A deep-learning model based on fusion images of chest radiography and X-ray sponge images supports human visual characteristics of retained surgical items detection

Kawakubo, Masateru
Department of Health Sciences, Faculty of Medical Sciences, Kyushu University

Waki, Hiroto
Department of Radiological Technology, Hyogo Medical University Hospital

Shirasaka, Takashi
Division of Radiology, Department of Medical Technology, Kyushu University Hospital

Kojima, Tsukasa
Division of Radiology, Department of Medical Technology, Kyushu University Hospital

他

出版情報：International Journal of Computer Assisted Radiology and Surgery, 2022-12-30. Springer
バージョン：
権利関係：

A deep-learning model based on fusion images of chest radiography and X-ray sponge images supports human visual characteristics of retained surgical items detection

[*,**]Masateru Kawakubo[1], [**]Hiroto Waki[2], Takashi Shirasaka[3,4], Tsukasa Kojima[3,5], Ryoji Mikayama[3], Hiroshi Hamasaki[3,5], Hiroshi Akamine[3,5], Toyoyuki Kato[3], Shingo Baba[6], Shin Ushiro[7,8], Kousei Ishigami[6].

1. Department of Health Sciences, Faculty of Medical Sciences, Kyushu University, Fukuoka, Japan
2. Department of Radiological Technology, Hyogo Medical University Hospital, Kobe, Japan
3. Division of Radiology, Department of Medical Technology, Kyushu University Hospital, Fukuoka, Japan
4. Department of Diagnostic Radiology, Faculty of Life Sciences, Kumamoto University, Kumamoto, Japan
5. Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
6. Department of Clinical Radiology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
7. Division of Patient Safety, Kyushu University Hospital, Fukuoka, Japan
8. Japan Council for Quality Health Care, Tokyo, Japan

***Corresponding author**:
Masateru Kawakubo, Ph.D., Assistant Professor
Department of Health Sciences, Faculty of Medical Sciences, Kyushu University
3-1-1 Maidashi, Higashi-ku, Fukuoka-shi, Fukuoka, 812-8582, Japan
Tel: +81-92-642-6728
Fax: +81-92-642-6728
E-mail: kawakubo.masateru.968@m.kyushu-u.ac.jp

** These authors contributed equally to this work as the first authors.

Brief title
Retained surgical item detection

Abstract

Purpose

Although a novel deep learning software was proposed using post-processed images obtained by the fusion between X-ray images of normal post-operative radiography and surgical sponge, the association of the retained surgical item detectability with human visual evaluation has not been sufficiently examined. In this study, we investigated the association of retained surgical item detectability between deep learning and human subjective evaluation.

Methods

A deep learning model was constructed from 2987 training images and 1298 validation images, which were obtained from post-processing of the image fusion between X-ray images of normal postoperative radiography and surgical sponge. Then, another 800 images were used, i.e., 400 with and 400 without surgical sponge. The detection characteristics of retained sponges between the model and a general observer with 10-year clinical experience were analyzed using the receiver operator characteristics.

Results

The following values from the deep learning model and observer were, respectively, derived: cutoff values of probability were 0.37 and 0.45; areas under the curves were 0.87 and 0.76; sensitivity values were 85 % and 61 %; and specificity values were 73 % and 92 %.

Conclusion

For the detection of surgical sponges, we concluded that the deep learning model has higher sensitivity while the human observer has higher specificity. These characteristics indicate that the deep learning system that is complementary to humans could support the clinical workflow in operation rooms for prevention of retained surgical items.

Keywords

Retained surgical item, Deep learning, Detection characteristic, Image fusion processing

Introduction

Retained surgical items (RSIs), also called "gossypiboma," can cause infection or damage to internal organs. To resolve such complications, operations are often required [1-5]. So far, the healthcare cost of RSIs has been reported to be more than $200,000 per incident [6]. Although the frequency of residual foreign body is approximately 1 in 10,000, some patient deaths due to RSI have been reported. Thus, it should be prevented because it is caused by serious human error in the background of risk factors such as higher amount of blood loss, participation of other medical teams in operations, and incorrect surgical count [7]. To prevent human error ~~caused by such unusual crisis situations~~, counts of surgical materials by operating room staff and visual evaluation with post-operative radiography are standard procedures in operating rooms [8]. Despite these precautions, RSIs still occur at least once per year in major hospitals [9]. This indicates the limit of the RSI detection procedure that depends on human subjective factors. Therefore, the precautions for RSI are required in the operating room independently of human subjective detection.

