Computer aided detection of surgical retained foreign object for prevention

Lubomir Hadjiiski[a)] and Theodore C. Marentis
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109

Amrita R. Chaudhury
Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan 48109

Lucas Rondon
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109

Nikolaos Chronis
Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan 48109

Heang-Ping Chan
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109

Purpose: Surgical retained foreign objects (RFOs) have significant morbidity and mortality. They are associated with approximately 1.5×10^9 annually in preventable medical costs. The detection accuracy of radiographs for RFOs is a mediocre 59%. The authors address the RFO problem with two complementary technologies: a three-dimensional (3D) gossypiboma micro tag, the μTag that improves the visibility of RFOs on radiographs, and a computer aided detection (CAD) system that detects the μTag. It is desirable for the CAD system to operate in a high specificity mode in the operating room (OR) and function as a first reader for the surgeon. This allows for fast point of care results and seamless workflow integration. The CAD system can also operate in a high sensitivity mode as a second reader for the radiologist to ensure the highest possible detection accuracy.

Methods: The 3D geometry of the μTag produces a similar two dimensional (2D) depiction on radiographs regardless of its orientation in the human body and ensures accurate detection by a radiologist and the CAD. The authors created a data set of 1800 cadaver images with the 3D μTag and other common man-made surgical objects positioned randomly. A total of 1061 cadaver images contained a single μTag and the remaining 739 were without μTag. A radiologist marked the location of the μTag using an in-house developed graphical user interface. The data set was partitioned into three independent subsets: a training set, a validation set, and a test set, consisting of 540, 560, and 700 images, respectively. A CAD system with modules that included preprocessing μTag enhancement, labeling, segmentation, feature analysis, classification, and detection was developed. The CAD system was developed using the training and the validation sets.

Results: On the training set, the CAD achieved 81.5% sensitivity with 0.014 false positives (FPs) per image in a high specificity mode for the surgeons in the OR and 96.1% sensitivity with 0.81 FPs per image in a high sensitivity mode for the radiologists. On the independent test set, the CAD achieved 79.5% sensitivity with 0.003 FPs per image in a high specificity mode for the surgeons and 90.2% sensitivity with 0.23 FPs per image in a high sensitivity mode for the radiologists.

Conclusions: To the best of the authors' knowledge, this is the first time a 3D μTag is used to produce a recognizable, substantially similar 2D projection on radiographs regardless of orientation in space. It is the first time a CAD system is used to search for man-made objects over anatomic background. The CAD system for the μTags achieved reasonable performance in both the high specificity and the high sensitivity modes. © *2015 American Association of Physicists in Medicine.* [http://dx.doi.org/10.1118/1.4907964]

1. INTRODUCTION

Surgical retained foreign objects (RFOs) have significant morbidity and mortality.[1] The associated combined medical and legal burden to the U.S. healthcare system is estimated at 1.5×10^9 annually.[2,3] The current approach to prevent RFOs, a combination of universal surgical count and elective radiographs when the count is off is flawed, as 70%–90% of confirmed gossypibomas cases have a correct count[4] and the detection accuracy of radiographs for RFOs is at best a mediocre 59%.[5] An example of a retained sponge on a clinical radiograph is shown in Fig. 1.

Even though surgical sponges carry x-ray visible features, they are malleable and have variable appearance. Moreover,

they are not standardized across vendors and their shape changes in the body. This can result in overlooking a gossypiboma, even if it is large and in plain sight. Although the detection accuracy of radiographs for RFOs is a mediocre 59%,[5] they are still considered cost-effective[4,6] and had been performed routinely after every operation in the Mayo Clinic for almost 70 yr.[5,7] Universal postoperative radiographs in open surgeries[4] and high risk cases, such as trauma, bariatric, or long operation,[8,9] have been recommended regardless of the count outcome.

To address the RFO problem, we are developing a computer aided detection (CAD) system that detects the three-dimensional (3D) gossypiboma micro tag, the μTag (Kalyspo LLC, Ann Arbor, MI[10]). The μTag is at present only at the experimental stage. We have previously reported a pilot study[11] in which a small data set of 346 cadaveric radiographs was used. The CAD system achieved 85.5% sensitivity at a false positive (FP) rate of 0.02 FPs/image in a high specificity mode for the surgeons in the operating room (OR) and 96% sensitivity at 0.73 FPs/image in a high sensitivity mode for the radiologists. In the current study, we continued the development of the CAD system with enlarged training and validation sets and performed a true performance assessment using a large independent test set.

2. METHODS

2.A. μTag design

Our goal was to produce a tag that is always visible on x-rays and a CAD system that detects the tag as a first or second reader. The μTag consists of high contrast features that greatly enhance the visibility of RFOs on radiographs. We have designed a 3D μTag that consists of four microspheres, 0.8 mm in diameter at the vertices of a tetrahedron (Fig. 2). The microspheres of the μTag are spaced apart by specific distances (2.1 mm between microspheres' centroids) so that they are easily distinguishable from naturally occurring high contrast features of the human anatomy, such as bone edges (Fig. 3).

The 3D geometry of the μTag provides at least three visible microspheres in any projection on the radiograph regardless of its orientation and location in the human body (Fig. 3). Even if two of the microspheres overlap, the μTag is still highly visible and recognizable [Fig. 3(b)]. The tag is not malleable and does not change shape.

2.B. Data set

With approval from the University of Michigan Anatomical Donations Program, we imaged the μTags placed over cadavers to simulate RFO in the patient body. Placing the μTag over the cadaver abdomen on the side farthest from the detector results in the greatest geometric blurring and scatter degradation of the μTags on the radiograph. This simulates the worst-case scenario, from an imaging point of view, so that we did not need to place μTags inside the cadavers. To ensure adequate sampling of RFO over a
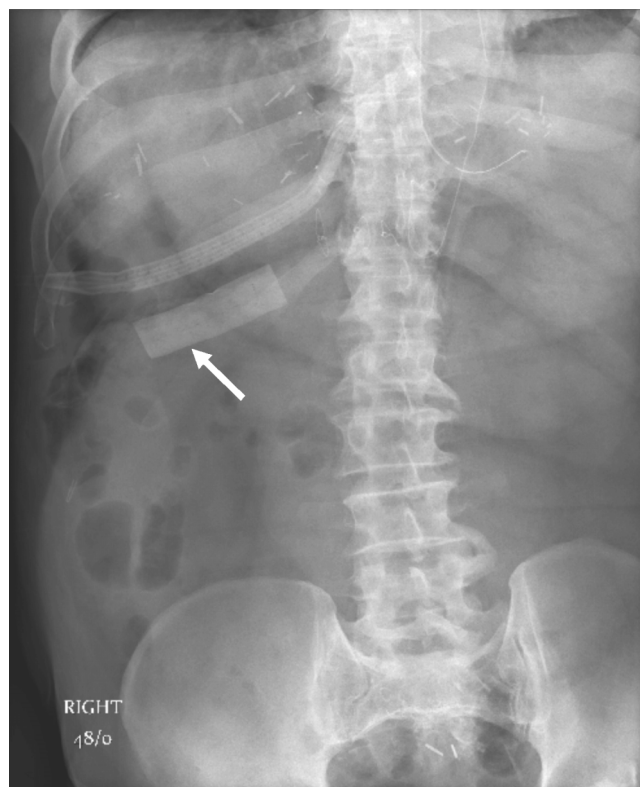

