Classification of Mammograms Using Convolutional Neural Network Based Feature Extraction

Taye Girma Debelee[1,3](✉), Mohammadreza Amirian[1], Achim Ibenthal[2], Günther Palm[1], and Friedhelm Schwenker[1]

[1] Institute of Neural Information Processing, Ulm University, 89069 Ulm, Germany
{taye.debelee,mohammadreza.amirian,guenther.palm, friedhelm.schwenker}@uni-ulm.de, tayegirma@gmail.com

[2] Adama Science and Technology University, Adama, Ethiopia
ibenthal@gmx.net

[3] Addis Abeba Science and Technology University, Addis Abeba, Ethiopia

Abstract. Breast cancer is the most common cause of death among women in the entire world and the second cause of death after lung cancer. The use of automatic breast cancer detection and classification might possibly enhance the survival rate of the patients through starting early treatment. In this paper, the convolutional Neural Networks (CNN) based feature extraction method is proposed. The features dimensionality was reduced using Principal Component Analysis (PCA). The reduced features are given to the K-Nearest Neighbors (KNN) to classify mammograms as normal or abnormal using 10-fold cross-validation. The experimental result of the proposed approach performed on Mammography Image Analysis Society (MIAS) and Digital Database for Screening Mammography (DDSM) datasets were found to be promising compared to previous studies in the area of image processing, artificial intelligence and CNN with an accuracy of 98.75% and 98.90% on MIAS and DDSM dataset respectively.

Keywords: Breast cancer · Mammogram · CNN
K-nearest neighbour · Feature extraction

1 Introduction

Breast cancer is one of the most prevalent types of cancer and the second cause of death among women [1,15,17,20,21]. In the USA breast cancer has been proven to be the second cause of death for women after lung cancer [1,19,21,30]. There are two severity of abnormalities associated with breast cancer cells: benign and malignant. The benign ones are cancerous cells that do not grow to neighboring tissues of the breast from where they originated and are no risk to life. The malignant ones, however, are cancerous cells that multiply to other parts of surrounding breast tissues from point of their origin and need to be treated as

F. Mekuria et al. (Eds.): ICT4DA 2017, LNICST 244, pp. 89–98, 2018.
https://doi.org/10.1007/978-3-319-95153-9_9

early as possible. Oliver explained the most common class of abnormalities that can indicate breast cancer [2,12]. The abnormality class includes geometrical asymmetries between left and right breast, normal architectural distortions of the breast tissue, presence of calcifications in the breast, and presence of masses. Breast cancer mass is a localized swelling or lump that appears to exist as benign or malignant in the breast [2,12].

According to the recent review made by Palazuelos et al. [20], breast cancer incidence rate has increased by 20% since 2008 and it is the most frequent diagnosed cancer. However, a study has shown that there is a decrease in mortality rate by 14% [20]. The reasons for decrease in mortality rate include: (a) improvement made on medical imaging technology for better visualization [18], and (b) the start of better treatment plan especially in developed countries [20] and (c) Computer Aided Detection and/Diagnosis (CAD). The CAD in turn enhanced the efficiency of radiologists and eased early detection of the breast abnormalities which indeed increased the survival rate of the patients through starting early treatment [21]. However, building CAD system using machine learning algorithm is not an easy task [5]. A set of features that well discriminate images from abnormal and normal must be extracted to build a well performing machine learning algorithms [5]. As cited in [5], X. Liu and J. Tang have explained how handcrafted method was used as a means of feature gathering and selection for mass lesion classification. But it is time consuming and also dependent on experts' knowledge.

As an alternative, deep learning became a choice to learn discriminant features directly from the data itself without special design of feature detectors [5]. CNN is one among many to become popular in the area of large size image processing. Based on the review made in [5,22], the success of CNN is proved to be promising in shape recognition, mass lesion classification using texture features and video recognition.