A computer-aided diagnosis (CAD) system for counting sponges using barcodes has also been reported [7]. Further, over the last 10 years, artificial intelligence-based image recognition technology has been applied rapidly to the medical field. In particular, deep learning-based models in chest radiography are reported for detection and stratification of pneumonia, as well as for detection of peripherally inserted central venous catheters [10, 11]. Moreover, Yamaguchi et al. recently proposed a novel CAD software for prevention of RSI with the use of deep learning [12]. In this study, a deep learning model was constructed using post-processed images obtained by the fusion between X-ray images of normal post-operative radiography and surgical sponge; the model showed 97.7% sensitivity and 83.8% specificity in the detection of retained surgical sponges. This result is considered to make a significant contribution to the development of an objective RSI detection system, but the association of RSI detectability with human visual evaluation has not been sufficiently examined. Therefore, the purpose of the present research study is to construct a deep learning-based RSI detection model using post-processed fusion images between standard chest radiography and

surgical sponge, and clarify the association between their RSI detectability and subjective evaluation by humans.

Methods

Training and validation data for deep learning

The training and validation data for deep learning was set by the fusion of images between standard chest radiography and surgical sponge X-ray (Fig. 1).

51 patients were enrolled for standard chest radiography. All cases were imaged posterior to anterior in a standing position, and any artificial objects such as cardiac devices or spinal instruments were removed. The imaging parameters were as follows: (1)source to detector distance = 200 cm; (2) tube voltage = 120 kVp; and (3) tube current = 160 mA. The exposure time was automatically controlled to keep the dose constant on the X-ray detector. The bit depth for all datasets was converted from 16 to 8 bits of portable network graphic files. The window level and width were rescaled to 256 gradients with the minimum and maximum values for each image intensity. The matrix size of all images was downsampled from 2816 × 2373 to 1024 × 1024 pixels through nearest neighbor interpolation. Further, the image was cropped into a 128 × 128 area with a 64-pixel overlap. However, the peripheral part of the chest radiograph was excluded because it is unlikely to be the chest surgical field. Consequently, the total number of cropped images was 8619 with 169 images per case (Fig. 1a).

The sponge image was the X-ray image of three surgical sponges soaked in water that were rolled/stretched and sandwiched between acrylic phantoms with a thickness of 10 cm. The imaging parameters were as follows: (1) source to detector distance = 100 cm; (2) tube voltage = 66 kVp; (3) tube current-time product = 1.6 mAs; (4) matrix size = 2832 × 2836 pixels. The matrix size of all images was downsampled to 1024 × 1024 pixels through nearest neighbor interpolation. Further, images of 128 × 128 pixels centered on the 150 points that were manually determined near the high X-ray absorption fibers of the surgical sponge were created. The 128 × 128 pixel area of only the acrylic phantom was used as background image. Consequently, 150 surgical sponge images for deep learning were obtained by subtracting the background image from the images of high X-ray absorption fibers (Fig. 1b). With and without surgical sponge images for deep learning were obtained

based on the flowchart shown in Fig. 1c. Randomly selected cropped radiography of 4285 images (out of 8619) were used for deep learning as images without surgical sponge. The remaining 4334 cropped images were fused with 150 surgical sponge images through image processing. First, one surgical sponge image was randomly selected for each cropped image. Next, the surgical sponge image was randomly rotated by 0, 90, 180, and 270 degrees. Finally, a cropped chest radiography with surgical sponge is obtained by the fusion using the following formula:

$$FX = CX + 0.3GX \qquad (1)$$

where FX is a cropped chest radiography with surgical sponge, CX is a cropped chest radiography, and GX is a surgical sponge image.

All images were randomly assigned to groups for training (with sponge: 3046 and without sponge: 2987) and validation (with sponge: 1288 and without sponge: 1298). To develop an established deep learning model, we used a deep learning platform called Neural Network Console (Sony Network Communications, Tokyo, Japan) and a specialized graphic processing unit (TITAN RTX 1080; Nvidia, Santa Clara, CA, USA) [13]. Consequently, the classification model for detecting surgical sponge was generated using the 100-epoch modified LeNet architecture with a batch size of 64, and Adam optimizer (shown in Fig. 2).