FIG. 1. A retained sponge on a radiograph (marked by the arrow) below a surgical drain.

variety of anatomic backgrounds and different locations of the abdomen, as well as random orientation of the μTag in space, imaging was performed as follows: the abdomen was divided into a 4×5 grid. The μTag was encapsulated in an opaque sphere that could be rolled and might stop at any random orientation. In addition, with IRB approval, we estimated the incidence of lines, tubes, and other man-made objects on intraoperative radiographs. A statistician determined the frequencies that the μTag and other foreign objects at the estimated incidence rates should be placed in each grid cell. These objects were then placed on the
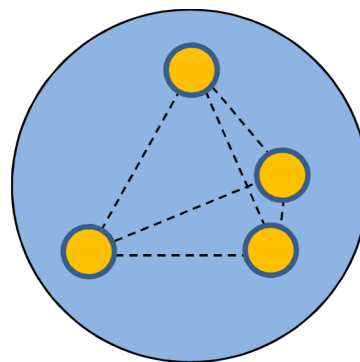

FIG. 2. The 3D μTag contained four microspheres arranged at the vertices of a tetrahedron embedded in a polymer matrix. The many possible configurations of the projected images of the μTag on 2D radiographs can be seen in Figs. 3, 4, 6, 8, and 9.
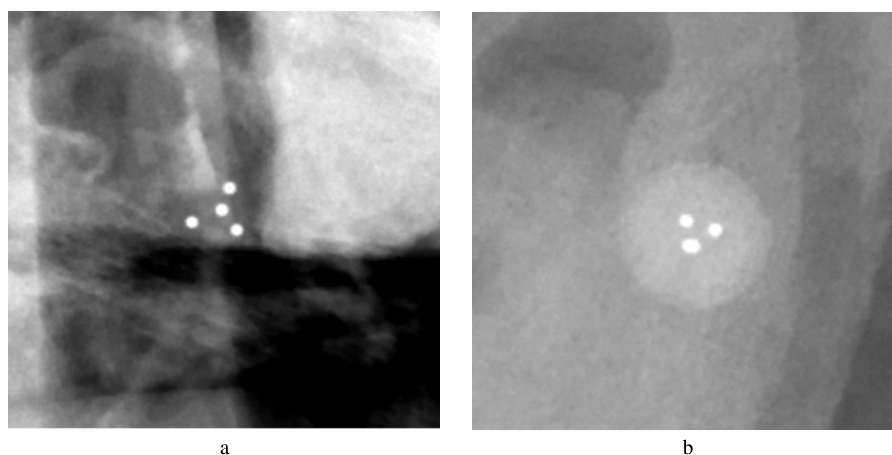
FIG. 3. (a) The 3D μTag imaged en-face with all four microspheres visible on anatomical background. (b) The 3D μTag rotated in a worst-case scenario. Two microspheres overlapped but the μTag was still highly visible and recognizable.

cadavers during image acquisition to simulate more realistic appearance on postoperative radiographs. Radiographs were acquired with a portable x-ray machine (Shimadzu) and an 8-megapixel flat-panel detector (Canon). The pixel pitch of the detector was 125×125 μm. An example of a cadaver image with μTag, tubes, and other man-made objects is shown in Fig. 4, in which the μTag can be seen overlapping with a tube [Fig. 4(b)].

We created a data set of 1800 cadaver images with the 3D μTag and other surgical objects commonly seen on clinical radiographs. A total of 1061 cadaver images contained a single μTag and the remaining 739 were without μTag. For the 1061 cadaver radiographs with μTag, a radiologist marked the location of the μTag with a bounding box using an in-house developed graphical user interface and also verified against the record of where the μTag was placed when a given cadaver radiograph was taken. These locations provide the

gold standard for evaluation of the CAD system performance. The data set was partitioned into three independent subsets: a training set, a validation set, and a test set. The sizes of the different partitions are presented in Table I. The CAD system was developed using the training and the validation sets. The performance of the CAD system was evaluated by the test set that was sequestered during training and validation.

2.C. CAD system design

Our current CAD system focuses on detection of μTags attached to sponges. The entire system is fully automated. A flow diagram of the CAD system is shown in Fig. 5. It consists of five processing modules: μTag enhancement module, μTag candidate labeling module, μTag segmentation module, feature analysis module, and feature classification module.
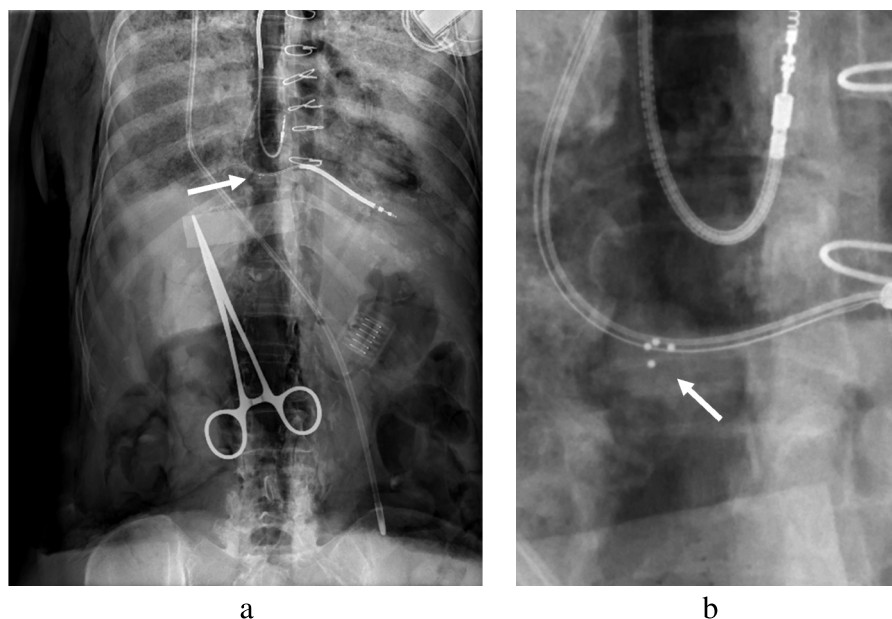

FIG. 4. (a) A typical image from our data set with multiple lines, tubes, and the μTag (marked by the arrow). (b) ROI showing the μTag overlapping with tubes.

