Received Date: 13-May-2016

Revised Date: 14-Oct-2016

Accepted Date: 09-Nov-2016

Article Type: Research Article

Computer-Aided Detection of Retained Surgical Needles from Postoperative Radiographs

Aunnasha Sengupta[1], Lubomir Hadjiiski[1], Heang-Ping Chan[1], Kenny Cha[1], Nikolaos Chronis[2], Theodore C. Marentis[1]

[1]Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109

[2]Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan 48109

Running Title: Computer-Aided Detection of Retained Surgical Needles from Postoperative Radiographs

Correspondence:

Aunnasha Sengupta
Department of Radiology
University of Michigan
1500 E. Medical Center Drive
MIB C476
Ann Arbor, MI 48109-5842
Telephone: (734) 647-7429
Fax: (734) 615-5513
E-mail:aunnie@umich.edu

Abstract

Purpose: Foreign objects, such as surgical sponges, needles, sutures and other surgical instruments, retained in the patient's body can have dire consequences in terms of patient mortality as well as legal and financial penalties. We propose computer-aided detection (CAD) on postoperative radiographs as a potential solution to reduce the chance of retained foreign objects (RFOs) after surgery, thus alleviating one of the major concerns for patient safety in the operation room. A CAD system can function as a second pair of eyes or a pre-screener for the surgeon and radiologist, depending on the CAD system design and the workflow. In this work, we focus on the detection of surgical needles on postoperative radiographs. As needles are frequently observed RFOs, a CAD system that can offer high sensitivity and specificity towards detecting surgical needles will be useful.

Methods: Our CAD system incorporates techniques such as image segmentation, image enhancement, feature analysis and curve fitting to detect surgical needles on radiographs. A data set consisted of 108 cadaver images with a total of 116 needles and 100 cadaver "normal" images without needles were acquired with a portable digital x-ray system. A reference standard was obtained by marking the needle locations using an in-house developed graphical user interface. The 108 cadaver images with the needles were partitioned into training set containing 53 cadaver images with 59 needles and a test set containing 55 cadaver images with 57 needles. All of the 100 cadaver normal images were reserved as a part of the test set and used to estimate the false positive detection rate. Two operating points were chosen from the CAD system such that it can be operated in two modes, one with higher specificity (mode I) and the other with higher sensitivity (mode II).

Results: For the training set, the CAD system with the rule-based classifier achieved a sensitivity of 74.6% with 0.15 false positives per image (FPs/image) in mode I and a sensitivity of 89.8 % with 0.36 FPs/image in mode II. For the test set, the CAD system achieved a sensitivity of 77.2% with 0.26 FPs/image in mode I and a sensitivity of 84.2% with 0.6 FPs/image in mode II. For comparison, the CAD system with the neural network classifier achieved a sensitivity of 74.6% with 0.08 FPs/image in mode I and a sensitivity of 88.1 % with 0.28 FPs/image in mode II for the training set, and a sensitivity of 75.4% with 0.23 FPs/image in mode I and a sensitivity of 86.0% with 0.57 FPs/image in mode II for the test set.

Conclusion: A novel CAD system has been developed for automated detection of needles inadvertently left behind in a patient's body from post-surgery radiographs. The pilot system

offers reasonable performance in both the high sensitivity and high specificity modes. This preliminary study shows the promise of CAD as a low-cost and efficient aid for reducing retained surgical needles in patients.

Keywords: computer aided detection (CAD), retained surgical needles, radiographs

1. INTRODUCTION

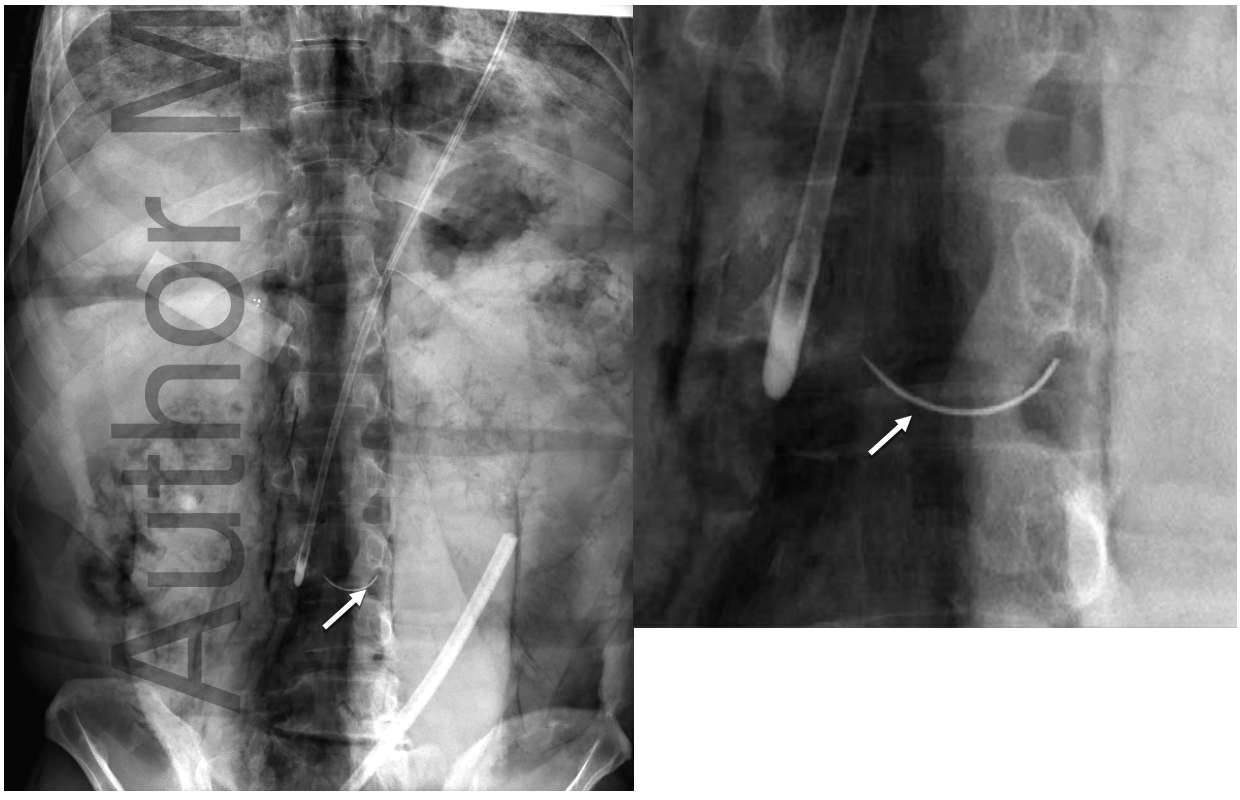
Foreign objects, such as surgical sponges, surgical needles, sutures and other surgical instruments, unintentionally left behind in a patient's body after surgery have rare occurrences but can lead to serious medical problems as well as legal and financial penalties for the responsible institution. Literature suggests that a retained foreign object (RFO) is likely to take place at least once in every 7000 surgeries[1], resulting in health concerns such as sepsis, fistula, visceral perforation or even death[2]. Fortunately, these ominous effects can be largely prevented.

Currently, the typical protocols to prevent such situations, involve manually counting the surgical instruments (surgical sponges, surgical needles, sutures, etc.) before and after the procedure to check for discrepancies, and examining post-operative x-ray images[1]. However, both approaches are heavily dependent on the accuracy of nurses, surgeons, and/or radiologists in counting or identifying RFOs from radiographs which can falter in demanding situations and make such protocols highly susceptible to human error. Therefore, there is a need for better methods to address these issues. An automated system that can detect such RFOs may be a viable solution.

We have previously addressed the RFO problem with two complementary technologies: a three-dimensional (3D) gossypiboma microtag (μTag) that is attached to and thus improves the visibility of sponges on radiographs and a computer-aided detection (CAD) system to detect the μTag[3]. The CAD system achieved very high sensitivity and specificity but the image analysis methods were specifically designed to detect microtag-sponges. In this study, we focus on designing a new CAD system for detection of needles on post-surgical radiographs, which is another important type of RFOs incorrectly counted most of the time[1], making it a pressing need to develop a more reliable method to complement the currently practiced protocols for needle accounting.

Several studies were performed previously to develop methods for the automatic detection and tracking of endovascular devices such as catheters and stent-grafts[4-7]. Although these devices are line-like objects, the techniques used for detecting such devices cannot be applied to the automatic detection of retained needles in radiographs, owing to the particular (elliptical) shape of the surgical needles.

There are two other previous studies[8, 9] on computer-aided detection (CAD) of retained surgical items like needles and sponges. While in one of the studies[8], the CAD system was only used for the detection of sponges, in the other study[9] a pattern recognition algorithm was implemented using a library of images of a surgical needle in various orientations. However, the test images used for evaluating this algorithm did not seem to include any anatomical noise, which is not the case in clinically obtained radiographic images. Additionally, the dataset used for the study was not partitioned into training and test sets, thus the evaluation result might be optimistically biased. In our current study, we proposed a different feature-extraction based approach, as it can be more versatile compared to the pattern recognition method used in the previous study[9]. We also used a training set and an independent test set of clinically relevant radiographs containing needles for the design and evaluation of the CAD system.



(a) (b)

Fig. 1. (a) A typical radiograph of a cadaver in our database with a needle. (b) Zoomed-in view of the radiograph from (a) showing the needle next to a tube. White arrows identify the needle location.

2. METHODS AND MATERIALS

To develop the CAD system for detecting surgical needles in radiographs, we first generated a data set of radiographs with needles commonly used in surgery. The data set was divided into a training set and a test set. We then designed image analysis methods to detect the needles and reduce false positives. The CAD system was trained with the training set and the performance was validated with the test set.

2.A. Data set

2.A.1 Surgical needles

Surgical needle sizes can vary in size between 6mm and 65mm depending on the procedures and the type of the tissue they are used for[10]. Of these, the two most commonly used needles were used to generate a set of radiographic images for this preliminary study. They are shown in Fig. 2 with their respective sizes summarized in Table 1.

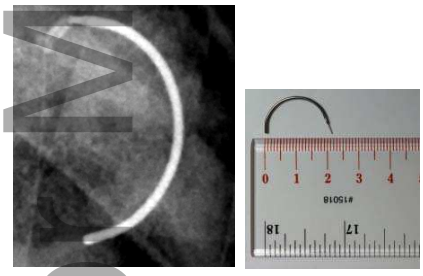
Table 1. Surgical Needle Types

Surgical Needle Type	Size(mm)
Type 1	22
Type 2	24

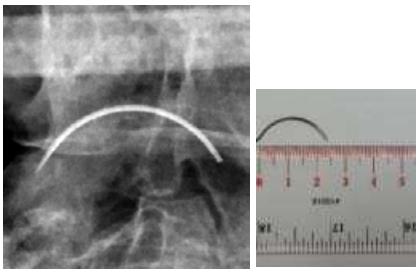
2.A.2 Data Sets

Because RFOs are rare events, we could not collect a large enough data set of post-surgery radiographs with needles from patient files within a reasonable time. With the approval of the University of Michigan Anatomical Donations Program, images of cadavers, with needles placed on top of them, were taken to obtain radiographs with simulated RFOs in the patient body. Placing the needles on the side farthest from the detector would result in the greatest geometric blurring and scatter degradation of the needles on the radiograph. This simulated the worst-case scenario, from an imaging point of view, so that we did not need to place needles inside the cadavers. All radiographs were acquired from cadavers with needles placed on the chest and abdomen regions because literature[2] suggests that the majority of needles were lost during surgery performed in the thoracoabdominal cavity. Furthermore, the

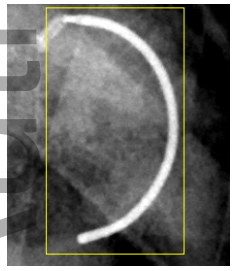
presence of a large number of anatomical structures in the abdominal region result in large contrast variations in the abdomen, which makes it a greater challenge for the CAD system to detect lost needles there than in other parts of the body. It is important to obtain an adequate sampling of surgical needles over a variety of anatomic backgrounds and different locations of the abdomen, as well as random orientation of the needles in space. Therefore, the imaging was performed as follows: the abdomen was divided into subregions in a 4 x 5 grid. A needle was stuck inside a piece of irregular-shaped foam that was rolled on the cadaver surface so that it would stop at a random orientation for each image. The cadaver was tilted and/or rotated relative to the x-ray incident beam to increase the variations of the projected anatomical background. 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 needles and other foreign objects at the estimated incidence rates should be placed in each grid cell. A Shimadzu portable x-ray machine coupled with an 8-megapixel flat-panel Cannon detector was used to image the cadavers with needles, sponges, tubes, sutures and other surgical instruments placed over them. The pixel pitch of the detector was 125×125 μm and the output digital radiographs have 12-bit (0 to 4095) gray levels.



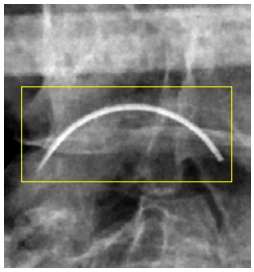
(a)



(b)



(c)



(d)

Fig. 2. Two most commonly used types of surgical needles used for the creation of the data set and their corresponding images on radiographic background (a) Needle Type 1, (b) Needle Type 2, (c) and (d) radiographic images of Type 1 and Type 2 needles, respectively, manually marked by radiologists with a bounding box, to be used as reference standard for evaluation of CAD performance.

One hundred and eight radiographs of 19 different cadavers with a total of 116 needles placed at different locations on their surface were acquired. Of the 108 cadaver radiographs, 53 radiographs with 59 needles were selected randomly for a training set and the remaining 55 radiographs with 57 needles were used for the test set. One hundred radiographs of cadavers without needles were also acquired and all were used as a part of the test set for the estimation of the false positive rate of the CAD system.

The location of the needles in each radiograph was manually marked by a radiologist with a bounding box, as shown in Fig. 2(c) and 2(d), using an in-house developed graphical user interface. The needles' true locations were used as a reference standard for the evaluation of the CAD system detection accuracy for both the training and test sets.