Detectability evaluation of surgical sponge with deep learning model

The test data for the deep learning model was obtained from the chest radiography dataset of National Institutes of Health (NIH) using the same workflow as the training and validation datasets [14]. First, 100 cases of chest radiography with no lesions or artifacts in the image were selected from the NIH dataset. Then, the matrix size of all images was resized to 1024 × 1024 pixels through nearest neighbor interpolation. From each case of chest radiography, images of 128 × 128 pixels centered on arbitrarily determined coordinates of each of 8 regions were cropped as shown in Fig. 3.

A total of 800 images in 100 cases were created. Randomly selected cropped radiography of 400 images (out of 800) were used for the testing phase of the deep learning model as the images without surgical sponge. The remaining 400 (out of 800) images were fused with 150 surgical sponge images

that were used for deep learning by the same workflow for creating the training and validation datasets. The 800 test images were input to the deep learning model for image classification, and output the probability of the presence of surgical sponge; the probability ranged from 0 (without sponge) to 1 (with sponge). The deep learning model evaluated the performance of surgical sponge detection by the receiver operator characteristic (ROC) analysis, which was performed on the detectability of surgical sponge of all 800 images and each of 8 regions of chest radiography.

Visual detectability evaluation of surgical sponge by observer

Two observers evaluated independently the detection performance of the surgical sponge on chest radiography. One of them was experienced with over 30 years experience of clinical scans of post-operative chest radiography as operating room staff. The other one was the general observer with approximately 10 years experience in a similar position. For the observer evaluation, 100 cases of chest radiography with the presence of surgical sponges were obtained by the following process; 400 cropped images for evaluation of the deep learning model were replaced with the same regions of the original NIH chest radiography (Fig. 4).

The continuous probability of the presence or absence of surgical sponges in each of the 8 regions was recorded (the probability ranged from 0 (without sponge) to 1 (with sponge)). It was given to the observers as prior information for the evaluation that the chest radiography was segmented into 8 regions and the presence or absence of high X-ray absorbing fibers of surgical sponge should be classified in each region. 100 images were displayed one by one for the two observers independently under a randomly sorted order with the general image contrast for chest on a general purpose 50-inch display. The observation time for each region was less than 30 sec, and no image processing was performed (such as windowing, zoom, or frequency processing).

Results

The areas under the ROC curves for detection of surgical sponge in all 800 images by the deep learning model, experienced, and general observer are shown in Fig. 5. The ROC values from the

deep learning model, experienced, and general observer were derived as follows: (1) cutoff values of probability were 0.37, 0.60, and 0.45, respectively; (2) areas under the ROC curves were 0.87, 0.89, and 0.76, respectively; (3) sensitivity values were 85%, 82%, and 61%, respectively; and (4) specificity values were 73%, 88%, and 92%, respectively.

Table 1 shows the ROC values at each of the 8 segments of chest radiography from the deep learning model, experienced, and general observer. For the lung regions (i.e., segments #1 to #5), the areas under the ROC curves and sensitivities of the deep learning model were greater than those of the general observer. For the heart (segment #6) and subdiaphragmatic (segments #7 and #8) regions, the areas under the ROC curves from the three methods were more than 0.9; thus, they were all equally high. For all regions excluding segment #6, the specificities from the deep learning model were smaller than those from the human observers.

Discussion

In this research study, we compared the detectability of surgical sponge in chest radiography of the deep learning model, experienced and general observer with ROC analysis. The areas under the ROC curves from the deep learning model and experienced observer were equivalently wider than that from the general observer. The deep learning model tended to have higher sensitivity for the detection of surgical sponge than the human observers. The human observers tended to have higher specificities for the detection of surgical sponge than the deep learning model. These results show that the deep learning model has high diagnostic ability in the presence of sponge in an image, whereas humans have high diagnostic ability in the absence of sponge in an image. In other words, the deep learning model reduces the rate of RSIs missed by human interpretation. Therefore, we believe that this research work has revealed that the deep learning model has complementary characteristics to humans for the detection of surgical items, and has shown its potential to reduce RSI.