TABLE I. Partition of the data set—number of images for training, validation, and test sets.

	No. of images with μTag	No. of images without μTag	Total no. of images
Training set	330	210	540
Validation set	321	239	560
Test set	410	290	700
Total	1061	739	1800

2.C.1. μTag *enhancement module*

The μTag enhancement module uses a combination of three filters to enhance the μTag and remove the background. Let the filters be denoted by F_1, F_2, and F_3, with filter kernels of $M_1 \times M_1$, $M_2 \times M_2$, and $M_3 \times M_3$, respectively, where $M_1 > M_2 \geq M_3$. The filters can be linear or nonlinear. If linear filters are chosen, the filters can be combined as follows to produce a single band-pass filter before convolution with the image to reduce processing time. The kernels of the filters are designed based on the following considerations. When the kernels are centered at one of the microspheres of a μTag

Input Image

↓

μTag enhancement module

↓

μTag candidate labeling module

↓

μTag segmentation module

↓

Feature analysis module

↓

Sufficient number of μTag candidates labeled?

No

Yes

↓

μTag classification module

↓

Output - detected μTag locations marked on the input image

FIG. 5. Block diagram of the CAD system for detection of μTags.

candidate, the difference of the filters F_1 and F_2 provided an estimate of the background image intensity around the microsphere.

Therefore, M_1 should be large enough to include the background, and M_2 should be of a size that can exclude the μTag microsphere from the background estimation. The filter F_3 is used to smooth or increase the signal intensity of the microspheres, depending on the selection of M_3. The contrast-to-noise ratio (CNR) enhancement can then be accomplished by convolving the combined band-pass filter with the image,

$$F(x,y) = F_3(x,y) - [F_1(x,y) - F_2(x,y)]. \qquad (1)$$

For simplicity, we chose to use uniform filters for F_1, F_2, and F_3 with $M_1 = 13$, $M_2 = 9$, $M_3 = 5$ pixels. The M_1, M_2, and M_3 size selection was based on the physical dimension of the μTag microspheres when imaged on the radiographs. The corresponding sizes of $M_1 = 1.625$ mm, $M_2 = 1.125$ mm, and $M_3 = 0.625$ mm are sufficiently large to cover the background around the μTag microsphere, the μTag microsphere itself, and the brightest central part of the μTag microsphere, respectively. Other filters such as Gaussian filters with appropriate kernels can be used as well.

2.C.2. μTag *candidate labeling module*

The μTag candidate labeling module identifies μTag candidates with an adaptive thresholding method based on histogram analysis, μTag segmentation, feature analysis, and an iterative optimization procedure for determination of the optimal threshold value. At the beginning of the automated optimization procedure, an initial threshold value is determined based on an analysis of the histogram of the CNR-enhanced image obtained from the μTag enhancement module. This threshold is then applied to the entire CNR-enhanced image to identify an initial set of potential μTag microsphere candidates that exceed the threshold. By using the μTag segmentation module (described in Sec. 2.C.3), the shape of every candidate object is refined. A set of features is then extracted from the refined objects and a rule-based classification is performed to reduce FP objects by the feature analysis module (described in Sec. 2.C.4). The optimization procedure stops when the number of remaining candidates falls within a predefined range or after a predefined number of iterations. Otherwise, a new gray-level threshold is determined automatically, which is applied again to the CNR-enhanced image and a new iteration starts. The iterative loop is shown in Fig. 5. Using the training set, we experimentally determined that a range of 25 to 100 objects at this stage will provide high sensitivity of detecting the true μTags and that a maximum of 10 iterations is enough to reach this targeted number of objects. The final set of objects selected by the iterative procedure is labeled as μTag microsphere candidates (TMSs). An average CNR (Ave_CNR) is calculated using the CNRs of the entire set of 25-100 TMSs and used in Sec. 2.C.5.
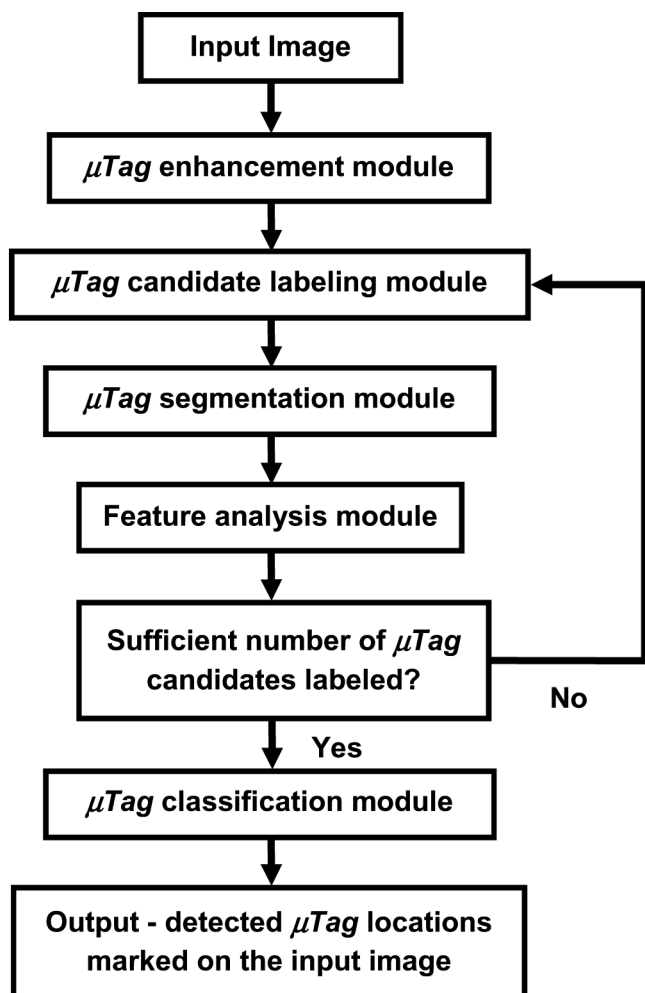
2.C.3. µTag segmentation module

The µTag segmentation module is guided by a cost function determined by the statistical variation of the pixel values in a small region around the TMS. The µTag microspheres have higher CNR values, on average, than other structures in their local area. A locally adaptive gray-level thresholding method is used to refine the candidates. The background mean and noise variations at the TMS location are estimated as the local mean and root-mean-square noise (RMSN) within a square kernel centered at the centroid of the TMS, excluding the pixels that contain the object itself and other TMS objects within the kernel region, on the CNR-enhanced image output from the µTag enhancement module. To automatically segment a refined TMS object, the module starts from the centroid pixel and includes an adjacent pixel, as determined by the eight-connectivity criterion,[12] as a part of the object if its value is larger than the background mean by a predefined multiple k of the RMSN, where k is the CNR threshold. A value of $k = 3$ was obtained experimentally using the training data set. Additional details related to the selection of k are presented in Sec. 4. The segmentation continues until no more adjacent pixels satisfy the conditions. The exact kernel size is not critical except that it should be sufficiently large to give a good estimate of the local background noise variations and was chosen to be 81×81 pixels.