2.B. CAD system design

The CAD system is designed to identify the location of a needle on the radiograph. The flowchart of the needle detection algorithm is shown in Fig 3. The parameters (the thresholds for the various decision rules) were selected empirically to achieve the best performance on the training set. Similarly, the weights of the feature classifiers were estimated with the training samples alone. The CAD system consists of the following image processing steps:

2.B.1 Image Enhancement

The purpose of image enhancement is to improve the contrast for the object of interest (in this case, needle) relative to its background. For the current application, it is achieved by a combination of three linear boxcar filters[3], namely, F_1, F_2 and F_3:

$$F(x,y) = F_3(x,y) - \{F_1(x,y) - F_2(x,y)\} \quad (1)$$

With proper selection of these boxcar filters, the difference between F_1 and F_2 can estimate the background intensity and F_3 can enhance the needle. The resulting filter, $F(x,y)$, is a band-pass filter that enhances the needle while removing the background

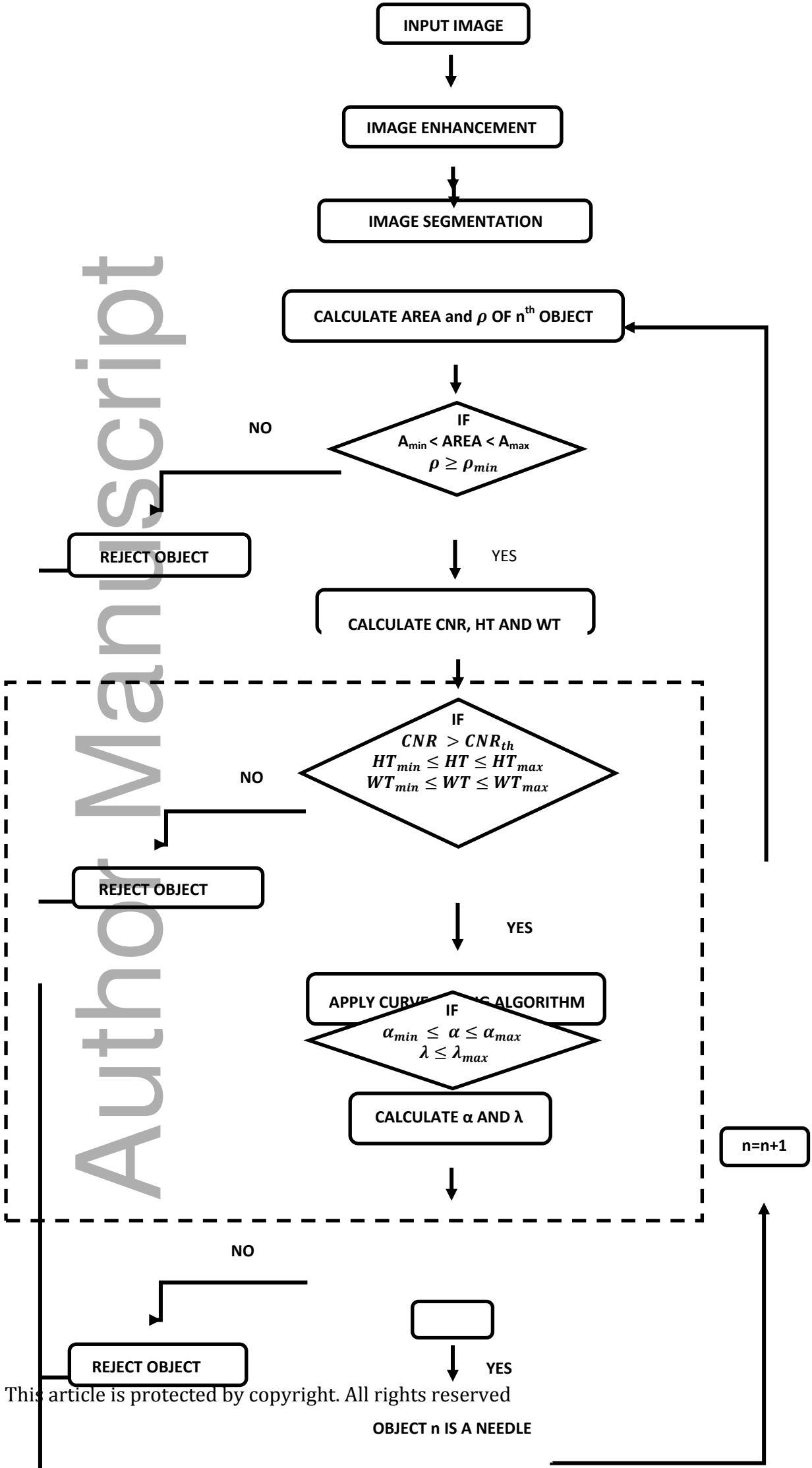
structures, thereby enhancing the contrast-to-noise ratio (CNR) of the needles. The relation between the kernel sizes of the filter, M_1, M_2 and M_3, is as follows: $M_1 > M_2 \geq M_3$. For this needle detection system, the filters were experimentally chosen by using the training set to be boxcar filters with kernel sizes: M_1 = 13 pixels (1.625 mm), M_2 = 9 pixels (1.125 mm) and M_3 = 5 pixels (0.625 mm).

After band-pass filtering, the mean gray level of the filtered image is 0. A constant gray level of 2048 was added to all pixels of the image to avoid cut-off of pixel values below 0.

2.B.2 Image Segmentation

Segmentation allows distinction between the objects of interest and the background. Global thresholding and region growing techniques were implemented to segment the objects of interest, which are the needle candidates in this study. The identified candidates can then be analyzed in the subsequent processing steps.

Gray level thresholding is first carried out to divide the filtered image into a "background" and a "foreground"[11]. By analysis of the histograms of the background-corrected images in the training set, we found that the majority of the needle pixels had gray levels greater than 135 above the mean of the filtered image. We therefore chose a global gray level threshold Th_s of 2183 (=2048+135) for pre-screening of the needle candidates, i.e., if a pixel has a value above Th_s, it is considered a candidate (foreground) pixel. The background pixels are assigned a gray level of 0, while the foreground pixels are set to 4095. The foreground pixels are then subjected to region growing.



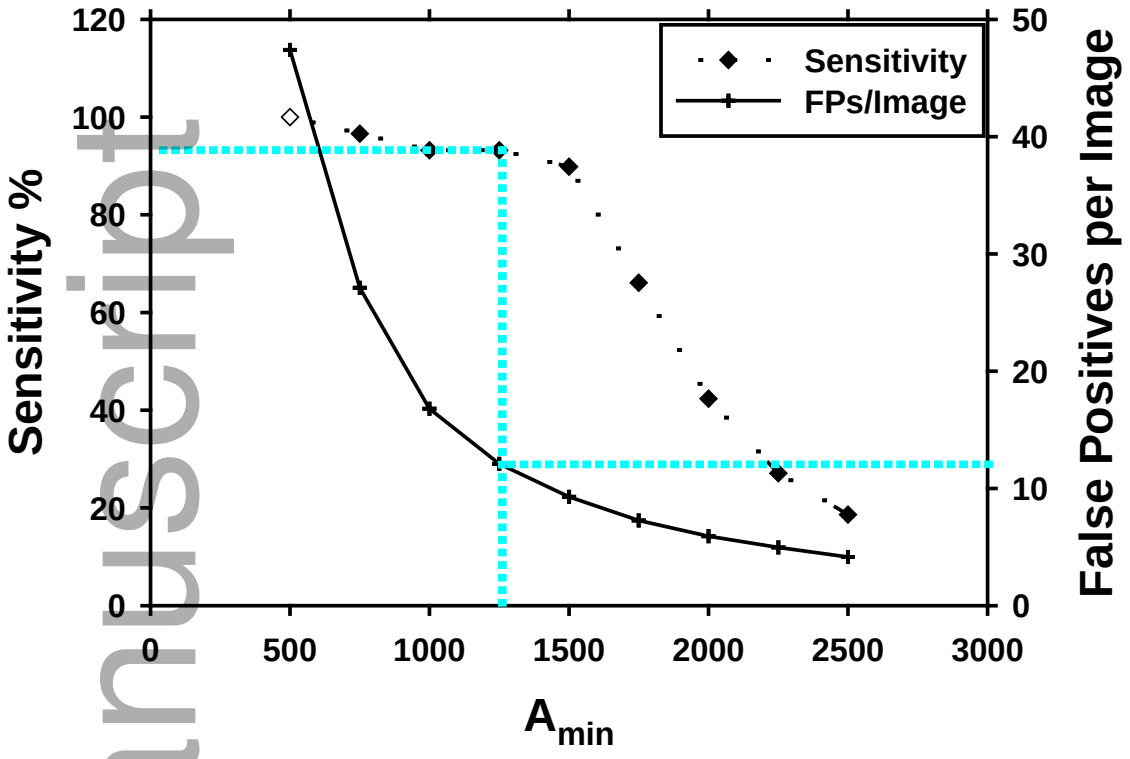


Fig. 4. Dependence of sensitivity and false positive rate (FPs/Image) on the minimum Area value A_{min} for objects at the pre-screening stage in the training set. The dashed vertical line on the plot indicates the chosen decision value for A_{min}.

The region growing technique[11] used here is initialized with a seed pixel. The algorithm determines if a neighboring pixel belongs to the same object by 8-connectivity and whether the pixel gray level satisfies a given criterion, i.e. having a gray level of 4095 in this case. After the region growing step, a number of needle candidate objects are obtained. Then the area (Area) of every candidate object is determined as the total number of connected pixels within the object. Because of the presence of higher contrast structures in human body and noisy background, many false positives (FPs) are included in the pool of needle candidates. Some of the FPs can be excluded by their size as follows. Based on the training set, we estimated the possible maximum and minimum area of the needles on the radiographs, and set the upper and lower bounds of Area to be A_{max} and A_{min}, respectively, on the object area. The objects with the area within the range A_{min} <Area < A_{max} are kept as needle candidates and the rest are removed as false positives. A_{max} is set at 6500 pixels (101.56 mm^2) to reject any object with an unusually large area, as compared to a needle. To

185 determine the value for A_{min}, we analyzed the tradeoffs between the sensitivity and the number of FPs per image, i.e., the FP rate, in our training set for a range of Area as shown in Fig. 4. It can be seen that the FPs increased sharply at small A_{min} values and the sensitivity fell rapidly at large A_{min} values. The value of A_{min} = 1300 pixels (20.31 mm^2) where the sensitivity was relatively stable at over 90% with a moderate FP rate of about 12 FPs/image
190 was selected as a compromise for this pre-screening stage. Further FP reduction is performed in the following step.

2.B.3. Feature Analysis

In order to further decide whether a particular candidate is a true needle, several
195 features such as object density, CNR and geometric dimensions are calculated and analyzed.

Object Density (ρ) is defined as follows:

$$\rho = \frac{\sum_{n=0}^{N_O} i(n)}{N_O} \qquad if\ i(n) > T_{cnr} \tag{2}$$

200 where *i(n)* is the gray level of pixel *n* in the CNR-enhanced image (see Section 2.B.1) and N_O is the total number of pixels in the object. The selected T_{cnr} threshold value of 2185 is very close to the gray level threshold Th_s of 2183 (Section 2.B.2) used for object segmentation, but slightly adjusted to better distinguish the object pixels from the background pixels.

If a selected candidate has a density ρ between a chosen range $\rho \geq \rho_{min}$, it is
205 considered a needle. ρ_{min} is determined by analyzing the trade-offs between sensitivity and the FP rate for a range of ρ_{min} values, as shown in Fig. 5. At ρ_{min} = 2250, we can reduce the FP rate from the previous stage (12.1 FPs/image) to 7.9 FPs/image, while keeping the sensitivity at over 90%. Note that the object density is calculated by using the enhanced image from Section 2.B.1.

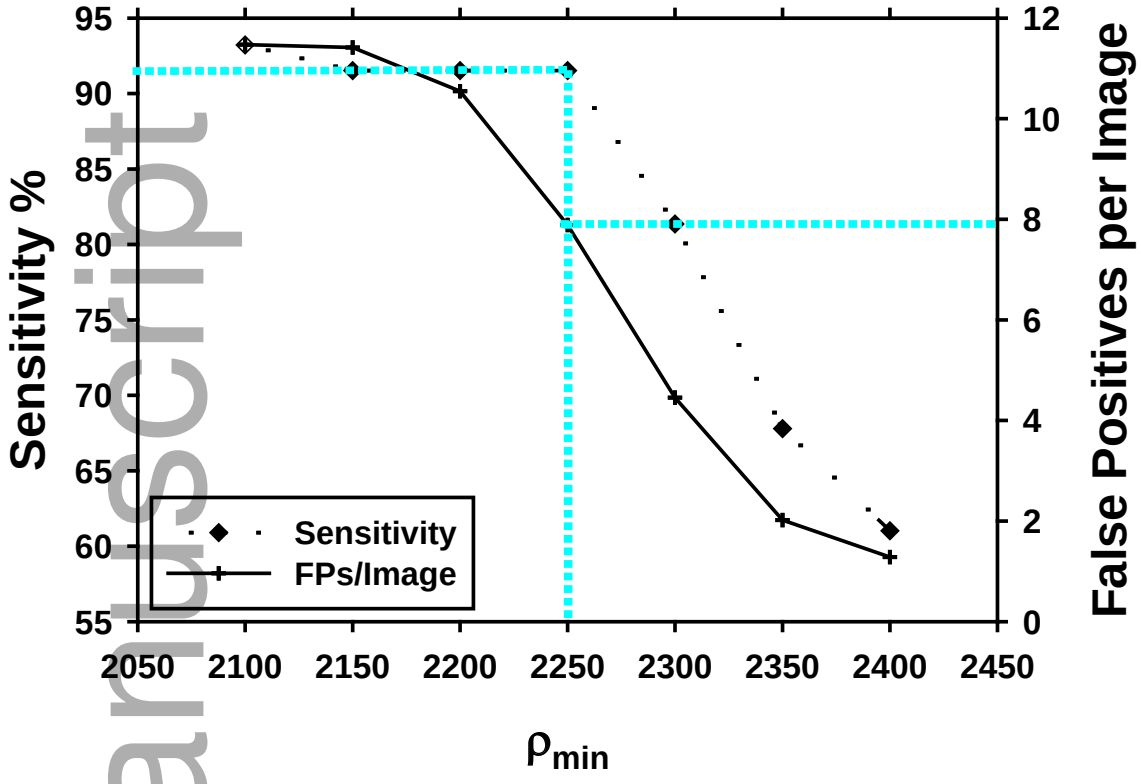


Fig. 5. Dependence of sensitivity and false positive rate (FPs/Image) on the minimum object density value ρ_{min} for objects at the pre-screening stage in the training set. The dashed vertical line on the plot indicates the selected decision value for ρ_{min} .

Since needles are radiopaque, the candidates that actually are needles would have higher pixel intensities than other candidates. Hence this would be a good way to judge the presence of a needle. However, this technique, when used by itself for processing cadaver images was found to create an unacceptably high number of false positives due to the large number of tissues, organs and surgical objects with high contrast appearance in the cadaver images. Therefore, additional techniques (described below) had to be implemented to identify the needles more accurately and reduce the false positives.