From the obtained results, the reason for the observers' low detectability of surgical sponges in the lung region (segments #1 to #5) appears to be the different levels of experience. Medical staff (such as surgeons and nurses) who assess post-surgery radiographs in the operating room are not

always familiar with image interpretation. In contrast, there are cases where post-operative image verification is sometimes performed without sufficient experience. For the medical staff in the operating room who do not specialize in diagnostic imaging, complexities can occur when evaluating lung regions. This is because the lung region is the one where bones with high X-ray absorption and lungs with low X-ray absorption coexist. Therefore, to detect lung lesions such as nodules, methods to perform X-ray energy subtraction imaging and bone removal processing using deep learning have been actively introduced [15-19]. Further, a recent study concluded that to prevent RSI events using a proactive multimodal approach that focuses on improving team communication and institutional support system, standardizing reports and implementing new technologies are most effective [20]. Our proposed method is a different approach from the ones aimed at improving human detection sensitivity using image processing. A recent publication reported that 57%, 35%, and <30% of surgeons, nurses/technologists, and anesthesiologists continually use X-ray for RSI detection, respectively, and that 26–50% of them considered X-ray to have poor effectiveness for RSI detection [21]. The results of this study indicate that deep learning has the potential to supplement an observer's experience in detecting surgical sponges. Therefore, the deep learning-based system proposed in this study is expected to improve the effectiveness of X-ray images in RSI detection and promote the use of X-ray images by medical staff in the operating room, thereby contributing to efficient surgical operation. Our proposed system expects to alert by the positive detection to the medical staff in operating room to any oversights and encourages them to check the images more carefully for detecting RSI. For the future, deploying deep learning system might lead to new clinical workflow which to seek the opinions of radiologists. For the heart (segment #6) and subdiaphragmatic (segments #7 and #8) regions, the detectability of surgical sponge by the deep learning model and observers was more than 0.9 (i.e., equally high for all three cases). This is because the contrast of the X-ray images in these areas was relatively simple compared to the lungs. In other words, both humans and the deep learning model were able to easily detect the X-ray absorption sponge fibers against organs such as the heart and liver, from the background signal in a wide and uniform area. Moreover, the results of this research may indicate the potential for automatic detection of surgical items in these

areas. The burden on medical staff is increasing due to the increase in the number of operations and complexity of surgical procedures. Thus, automatic detection of surgical items by deep learning can be expected to reduce that burden and prevent the occurrence of RSI accidents.

There are some limitations to this study. First, the surgical sponge images were fused with standard chest radiography rather than post-operative chest radiography. The purpose of this study was to investigate the characteristics of the detection of surgical sponge through deep learning and human observers. We consider that it would be a bias for observer evaluation to use post-operative images obtained at the hospital where they work. However, it was not possible to have access to a database of post-operative images, and thus, we used the generally accessible NIH chest radiography database. Second, this study has been limited to the detection of surgical sponges in the chest area. In the future, we should plan to further investigate the detection of several surgical items in post-operative images.

Conclusions

In this research, we constructed a deep learning-based RSI detection model using post-processed fusion images between standard chest radiography and surgical sponge. Then, we investigated the characteristics of the detection of surgical sponge with the deep learning model and human observers. It was clarified from the results that deep learning has complementary characteristics to humans for surgical sponge detection, and could support RSI detection by humans.

Statements and Declarations

Ethics approval and consent to participate

This retrospective study was approved by the institutional review board in Kyushu University and conducted in accordance with the 1964 Declaration of Helsinki.

Consent for publication

The review board waived the requirement for written informed consent. The manuscript does not contain any individual person's data in any form.

Availability of data and materials

Not applicable.

Competing interests

This study has a pending patent in Japan (2020–094615).

Funding

This work was supported by the Japanese Grant of The Clinical Research Promotion Foundation (2020).