2.C.4. Feature analysis module

The feature analysis module extracts features from the refined TMSs and applies rule-based classification based on the extracted features to differentiate the FP candidates from true microspheres.

2.C.4.a. Feature extraction. The following features[13] were extracted from every segmented TMS.

Area (AR). The area is calculated as the number of pixels segmented for the TMS.

Contrast relative to the neighboring background (C). The contrast is calculated as the peak pixel value within the segmented TMS region above the background mean.

Eccentricity (e). In order to calculate the eccentricity, the moments and the second moments have to be derived first. The moments M_0, M_x, and M_y are calculated as follows:

$$M_0 = \sum_i g_i, \tag{2}$$

$$M_x = \sum_i g_i x_i / M_0, \tag{3}$$

$$M_y = \sum_i g_i y_i / M_0, \tag{4}$$

where g_i is the gray-level value of pixel i. (x_i, y_i) are the co-ordinates of the pixel i. The second moments are calculated as

$$M_{xx} = \sum_i g_i (x_i - M_x)^2 / M_0, \tag{5}$$

$$M_{yy} = \sum_i g_i (y_i - M_y)^2 / M_0, \tag{6}$$

$$M_{xy} = \sum_i g_i (x_i - M_x)(y_i - M_y) / M_0. \tag{7}$$

The summations are over all pixels within the segmented TMS.

The **major axis 2a** and the **minor axis 2b** of the ellipse that characterizes the object shape can be derived from the second moments as

$$2a = \sqrt{2 \left[M_{xx} + M_{yy} + \sqrt{(M_{xx} - M_{yy})^2 + 4M_{xy}^2} \right]}, \tag{8}$$

$$2b = \sqrt{2 \left[M_{xx} + M_{yy} - \sqrt{(M_{xx} - M_{yy})^2 + 4M_{xy}^2} \right]}. \tag{9}$$

The eccentricity of the ellipse is given by

$$e = \frac{\sqrt{a^2 - b^2}}{a}. \tag{10}$$

The eccentricity values are in a range of 0 to 1. For a circle, $e = 0$ ($a = b$), and for a line, $e = 1$ ($a > 0$; $b = 0$).

These features were proven to be effective descriptors for differentiating true microcalcifications from other FP objects on mammograms in our previous study.[13] There are substantial resemblances between the geometric shapes of µTag microspheres and microcalcifications, and therefore, these features are good candidates as descriptors of the TMSs.

2.C.4.b. Rule-based classification. A rule-based classification is used to combine the image features and to determine whether a TMS is a true positive (TP) µTag microsphere or a FP.

Rule 1:

$$\text{TMS} = \begin{cases} \text{keep} & \text{Th}_{A\min} \leq \text{AR} \leq \text{Th}_{A\max} \\ \text{remove} & \text{otherwise} \end{cases}. \tag{11}$$

This rule imposes constraints on the area of the TMS candidate (AR). The minimum ($\text{Th}_{A\min}$) and maximum ($\text{Th}_{A\max}$) area thresholds were estimated based on the physical dimensions and possible variations of the imaged µTag microsphere sizes to be $\text{Th}_{A\min} = 25$ pixels and $\text{Th}_{A\max} = 100$ pixels, respectively. If the area of the candidate is less than 25 pixels, it is excluded because the area is too small to correspond to a µTag microsphere and it is likely to be noise. If the area of the candidate is more than 100 pixels, it is excluded because the area is too large to correspond to a µTag microsphere or two overlapping microspheres and it is likely to be a different large object.

Rule 2:

$$\text{TMS} = \begin{cases} \text{keep} & C \leq K_C * \text{RMSN} \\ \text{remove} & C > K_C * \text{RMSN} \end{cases}. \tag{12}$$

This rule imposes constraint on the contrast of the TMS. The constant K_C was selected as $K_C = 15$ based on the analysis of the µTag microspheres on the training radiographs. If the contrast is higher than 15 times the RMSN, then the TMS is likely a dense surgical tool or other objects and removed.

Rule 3:

$$\text{TMS} = \begin{cases} \text{keep} & 2a \leq \text{Th}_{\text{Ma}} \text{ OR } e < \text{Th}_e \\ \text{remove} & 2a > \text{Th}_{\text{Ma}} \text{ AND } e \geq \text{Th}_e \end{cases}. \tag{13}$$

This rule removes linear objects. The major axis threshold Th_{Ma} was chosen to be 15 pixels. The eccentricity threshold Th_e was selected to be 0.95.

A TMS that satisfies all three rules will be kept as input to the next module.

2.C.5. μTag *classification module*

The μTag classification module consists of a multilayer neural network classifier of the μTag microspheres and a procedure for μTag identification and reduction of false positives.

2.C.5.a. A multilayer neural network classification of the μTags. A convolution neural network (CNN) classifier[14–16] is applied to the TMSs that are kept after rule-based classification to further reduce the FPs. The structure and the parameters of the CNN were determined in our previous studies,[15–17] which were found to be effective for reducing FPs in tasks of detecting compact-shaped objects. The CNN consists of input group of neurons, two convolution hidden layers, and one output neuron. The input group of neurons (16×16 neurons) correspond to two dimensional (2D) receptive field that can cover 16×16 pixel region of interest (ROI) centered at a TMS. The first convolution hidden layer consists of 14 groups of neurons. The second convolution hidden layer consists of ten groups of neurons. All neurons in the hidden layers are fully connected. The kernel sizes of the first group of filters between the input and the first hidden layer are 5×5 pixels, and those of the second group of filters between the first and second hidden layers are 7×7 pixels. A logistic sigmoid function was chosen as the activation function for both the hidden neurons and output neurons. The CNN was trained to have an output node score S_{CNN} between 0 (FP) and 1 (TP). Backpropagation training was used to adjust the connection weights of the CNN by minimizing the sum-of-squares error (SSE) function, which leads to a CNN output S_{CNN} that can be interpreted as a probability[18] of the input ROI containing a true μTag microsphere. A decision threshold Th_{CNN} can be applied to the S_{CNN} and the ROIs with S_{CNN} values above the Th_{CNN} will be kept and ROIs with S_{CNN} values below the Th_{CNN} will be eliminated as FPs,

$$TMS = \begin{cases} keep & S_{CNN} \geq Th_{CNN} \\ remove & S_{CNN} < Th_{CNN} \end{cases}. \qquad (14)$$