CNR is calculated next, which is defined by the following equation:

$$CNR = \frac{\mu_{sig}}{\sigma_{Noise}} \quad , \tag{3}$$

where

$$\mu_{sig} = \bar{O} - \bar{B} \tag{4}$$

$$\bar{O} = \frac{\sum_{n=0}^{N_O} i(n)}{N_O} \qquad if\ i(n) > T_{cnr} \tag{5}$$

$$\bar{B} = \frac{\sum_{n=0}^{N_B} i(n)}{N_B} \qquad if\ i(n) \leq T_{cnr}, \tag{6}$$

where $i(n)$ is the intensity of pixel n, N_O is the total number of pixels in the object $\bar{O}$, N_B is the total number of pixels in the background region $\bar{B}$, and μ_{sig} is the difference between the mean pixel values of the object $\bar{O}$ and its background $\bar{B}$. σ_{Noise} is the standard deviation of the background noise, which is calculated as follows:

$$\sigma_{Noise} = \sqrt{\left(\frac{\sum_{n=0}^{N_B}(i(n)-\bar{B})^2}{N_B}\right)} \tag{7}$$

The needles or other metallic objects would have higher CNR values, as compared to tissues and muscles in the background. Therefore, if we impose a constraint on the CNR as a decision rule CNR > $CNR_{th,}$ where CNR_{th} is a CNR threshold, this decision rule will function as a second filter, eliminating FP objects, such as those from bones and muscles. Fig. 6 shows the dependence of the true and false positives as a function of CNR values for the training set. It can be seen that, when CNR_{th} is set to 3.45, the FP rate can be reduced to 2.9. This is 63.3% ((7.9-2.9)/7.9) lower than the FP rate obtained after previous stage while the sensitivity is 89.8%.

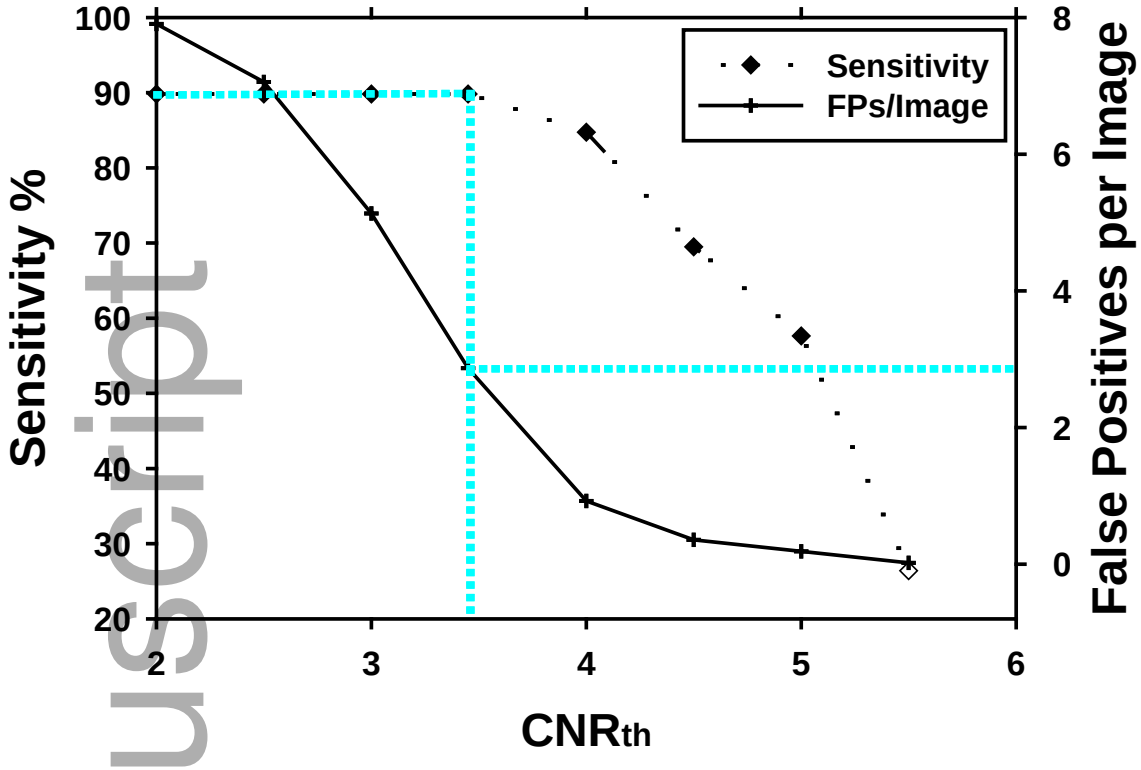
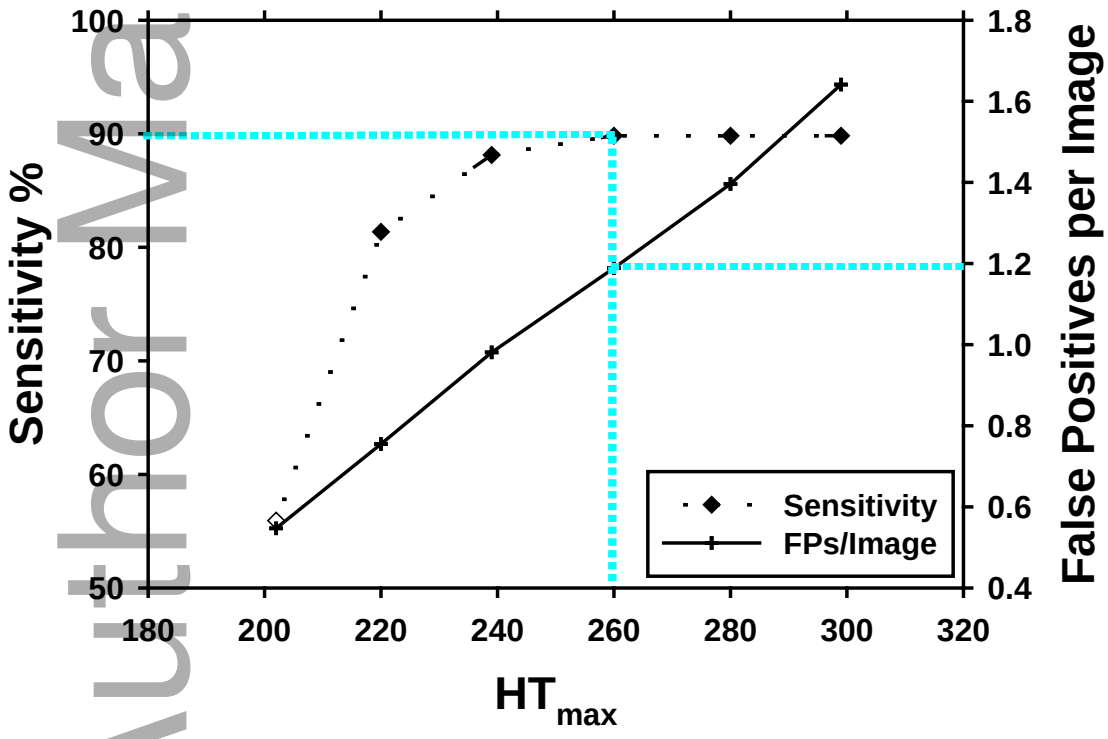


Fig. 6. Dependence of sensitivity and false positive rate (FPs/Image) on the CNR threshold for objects in the training set. The dashed vertical line on the plot indicates the chosen decision threshold CNR_{th}.

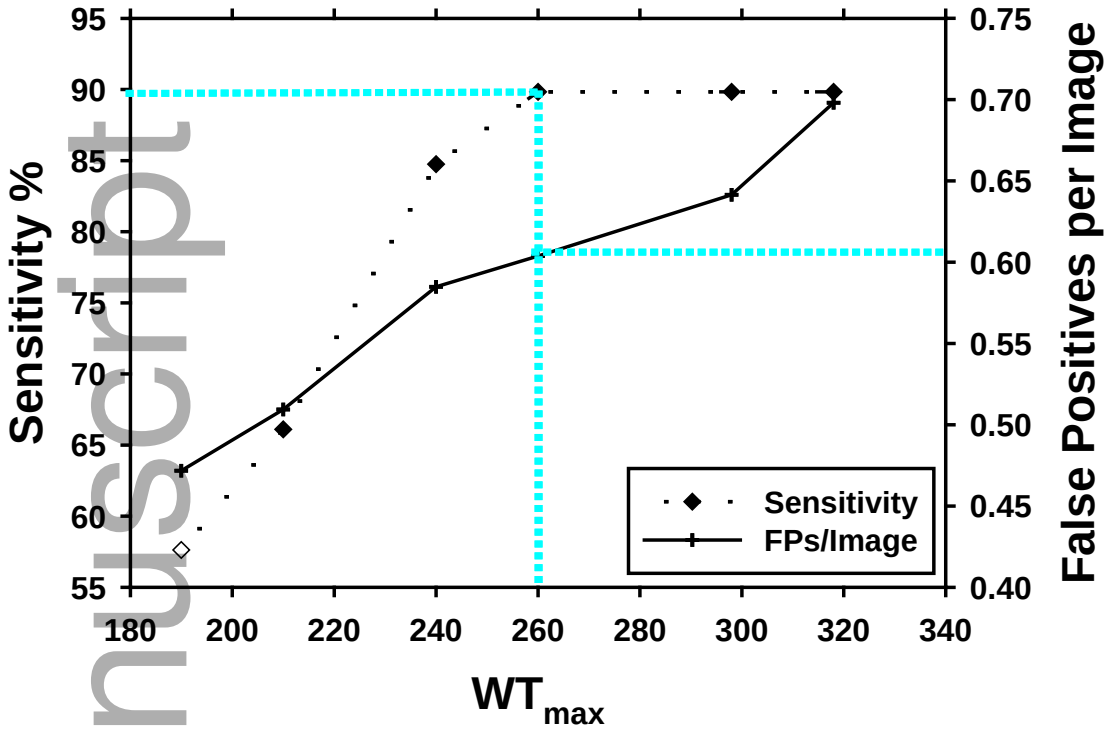
To further reduce the number of FPs, the geometric parameters of the chosen candidate was taken into consideration. The geometric parameters are denoted by *the height* (HT) *and width* (WT) of a virtual box, as shown in Fig. 7, enclosing the entire needle candidate. Constraining these geometric parameters within a range defined as $HT_{min}<HT<HT_{max}$ and $WT_{min}<WT<WT_{max}$, will result in elimination of a substantial number of false positive bone-containing objects of sizes different from needles but having CNR and object density values comparable to that of the needles. Based on the training set, the values for HT_{min} and WT_{min} were chosen to be 70 pixels (8.75 mm) to remove very small objects as compared to needles. The values of HT_{max} and WT_{max} were chosen by examining the dependence of the sensitivity and FP rate on their values, as shown in Fig. 8. The optimal values for HT_{max} and WT_{max} were chosen to be both 260 pixels (32.50 mm). With these HT_{min}, HT_{max}, WT_{min}, and WT_{max} values, the FP rate was reduced from 2.9 (after the CNR stage) to 0.6, at a sensitivity of 89.8% for the training set.

(a) (b)

Fig. 7. Examples of how the CAD system estimates the HT and WT values of the needle by defining a rectangular box to enclose the needle.



(a)



(b)

Fig. 8. Dependence of sensitivity and false positive rate (FPs/Image) on (a) HT_{max} and (b) WT_{max} values for objects in the training set. The dashed vertical lines on the plots indicate the chosen decision thresholds for HT_{max} and WT_{max}.

After the three different types of features and decision rules based on the object size and shape, the FP rate was reduced substantially by about 96% (FP rate decreased from 12.1 to 0.6 FPs/image in the training set) compared to the pre-screening stage. However, some other FP objects in the radiographs still remained because they met all the conditions imposed by the decision rules described above. We designed another FP reduction method to exploit the rather distinctive shape of the needle. The various projections of the needles could be thought of as parts of an ellipse or hyperbola. Thus, by fitting an ellipse or a hyperbola to the object, one can determine whether the detected object is a needle by estimating its goodness of fit to an ellipse or a hyperbola.

2.B.4. Curve Fitting Algorithm

While applying the curve fitting algorithm, it is unknown whether the needle conforms better to a hyperbola, ellipse or parabola. However, for simplicity, our aim is to

force an ellipse to fit to the candidate object shape. Therefore, the algorithm fits the general equation for a conic section[12] to the pixels of the binary segmented object obtained in Section 2.B.2:

$$G\ (x,y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0, \quad (8)$$

where *A, B, C, D, E,* and *F* are coefficients of the conic section.

This equation describes a hyperbola if the determinant, $B^2 - 4AC$ is positive, an ellipse if it is negative, and a parabola if it is 0. We used two techniques the least squares curve fitting[12] and the nonlinear simplex optimization[13] based fitting to determine the values of the coefficients *A, B, C, D, E,* and *F* for the conic section with the best fit to a set of data points (object pixels on the image) as follows.

Least Squares Ellipse Fitting:

In the case of least squares ellipse fitting, it is necessary first to select an appropriate error function, and then to estimate an aggregate error. The error function should be such that it is 0 for object pixels with coordinates lying on the curve and non-zero for those located away from the curve. Eq. (8) is used as an error function. Without loss of generality, we can set the coefficient F to 1.

Error Function: $G(x,y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + 1$ (9)

Based on $G(x,y)$ the pointwise error and the aggregate error are defined as:

Pointwise Error: $\xi_i = G(x_i, y_i) = Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + 1$ (10)

Aggregate Error: $E = \sum_{i=0}^{N_O} \xi_i^2$ (11)

where, N_O is the total number of pixels in the object.

