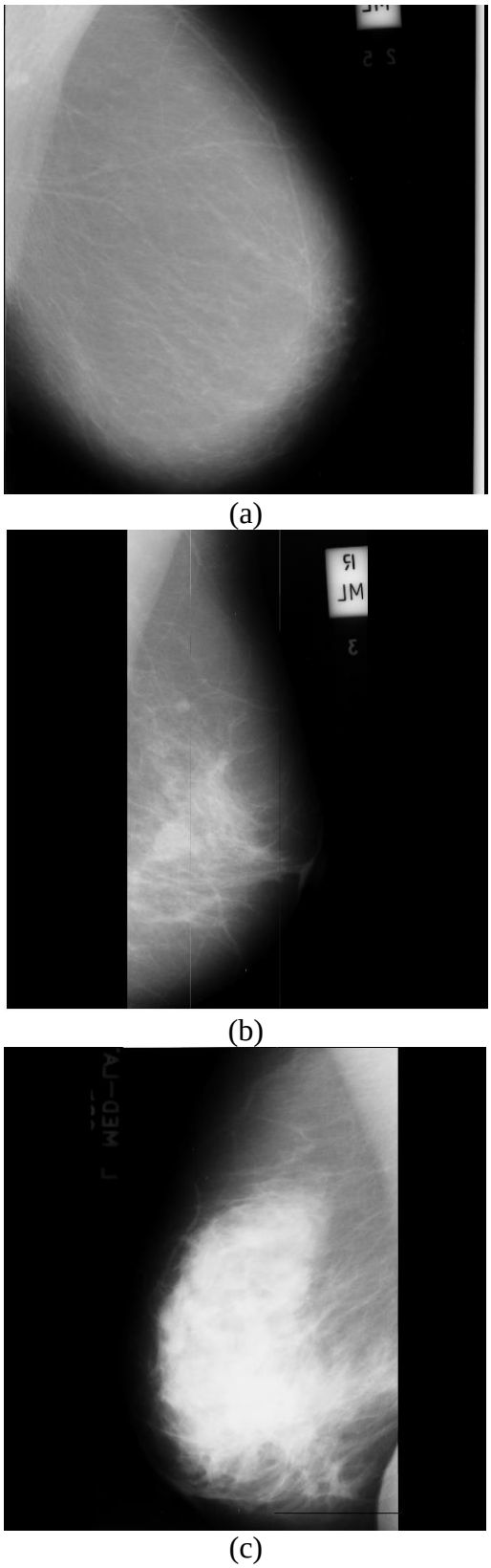

Figure 4. MIAS database sample images, a) normal, b) benign, and c) malignant

3. Results

To evaluate the performance of the proposed algorithm, an experimental investigation was conducted using MIAS database [12]. We used 288 images of the database. 208 of them were normal and 80 included masses and calcifications. Forty-seven images from these 80 abnormal images showed benign lesion and 33 malignant lesion. Table 1 illustrates the two-class confusion matrix to evaluate the proposed algorithm.

Table 1. Confusion matrix definition to evaluate the proposed algorithm

		Predicted classes	
Real classes		Abnormal	Normal
	Abnormal	True positive	False negative
	Normal	False positive	True negative

Where, true positive (TP) means the classification result is abnormal in the presence of a abnormal person. False negative (FN) implies that the classification result is normal, while the person is abnormal. False positive (FP) means the classification result is abnormal while the person is normal, and true negative (TN) indicates that the classification result is normal in the presence of normal person.

According to the above definitions, the equations related to specificity (accuracy of negative class), sensitivity (accuracy of positive class), accuracy of recognizing both negative and positive classes, FPR, and FNR could be defined as

$$Specificity = TN/(TN + FP) \quad (6)$$

$$Sensitivity = TP/(TP + FN) \quad (7)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

$$FNR = 1 - Sensitivity \quad (9)$$

$$FPR = 1 - Specificity \quad (10)$$

Minimizing FNR can be considered as the main goal in normal-abnormal breast cancer detection systems. Labelling abnormal persons as normal may jeopardize their life. Wrong prediction of a normal persons may just lead them to further ensuring treatments. To this end, cost-error parameter of the AdaBoost ensemble is set to give more priority to minimizing FNR. Table 2

shows the result of applying the proposed method to the MIAS dataset using 18-fold cross validation strategy. As demonstrated in Table 2, the FPR of the proposed system is 0, meaning that all abnormal persons are predicted abnormal.

Table 2. Confusion matrix of the proposed algorithm for normal-abnormal classification

		Predicted class	
Real class		Abnormal	Normal
	Abnormal	95.83333	4.166667
	normal	0	100

In order to differentiate between benign and malignant cases in the abnormal class, another AdaBoost classifier is used, as demonstrated in Figure 2. The percentage of its experimental results is shown in Table 3.

Table 3. Confusion matrix of the proposed algorithm for benign-malign classification

		Predicted class	
Real class		Benign	Malignant
	Benign	75	25
	Malign	25	75

4. Discussion

The comparison between the results of the proposed algorithm and other prior systems is presented in Table 4. In order to have a good comparison, recent methods applied on MIAS database are considered. The first and second rows show the two steps of the proposed method. As stated before, the specificity, sensitivity and overall accuracy of the proposed method in its first step is100%, 95.83%, and 97.91%, respectively. These values are higher than most of the available methods. The results of the second step is less important than the first step. The accuracy of the proposed system in the this step is 75%.