However, in order to preserve some TMSs with high CNR, but eliminated by the CNN, rule (14) was modified as follows:

TMS
$$= \begin{cases} keep & S_{CNN} \geq Th_{CNN} \\ keep & S_{CNN} < Th_{CNN} \text{ AND } CNR > K_{SNR} * Ave\_CNR \\ remove & S_{CNN} < Th_{CNN} \text{ AND } CNR \leq K_{SNR} * Ave\_CNR \end{cases},$$
$$(15)$$

where K_{SNR} is a parameter that is used as a second decision threshold. Two pairs of decision thresholds were selected by using the training and validation data sets: one for the high specificity operating point and the second one for the high sensitivity operating point. The decision thresholds for the high specificity operating point are $Th_{CNN} = 0.65$ and $K_{SNR} = 1.6$. The decision thresholds for the high sensitivity operating point are $Th_{CNN} = 0.3$ and $K_{SNR} = 1$. More details on the selection of these thresholds are described in Sec. 3 below.

2.C.5.b. μTag identification and false positive reduction.
2.C.5.b.i. μTag identification. At this stage, the potential μTags are identified by a regional object grouping procedure, which is based on the geometric properties and physical dimensions of the μTag. A region with higher concentration of the TMSs is selected first as a starting region to find the μTags. The growing procedure is based on a dynamic object grouping algorithm, which searches for a new TMS in the neighborhood and updates the μTag centroid after each new candidate is added. A potential TMS is included as a member of the μTag if it is within a preselected distance threshold of 30 pixels from the current μTag centroid. The μTag candidates are attempted to be grown at all potential regions with high concentration of TMSs on the radiograph.

2.C.5.b.ii. False positive reduction. After grouping, groups containing more than one TMSs are considered to be μTag candidates. The remaining TMSs that are not found to be members of any potential μTags are excluded as FPs.

In order to further reduce the FPs due to sutures or other line structures, a ridge feature is designed to estimate the standard deviation of pixel values within a rectangular region between every two TMS members of a μTag candidate. The rectangular region is positioned between the centroids of the TMSs, and the midline of the rectangular region coincides with the line connecting the centroids of the TMSs. The length of the rectangular region is the distance between the centroids of the TMSs and its width is 5 pixels. The width of 5 pixels for the rectangle is selected to cover the main region between the two μTag microspheres. If the TMSs are true μTag microspheres, the standard deviation within the rectangular region will be large because the background pixels between the μTag microspheres have lower pixel values compared to the μTag microsphere pixels. If the TMSs are FPs, the pixel values between the TMSs are more likely to be comparable to the pixel values of the TMSs, which will result in a smaller standard deviation and thus a smaller value of the ridge feature. After exhaustive pairing of the TMSs in a given μTag candidate, the μTag candidate is kept as a TP if the majority of μTag microspheres pairs are considered true microspheres.

2.D. Evaluation methods

The detection performance of the μTag CAD system was compared to the true μTag locations marked by the radiologist. The μTag CAD system output is provided as the smallest rectangular bounding boxes enclosing the detected μTags. The scoring program automatically calculated the intersection of the areas enclosed by a CAD detected rectangle and the gold standard rectangle. If the ratio of the intersection of the rectangles is greater than 10% of the gold standard rectangle area, the μTag is considered to be detected. If a gold standard rectangle intersects with more than one CAD
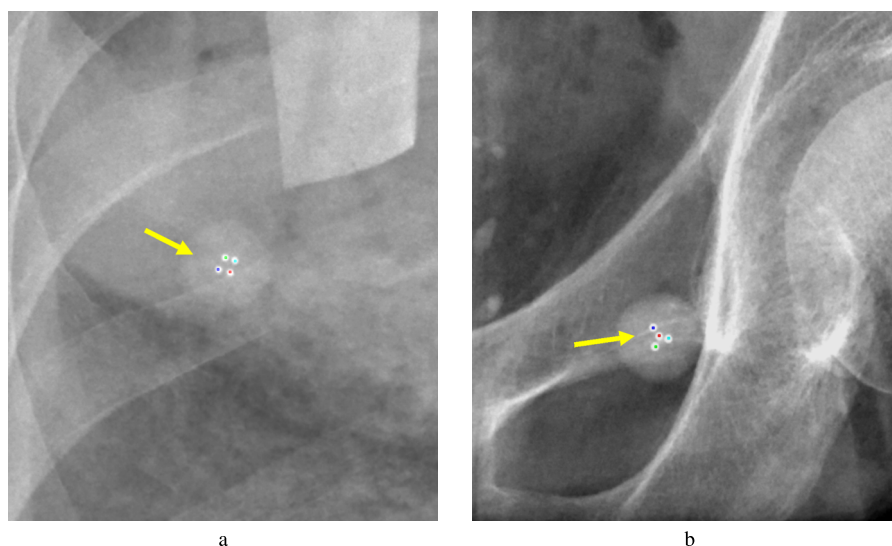
FIG. 6. CAD detection of μTags on the radiographs (marked by the arrow): the μTag was placed on the ribs (a) and on the pelvis (b). The four microspheres are distinctly visible. CAD detections are shown as dots overlaying the microspheres.

detected rectangle, only one TP finding is recorded. The μTag detection performance of the CAD system was evaluated by the free response receiver operating characteristic (FROC) analysis. The area under the curve (AUC) of the FROC curves was estimated by using the method by Bunch *et al.*[19,20] The detection sensitivity was calculated relative to the number of cadaver images with μTag. The average FP detection rate was determined by counting the CAD detections in the cadaver images without μTag (the negative cases) relative to the number of cadaver images without μTag. For comparison purpose, the average FP detection rate was also determined by counting the CAD detections in both the cadaver images with and without μTag relative to the total number of cadaver images with and without μTag. The detection performance of the μTag CAD system was analyzed for the training and the validation sets during the development of the CAD system. After all the parameters were selected and frozen, the CAD system was applied to the independent test set to predict its performance in unknown cases.