In order to minimize the aggregate error with the least squares minimization, first the partial derivatives of Eq. (11) with respect to *A, B, C, D* and *E* are obtained and are equated to zero. The following equations are derived as a result:

$$A\sum x_i^4 + B\sum x_i^3 y_i + C\sum x_i^2 y_i^2 + D\sum x_i^3 + E\sum x_i^2 y_i + \sum x_i^2 = 0 \quad (12)$$

$$A\sum x_i^3 y_i + B\sum x_i^2 y_i^2 + C\sum x_i y_i^3 + D\sum x_i^2 y_i + E\sum x_i y_i^2 + \sum x_i y_i = 0 \quad (13)$$

$$A\sum x_i^2 y_i^2 + B\sum x_i y_i^3 + C\sum y_i^4 + D\sum x_i y_i^2 + E\sum y_i^3 + \sum y_i^2 = 0 \quad (14)$$

$$A\sum x_i^3 + B\sum x_i^2 y_i + C\sum x_i y_i^2 + D\sum x_i^2 + E\sum x_i y_i + \sum x_i = 0 \quad (15)$$

$$A\sum x_i^2 y_i + B\sum x_i y_i^2 + C\sum y_i^3 + D\sum x_i y_i + E\sum y_i^2 + \sum y_i = 0 \quad (16)$$

The solution to these five equations yields a conic which will minimize the error function. Since there are five linear equations with five unknowns, they can be solved using the matrix method, as shown below:

$$\begin{bmatrix} \sum x_i^4 & \sum x_i^3 y_i & \sum x_i^2 y_i^2 & \sum x_i^3 & \sum x_i^2 y_i \\ \sum x_i^3 y_i & \sum x_i^2 y_i^2 & \sum x_i y_i^3 & \sum x_i^2 y_i & \sum x_i y_i^2 \\ \sum x_i^2 y_i^2 & \sum x_i y_i^3 & \sum y_i^4 & \sum x_i y_i^2 & \sum y_i^3 \\ \sum x_i^3 & \sum x_i^2 y_i & \sum x_i y_i^2 & \sum x_i^2 & \sum x_i y_i \\ \sum x_i^2 y_i & \sum x_i y_i^2 & \sum y_i^3 & \sum x_i y_i & \sum y_i^2 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \\ E \end{bmatrix} = \begin{bmatrix} \sum x_i^2 \\ \sum x_i y_i \\ \sum y_i^2 \\ \sum x_i \\ \sum y_i \end{bmatrix} \quad (17)$$

Matrix X Matrix M Matrix Y

Matrices X and Y in Eq. (17) can be obtained by using the coordinates of every pixel within the candidate object being fit. Thus by solving Eq. (17), the coefficients A, B, C, D and E can be estimated as follows to satisfy Eq. (9) :

$$M = X^{-1} \times Y \quad (18)$$

Nonlinear Simplex Optimization Ellipse Fitting:

This optimization method, authored by Nedler and Mead[13], allows the minimization of a function of *k* variables. It begins with a *simplex*[14] characterized by *k* +1 vertices in an *k*-dimensional space of the variables, which are associated with the function values calculated at those vertices and estimation of the corresponding errors in these vertices. With every iteration, a new *simplex* is formed such that the estimated errors at its vertices, slowly descend to the minimum[14].

In our application, once the new simplex with the corresponding new set of coefficients *A, B, C, D* and *E* was generated for a given iteration, the determinant, $B^2 - 4AC$ was calculated to estimate whether the candidate conic in question conformed to an ellipse. If the determinant was positive or equal to zero i.e. the conic was a hyperbola or a parabola, this specific simplex solution was penalized by assignment of a large error and excluded from the set that could generate simplexes. Therefore, by constraining the simplex procedure with this additional condition it was possible to force an ellipse to fit the segmented object candidate data points.

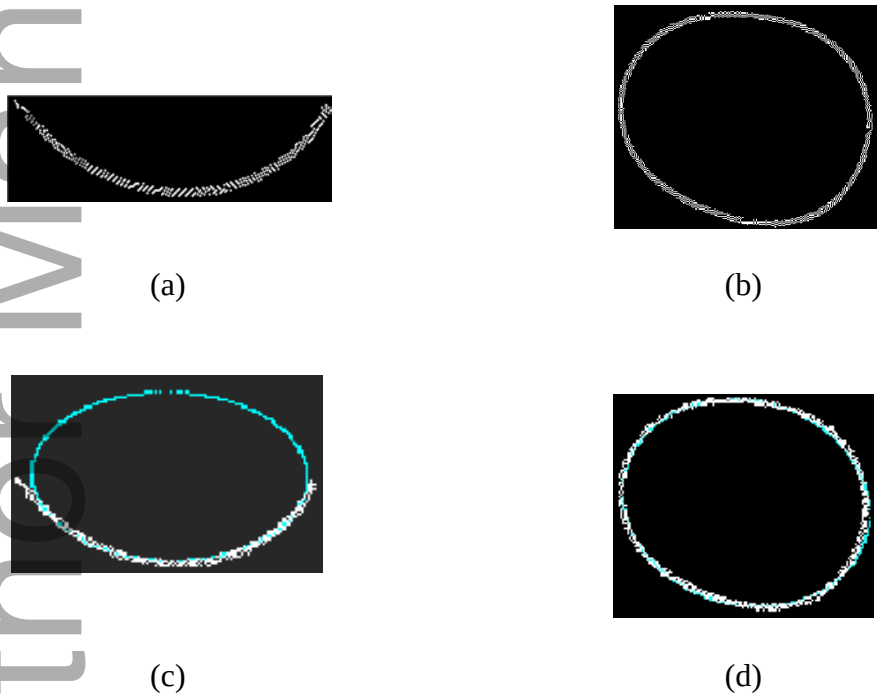
2.B.5. Goodness of fit features

The next step is to estimate how well the conic curve fit the object in question. The goodness of fit was determined by two functions.

Overlap function:

This function basically calculates α that describes what percentage of the fitted ellipse coincides with the object pixels. This allows separation of needles from other foreign objects such as sutures and pacemakers placed on the cadavers. As we can see from Fig. 9(d), for pacemakers, the segmented pacemaker object (Fig. 9(b)) almost completely overlaps with the

fitted ellipse. Hence, the overlap percentage, α, would be close to 100% for these false positives, while for needles (Fig. 9(a)), the overlap percentage is typically in the range of 10-65%, as shown in Fig. 9(c). Fig. 9(a) is the filtered, thresholded and cropped section of the image in Fig. 7(b). We set the criterion that, if $\alpha_{min} \leq \alpha \leq \alpha_{max}$, the candidate is considered a needle. The thresholds for α_{min} and α_{max} are determined by analysis of the false positive and true positive rate for different α values for all objects after the feature analysis step in the training set, as shown in Fig. 10(a). The plot shows that the peak in the graph corresponds to the objects with α values in the range of 7-65%, which basically represents all the needles. Therefore, α_{min} and α_{max} were set at 7% and 65%, respectively. By imposing the lower and upper bounds of the overlap function as described above, the number of FPs per image was reduced from 0.6 (after the width stage) to. 0.35 (FPs/Image (α_{max}) - FPs/Image (α_{min}) = 0.52-0.17)



(a) (b)

(c) (d)

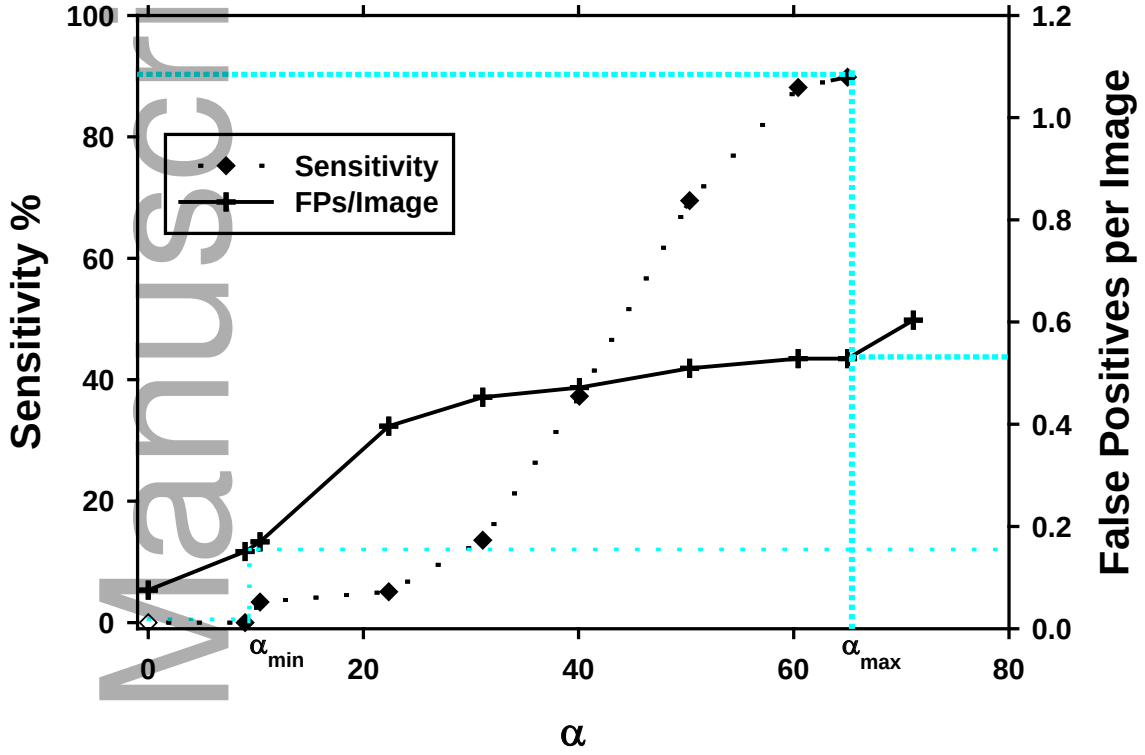
Fig. 9. Candidates representing a needle (the needle from Fig. 8b) (a), and a pacemaker (b) after enhancement and segmentation. (c) Ellipse fitted (in blue) to the needle (in white) ($\alpha = 49.4\%$). (d) Ellipse fitted to the pacemaker candidate ($\alpha = 99.6\%$).

Normal Distance function:

The normal distance function[12] computes the perpendicular distance λ, from every pixel location in the candidate to the fitted ellipse. It is calculated as the ratio of *G(x, y)* (Eq. (9)) to the magnitude of the gradient of *G(x, y)*:

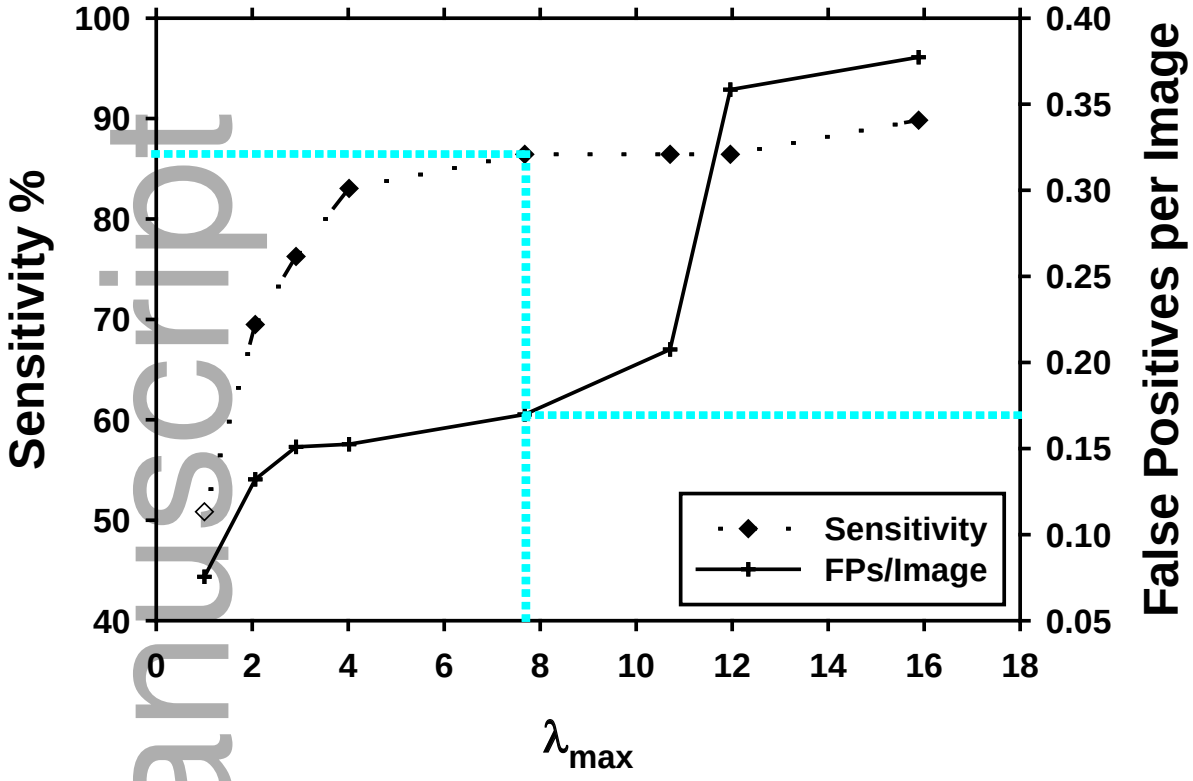
$$\lambda = \sum_{n=0}^{N_O} \left(\frac{G(x,y)}{|\nabla G|}\right)^2 \quad , \tag{19}$$

where



(a)

$$|\nabla G| = \sqrt{\left(\frac{\partial G(x,y)}{\partial x}\right)^2 + \left(\frac{\partial G(x,y)}{\partial y}\right)^2} \tag{20}$$



(b)

Fig. 10. (a) Dependence of the sensitivity and FP rate (FPs/Image) on α, α_{min} = 7% and α_{max} = 65% were the values chosen from the training set. (b) Dependence of the sensitivity and FP rate on λ_{max} from the training set. The dashed vertical line indicates the corresponding selected thresholds.

If the calculated λ is such that $\lambda \leq \lambda_{max}$, the selected candidate is labelled as a detected needle, otherwise the candidate is dismissed. The dependence of sensitivity and FP rate on λ_{max} in the training set is plotted in Fig. 10(b). At a value of λ_{max} = 8.0 pixels (0.16 mm^2), the FP rate was reduced from 0.35 (after applying the overlap function) to 0.17 with a sensitivity of 86.4%. Furthermore, the choice of threshold for λ can be used to adjust the operating point of the CAD system to achieve different sensitivity and specificity values. The specific values for this threshold are discussed in Section 3.

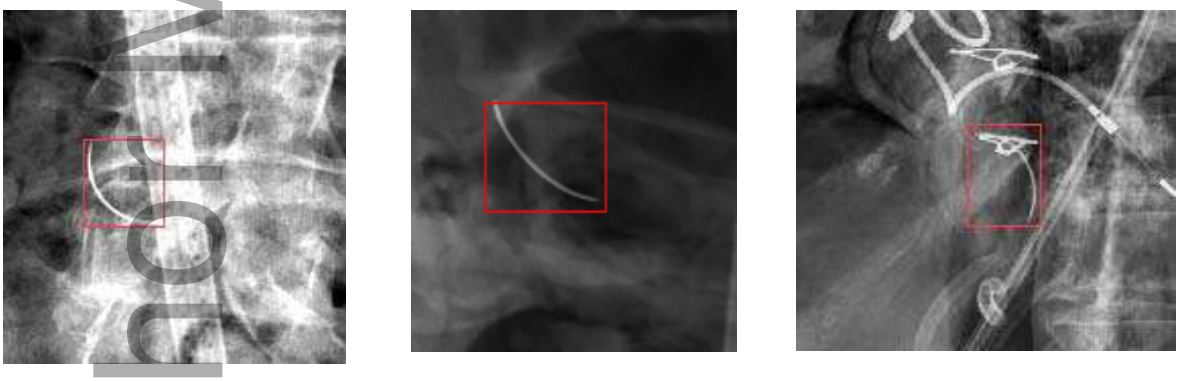
2.B.6. Neural network, linear discriminant analysis (LDA) and random forest classifiers

In addition to the rule-based classifier described above, we evaluated the performance of three other classifiers, neural network, random forest and linear discriminant analysis (LDA), to merge the extracted features and reduce false positives. For these classifiers, we first used the rule-based method with the fixed threshold values (as described in the previous section) up to the CNR stage to reduce some of the false positives. We then trained the classifiers with the five extracted features, CNR, HT, WT, α and λ, as input predictor variables using the training set.