2 Convolutional Neural Networks

Convolutional Neural Networks are a biologically inspired variants of ordinary Neural Networks like multilayer perceptrons (MLP) which reduces the computational time and translational invariance [5,26]. Its major components are convolutional layer, pooling layer, normalization layer and fully connected layer [5,26]. And so far three different attempts are made to gain improvement in CNN architecture and achieve better accuracy [22]. In [28] using smaller window size and stride were made possible. In [27] training and testing the CNN over the whole image at multiple scales is achieved. As a result of smaller convolutional filters, authors in [22] have increased the depth of the convNet through increasing convolutional layers, and among them convNet with 16 and 19 layers achieved the best performance. The architectural overview of CNN is indicated in Fig. 1 and its detail can be found in [31] and Sect. 2.1 of [22]. Figure 1 depicts the basic CNN architecture where two convolutional and two pooling layers are stacked one after the other and then applied to the original images. In this architecture the extracted features are given to the fully connected layers as input to make classification. However, in our case the extracted features are fed to KNN after dimensionality reduction. The convolutional layer is responsible to create the feature mapping from the original image, and as indicated in Fig. 1, four feature maps are generated after convolutional operation. The convolutional operation is followed by the down-sampling operation using max-pooling layer to avoid non-maximal values which in turn reduces the computation to the subsequent upper layers [26,29].

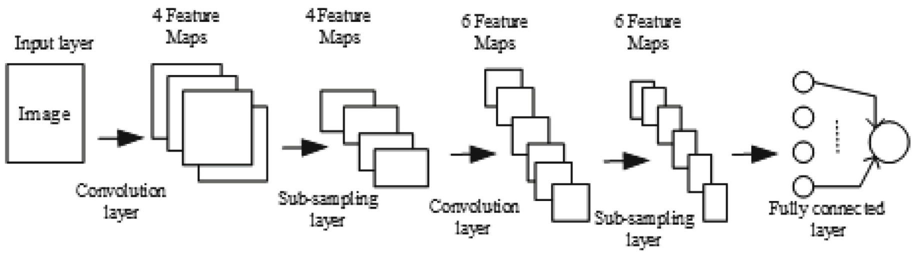

Fig. 1. Basic layers (convolutional, pooling, fully-connected) build CNN architecture from [26]. Convolutional layer is followed by pooling for down-sampling and in between, rectified linear unit (ReLU) is used with element wise activation function like $\max(0, \mathrm{x})$ without changing the size of the volume [22].

3 Related Work

There are many proposed and implemented approaches related to classification of mammograms using different techniques on MIAS and DDSM datasets. However, in this paper an attempt is made to briefly show the current state-of-the-arts in the area of image processing techniques, artificial intelligence technique and CNN.

In [4], the pixel-based approach is proposed to classify the mammograms taken from MIAS database as tumor and non-tumor. The computed Gabor feature pool from the mammograms are given to the support vector machine (SVM) classifier and obtained an accuracy of 80%.

In [10], a Particle Swarm Optimized Wavelet Neural Network (PSOWNN) for classification and Laws Texture Energy measures for feature extraction are proposed. A privately collected 216 mammograms are used for the experiment. Features are extracted from region of interest (ROI) after segmentation. The aim of the paper is to classify mammograms as normal or abnormal and the proposed approach achieved an accuracy of 93.68%, sensitivity of 94.14%, specificity of 92.10% and area under the receiver operating characteristic curve(AUC) of 0.968.

In [9], a GLCM texture features extracted from MIAS database are used in the proposed method. These features are given to Radial Basis Function Neural Network(RBFNN) as an input to classify the mammograms as normal or abnormal. The method achieved an accuracy of 93.98%.

In [8], 100 mammograms from MIAS database are used in their experiment to classify the mammograms as normal, benign and malignant. Before ANN classifier, the median filter and seeded region growing algorithm are applied to mammograms to remove noise and artifacts. The rough-set theory based feature selection algorithm is applied to the extracted 16 texture properties so as to reduce the number of features to 5. And using these features, the ANN classifier achieved a sensitivity of 98.6%, specificity of 89.3% and accuracy of 96%.

The mammograms from MIAS database is used in [7] as source image dataset during classification of mammograms as normal and cancerous. The mammograms are first enhanced and noise removal technique has been applied before extracting the wavelet coefficients using generalized Gaussian density model. Two classifiers are used in this paper and the accuracy achieved by Neural Network with a Bayesian Back Propagation(NNBBP) algorithm is 97.08% and 95.42% by Artificial Neuro-Fuzzy Inference system (ANFIS).

In [6], high-level and middle-level features are extracted from images in DDSM database using pre-trained CNN model at two different layers. The aim of the paper is to classify the breast mass as bengin and malignant using SVM as a classifier, which achieved an accuracy of 96.7%.

In [5], the performance of classifying mammography mass lesion as benign and malignant has been increased from 0.787 to 0.822 in terms of area under ROC. The features are extracted from the BCDR database, particularly BCDR-F03, using the pre-trained CNN model and given as input to SVM classifier. In this paper, data augmentation operation like flipping and rotation are applied to the original images to achieve balanced dataset for each class.

