

# Mammographic Mass Segmentation with Online Learned Shape and Appearance Priors

Menglin Jiang<sup>1</sup>, Shaoting Zhang<sup>2(✉)</sup>, Yuanjie Zheng<sup>3</sup>,  
and Dimitris N. Metaxas<sup>1</sup>

<sup>1</sup> Department of Computer Science, Rutgers University, Piscataway, NJ, USA

<sup>2</sup> Department of Computer Science, UNC Charlotte, Charlotte, NC, USA  
szhang16@uncc.edu

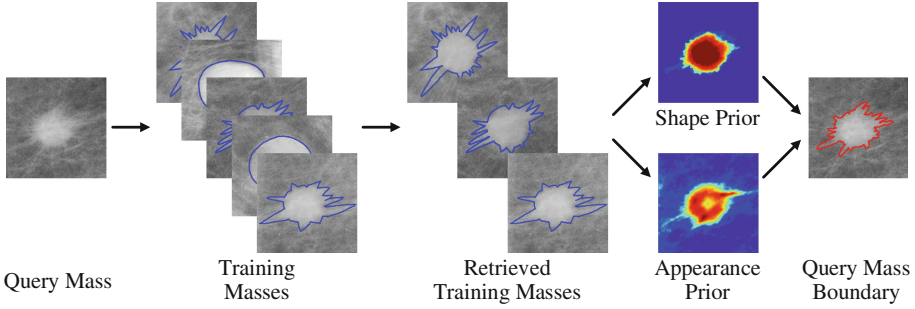
<sup>3</sup> School of Information Science and Engineering, Shandong Normal University,  
Jinan, China

**Abstract.** Automatic segmentation of mammographic mass is an important yet challenging task. Despite the great success of shape prior in biomedical image analysis, existing shape modeling methods are not suitable for mass segmentation. The reason is that masses have no specific biological structure and exhibit complex variation in shape, margin, and size. In addition, it is difficult to preserve the local details of mass boundaries, as masses may have spiculated and obscure boundaries. To solve these problems, we propose to learn online shape and appearance priors via image retrieval. In particular, given a query image, its visually similar training masses are first retrieved via Hough voting of local features. Then, query specific shape and appearance priors are calculated from these training masses on the fly. Finally, the query mass is segmented using these priors and graph cuts. The proposed approach is extensively validated on a large dataset constructed on DDSM. Results demonstrate that our online learned priors lead to substantial improvement in mass segmentation accuracy, compared with previous systems.

## 1 Introduction

For years, mammography has played a key role in the diagnosis of breast cancer, which is the second leading cause of cancer-related death among women. The major indicators of breast cancer are mass and microcalcification. Mass segmentation is important to many clinical applications. For example, it is critical to diagnosis of mass, since morphological and spiculation characteristics derived from segmentation result are strongly correlated to mass pathology [2]. However, mass segmentation is very challenging, since masses vary substantially in shape, margin, and size, and they often have obscure boundaries [7].

During the past two decades, many approaches have been proposed to facilitate mass segmentation [7]. Nevertheless, few of them adopt shape and appearance priors, which provide promising directions for many other biomedical image segmentation problems [12, 13], such as segmentation of human lung, liver, prostate, and hippocampus. In mass segmentation, the absence of the study



**Fig. 1.** Overview of our approach. The blue lines around training masses denote radiologist-labeled boundaries, and the red line on the rightmost image denotes our segmentation result.

of shape and appearance priors is mainly due to two reasons. First, unlike the aforementioned organs/objects, mammographic masses have no specific biological structure, and they present large variation in shape, margin, and size. Naturally, it is very hard to construct shape or appearance models for mass [7]. Second, masses are often indistinguishable from surrounding tissues and may have greatly spiculated margins. Therefore, it is difficult to preserve the local details of mass boundaries.