For the neural network, we used a multilayer perceptron with 5 input neurons, two hidden layers (with 5 hidden nodes in the first layer and 2 hidden nodes in the second layer) and one output neuron. It was trained with back-propagation for 1000 iterations. We selected the smallest neural network that could reach training AUC of 0.99.

For the random forest classifier the best training result (highest sensitivity with the lowest false positive rate) was obtained for 200 decision trees with a minimum leaf size of 2 and therefore we chose this configuration. We experimentally determined these parameters using the training set and then applied the selected parameters to the test set.

2.C. Evaluation of the CAD system



(a) (b) (c)

Fig. 11. Needles of various shapes and orientations with different background that were detected by the CAD system.

The bounding box of a needle is used as reference standard to determine whether an object labelled by the CAD system is a true positive (TP) or an FP using a scoring program[3]. For the rule-based classifier, λ was used as a decision variable. For the random forest, LDA and neural network classifiers, the output score from the classifier was used as a decision variable to obtain the FROC curves as well as the operating points for mode I and II. At a given

decision threshold value, each detected object is first enclosed by a bounding box. If the centroid of the bounding box of a CAD detected object fell within a reference standard box, then the detected object is a TP, otherwise an FP. The entire CAD system was developed using the training set.

After all parameters are fixed, the performance of the CAD system was evaluated on the test set. Free response receiver operating characteristic (FROC) analysis was used to evaluate the performance of the CAD system on both the training and test sets over a range of decision thresholds. The sensitivity for the training and test sets were estimated by using the 53 training and the 55 test cadaver images with needles, respectively. The corresponding FP rates for the test curve were estimated by using the 100 normal images without needles.

3. RESULTS

Examples of the detected needles with different orientations, locations and with different levels of visibility are shown in Figs. 1, 7 and 11. The FROC curves using the rule-based classifier for the training and the test sets are shown in Fig. 12. Two hypothetical operating points are marked along the FROC curves for the rule-based classifier, mode I with higher specificity and mode II with higher sensitivity. The decision thresholds for the two modes were selected using the training set:

1. Mode I: λ_{max} = 2.74 pixels (0.0428 mm^2)
2. Mode II: λ_{max} = 15.9 pixels (0.2484 mm^2)

The detection results for both modes are summarized in Table 2.

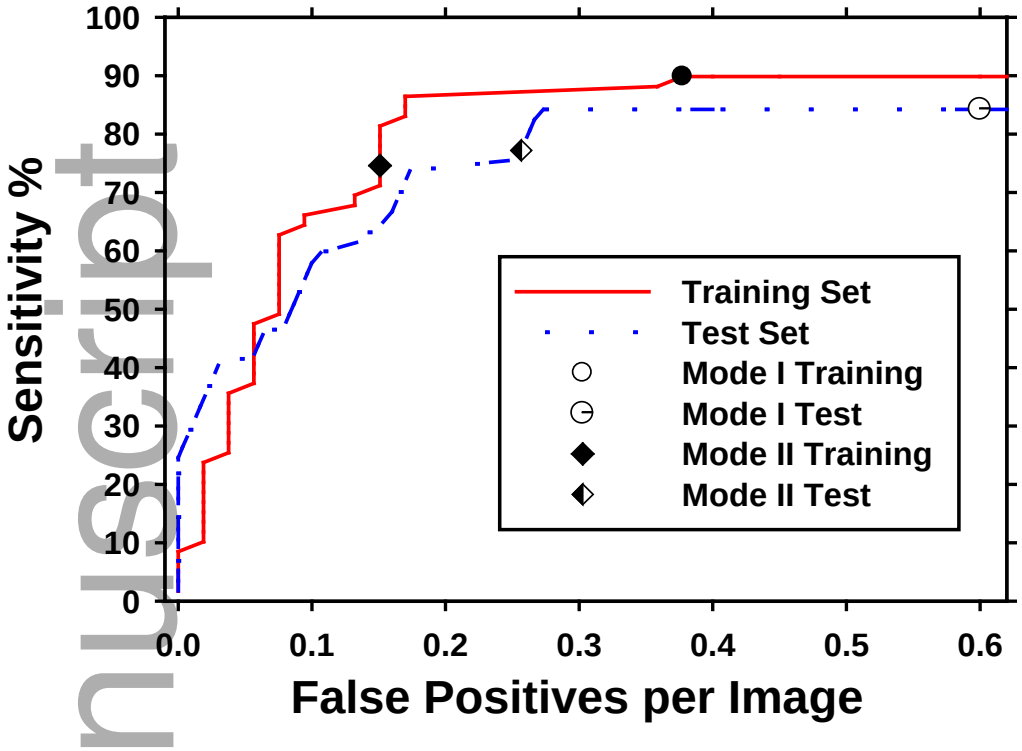


Fig. 12. FROC curves for training and test sets for rule based classifier.

405 The operating points for mode I with higher specificity and for mode II with higher sensitivity for the neural network classifier are shown in Fig. 13. The operating points were selected in the same way as for the rule-based classifier. The detection results for both modes, together with the results for the random forest and LDA classifiers are summarized in Table 2.

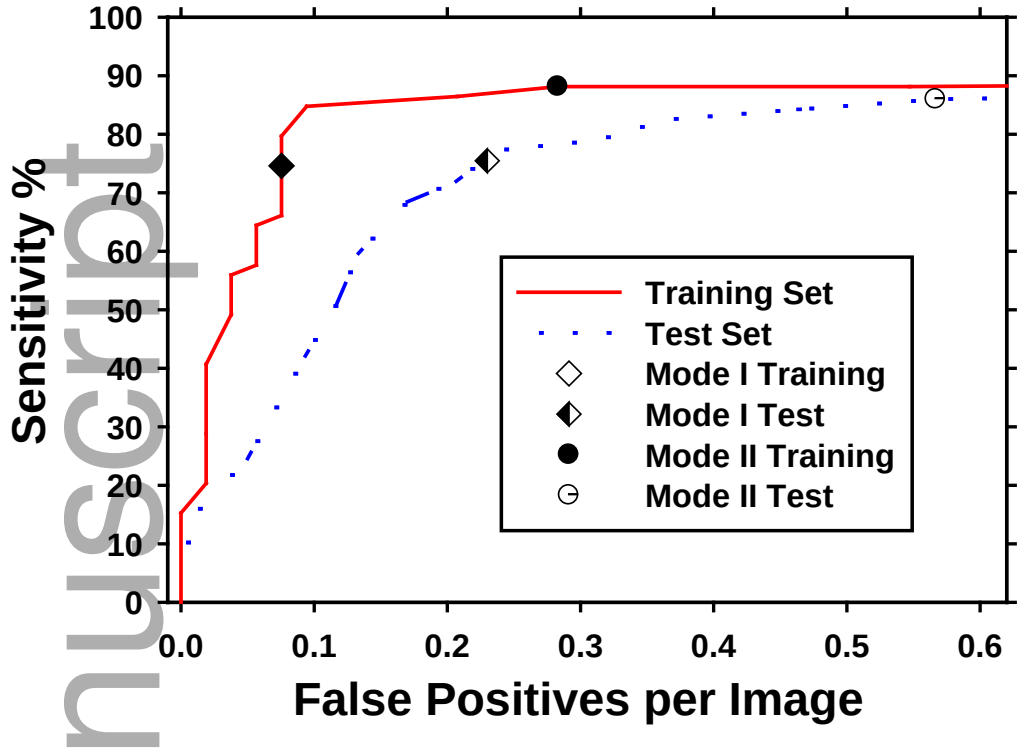


Fig. 13. FROC curves for training and test sets for neural network classifier.

410

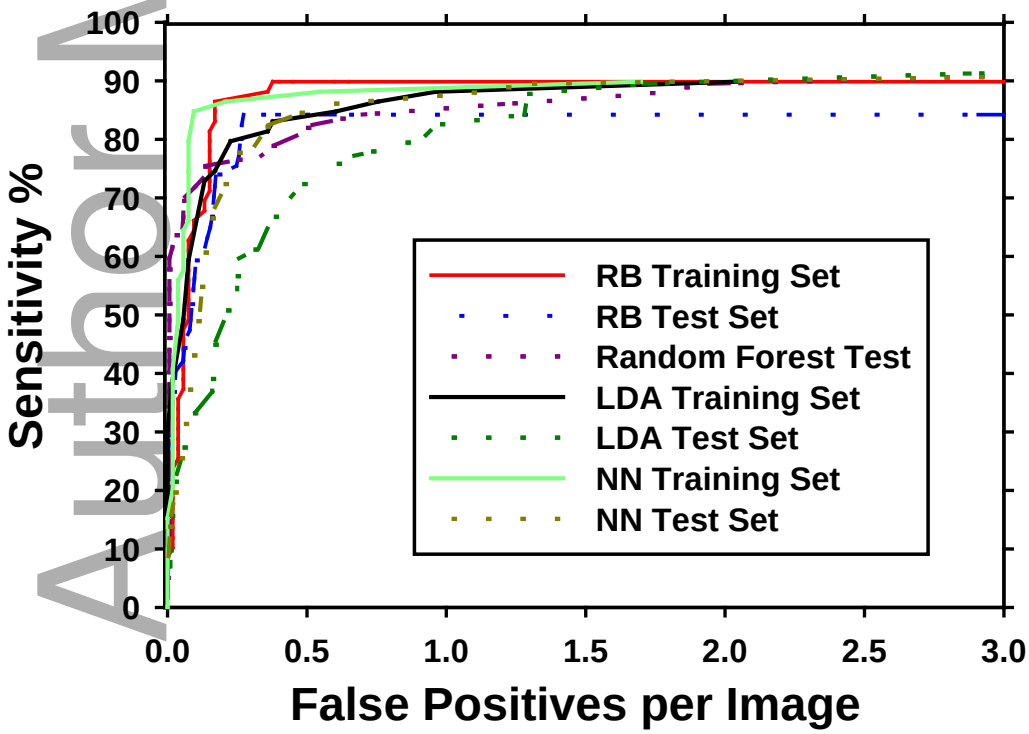


Fig. 14. Comparison of the FROC curves for the different classifiers

The FROC curves for the different classifiers: rule-based, random forest, LDA, and neural network are compared in Fig. 14. The FROC curve for the random forest training set is not shown in this figure, as no false positives were produced for the training set and all needles at that stage were detected with an overall sensitivity of 89.8%. The neural network classifier had the best performance, with a higher sensitivity of 86.0% with the test set and a slightly lower FP rate of 0.57 FPs/Image with the normal set.

Table 2. Results for needle detection on cadaver radiographs.

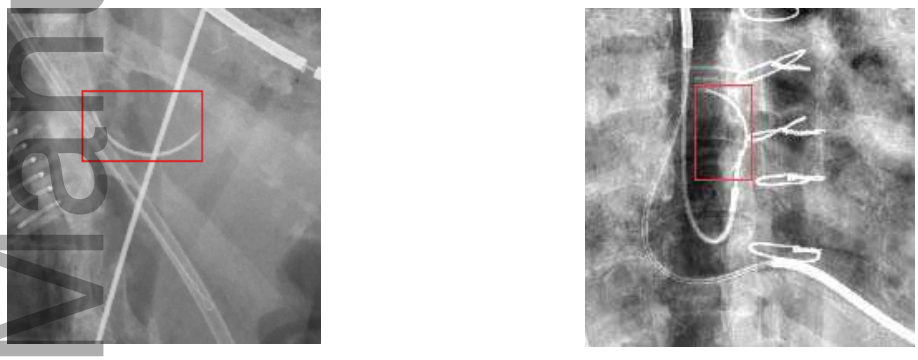
Classifier		Training Set		Test Set	
		Sensitivity	False positives/ Image	Sensitivity	False positives/ Image
Rule Based	Mode I	74.6	0.15	77.2	0.26
	Mode II	89.8	0.36	84.2	0.60
Random Forest	Mode I	89.8	0.0	75.4	0.21
	Mode II	89.8	0.0	75.4	0.21
LDA	Mode I	79.7	0.22	71.9	0.48
	Mode II	86.4	0.83	86.0	1.28
Neural Networks	Mode I	74.6	0.08	75.4	0.23
	Mode II	88.1	0.28	86.0	0.57

4. DISCUSSION

Although there have been previous studies[4-7] to detect line-like objects, those techniques could not be directly applied for the automatic detection of needles. For example, for the detection of lasso catheters[5] , a similar idea of ellipse fitting was used, but the ellipse fitting

was done by first detecting the electrodes as 'blobs' and then checking whether these 'blobs' could fit an ellipse. However, we could not apply the same 'blob' detection technique directly to surgical needles. Moreover, the higher radio-opacity of the needles as compared to these devices, prompted us to investigate other features like contrast-to-noise ratio, object density etc. as we discussed in previous sections of this study. This type of feature analysis reduced some of the false positives.

The cadaver images with the needles and the additional surgical instruments, sponges and medical tubes on the cadaver simulated a challenging data set. The preliminary results indicated that the CAD system performed well in both modes I and II. The system was able to detect most of the surgical needles with relatively low FP rates. The CAD systems with the rule-based or the neural network classifier performed similarly and slightly better than the CAD systems with the random forest or LDA classifier.



(a) (b)

Fig. 15. Example of needle missed by both the rule based and the neural network based CAD systems.

The two operating points of the CAD system are selected to illustrate the potential application of the CAD system in different clinical situations. Mode I with higher specificity can be used by the surgeon in the operating room, while mode II with higher sensitivity can be used by the radiologist. For example, the CAD system with the neural network classifier operated in mode I offers 75.4% sensitivity with 0.23 FPs/image on the test set, or approximately only 1 FP in every 4 to 5 cases. This will allow the surgeon to close the patient in about 75% of the time for whom no needle (or false positive) on the radiographs is detected. For the 23% of the radiographs with a CAD mark, the surgeon can take a quick look at the radiograph to determine if the marked object is a real needle or not. The surgeon

can easily rule out FPs such as bones, sutures or tube which may be mistakenly recognized by the system as needles. In an occasional doubtful situation, where the surgeon is unable to distinguish between a needle and a false positive from the objects marked on the radiographs by the CAD system, the surgeon will have the opportunity to verify his conclusions by inspecting the patient before closing the patient. With the CAD sensitivity at about 75%, the surgeon will be able to remove about 75% RFO needles from the patients before closing. Therefore, the CAD system can function as a first reader in the OR with surgeon spending minimal time checking the CAD marks only. In mode II the CAD system will operate at 86.0% sensitivity with 0.57 FPs/image. It may be used by the radiologist as a second reader that will complement their own reading. The combined sensitivity of radiologist and the CAD system will be most likely higher than the radiologist alone[15-18]. A radiologist can easily rule out FPs such as bones, sutures or tubes that are identified by the CAD system as needles.