3. RESULTS

Examples of CAD detection of the μTags with different orientations and locations on the radiographs and with different levels of visibility are shown in Figs. 6 and 8. The examples shown in these figures and in Figs. 3 and 4 were all detected successfully. The FROC curves for the training and the validation sets are plotted in Fig. 7.

The AUC of the FROC curves for the training and the validation sets was 0.962 and 0.961, respectively (Table II). To simulate the situations that the CAD system may be operated at a high sensitivity or at a high specificity condition, we chose two operating points along the FROC curves. The decision threshold ($Th_{CNN} = 0.3$, $K_{SNR} = 1$) for the high sensitivity operating point for the CAD system was selected based on the training set as the point with the highest sensitivity, i.e., the sensitivity of 96.1% at 0.81 FPs/image (Fig. 7). Using the

selected decision threshold, the corresponding high sensitivity operating point on the validation set was found to have 93.5% sensitivity at 0.44 FPs/image. The decision threshold for the high specificity operating point was selected based on both the training and validation sets as the point with very low FPs/image rate and the highest possible sensitivity. The decision threshold ($Th_{CNN} = 0.65$, $K_{SNR} = 1.6$) that yielded a sensitivity of 81.5%, at 0.014 FPs/image for the training set, and a sensitivity of 80.1% at 0.004 FPs/image for the validation set was chosen (Fig. 7).

Once the two sets of decision thresholds were fixed, the CAD system was applied to the test set. In order to preserve the
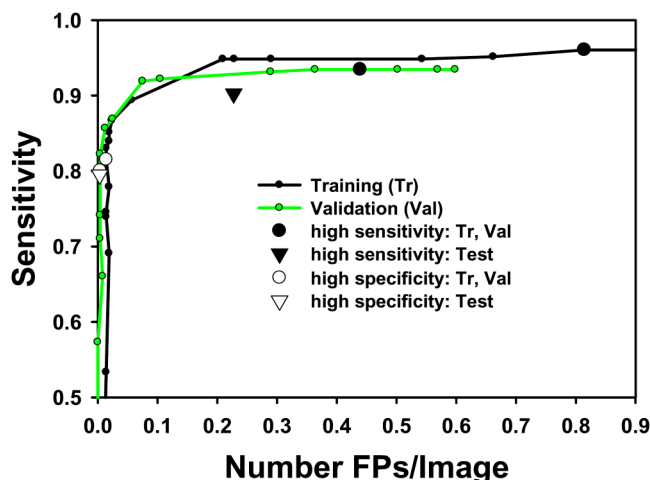

FIG. 7. CAD detection results for the training, validation, and test sets. The FROC curves of the CAD detection results for the training and validation sets are shown. The high sensitivity operating points for the training and validation sets are marked with large black circles. The corresponding high sensitivity operating point for the test set is marked with a black triangle. The high specificity operating points for the training and validation sets are marked with large white circles. The corresponding high specificity operating point for the test set is marked with a white triangle. The FP rates for these FROC curves were determined using the cadaver images without μTag.
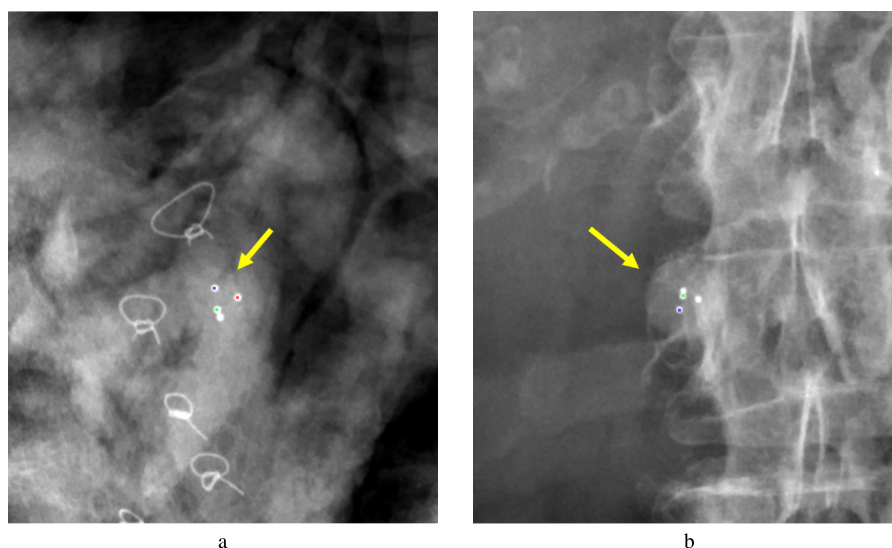
FIG. 8. CAD detection of μTags on the radiographs (marked by the arrow): the μTag was placed on soft tissue near sutures (a) and on the spine (b). Two of the μTag microspheres overlapped. The bone overlapped with one of the microspheres and made it difficult to detect the μTag (b).

independence of the test set, only two runs were performed; one used the decision threshold for the high sensitivity operating point and the other the decision threshold for the high specificity operating point. The high sensitivity operating point of the CAD system for the test set achieved 90.2% sensitivity with 0.23 FPs/image and the high specificity operating point achieved 79.5% sensitivity with 0.003 FPs/image (Fig. 7). The detection results are summarized in Table II.

The FP rates estimated using both the images with and without μTag (Table III) were slightly lower than the FP rates estimated using the images without μTag alone (Table II).

4. DISCUSSION

The CAD system for the μTags achieved reasonable sensitivity and specificity in the three data sets. The CAD system was able to detect the μTags overlapped with different anatomical backgrounds and locations on the radiographs. Examples of CAD detections are shown in Figs. 6 and 8. In a number of cases, the μTag happened to be rotated in a way that all four microspheres were distinctly visible on the radiograph. In such cases, the μTag was detected successfully regardless of the type of anatomical background such as ribs [Fig. 6(a)], pelvis [Fig. 6(b)], spine [Fig. 3(a)], or other objects [Fig. 4(b)]

with which the μTag was superimposed. However, in most cases, two of the μTag microspheres overlapped to certain degree due to different orientations of the μTag. The overlap could range from minimal [Figs. 8(a) and 8(b)] to a worst-case scenario where the microspheres almost completely overlap [Fig. 3(b)]. Nevertheless, the design of the μTag ensures that three microspheres will be well distinguishable on different anatomical backgrounds and the CAD system was able to detect the μTag in most of the cases.

Even more challenging is the situation presented in Fig. 8(b), where the bone obscured some of the microspheres and made it difficult to detect. The μTags were located on the spine [Fig. 8(b)] and two of the microspheres were partially overlapping. However, the CAD system was able to detect such μTags in most cases.