To solve the above problems, we propose to incorporate image retrieval into mammographic mass segmentation, and learn “customized” shape and appearance priors for each query mass. The overview of our approach is shown in Fig. 1. Specifically, during the offline process, a large number of diagnosed masses form a training set. SIFT features [6] are extracted from these masses and stored in an inverted index for fast retrieval [11]. During the online process, given a query mass, it is first matched with all the training mass through Hough voting of SIFT features [10] to find the most similar ones. A similarity score is also calculated to measure the overall similarity between the query mass and its retrieved training masses. Then, shape and appearance priors are learned from the retrieved masses on the fly, which characterize the global shape and local appearance information of these masses. Finally, the two priors are integrated in a segmentation energy function, and their weights are automatically adjusted using the aforesaid similarity score. The query mass is segmented by solving the energy function via graph cuts [1].

In mass segmentation, our approach has several advantages over existing online shape prior modeling methods, such as atlas-based methods [12] and sparse shape composition (SSC) [13]. First, these methods are generally designed for organs/objects with anatomical structures and relatively simple shapes. When dealing with mass segmentation, some assumptions of those methods, such as correspondence between organ landmarks, will be violated and thus the results will become unreliable. On the contrary, our approach adopts a retrieval method dedicated to handle objects with complex shape variations. Therefore, it

could get effective shape priors for masses. Second, since the retrieved training masses are similar to the query mass in terms of not only global shape but also local appearance, our approach incorporates a novel appearance prior, which complements shape prior and preserves the local details of mass boundaries. Finally, the priors' weights in our segmentation energy function are automatically adjusted using the similarity between the query mass and its retrieved training masses, which makes our approach even more adaptive.

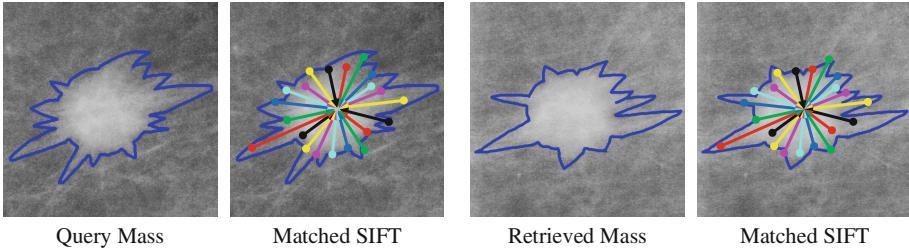
## 2 Methodology

In this section, we first introduce our mass retrieval process, and then describe how to learn shape and appearance priors from the retrieval result, followed by our mass segmentation method. The framework of our approach is illustrated in Fig. 1.

**Mass Retrieval Based on Hough Voting:** Our approach characterizes mammographic images with densely sampled SIFT features [6], which demonstrate excellent performance in mass retrieval and analysis [4]. To accelerate the retrieval process, all the SIFT features are quantized using bag-of-words (BoW) method [4, 11], and the quantized SIFT features extracted from training set are stored in an inverted index [4]. The training set  $\mathcal{D}$  comprises a series of samples, each of which contains a diagnosed mass located at the center. A training mass  $d \in \mathcal{D}$  is represented as  $d = \{(v_j^d, \mathbf{p}_j^d)\}_{j=1}^n$ , where  $n$  is the number of features extracted from  $d$ ,  $v_j^d$  denotes the  $j$ -th quantized feature (visual word ID), and  $\mathbf{p}_j^d = [x_j^d, y_j^d]^T$  denotes the *relative* position of  $v_j^d$  from the center of  $d$  (the coordinate origin is at mass center). The query set  $\mathcal{Q}$  includes a series of query masses. Note that query masses are not necessarily located at image centers. A query mass  $q \in \mathcal{Q}$  is represented as  $q = \{(v_i^q, \mathbf{p}_i^q)\}_{i=1}^m$ , where  $\mathbf{p}_i^q = [x_i^q, y_i^q]^T$  denotes the *absolute* position of  $v_i^q$  (the origin is at the upper left corner of the image since the position of the mass center is unknown).