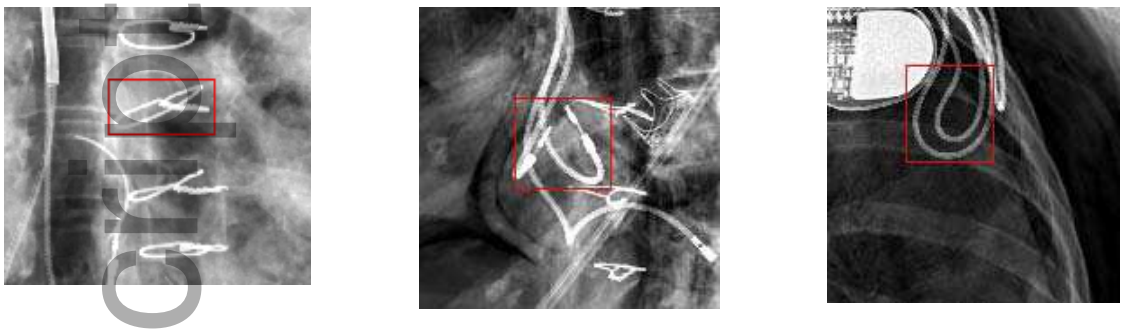
One of the most challenging problems faced by the CAD system was caused by the presence of other objects in the radiograph, such as pacemakers, tubes, sutures etc. In fact when the curve fitting algorithm was applied to these objects, it was found that they conform to the shape of an ellipse with very small error. However, it was observed that the needle overlapped with the ellipse by only 10-65%, while the sutures and pacemakers overlapped more than 90%. Thus, a decision based on the overlap function helped eliminate several such false positives, improving the specificity of the CAD system even in these challenging situations.

In this preliminary study, we used two types of needles, which are frequently used in clinical practice and have similar sizes (22 mm and 24 mm). While this is a current limitation of the system, in the future, we plan to increase the data set to include radiographs of needles with a wide range of sizes, which will allow us to identify any deficiencies or limitations the current system may have and design new methods to improve its performance.

Fig. 15 shows examples of needles missed by the CAD system with the neural networks as well as with the rule based system. Due to the presence of other overlapping structures like bones and medical instruments, the shapes of the segmented objects are distorted and the fitting of these shapes to an ellipse was poor with large errors. In the future when we generate additional training samples, we may try to include additional shape analysis to differentiate the different types of high-contrast objects. Another possible strategy is to separate the overlapping objects before fitting the ellipse.

Fig. 16 shows some of the false positives detected by the rule based but omitted by the neural network based CAD system. They satisfied all the conditions that were designed to

detect the needles. Nevertheless, most of these false positives can be easily recognized and ignored by the physicians.



(a) (b) (c)

Fig. 16. Examples of false positives detected by the rule based but omitted by the neural network based CAD system

This study has several limitations. First, the data set is relatively small and includes only two types of needles so that it does not cover all possible scenarios that may occur in the operating room. Second, we could only use cadavers to generate the training and test radiographs, rather than collecting real human subject images, although the cadaver may be more challenging because of the various procedures that might have been performed near and after the end of the life and the various objects left behind in the cadaver body. Third, there are other possible types of RFOs and this study only focused on surgical needles. Fourth, the sensitivity and specificity are still relatively low, including the potential operating points for mode I and mode II, and further improvement is needed when a larger data set can be generated.

5. Conclusion

To the best of our knowledge, this was the first time a CAD system is developed to detect needles from post-operative radiographs with realistic anatomical background structures[8, 9]. We demonstrated that the CAD system can be operated in different modes by properly selecting the operating points; for example, a high specificity mode as a first reader for the surgeon in the operating room and a high sensitivity mode for the radiologist later as a second reader. With such a computer-assisted approach, it is possible to substantially reduce the number of RFOs unintentionally left behind in a patient's body, thereby reducing the morbidity and mortality as well as the patient care costs associated with RFOs.

ACKNOWLEDGMENTS

This work has been supported in part by a University of Michigan Fostering Innovation Grant No. FIG 530-R10 and a Michigan Initiative for Innovation and Entrepreneurship Grant. T. C. Marentis and N. Chronis are cofounders of Kalyspo, LLC.

The authors for this research article have no relevant conflicts of interest to disclose.

REFERENCES

[1]N. N. Egorova, A. Moskowitz, A. Gelijns, A. Weinberg, J. Curty, B. Rabin-Fastman, H. Kaplan, M. Cooper, D. Fowler, J. C. Emond and G. Greco, "Managing the prevention of retained surgical instruments: what is the value of counting?," Annals of Surgery **247**, 13-8 (2008).

[2]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," New England Journal of Medicine **348**, 229-235 (2003).

[3]L. Hadjiiski, T. C. Marentis, A. R. Chaudhury, L. Rondon, N. Chronis, and H. P. Chan, "Computer aided detection of surgical retained foreign object for prevention," Med. Phys. **42**, 1213-1222 (2015).

[4]D. Volpi, M.H. Sarhan, R. Ghotbi, N. Navab, D. Mateus and S. Demirci, "Online tracking of interventional devices for endovascular aortic repair," International journal of computer assisted radiology and surgery, **10(6)**, 773-781 (2015).

[5]Y. Ma, N. Gogin, P. Cathier, R. J. Housden, G. Gijsbers, M. Cooklin, M. O'Neill, J. Gill, C.A. Rinaldi, R. Razavi and K.S. Rhode, "Real-time x-ray fluoroscopy-based catheter detection and tracking for cardiac electrophysiology interventions," Medical Physics, **40**, 071902 (2013)

[6]F. Milletari, N. Navab and P. Fallavollita, "Automatic detection of multiple and overlapping EP catheters in fluoroscopic sequences," International Conference on Medical Image Computing and Computer-Assisted Intervention, 371-379 (2013).

[7]F. Milletari, V. Belagiannis, N. Navab, and P. Fallavollita, "Fully automatic catheter localization in C-arm images using l1-sparse coding," Medical Image Computing and

Computer-Assisted Intervention: MICCAI, International Conference on Medical Image Computing and Computer-Assisted Intervention, **17**, 570-577 (2013).

[8]G. Agam, L. Gan, M. Moric, and V. Gluncic, "Automated identification of retained surgical items in radiological images," in Medical Imaging 2015: Pacs and Imaging Informatics: Next Generation and Innovations, edited by T. S. Cook and J.Zhang (2015), Vol. 9418.

[9]B. Asiyanbola, C. W. Chao, J. S. Lewin, and R. Etienne-Cummings, "Modified Map-Seeking Circuit: Use of Computer Aided Detection in Locating Postoperative Retained Foreign Bodies," Journal of Surgical Research **175** (2), E47E52 (2012).

[10]M. D. Macilquham, R. G. Riley, and P. Grossberg, "Identifying lost surgical needles using radiographic techniques," AORN Journal, **78**, 73-78 (2003).

[11]G. Dougherty, *Digital image processing for medical applications.* (Cambridge University Press, New York, 2009).

[12]G. J. Agin, Ph.D. Thesis, Carnegie-Mellon University, Pittsburgh (1981).

[13]J. A. Nelder and R. Mead, "A simplex method for function minimization," The Computer Journal, **7**, 308-313 (1965).

[14]J. C. Lagarias, J. A. Reeds, M. H. Wright and P. E. Wright, "Convergence properties of the Nelder--Mead simplex method in low dimensions," SIAM Journal on Optimization, **9**, 112-147 (1998).

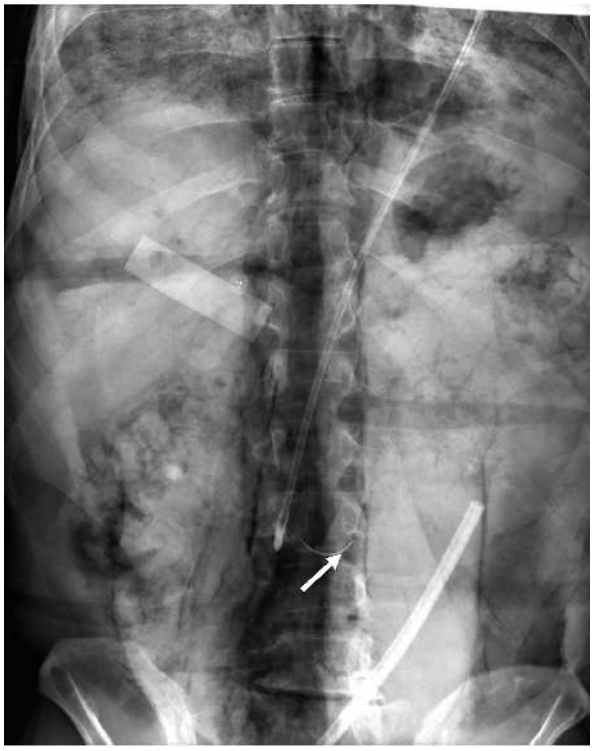
[15]N. Petrick, B. Sahiner, S. G. Armato III, A. Bert, L. Correale, S. Delsanto, M. T. Freedman, D. Fryd, D. Gur, L. Hadjiiski, Z. Huo, Y. Jiang, L. Morra, S. Paquerault, V. Raykar, F. Samuelson, R. M. Summers, G. Tourassi, H. Yoshida, B. Zheng, C. Zhou and H. P. Chan, "Evaluation of computer-aided detection and diagnosis systems[a)]," Medical Physics **40**, 087001 (2013).

[16]R. F. Brem and J. M. Schoonjans, "Radiologist detection of microcalcifications with and without computer-aided detection: a comparative study," Clinical Radiology **56**, 150-154 (2001).

[17]J. C. Dean and C. C. Ilvento, "Improved cancer detection using computer-aided detection with diagnostic and screening mammography: prospective study of 104 cancers," American Journal of Roentgenology **187**, 20-28 (2006).

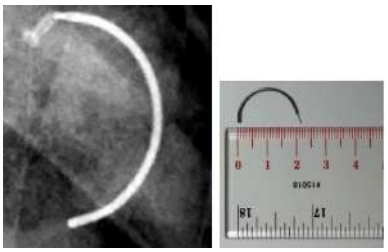
[18]D. Regge, P. D. Monica, G. Galatola, C. Laudi, A. Zambon, L. Correale, L., R. Asnaghi, B. Barbaro, C. Borghi, D. Campanella, M. C. Cassinis, R. Ferrari, A. Ferraris, C. Hassan, F. Iafrate, G. Iussich, A. Laghi, R. Massara, E. Neri, L. Sali, S. Venturini and G. Gandini, "Efficacy of computer-aided detection as a second reader for 6–9-mm

575 lesions at CT colonography: multicenter prospective trial," Radiology **266**, 168-176 (2013).