4 Materials and Methods

4.1 Dataset

The proposed approach is applied on benchmark databases that many researchers have used in previous breast cancer image analysis. The two databases are DDSM [13] and MIAS [14,16]. There are 322 images in the MIAS database with information that include character of background tissue, class of abnormalities, severity of abnormality, central coordinate of abnormality, and radius of circle enclosing abnormality [11].

The images used were 112 abnormal and 208 normal from MIAS database. DDSM database contains 2620 cases and 43 volumes with images taken in MLO and craniocaudal(CC) views. Similar number of images were used from DDSM database.

4.2 Preprocessing

Preprocessing stage plays an important role in many image processing application. To reduce the impact of dark parts in the borders of the mammograms, cropping was performed. This stage also removes the background and artifacts

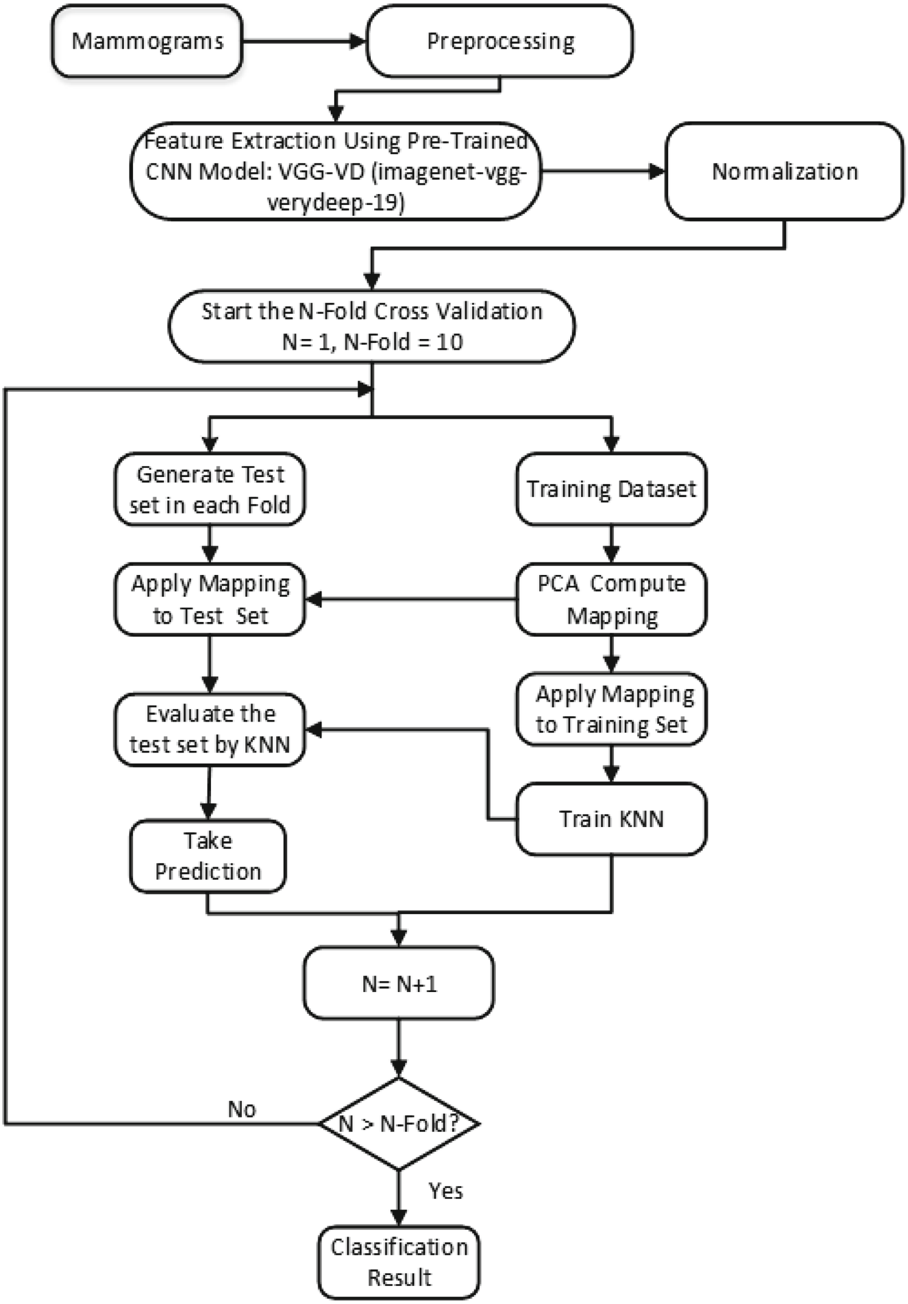

Fig. 2. Framework of the proposed approach. Mammograms from MIAS and DDSM is preprocessed and features are extracted using the pre-trained CNN model and normalized. Finally PCA is used for features dimensionality reduction inside 10 fold cross validation to train and evaluate the performance of KNN