Given a query mass  $q$ , it is matched with all the training masses. In order to find similar training masses with different orientations or sizes, all the training masses are *virtually* transformed using 8 rotation degrees (from 0 to  $7\pi/4$ ) and 8 scaling factors (from 1/2 to 2). To this end, we only need to re-calculate the positions of SIFT features, since SIFT is invariant to rotation and scale change [6].

For the given query mass  $q$  and any (transformed) training mass  $d$ , we calculate a similarity map  $\mathcal{S}_{q,d}$ , a similarity score  $s_{q,d}$ , and the position of the query mass center  $\mathbf{c}_{q,d}$ .  $\mathcal{S}_{q,d}$  is a matrix of the same size of  $q$ , and its element at position  $\mathbf{p}$ , denoted as  $\mathcal{S}_{q,d}(\mathbf{p})$ , indicates the similarity between the region of  $q$  centered at  $\mathbf{p}$  and  $d$ . The matching process is based on generalized Hough voting of SIFT features [10], which is illustrated in Fig. 2. The basic idea is that the features should be quantized to the same visual words and be spatially consistent (i.e., have similar positions relative to mass centers) if  $q$  matches  $d$ . In particular, given a pair of matched features  $v_i^q = v_j^d = v$ , the *absolute* position of the query mass center, denoted as  $\mathbf{c}_i^q$ , is first computed based on  $\mathbf{p}_i^q$  and  $\mathbf{p}_j^d$ . Then  $v_i^q$  updates the similarity map  $\mathcal{S}_{q,d}$ . To resist gentle nonrigid deformation,



**Fig. 2.** Illustration of our mass matching algorithm. The blue lines denote the mass boundaries labeled by radiologists. The dots indicate the positions of the matched SIFT features, which are spatially consistent. The arrows denote the *relative* positions of the training features to the center of the training mass. The center of the query mass is localized by finding the maximum element in  $\mathbf{S}_{q,d}$ .

$v_i^q$  votes in favor of not only  $\mathbf{c}_i^q$  but also the neighbors of  $\mathbf{c}_i^q$ .  $\mathbf{c}_i^q$  earns a full vote, and each neighbor gains a vote weighed by a Gaussian factor:

$$\mathbf{S}_{q,d}(\mathbf{c}_i^q + \delta\mathbf{p}) = \frac{\text{idf}^2(v)}{\text{tf}(v,q)\text{tf}(v,d)} \exp\left(-\frac{\|\delta\mathbf{p}\|^2}{2\sigma^2}\right), \quad (1)$$

where  $\delta\mathbf{p}$  represents the displacement from  $\mathbf{c}_i^q$  to its neighbor,  $\sigma$  determines the order of  $\|\delta\mathbf{p}\|$ ,  $\text{tf}(v,q)$  and  $\text{tf}(v,d)$  are the term frequencies (TFs) of  $v$  in  $q$  and  $d$  respectively, and  $\text{idf}(v)$  is the inverse document frequency (IDF) of  $v$ . TF-IDF reflects the importance of a visual word to an image in a collection of images, and is widely adopted in BoW-based image retrieval methods [4, 10, 11]. The cumulative votes of all the feature pairs generate the similarity map  $\mathbf{S}_{q,d}$ . The largest element in  $\mathbf{S}_{q,d}$  is defined as the similarity score  $s_{q,d}$ , and the position of the largest element is defined as the query mass center  $\mathbf{c}_{q,d}$ .

After computing the similarity scores between  $q$  and all the (transformed) training masses, the top  $k$  most similar training masses along with their diagnostic reports are returned to radiologists. These masses are referred to as the retrieval set of  $q$ , which is denoted as  $\mathcal{N}_q$ . The average similarity score of these  $k$  training masses is denoted as  $\omega = (1/k) \sum_{d \in \mathcal{N}_q} s_{q,d}$ . During the segmentation of  $q$ ,  $\omega$  measures the confidence of our shape and appearance priors learned from  $\mathcal{N}_q$ .