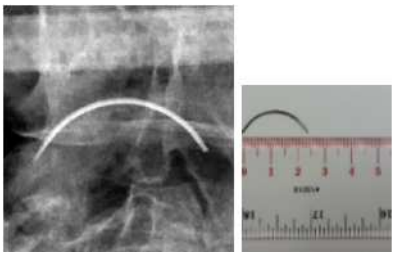


mp_12011_f1a.tif

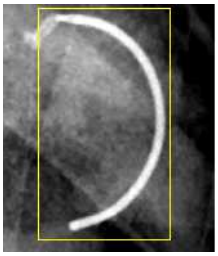
mp_12011_f1b.tif



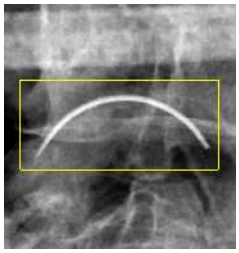
mp_12011_f2a.tif



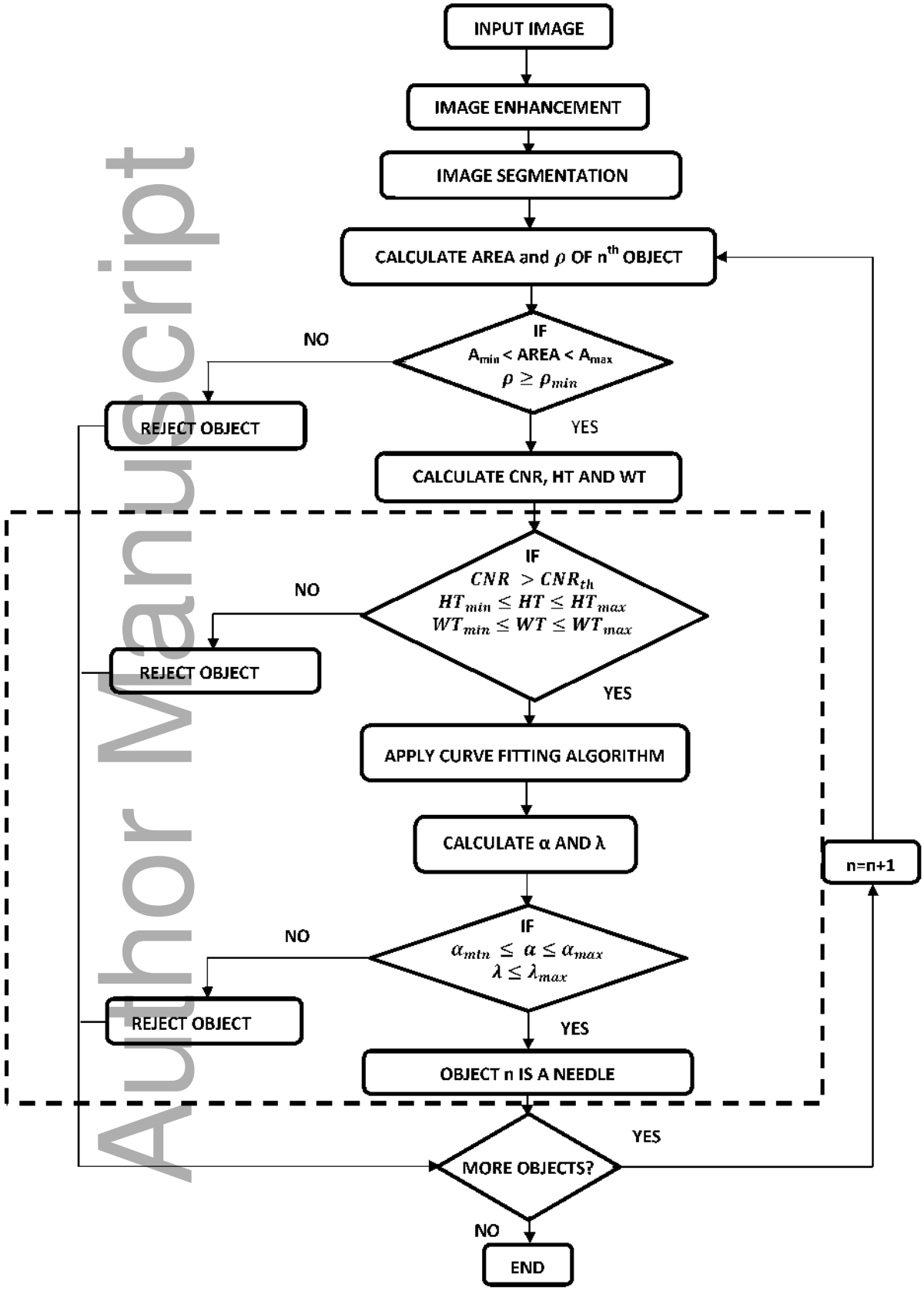
mp_12011_f2b.tif



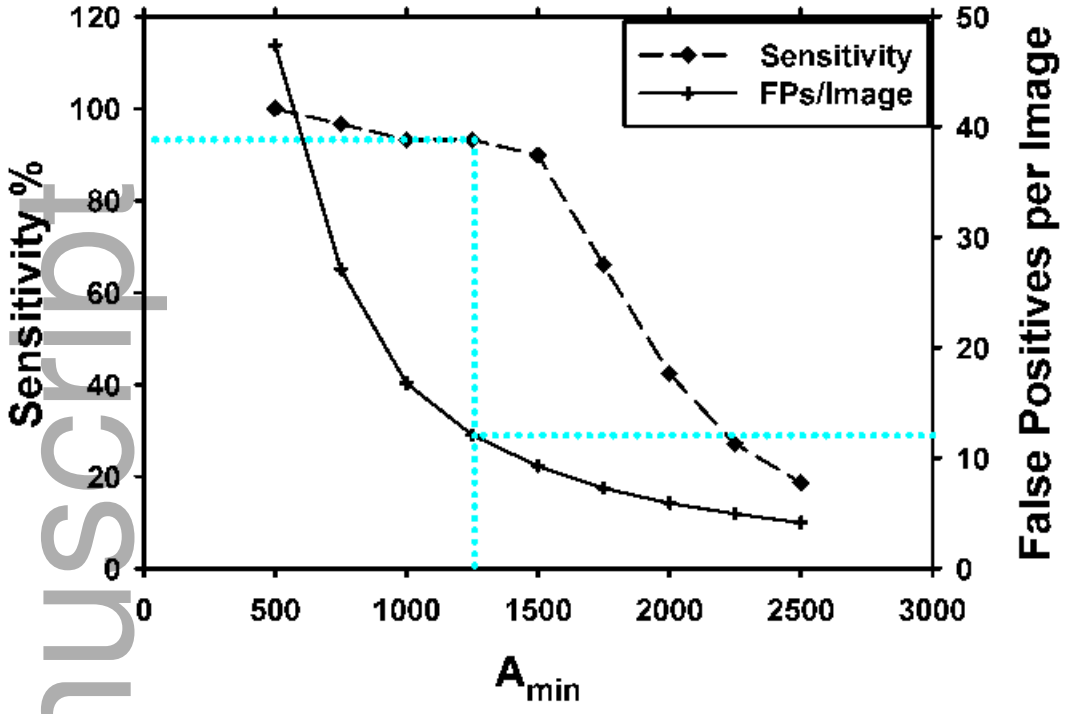
mp_12011_f2c.tif



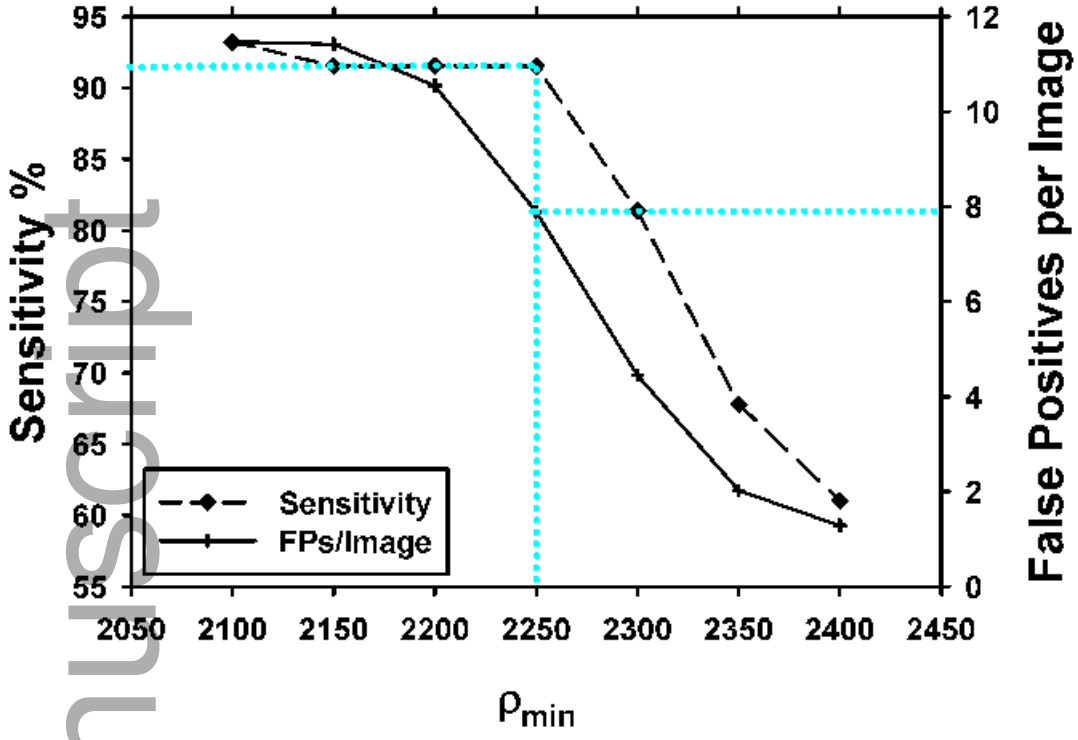
mp_12011_f2d.tif



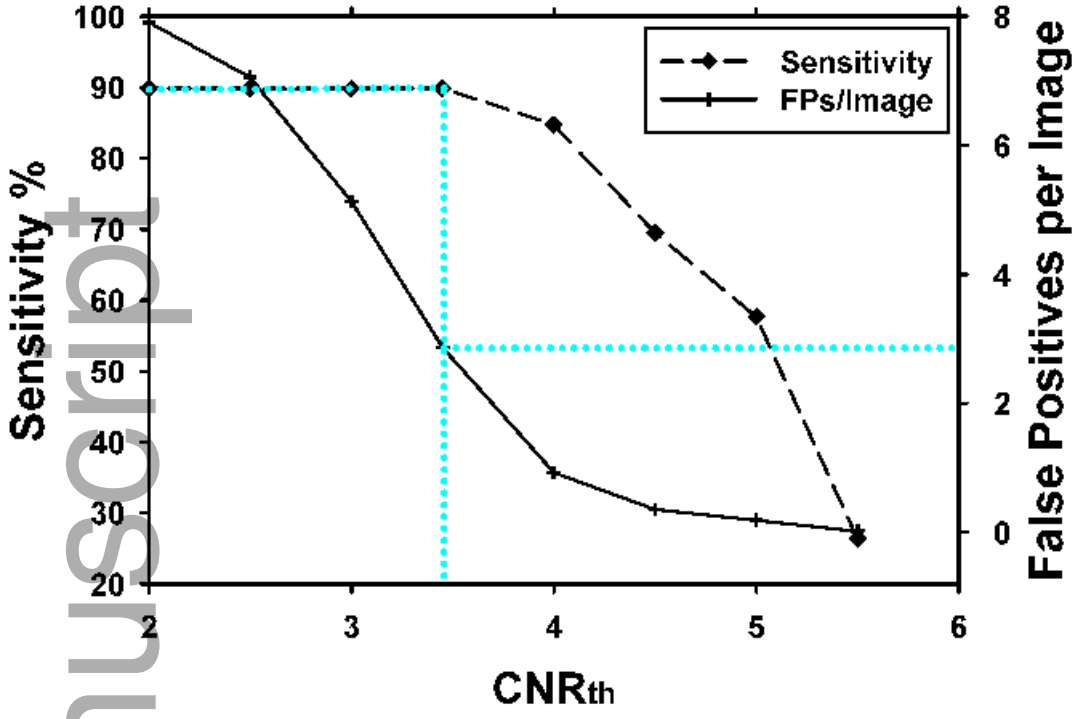
mp_12011_f3.tif



mp_12011_f4.tif

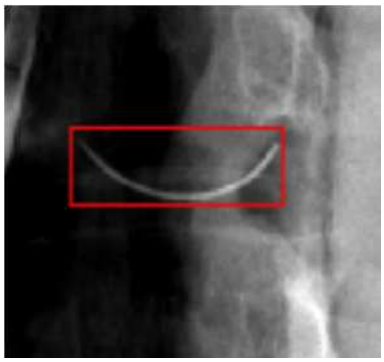


mp_12011_f5.tif

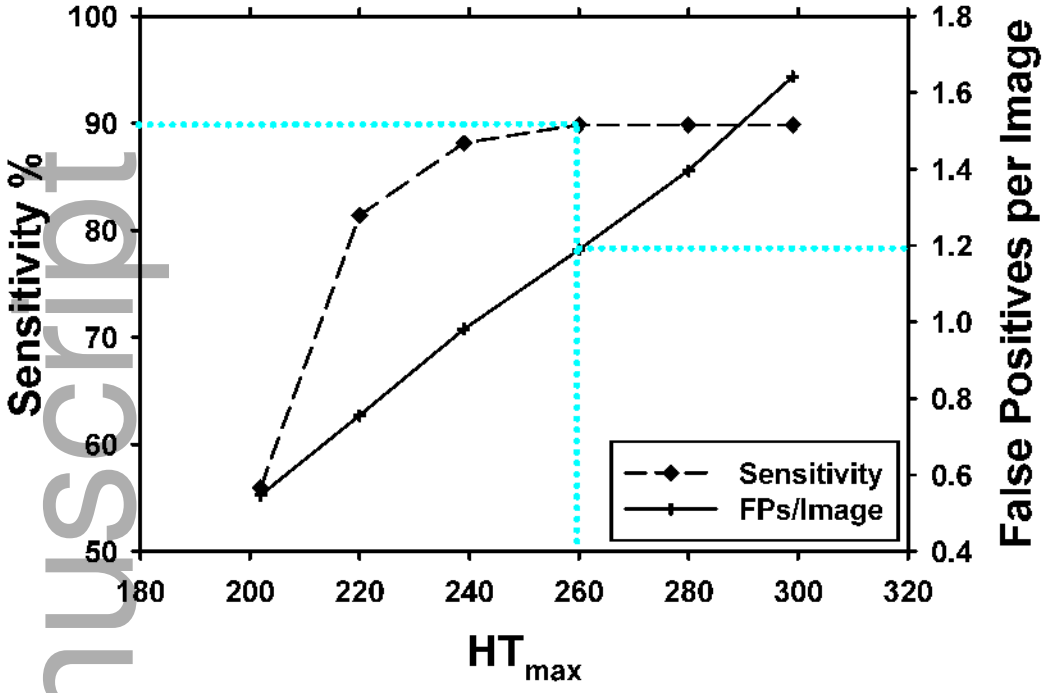


mp_12011_f6.tif

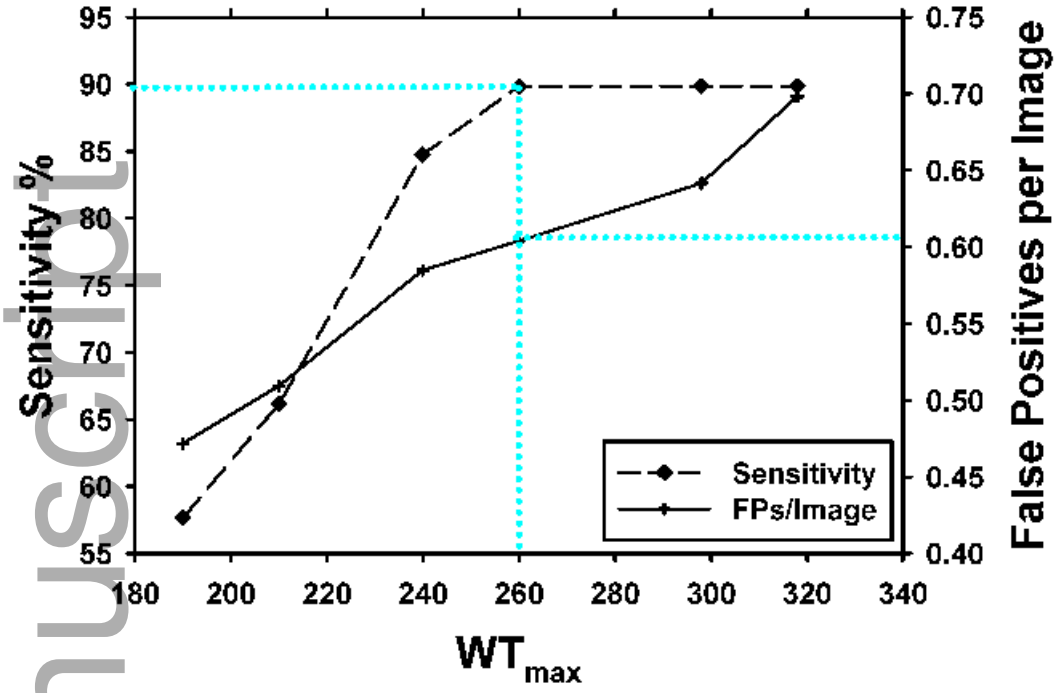
mp_12011_f7a.tif



mp_12011_f7b.tif

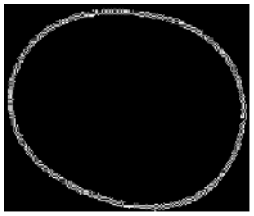


mp_12011_f8a.tif

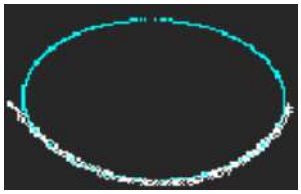


mp_12011_f8b.tif

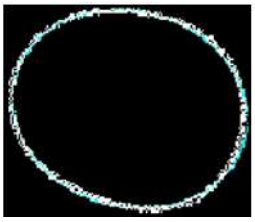
mp_12011_f9a.tif