on the original images [10]. For cropping, an algorithm that finds the first column in the left of the mammograms in which the sum of the pixels is greater than the threshold value is written. Then from this point the algorithm finds the first column in the right where the sum of the pixels is less than the given threshold value. The same cropping algorithm is applied to all mammograms used from MIAS and DDSM databases. Then, image enhancement, noise removal, image scaling and histogram equalization are applied to improve the quality of the original mammograms. In Fig. 3, the original and contrast enhanced image after cropping is given.

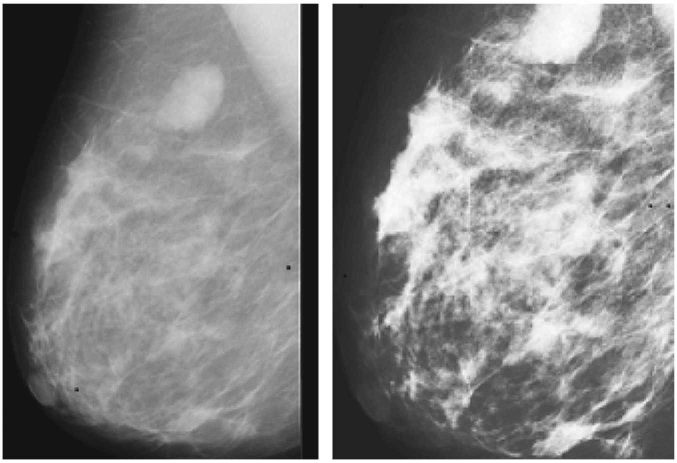

Fig. 3. Abnormal original image-MIAS (mdb015.pgm) (left), cropped and contrast ehnanced image (right) from [14]

4.3 Feature Extraction and Selection

The major steps in CAD system are preprocessing, segmentation, feature extraction, feature selection and classification. Even so, the role of feature extraction and selection are very significant on the performance of any classifier. After Preprocessing and image segmentation, set of features are required for each image that well represents the normal and abnormal mammograms. In this paper the pre-trained CNN model in [22] is used to extract features from the dataset of MIAS [3] and DDSM.

Once the features are extracted, classification can be performed as the next step without feature selection. All extracted features don't mean always important for better classification accuracy [23]. Feature selection is a search method that can be used to generate the subset of features [24]. It reduces the redundant features and computational cost [23,24]. For this purpose, PCA is used in dimensionality reduction.

4.4 Classification

Classification is one of the major decision making step of image recognition to separate mammogram whether cancerous or not. Among many of the classifiers, Naive Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Artificial Neural Networks (ANN), etc. are already introduced. In this paper KNN is used as a classifier with an aim of separating the mammograms as normal and abnormal.

KNN is one of the simplest and most important non-parameter algorithms among the supervised learning algorithms. It memorizes the training set and then predicts the label of any new instance based on the labels of its closest neighbors in the training data set. The classification performance of KNN classifier is evaluated based on accuracy as indicated in Eq. 1 [6,24].

5 Result and Discussion

The experiment was done to evaluate the performance of KNN with K value of 3 on MIAS and DDSM databases. The feature vectors extracted is given as input to the KNN classifier after PCS dimensionality reduction and trained using 10-fold cross-validation. The performance of a classifier was evaluated using accuracy, sensitivity and specificity as evaluation parameter. Given True Positive(TP), True Negative(TN), False Negative(FN) and False Positive(FP), the accuracy, sensitivity and specificity can be defined as:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{1}$$

$$sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$specificity = \frac{TN}{TN + FP} \tag{3}$$

The experimental result shows that KNN has scored a higher performance in terms of classification accuracy. The classification accuracy of KNN was 98.75% and 98.90% for MIAS and DDSM databases respectively. The 2D visualization of the data points after KNN is given in Fig. 4. It shows the classification boundary with misclassified data points. The data points near the boundary lines were checked to which image groups they belong to confirm that they are well classified or misclassified. The result of this study was also compared with previous research outputs as indicated in Table 1 and this study scored a better classification accuracy, sensitivity and specificity. The comparison was conducted in terms of time required to build the model during training and testing, and the result indicates that, for both datasets, computational time is insignificant.