Note that our retrieval method could find training masses which are similar in local appearance and global shape to the query mass. A match between a query feature and a training feature assures that the two local patches, from where the SIFT features are extracted, have similar appearances. Besides, the spatial consistency constraint guarantees that two matched masses have similar shapes and sizes. Consequently, the retrieved training masses could guide segmentation of the query mass. Moreover, due to the adoption of BoW technique and inverted index, our retrieval method is computationally efficient.

**Learning Online Shape and Appearance Priors:** Given a query mass  $q$ , our segmentation method aims to find a foreground mask  $\mathbf{L}_q$ .  $\mathbf{L}_q$  is a binary matrix

of the size of  $q$ , and its element  $\mathbf{L}_q(\mathbf{p}) \in \{0, 1\}$  indicates the label of the pixel at position  $\mathbf{p}$ , where 0 and 1 represent background and mass respectively. Each retrieved training mass  $d \in \mathcal{N}_q$  has a foreground mask  $\mathbf{L}_d$ . To align  $d$  with  $q$ , we simply copy  $\mathbf{L}_d$  to a new mask of the same size of  $\mathbf{L}_q$  and move the center of  $\mathbf{L}_d$  to the query mass center  $\mathbf{c}_{q,d}$ .  $\mathbf{L}_d$  will hereafter denote the aligned foreground mask.

Utilizing the foreground masks of the retrieved training masses in  $\mathcal{N}_q$ , we could learn shape and appearance priors for  $q$  on the fly. Shape prior models the spatial distribution of the pixels in  $q$  belonging to a mass. Our approach estimates this prior by averaging the foreground masks of the retrieved masses:

$$\begin{aligned} p_S(\mathbf{L}_q(\mathbf{p}) = 1) &= \frac{1}{k} \sum_{d \in \mathcal{N}_q} \mathbf{L}_d(\mathbf{p}), \\ p_S(\mathbf{L}_q(\mathbf{p}) = 0) &= 1 - p_S(\mathbf{L}_q(\mathbf{p}) = 1). \end{aligned} \quad (2)$$

Appearance prior models how likely a small patch in  $q$  belongs to a mass. In our approach, a patch is a small region from where a SIFT feature is extracted, and it is characterized by its visual word (quantized SIFT feature). The probability of word  $v$  belonging to a mass is estimated on  $\mathcal{N}_q$ :

$$\begin{aligned} p_A(\mathbf{L}_q(\mathbf{p}_v) = 1) &= \frac{n_v^f}{n_v}, \\ p_A(\mathbf{L}_q(\mathbf{p}_v) = 0) &= 1 - p_A(\mathbf{L}_q(\mathbf{p}_v) = 1), \end{aligned} \quad (3)$$

where  $\mathbf{p}_v$  is the position of word  $v$ ,  $n_v$  is the total number of times that  $v$  appears in  $\mathcal{N}_q$ ,  $n_v^f$  is the number of times that  $v$  appears in foreground masses.

It is noteworthy that our shape and appearance priors are complementary. In particular, shape prior tends to recognize mass centers, since the average foreground mask of the retrieved training masses generally has large scores around mass centers. Appearance prior, on the other hand, tends to recognize mass edges, as SIFT features extracted from mass edges are very discriminative [4]. Examples of shape and appearance priors are provided in Fig. 1.

**Mass Segmentation via Graph Cuts with Priors:** Our segmentation method computes the foreground mask  $\mathbf{L}_q$  by minimizing the following energy function:

$$\begin{aligned} E(\mathbf{L}_q) &= \lambda_1 E_I(\mathbf{L}_q) + \lambda_2 \omega E_S(\mathbf{L}_q) + \lambda_3 \omega E_A(\mathbf{L}_q) + E_R(\mathbf{L}_q) \\ &= -\lambda_1 \sum_{\mathbf{p}} \ln p_I(\mathbf{I}_q(\mathbf{p}) | \mathbf{L}_q(\mathbf{p})) - \lambda_2 \omega \sum_{\mathbf{p}} \ln p_S(\mathbf{L}_q(\mathbf{p})) \\ &\quad - \lambda_3 \omega \sum_{\mathbf{p}} \ln p_A(\mathbf{L}_q(\mathbf{p})) + \sum_{\mathbf{p}, \mathbf{p}'} \beta(\mathbf{L}_q(\mathbf{p}), \mathbf{L}_q(\mathbf{p}')), \end{aligned} \quad (4)$$