Author contributions

Study conception and design: Masateru Kawakubo, Hiroto Waki, Takashi Shirasaka, Tsukasa Kojima, Ryoji Mikayama, Hiroshi Hmasaki;

Acquisition of data: Masateru Kawakubo, Hiroto Waki, Hiroshi Akamine, Toyoyuki Kato, Shingo Baba;

Analysis and interpretation of data: Masateru Kawakubo, Hiroto Waki, Takashi Shirasaka, Tsukasa Kojima, Ryoji Mikayama, Hiroshi Hmasaki;

Drafting of manuscript: Masateru Kawakubo, Takashi Shirasaka, Tsukasa Kojima, Ryoji Mikayama, Hiroshi Hmasaki;

Critical revision: Masateru Kawakubo, Hiroto Waki, Toyoyuki Kato, Dhingo Baba, Shin Ushiro, Kousei Ishigami.

Acknowledgements

We thank Editage (www.editage.com) for the English language editing.

References

1. Rajagopal A, Martin J (2002) Gossypiboma -"A Surgeon's Legacy": Report of a Case and Review of the Literature. Dis Colon Rectum 45:119–120. https://doi.org/10.1007/s10350-004-6124-1

2. Stawicki SP, Evans DC, Cipolla J, Seamon MJ, Lukaszczyk JJ, Prosciak MP, Torigian DA, Doraiswamy VA, Yazzie NP, Gunter Jr. OL, Steinberg SM (2009) Retained Surgical Foreign Bodies: A Comprehensive Review of Risks and Preventive Strategies. Scand J Surg 98:8–17. https://doi.org/10.1177/145749690909800103

3. Gencosmanoglu R, Inceoglu R (2003) An unusual cause of small bowel obstruction: Gossypiboma—case report. BMC Surg 3:6. https://doi.org/10.1186/1471-2482-3-6

4. Jason RS, Chisolm A, Lubetsky HW (1979) Retained surgical sponge simulating a pancreatic mass. J Natl Med Assoc 71:501–503.

5. Sun HS, Chen SL, Kuo CC, Wang S, Kao Y (2007) Gossypiboma—Retained Surgical Sponge. J Chin Med Assoc 70:511–513. https://doi.org/10.1016/S1726-4901(08)70051-0

6. Regenbogen SE, Greenberg CC, Resch SC, Kollengode A, Cima RR, Zinner MJ, Gawande AA (2009) Prevention of retained surgical sponges: A decision-analytic model predicting relative cost-effectiveness. Surgery 145:527–535. https://doi.org/10.1016/j.surg.2009.01.011

7. Moffatt-Bruce SD, Cook CH, Steinberg SM, Stawicki SP (2014) Risk factors for retained surgical items: a meta-analysis and proposed risk stratification system. J Surg Res 190:429–436. https://doi.org/10.1016/j.jss.2014.05.044

8. Cima RR, Kollengode A, Garnatz J, Storsveen A, Weisbrod C, Deschamps C (2008) Incidence and Characteristics of Potential and Actual Retained Foreign Object Events in Surgical Patients. J Am Coll Surg 207:80–87. https://doi.org/10.1016/j.jamcollsurg.2007.12.047

9. Mahran MA, Toeima E, Morris EP (2013) The recurring problem of retained swabs and instruments. Best Pract Res Clin Obstet Gynaecol 27:489-495. https://doi.org/10.1016/j.bpobgyn.2013.03.001

10. Yu D, Zhang K, Huang L, Zhao B, Zhang X, Guo X, Li M, Gu Z, Fu G, Hu M, Ping Y, Sheng Y, Liu Z, Hu X, Zhao R (2020) Detection of peripherally inserted central catheter (PICC) in chest X-ray

images: A multi-task deep learning model. Comput Methods Programs Biomed 197:105674. https://doi.org/10.1016/j.cmpb.2020.105674