The feature vector after dimensionality reduction is also given to RBF and SVM. The performance of both classifiers were evaluated in terms of sensitivity, specificity and accuracy. They achieved better performance compared to the state-of-the-art.

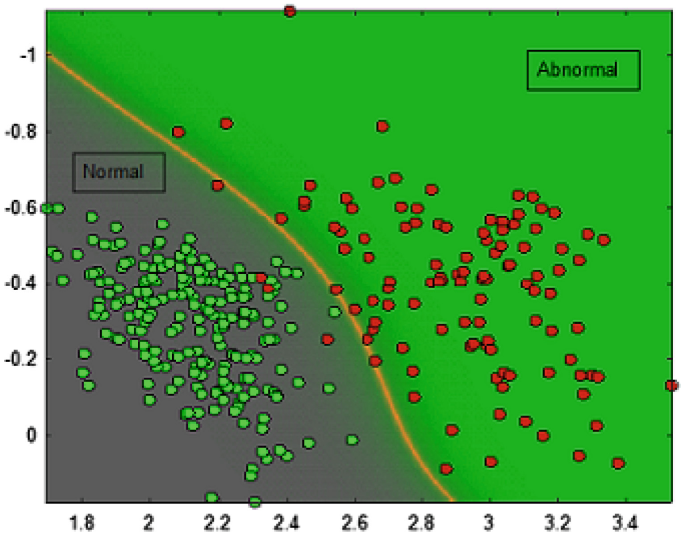

Fig. 4. 2D visual observation of the data points and the decision boundary of a polynomial of degree 3. The red data points represent data points from abnormal mammograms and yellow represents data points from normal mammograms (Color figure online)

Table 1. Comparison of proposed approach with previous studies

Author, year	Dataset	Accuracy(%)/ AUC	Features	Classifier
J. Torrents-Barrena et al. 2014	MIAS	80.00	Gabor	SVM
Y. A. Reyad and M. A. Berbar 2014	DDSM	98.63	Statistical and LBP	SVM
W. Xie et al. 2016	MIAS, DDSM	96.02,95.73	-	ELM
A. Alqoud and M. A. Jaffar 2016	MIAS	95.97	Gabor	ANN
A. Alqoud and M. A. Jaffar 2016	MIAS	96.82	LBP	ANN
A. Alqoud and M. A. Jaffar 2016	MIAS	98.72	Gabor and LBP	ANN
J. Dheeba et al. 2016	Private dataset	93.68	-	PSOWNN
Mellisa Pratiwi et al. 2015	MIAS	93.98	Texture	RBFNN
W. Peng et al. 2016	MIAS	96.00	Texture	ANN
Hela Mahersia et al. 2016	MIAS	97.08,95.42	-	NNBBP, ANFIS
Zhicheng Jiao et al. 2016	DDSM	96.70	CNN based features	SVM
John Arevalo et al. 2016	BCDR	-/0.822	CNN based features	SVM
Proposed approach	MIAS, DDSM	98.75,98.90	CNN based features	KNN

6 Conclusion

In this paper, a CNN based feature extraction method was proposed. The experiment was done to evaluate the distinguishing power of the features using KNN classifier. From the experimental result shown in previous section, the CNN based Feature Extraction technique for automatic breast cancer detection and classification is possible even without segmentation for region of interests. The results achieved in this study for the two popular datasets, MIAS and DDSM are promising and encourage for further improvement with better segmentation algorithm.