where  $\mathbf{I}_q$  denotes the intensity matrix of  $q$ ,  $\mathbf{I}_q(\mathbf{p})$  represents the value of the pixel at position  $\mathbf{p}$ .  $E_I(\mathbf{L}_q)$ ,  $E_S(\mathbf{L}_q)$ ,  $E_A(\mathbf{L}_q)$  and  $E_R(\mathbf{L}_q)$  are the energy terms related to intensity information, shape prior, appearance prior, and regularity constraint, respectively.  $\beta(\mathbf{L}_q(\mathbf{p}), \mathbf{L}_q(\mathbf{p}'))$  is a penalty term for adjacent pixels with different labels.  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  are the weights for the first three energy terms, and the last term has an implicit weight 1.

In particular, the intensity energy  $E_I(\mathbf{L}_q)$  evaluates how well  $\mathbf{L}_q$  explains  $\mathbf{I}_q$ . It is derived from the total likelihood of the observed intensities given certain labels. Following conventions in radiological image segmentation [5, 12], the foreground likelihood  $p_I(\mathbf{I}_q(\mathbf{p}) | \mathbf{L}_q(\mathbf{p}) = 1)$  and background likelihood  $p_I(\mathbf{I}_q(\mathbf{p}) | \mathbf{L}_q(\mathbf{p}) = 0)$  are approximated by Gaussian density function and Parzen window estimator respectively, and are both learned on the entire training set  $\mathcal{D}$ . The shape energy  $E_S(\mathbf{L}_q)$  and appearance energy  $E_A(\mathbf{L}_q)$  measure how well  $\mathbf{L}_q$  fits the shape and appearance priors. The regularity energy  $E_R(\mathbf{L}_q)$  is employed to promote smooth segmentation. It calculates a penalty score for every pair of neighboring pixels  $(\mathbf{p}, \mathbf{p}')$ . Following [1, 12], we compute this score using:

$$\beta(\mathbf{L}_q(\mathbf{p}), \mathbf{L}_q(\mathbf{p}')) = \frac{\mathbb{1}(\mathbf{L}_q(\mathbf{p}) \neq \mathbf{L}_q(\mathbf{p}'))}{2 \|\mathbf{p} - \mathbf{p}'\|} \exp\left(-\frac{(\mathbf{I}_q(\mathbf{p}) - \mathbf{I}_q(\mathbf{p}'))^2}{2\zeta^2}\right), \quad (5)$$

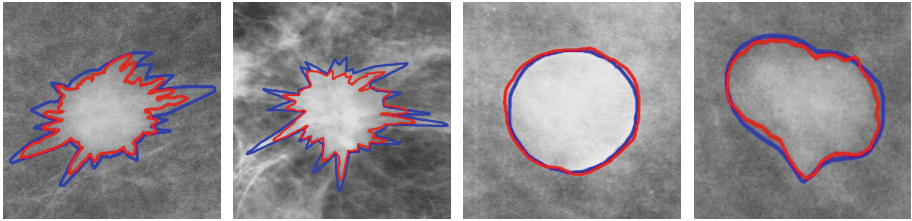
where  $\mathbb{1}$  is the indicator function, and  $\zeta$  determines the order of intensity difference. The above function assigns a positive score to  $(\mathbf{p}, \mathbf{p}')$  only if they have different labels, and the score will be large if they have similar intensities and short distance. Similar to [12], we first plug in Eqs. (2), (3) and (5) to energy function Eq. (4), then convert it to the sum of unary potentials and pairwise potentials, and finally minimize it via graph cuts [1] to obtain the foreground mask  $\mathbf{L}_q$ .

Note that the overall similarity score  $\omega$  is utilized to adjust the weights of the prior-related energy terms in Eq. (4). As a result, if there are similar masses in the training set, our segmentation method will rely on the priors. Otherwise,  $\omega$  will be very small and Eq. (4) automatically degenerates to traditional graph cuts-based segmentation, which prevents ineffective priors from reducing the segmentation accuracy.