5. Conclusion

A two-layer classification system was proposed in this paper for classification of breast tumor images. The system is comprised of two individual AdaBoost classifiers. The first one is separating the input mammographic images into two categories of normal and abnormal. The second classifier is applied on the abnormal class of previous stage and divides them into two subgroups of benign and malignant. The proposed system is applied on publicly available MIAS database. Specificity, sensitivity, and accuracy of the first classifier are 100%, 95.83%, and 97.91%, respectively. These values are 75% for the second system.

Table 4. Comparison of the proposed computer-aided diagnosis (CAD) system with other CAD systems

Method	Year	Database	Feature extraction Technique	Classification technique	Sensitivity (%)	Specificity (%)	Accuracy (%)
The proposed method (normal-abnormal classes)		Mammographic Image Analysis Society (MIAS)	Skewness of area, compactness, fractal, central moment, eccentricity, spread, and average gray level	AdaBoost	95.83	100	97.91
The proposed method (benign-malign classes		MIAS	Skewness of area, compactness, fractal, central moment, eccentricity, spread, and average gray level	AdaBoost	75	75	75
Pak-Rashidi-Alikhassi [9]	2015	MIAS	Skewness of area, compactness, fractal, central moment, eccentricity, spread, and average gray level	AdaBoost	87.15	93.58	91.43
Polar Complex Exponential Transform (PCET) [8]	2016	MIAS	Real and imaginary part of orthogonal moment PCET as a texture descriptors	Adaptive Differential Evolution Wavelet Neural Network (ADEWNN)	95.1	94.35	94.64
Phase Corrected PCET [8]	2016	MIAS	Magnitude orthogonal moment PCET as a texture descriptors	ADEWNN	98.19	97.19	97.96
Rouhi et al. [2]	2015	MIAS	Intensity histogram, shape, and texture	Multi-layer neural network (MLP)	92.70	90.54	90.16
Tahmasbi et al. [3]	2011	MIAS	Magnitude of Zernike moments as a shape and margin descriptors	Multi-layer neural network (MLP)	100	94.50	96.43
Saki et al. [13]	2013	MIAS	AGL, Con, SpI, NRL derivatives, Zernike moments	Opposite Weight Back Propagation per Epoch (OWBPE)	90.10	88.06	89.28
RamirezVillegas et al. [14]	2012	MIAS	Statistical + wavelet packet energy and Tsallis entropy parameterization	Support Vector machines	91.67	100	93.75

Acknowledgements

The authors gratefully acknowledge financial support from Gonabad University of Medical Sciences through grant number p.1.482.

References

1. American Cancer Society (2016). Breast cancer: Facts and figures 2016. Atlanta: ACS.
2. Rouhi R, Jafari M, Kasaei S, Keshavarzian P. Benign and malignant breast tumors classification based on region growing and CNN segmentation. Expert Systems with Applications. 2015 Feb 15;42(3):990-1002.
3. Tahmasbi A, Saki F, Shokouhi SB. Classification of benign and malignant masses based on Zernike moments. Computers in biology and medicine. 2011 Aug 31;41(8):726-35.
4. Sickles EA, D'Orsi CJ, Bassett LW, Appleton CM, Berg WA, Burnside ES. ACR BI-RADS® Atlas, Breast imaging reporting and data system. Reston, VA: American College of Radiology. 2013:39-48.
5. Kwon S, Lee H, Lee S. Image enhancement with Gaussian filtering in time-domain microwave imaging system for breast cancer detection. Electronics Letters. 2016 Jan 22;52(5):342-4.
6. Tan M, Zheng B, Leader JK, Gur D. Association between changes in mammographic image features and risk for near-term breast cancer development. IEEE transactions on medical imaging. 2016 Jul;35(7):1719-28.
7. Albarqouni S, Baur C, Achilles F, Belagiannis V, Demirci S, Navab N. Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. IEEE transactions on medical imaging. 2016 May;35(5):1313-21.
8. Singh SP, Urooj S, Lay-Ekuakille A. Breast Cancer Detection Using PCPCET and ADEWNN: A Geometric Invariant Approach to Medical X-Ray Image Sensors. IEEE Sensors Journal. 2016 Jun 15;16(12):4847-55.
9. Pak F, Kanan HR, Alikhassi A. Breast cancer detection and classification in digital mammography based on Non-Subsampled Contourlet Transform (NSCT) and Super Resolution. Computer methods and programs in biomedicine. 2015 Nov 30;122(2):89-107.
10. Tax DM, Duin RP. Using two-class classifiers for multiclass classification. InPattern Recognition, 2002. Proceedings. 16th International Conference on 2002 (Vol. 2, pp. 124-127). IEEE.
11. Zhu J, Zou H, Rosset S, Hastie T. Multi-class adaboost. Statistics and its Interface. 2009 Jan 1;2(3):349-60.
12. MIAS database, http:// peipa.essex.ac.uk /info/mias.html, Access Date. 10 December 2015.
13. Saki F, Tahmasbi A, Soltanian-Zadeh H, Shokouhi SB. Fast opposite weight learning rules with application in breast cancer diagnosis. Computers in biology and medicine. 2013 Jan 1;43(1):32-41.
14. Ramirez-Villegas JF, Ramirez-Moreno DF. Wavelet packet energy, Tsallis entropy and statistical parameterization for support vector-based and neural-based classification of mammographic regions. Neurocomputing. 2012 Feb 1;77(1):82-100.
15. Schapire RE. Explaining adaboost. InEmpirical inference 2013 (pp. 37-52). Springer Berlin Heidelberg.
16. Freund Y, Schapire R, Abe N. A short introduction to boosting. Journal-Japanese Society For Artificial Intelligence. 1999 Sep 1;14(771-780):1612.