The μTags missed by the CAD system [false negatives (FNs)] were caused by overlying structures, low pixel intensity of the radiograph, and μTags located along the edge of the image and only partially visible. An example of a missed μTag is shown in Fig. 9. The μTag overlapped with a suture that obscured the microspheres. In addition, two of the microspheres perfectly overlapped, which also contributed to the subtle appearance of the μTag.

An example of FP detection is presented in Fig. 10. The turning points of the suture mimicked closely the microspheres,

TABLE II. Performance of the CAD system on the training, validation, and test sets for the high sensitivity and high specificity operating points. The FP rates were determined using the images without μTag.

	High sensitivity operating point		High specificity operating point		
	Sensitivity (%)	Specificity (FPs/image)	Sensitivity (%)	Specificity (FPs/image)	AUC
Training set	96.1	0.81	81.5	0.014	0.962
Validation set	93.5	0.44	80.1	0.004	0.961
Test set	90.2	0.23	79.5	0.003	—

TABLE III. Performance of the CAD system on the training, validation, and test sets for the high sensitivity and high specificity operating points. The FP rates were determined using both the cadaver images with and without μTag.

	High sensitivity operating point		High specificity operating point	
	Sensitivity (%)	Specificity (FPs/image)	Sensitivity (%)	Specificity (FPs/image)
Training set	96.1	0.66	81.5	0.011
Validation set	93.5	0.44	80.1	0.009
Test set	90.2	0.20	79.5	0.001

which confused the CAD system. However, most of this type of FPs was handled successfully by the multistage FP reduction procedure.

The CNR threshold k used for TMS object segmentation was determined experimentally, by using the training set, and it was chosen as the threshold that provides a high sensitivity while maintaining a low FP rate ($k = 3$) (Table IV).

The two operating points of the CAD system are intended for two situations in clinical use. The high specificity operating point can be used by the surgeon while the high sensitivity operating point can be used by the radiologist. The high specificity CAD system, such as 79.5% sensitivity with 0.003 FPs/image (achieved on the test set), can be used for intra or perioperative detection of μTags. This will allow the surgeon to close the patient more than 99.7% of the time, for whom no μTag was detected, and send the radiographs to the radiologist for confirmation as in current clinical practice. Therefore, the CAD system can function as a first reader in

the operating room, a first for CAD, with the radiologist as backup. The high sensitivity CAD system, such as 90.2% sensitivity with 0.23 FPs/image, helps the radiologist identify the majority of μTags, if any. A radiologist can easily rule out FPs because of the easily recognizable μTag shape.

One limitation of the study is the size of the data set. Even though the data set is relatively large, adding more new cases to the training set that cover a wide range of clinical situations and patient population will further improve the training and thus the performance of the CAD system, especially in the difficult situations such as μTag overlapping with dense tissues or bony structures or the presence of structures with multiple closely spaced dense spots such as sutures that can mimic the μTag. Another limitation is that the μTags were placed outside the cadavers so that the contrast of the microspheres may be somewhat different from those actually embedded inside the human body. However, since the RFO cases are rare, it will be difficult and cost-prohibitive to collect a large enough sample with RFOs inside patients or cadavers
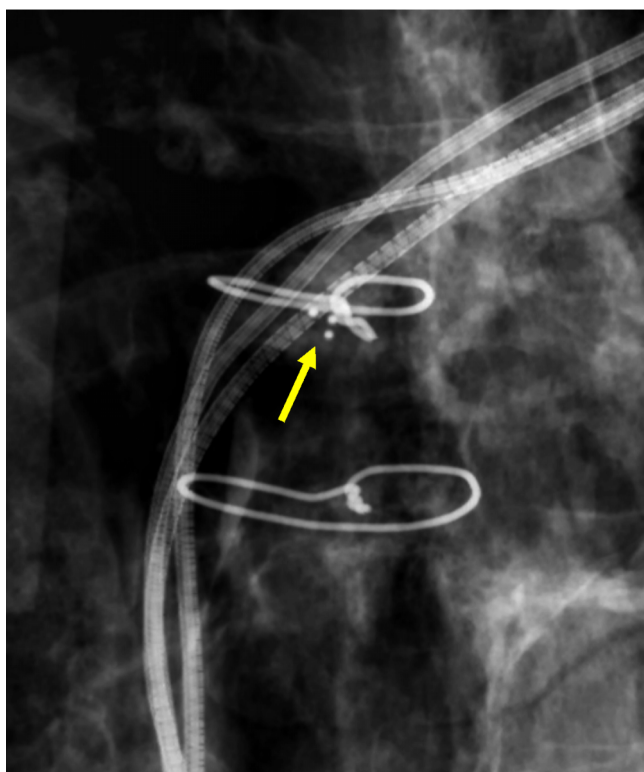

FIG. 9. μTag missed by the CAD system. The μTag is marked by the arrow. The μTag was placed on a suture which obscured some of the microspheres. Two of the microspheres perfectly overlapped and seen as one.
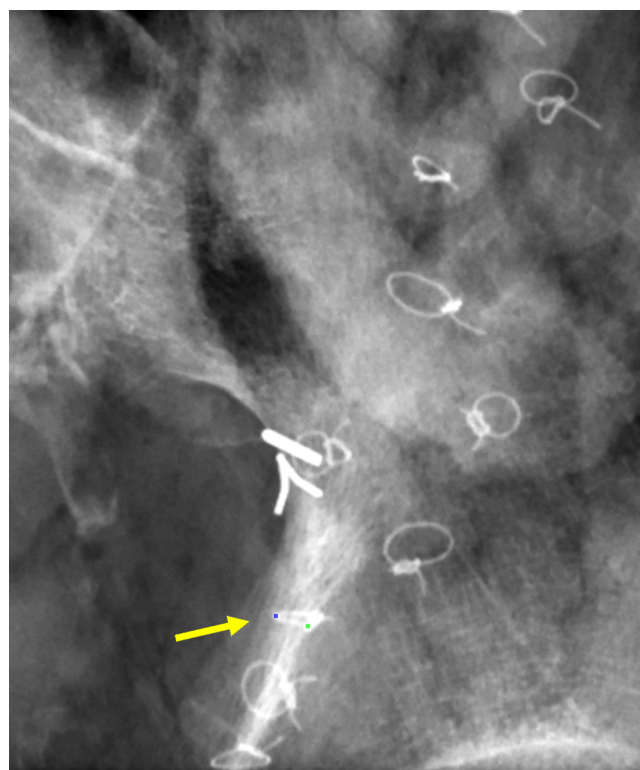

FIG. 10. False positive CAD detection. The CAD system detected the turning points of the suture (marked by the arrow) which mimic microspheres.