### 3 Experiments

In this section, we first describe our dataset, then evaluate the performance of mass retrieval and mass segmentation using our approach.

**Dataset:** Our dataset builds on the digital database for screening mammography (DDSM) [3], which is currently the largest public mammogram database. DDSM is comprised of 2,604 cases, and every case consists of four views, i.e., LEFT-CC, LEFT-MLO, RIGHT-CC and RIGHT-MLO. The masses have diverse shapes, margins, sizes, breast densities as well as patients' ages. They also have radiologist-labeled boundaries and diagnosed pathologies. To build our dataset, 2,340 image regions centered at masses are extracted. Our approach and the comparison methods are tested five times. During each time, 100 images are randomly selected to form the query set  $\mathcal{Q}$ , and the remaining 2,240 images form the training set  $\mathcal{D}$ .  $\mathcal{Q}$  and  $\mathcal{D}$  are selected from different cases in order to avoid positive bias. Below we report the average of the evaluation results during five tests.



**Fig. 3.** Our segmentation results on four masses, which are represented by red lines. The blue lines denote radiologist-labeled mass boundaries. The left two masses are malignant (cancer), and the right two masses are benign.

**Evaluation of Mass Retrieval:** The evaluation metric adopted here is retrieval precision. In our context, precision is defined as the percentage of retrieved training masses that have the same shape category as that of the query mass. All the shape attributes are divided as two categories. The first category includes “irregular”, “lobulated”, “architectural distortion”, and their combinations. The second category includes “round” and “oval”. We compare our method with a state-of-the-art mass retrieval approach [4], which indexes quantized SIFT features with a vocabulary tree. The precision scores of both methods change slightly as the size of retrieval set  $k$  increases from 1 to 30, and our method systematically outperforms the comparison method. For instance, at  $k = 20$ , the precision scores of our method and the vocabulary tree-based method are  $0.85 \pm 0.11$  and  $0.81 \pm 0.14$ , respectively. Our precise mass retrieval method lays the foundation for learning accurate priors and improving mass segmentation performance.

**Evaluation of Mass Segmentation:** Segmentation accuracy is assessed by area overlap measure (AOM) and average minimum distance (DIST), which are two widely used evaluation metrics in medical image segmentation. Our method is tested with three settings, i.e., employing shape prior, appearance prior, and both priors. The three configurations are hereafter denoted as “Ours-Shape”, “Ours-App”, and “Ours-Both”. For all the configurations, we set  $k$  to the same value in mass retrieval experiments, i.e.  $k = 20$ .  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  are tuned through cross validation.

Three classical and state-of-the-art mass segmentation approaches are implemented for comparison, which are based on active contour (AC) [8], convolution neural network (CNN) [5], and traditional graph cuts (GC) [9], respectively. The key parameters of these methods are tuned using cross validation. The evaluation results are summarized in Table 1. A few segmentation results obtained by Ours-Both are provided in Fig. 3 for qualitative evaluation.

The above results lead to several conclusions. First, our approach could find visually similar training masses for most query masses and calculate effective shape and appearance priors. Therefore, Ours-Shape and Ours-App substantially surpass GC. Second, as noted earlier, the two priors are complementary: shape prior recognizes mass centers, whereas appearance prior is vital to keep the