References

1. American Cancer Society: Breast Cancer Facts and Figures 2015–2016. http://www.cancer.org/acs/groups/content/@research/documents/document/acspc-046381.pdf
2. Sampat, M.P., Markey, M.K., Bovik, A.C.: Computer-aided detection and diagnosis in mammography. In: Handbook of Image and Video Processing. Elsevier Academic Press, San Francisco (2005)
3. Vedaldi, A., Lenc, K.: MatConvNet-convolutional neural networks for MATLAB. In: Proceedings of the 23rd ACM International Conference on Multimedia, Brisbane, Australia, pp. 689–692 (2015)
4. Torrents-Barrena, J., Puig, D., Ferre, M., Melendez, J., Diez-Presa, L., Arenas, M., Marti, J.: Breast masses identification through pixel-based texture classification. In: Fujita, H., Hara, T., Muramatsu, C. (eds.) IWDM 2014. LNCS, vol. 8539, pp. 581–588. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07887-8_81
5. Arevalo, J., González, F.A., Ramos-Pollán, R., Oliveira, J.L., Lopez, M.A.G.: Representation learning for mammography mass lesion classification with convolutional neural networks. Comput. Methods Programs Biomed. **127**, 248–257 (2016)
6. Jiao, Z., Gao, X., Wang, Y., Li, J.: A deep feature based framework for breast masses classification. Neurocomputing **197**, 221–231 (2016)
7. Mahersia, H., Boulehmi, H., Hamrouni, K.: Development of intelligent systems based on Bayesian regularization network and neuro-fuzzy models for mass detection in mammograms: a comparative analysis. Comput. Methods Programs Biomed. **126**, 46–62 (2016)
8. Peng, W., Mayorga, R.V., Hussein, E.M.A.: An automated confirmatory system for analysisof mammograms. Comput. Methods Programs Biomed. **125**, 134–144 (2016)
9. Pratiwi, M., Alexander, Harefa, J., Nandai, S.: Mammograms classification using gray-level co-occurrence matrix and radial basis function neural network. Procedia Comput. Sci. **59**, 83–91 (2015)
10. Dheeba, J., Albert Singh, N., Tamil Selvi, S.: Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach. J. Biomed. Inform. **49**, 45–52 (2014)
11. Pereira, E.T., Eleutério, S.P., Marques de Carvalho, J.: Local binary patterns applied to breast cancer classification in mammographies. RITA. **21**(2), 32–46 (2014)
12. Mammographic Image Analysis: Signs of Diseases. http://www.mammoimage.org/signs-of-disease/
13. University of South Florida Digital Mammography. http://marathon.csee.usf.edu/Mammography/Database.html
14. The mini-MIAS database of mammograms. http://peipa.essex.ac.uk/info/mias.html
15. Kanadam, K.P., Chereddy, S.R.: Mammogram classification using sparse-ROI: a novel representation to arbitrary shaped masses. Expert Syst. Appl. **57**, 204–213 (2016)
16. Dong, M., Lu, X., Ma, Y., Guo, Y., Ma, Y., Wang, K.: An efficient approach for automated mass segmentation and classification in mammograms. J. Digit. Imaging **28**(5), 613–625 (2015)
17. Rouhi, R., Jafari, M.: Classification of benign and malignant breast tumors based on hybrid level set segmentation. Expert Syst. Appl. **46**, 45–59 (2016)

18. Prasad, S.N., Houserkova, D.: The role of various modalities in breast imaging. Biomed. Pap. Med. Fac. Univ. Palacky Olomouc Czech Repub. **151**(2), 209–218 (2007)
19. Tirona, M.T.: Breast cancer screening update. Am. Fam. Physician **87**(2), 274–278 (2011)
20. Palazuelos, G., Trujillo, S., Romero, J.: Breast tomosynthesis: the new age of mammography. Rev. Colomb. Radiol. **25**(2), 3926–3933 (2014)
21. Dubey, R.B., Hanmandlu, M., Gupta, S.K.: A comparison of two methods for the segmentation of masses in the digital mammograms. Comput. Med. Imaging Graph. **34**, 185–191 (2010)
22. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556(2014)
23. Xie, W., Li, Y., Ma, Y.: Breast mass classification in digital mammography based on extreme learning machine. Neurocomputing **173**, 930–941 (2016)
24. Rouhi, R., Jafari, M., Kasaei, S., Keshavrzian, P.: Benign and malignant breast tumor classification based on region growing and CNN segmentation. Expert Syst. Appl. **42**, 990–1002 (2015)
25. Shalev-Shwartz, S., Ben-David, S.: Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, New York (2014)
26. Convolutional Neural Networks (LeNet). http://deeplearning.net/tutorial/lenet.html
27. Howard, A.G.: Some improvements on deep convolutional neural network based image classification. CoRR abs/1312.5402(2014)
28. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. CoRR abs/1311.2901 (2014)
29. Convolutional Neural Networks for Visual Recognition. http://cs231n.github.io/convolutional-networks/
30. Bonafede, M.M., Kalra, V.B., Miller, J.D., Fajardo, L.L.: Value analysis of digital breast tomosynthesis for breast cancer screening in a commercially-insured US population. ClinicoEconom. Outcomes Res. **7**, 53–63 (2015)
31. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. **25**, 1097–1105 (2012)