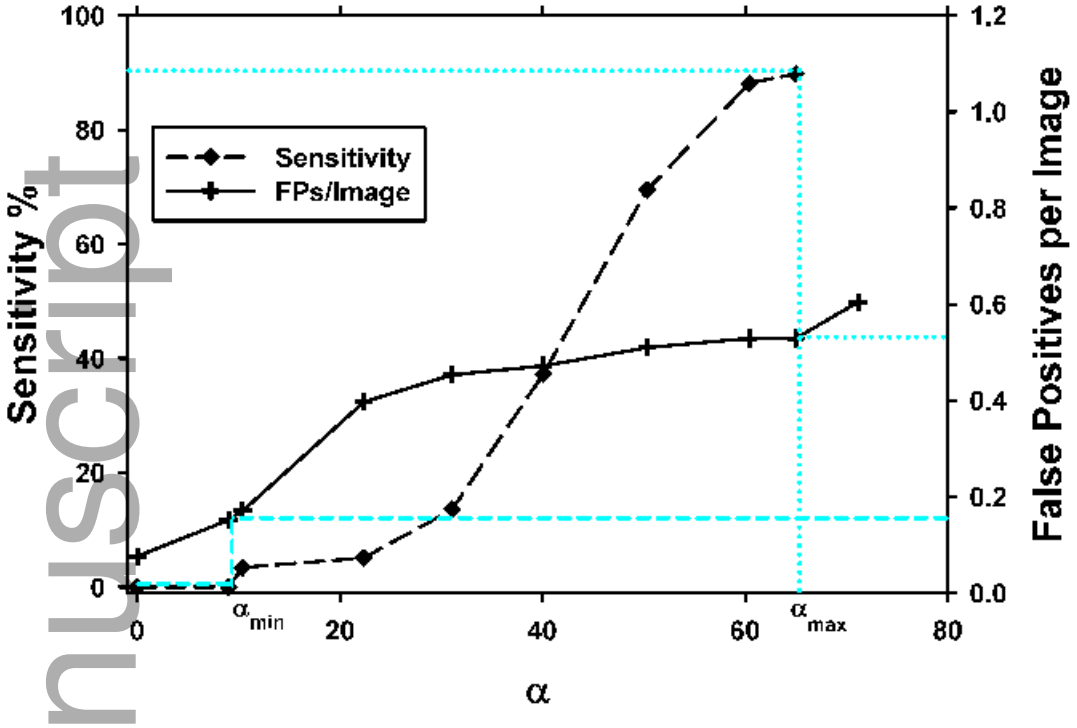
mp_12011_f9b.tif



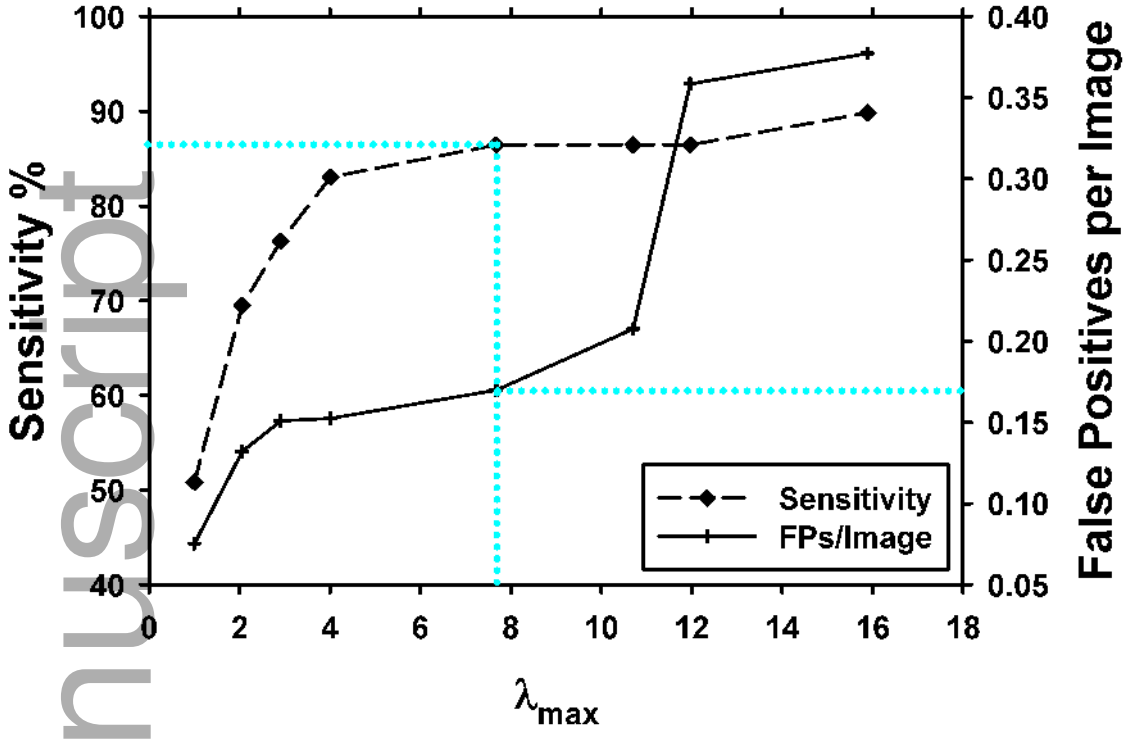
mp_12011_f9c.tif



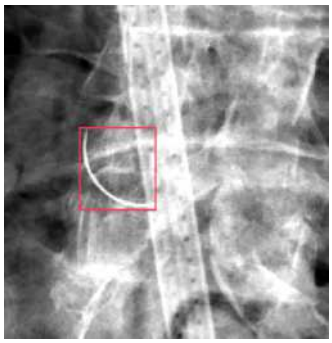
mp_12011_f9d.tif



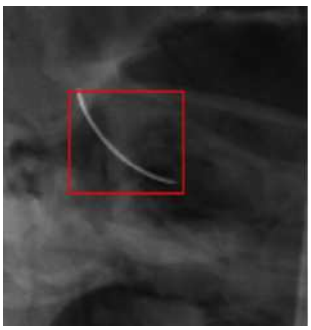
mp_12011_f10a.tif



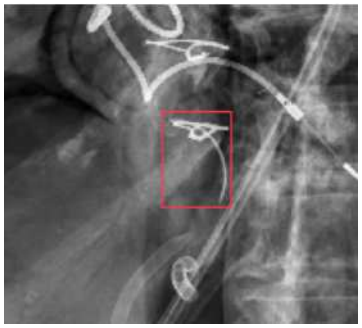
mp_12011_f10b.tif



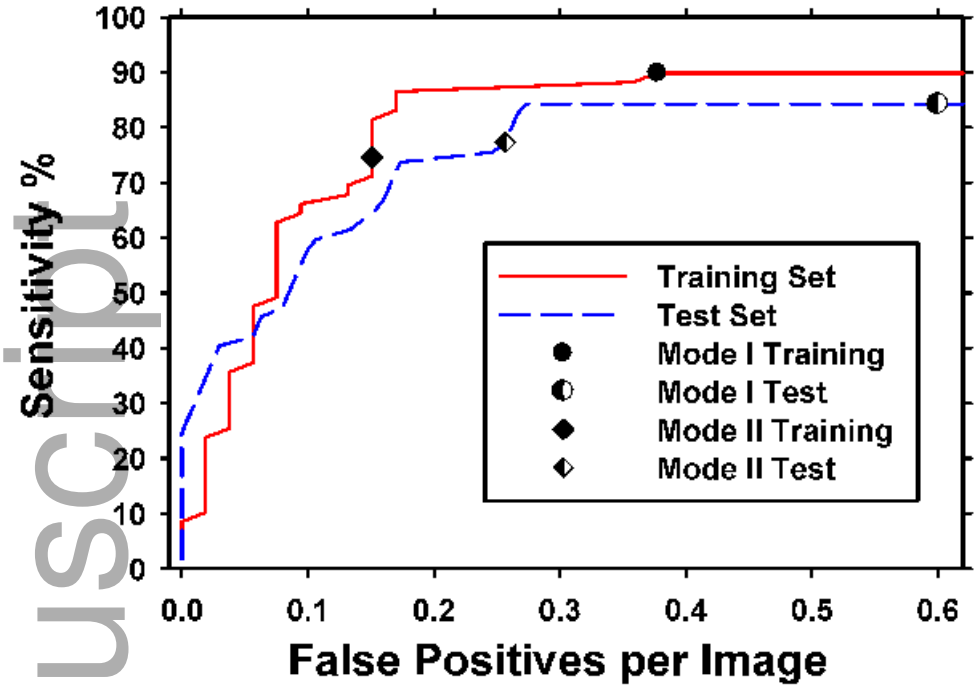
mp_12011_f11a.tif



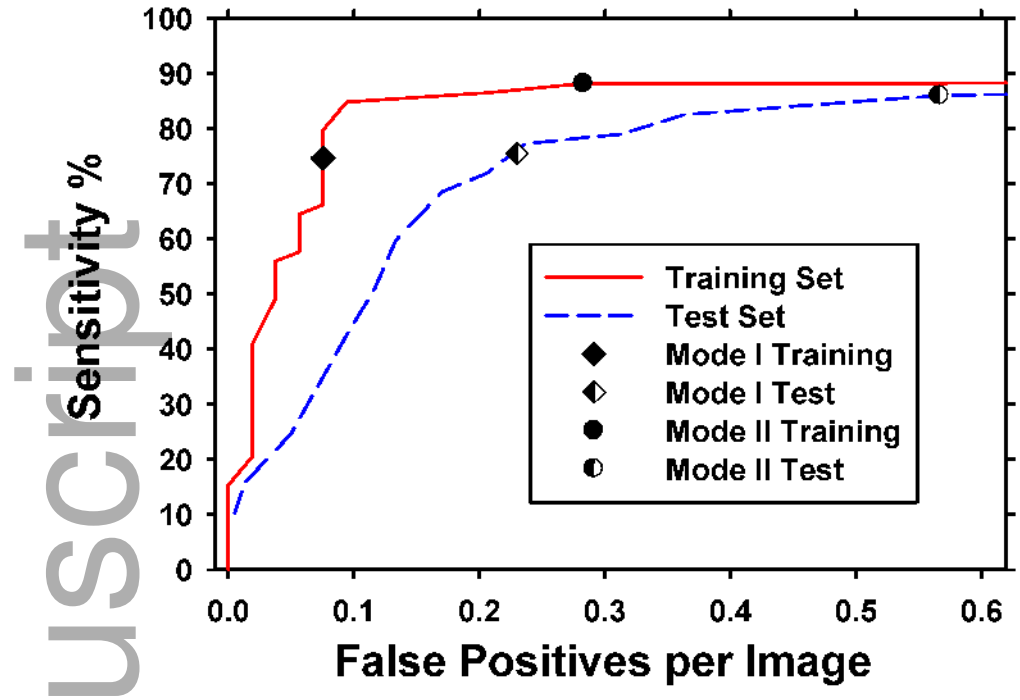
mp_12011_f11b.tif



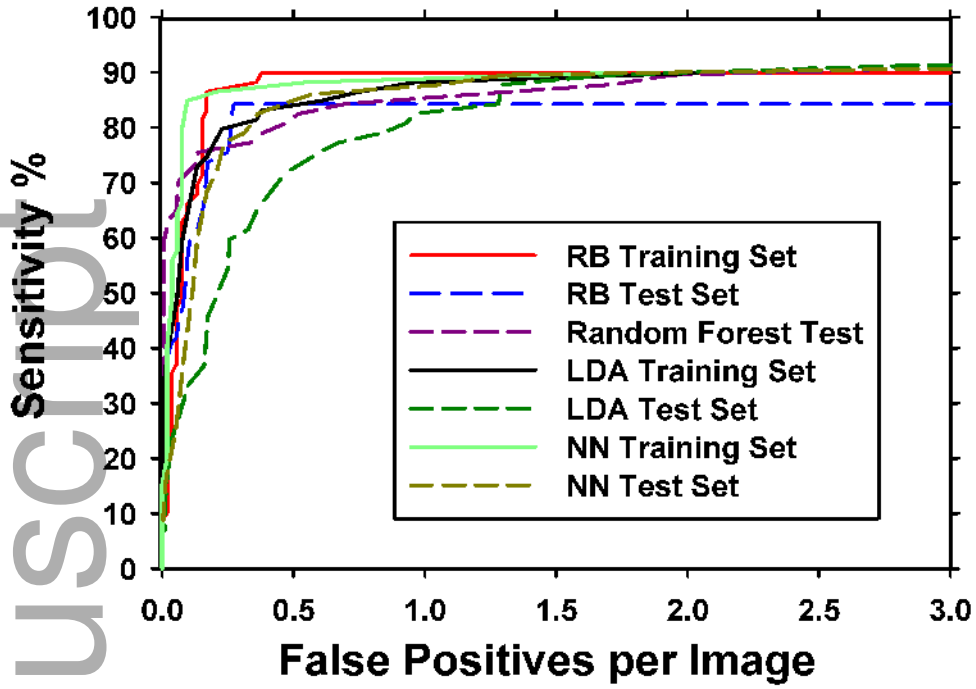
mp_12011_f11c.tif



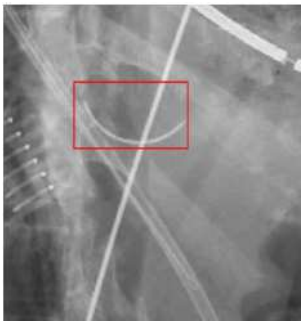
mp_12011_f12.tif



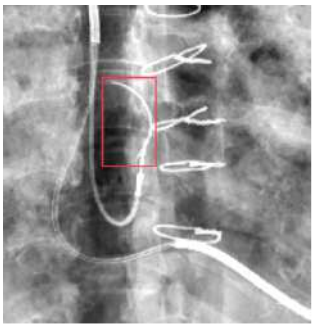
mp_12011_f13.tif



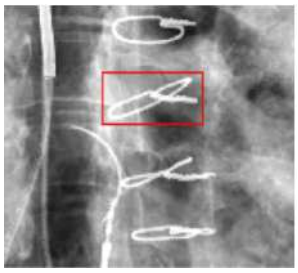
mp_12011_f14.tif



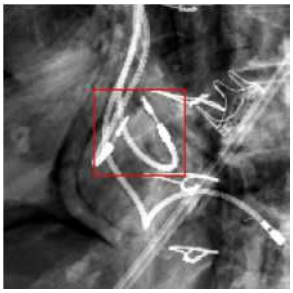
mp_12011_f15a.tif



mp_12011_f15b.tif



mp_12011_f16a.tif



mp_12011_f16b.tif

mp_12011_f16c.tif