**Table 1.** AOM and DIST (unit is mm) scores of the evaluated methods

	AC [8]	CNN [5]	GC [9]	Ours-Shape	Ours-App	Ours-Both
AOM	$0.78 \pm 0.12$	$0.73 \pm 0.17$	$0.75 \pm 0.14$	$0.81 \pm 0.13$	$0.80 \pm 0.10$	<b><math>0.84 \pm 0.09</math></b>
DIST	$1.09 \pm 0.43$	$1.36 \pm 0.62$	$1.24 \pm 0.54$	$0.97 \pm 0.49$	$1.01 \pm 0.45$	<b><math>0.88 \pm 0.47</math></b>

local details of mass boundaries. Thus, by integrating both priors, Ours-Both further improves the segmentation accuracy. Third, detailed results show that for some “uncommon” query masses, which have few similar training masses, the overall similarity score  $\omega$  is very small and the segmentation results of Ours-Both are similar to those of GC. That is, the adaptive weights of priors successfully prevent ineffective priors from backfiring. Finally, Ours-Both outperforms all the comparison methods especially for masses with irregular and spiculated shapes. Its segmentation results have a close agreement with radiologist-labeled mass boundaries, and are highly consistent with mass pathologies.

## 4 Conclusion

In this paper, we leverage image retrieval method to learn query specific shape and appearance priors for mammographic mass segmentation. Given a query mass, similar training masses are found via Hough voting of SIFT features, and priors are learned from these masses. The query mass is segmented through graph cuts with priors, where the weights of priors are automatically adjusted according to the overall similarity between query mass and its retrieved training masses. Extensive experiments on DDSM demonstrate that our online learned priors considerably improve the segmentation accuracy, and our approach outperforms several widely used mass segmentation methods or systems. Future endeavors will be devoted to distinguishing between benign and malignant masses using features derived from mass boundaries.

## References

1. Boykov, Y.Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(11), 1222–1239 (2001)
2. D’Orsi, C.J., Sickles, E.A., Mendelson, E.B., et al.: *ACR BI-RADS Atlas, Breast Imaging Reporting and Data System*, 5th edn. American College of Radiology, Reston (2013)
3. Heath, M., Bowyer, K., Kopans, D., Moore, R., Kegelmeyer, W.P.: The digital database for screening mammography. In: *Proceeding IWDM*, pp. 212–218 (2000)
4. Jiang, M., Zhang, S., Li, H., Metaxas, D.N.: Computer-aided diagnosis of mammographic masses using scalable image retrieval. *IEEE Trans. Biomed. Eng.* **62**(2), 783–792 (2015)
5. Lo, S.B., Li, H., Wang, Y.J., Kinnard, L., Freedman, M.T.: A multiple circular paths convolution neural network system for detection of mammographic masses. *IEEE Trans. Med. Imaging* **21**(2), 150–158 (2002)



6. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
7. Oliver, A., Freixenet, J., Martí, J., Pérez, E., Pont, J., Denton, E.R.E., Zwiggelaar, R.: A review of automatic mass detection and segmentation in mammographic images. *Med. Image Anal.* **14**(2), 87–110 (2010)
8. Rahmati, P., Adler, A., Hamarneh, G.: Mammography segmentation with maximum likelihood active contours. *Med. Image Anal.* **16**(6), 1167–1186 (2012)
9. Saidin, N., Sakim, H.A.M., Ngah, U.K., Shuaib, I.L.: Computer aided detection of breast density and mass, and visualization of other breast anatomical regions on mammograms using graph cuts. *Comput. Math. Methods Med.* **2013**(205384), 1–13 (2013)
10. Shen, X., Lin, Z., Brandt, J., Avidan, S., Wu, Y.: Object retrieval and localization with spatially-constrained similarity measure and k-NN re-ranking. In: *Proceeding CVPR*, pp. 3013–3020 (2012)
11. Sivic, J., Zisserman, A.: Video Google: a text retrieval approach to object matching in videos. In: *Proceeding ICCV*, pp. 1470–1477 (2003)
12. van der Lijn, F., den Heijer, T., Breteler, M.M.B., Niessen, W.J.: Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. *NeuroImage* **43**(4), 708–720 (2008)
13. Zhang, S., Zhan, Y., Zhou, Y., Uzunbas, M., Metaxas, D.N.: Shape prior modeling using sparse representation and online dictionary learning. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) *MICCAI 2012*. LNCS, vol. 7512, pp. 435–442. Springer, Heidelberg (2012). doi:[10.1007/978-3-642-33454-2\\_54](https://doi.org/10.1007/978-3-642-33454-2_54)