TABLE IV. Effect of the CNR threshold (k) on the performance of the CAD system using the training data set.

k	1	2	2.5	3	3.5	4	5
Sensitivity (%)	93.9	95.8	96.4	96.1	94.8	93.9	87.0
FPs/image	0.51	1.18	1.35	0.81	0.84	0.50	0.10

in a reasonable time. We believe that our approach would be a practical and efficient way to collect data for development of such a CAD system.

In our future work, we will continue to train the CAD system with a larger and more general data set to represent all possible variations of the patient population. We will also optimize the design of the μTag to make it more distinguishable from tissues and structures in the human body. The CAD system may also need to be updated to match the design of the μTag such as designing new enhancement filters, new morphological features, and different types of classifiers for FP reduction. Furthermore, the current CAD system focuses on detection of sponges or towels labeled with μTags. We will continue to develop different CAD systems for detection of other types of RFOs such as needles or surgical tools.

5. CONCLUSIONS

To the best of our knowledge, this is the first time a 3D μTag is used to produce a recognizable, substantially similar 2D projection on radiographs regardless of orientation in space. It is the first time a CAD system is used to search for man-made objects over anatomic background. The use of a lock-and-key, device-software combination mechanism is novel. The *a-priori* knowledge of the μTag design specifications, such as contrast, size, shape, and spacing of the high contrast features, and their spatial grouping allows for narrower detection criteria that will confer increased detection accuracy. This offers the opportunity to have the CAD function as a first reader.

The potential to prevent the vast majority RFOs with a single system in a cost-effective manner has never existed before. The possibility to eliminate the count altogether and save operating room time without compromising patient safety would be a major step forward in clinical practice as time is one of the most valuable resources to a hospital. A reader study is under way. We are imaging healthy volunteers toward a noncadaveric data set for further validation of the CAD system. An operating room pilot is being designed to test the complete RFO prevention technology.

ACKNOWLEDGMENTS

[a] Author to whom correspondence should be addressed. Electronic mail: lhadjisk@umich.edu; Telephone: (734) 647-7428; Fax: (734) 615-5513.

[1] V. C. Gibbs, F. D. Coakley, and H. D. Reines, "Preventable errors in the operating room: Retained foreign bodies after surgery–Part I," Curr. Probl. Surg. **44**, 281–337 (2007).

[2] J. Kennedy, "Prevention of retained foreign objects," http://www.wshmma.org/events/2010/Prevention_of_Retained_Foreign_Objects.pdf (2009).

[3] M. Krolikowski and M. Raza, "AutoCount–Automated surgical tool counting system," http://www.ideatoproduct.org/global/images/stories/Presentations_09_temp/SaturdayTES/AutoCountIIT.pdf (2009).

[4] A. A. Gawande, D. M. Studdert, E. J. Orav, T. A. Brennan, and M. J. Zinner, "Risk factors for retained instruments and sponges after surgery," N. Engl. J. Med. **348**, 229–235 (2003).

[5] R. R. Cima, A. Kollengode, J. Garnatz, A. Storsveen, C. Weisbrod, and C. Deschamps, "Incidence and characteristics of potential and actual retained foreign object events in surgical patients," J. Am. Coll. Surg. **207**, 80–87 (2008).

[6] L. A. Dossett, R. S. Dittus, T. Speroff, A. K. May, and B. A. Cotton, "Cost-effectiveness of routine radiographs after emergent open cavity operations," Surgery **144**, 317–321 (2008).

[7] R. R. Cima, A. Kollengode, J. Clark, S. Pool, C. Weisbrod, G. J. Amstutz, and C. Deschamps, "Using a data-matrix-coded sponge counting system across a surgical practice: Impact after 18 months," Jt. Comm. J. Qual. Patient Saf./Jt. Comm. Resour. **37**, 51–58 (2011).

[8] W. Rappaport and K. Haynes, "The retained surgical sponge following intraabdominal surgery–A continuing problem," Arch. Surg. **125**, 405–407 (1990).

[9] P. R. Lauwers and R. H. Van Hee, "Intraperitoneal gossypibomas: The need to count sponges," World J. Surg. **24**, 521–527 (2000).

[10] Kalyspo LLC, "The solution to retained foreign objects," www.kalyspo.com (2014), retrieved 12/22/2014.

[11] T. C. Marentis, L. Hadjiiski, L. Rondon, A. R. Chaudhury, N. Chronis, and H. Chan, "Surgical retained foreign object (RFO) prevention by computer aided detection (CAD)," Proc. SPIE **9035**, 903529 (2014).

[12] W. K. Pratt, *Digital Imaging Processing*, 2 ed. (Wiley, New York, NY, 1991).

[13] H.-P. Chan, B. Sahiner, K. L. Lam, N. Petrick, M. A. Helvie, M. M. Goodsitt, and D. D. Adler, "Computerized analysis of mammographic microcalcifications in morphological and texture feature space," Med. Phys. **25**, 2007–2019 (1998).

[14] H.-P. Chan, S. C. B. Lo, B. Sahiner, K. L. Lam, and M. A. Helvie, "Computer-aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network," Med. Phys. **22**, 1555–1567 (1995).

[15] M. N. Gurcan, H.-P. Chan, B. Sahiner, L. Hadjiiski, N. Petrick, and M. A. Helvie, "Optimal neural network architecture selection: Improvement in computerized detection of microcalcifications," Acad. Radiol. **9**, 420–429 (2002).

[16] J. Ge, B. Sahiner, L. M. Hadjiiski, H.-P. Chan, J. Wei, M. A. Helvie, and C. Zhou, "Computer aided detection of clusters of microcalcifications on full field digital mammograms," Med. Phys. **33**, 2975–2988 (2006).

[17] M. N. Gurcan, B. Sahiner, H.-P. Chan, L. M. Hadjiiski, and N. Petrick, "Selection of an optimal neural network architecture for computer-aided detection of microcalcifications–Comparison of automated optimization techniques," Med. Phys. **28**, 1937–1948 (2001).

[18] C. M. Bishop, *Neural Networks for Pattern Recognition* (Clarendon, Oxford, 1995).

[19] P. C. Bunch, J. F. Hamilton, G. K. Sanderson, and A. H. Simmons, "A free response approach to the measurement and characterization of radiographic observer performance," Proc. SPIE **127**, 124–135 (1977).

[20] P. C. Bunch, J. F. Hamilton, G. K. Sanderson, and A. H. Simmons, "A free response approach to the measurement and characterization of radiographic observer performance," J. Appl. Photogr. Eng. **4**, 166–171 (1978).