11. El Asnaoui K, Chawki Y, Idri A (2021) Automated Methods for Detection and Classification Pneumonia Based on X-Ray Images Using Deep Learning. Stud Big Data. Springer, Cham:257–284. https://doi.org/10.1007/978-3-030-74575-2_14
12. Yamaguchi S, Soyama A, Ono S, Hamauzu S, Yamada M, Fukuda T, Hidaka M, Tsurumoto T, Uetani M, Eguchi S (2021) Novel Computer-Aided Diagnosis Software for the Prevention of Retained Surgical Items. J Am Coll Surg 233:686–696. https://doi.org/10.1016/j.jamcollsurg.2021.08.689
13. Sony Neural Network Console. https://dl.sony.com/ Accessed 4 Dec 2020
14. Wang X, Peng Y, Lu Z, Lu Z, Bagheri M, Summers RM (2017) ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly- Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR 3462–3471. https://doi.org/10.1109/CVPR.2017.369
15. Vock P, Szucs-Farkas Z (2009) Dual energy subtraction: Principles and clinical applications. Eur J Radiol 72:231–237. https://doi.org/10.1016/j.ejrad.2009.03.046
16. Kuhlman JE, Collins J, Brooks GN, Yandow DR, Broderick LS (2006) Dual-Energy Subtraction Chest Radiography: What to Look for beyond Calcified Nodules. RadioGraphics 26:79–92. https://doi.org/10.1148/rg.261055034
17. Yang W, Chen Y, Liu Y, Zhong L, Qin G, Lu Z, Feng Q, Chen W (2017) Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. Med Image Anal 35:421–433. https://doi.org/10.1016/j.media.2016.08.004
18. Zhou Z, Zhou L, Shen K (2020) Dilated conditional GAN for bone suppression in chest radiographs with enforced semantic features. Med Phys 47:6207–6215. https://doi.org/10.1002/mp.14371
19. Han L, Lyu Y, Peng C, Zhou SK (2022) GAN-based disentanglement learning for chest X-ray rib suppression. Med Image Anal 77:102369. https://doi.org/10.1016/j.media.2022.102369
20. Weprin S, Crocerossa F, Meyer D, Maddra K, Valancy D, Osardu R, Kang HS, Moore RH,

Carbonara U, J Kim F, Autorino R. Risk factors and preventive strategies for unintentionally retained surgical sharps: a systematic review. Patient Saf Surg. 2021 Jul 12;15(1):24. doi: 10.1186/s13037-021-00297-3. PMID: 34253246; PMCID: PMC8276389.

21. Weprin SA, Meyer D, Li R, Carbonara U, Crocerossa F, Kim FJ, Autorino R, Speich JE, Klausner AP. Incidence and OR team awareness of "near-miss" and retained surgical sharps: a national survey on United States operating rooms. Patient Saf Surg. 2021 Apr 3;15(1):14. doi: 10.1186/s13037-021-00287-5. PMID: 33812376; PMCID: PMC8019169.

Table captions

Table 1. Receiver operator characteristic values at each segment of chest radiography

(AUC: area under the curve).

Figure captions

Fig. 1 Dataset construction for deep learning by fusion of images between standard chest radiography and surgical sponge X-ray

(a) The image was cropped in a 128 × 128 pixel area with a 64-pixel overlap (magenta and green matrices). However, the peripheral part of the chest radiograph was excluded because it is unlikely to be the chest surgical field (dark gray matrices). The total number of cropped images was 8619 with 169 images per case (lower). (b) Images of 128 × 128 pixels centered on the 150 points that are manually determined near the high X-ray absorption fibers of the surgical sponge were created (magenta areas). The 128 × 128 pixel area of only the acrylic phantom was used as background image (green areas). A total of 150 surgical sponge images for deep learning were obtained by subtracting the background image from the images of high X-ray absorption fibers (lower). (c) Randomly selected cropped radiography of 4285 (out of 8619) were used for deep learning as the images without surgical sponge (black flow). The other 4334 (out of 8619) cropped images were fused with 150 surgical sponge images by image processing (red flow).

Fig. 2 Modified LeNet architecture for detection model of surgical sponge

Fig. 3 Segments of the chest radiography

All chest radiographs were segmented into 8 regions. Segments #1 to #5 are lung regions, segment #6 is heart region, and segments #7 and #8 are subdiaphragmatic regions.

Fig. 4 Radiography of chest and surgical sponge for observer evaluations

The area in original chest radiography (white dashed matrix) is replaced by the same area of that with sponge image (green matrix). The size of that with sponge image is 128 × 128 pixels. The sponge fiber is described by green lines for readability, but the actual color is grayscale.

Fig. 5 Receiver operator characteristics curves for detection of surgical sponges

The curves in the color of red, blue, and black are indicated as the detectability of surgical sponges in all regions by the deep learning model, experienced, and general observer, respectively.

Figure

Fig. 1

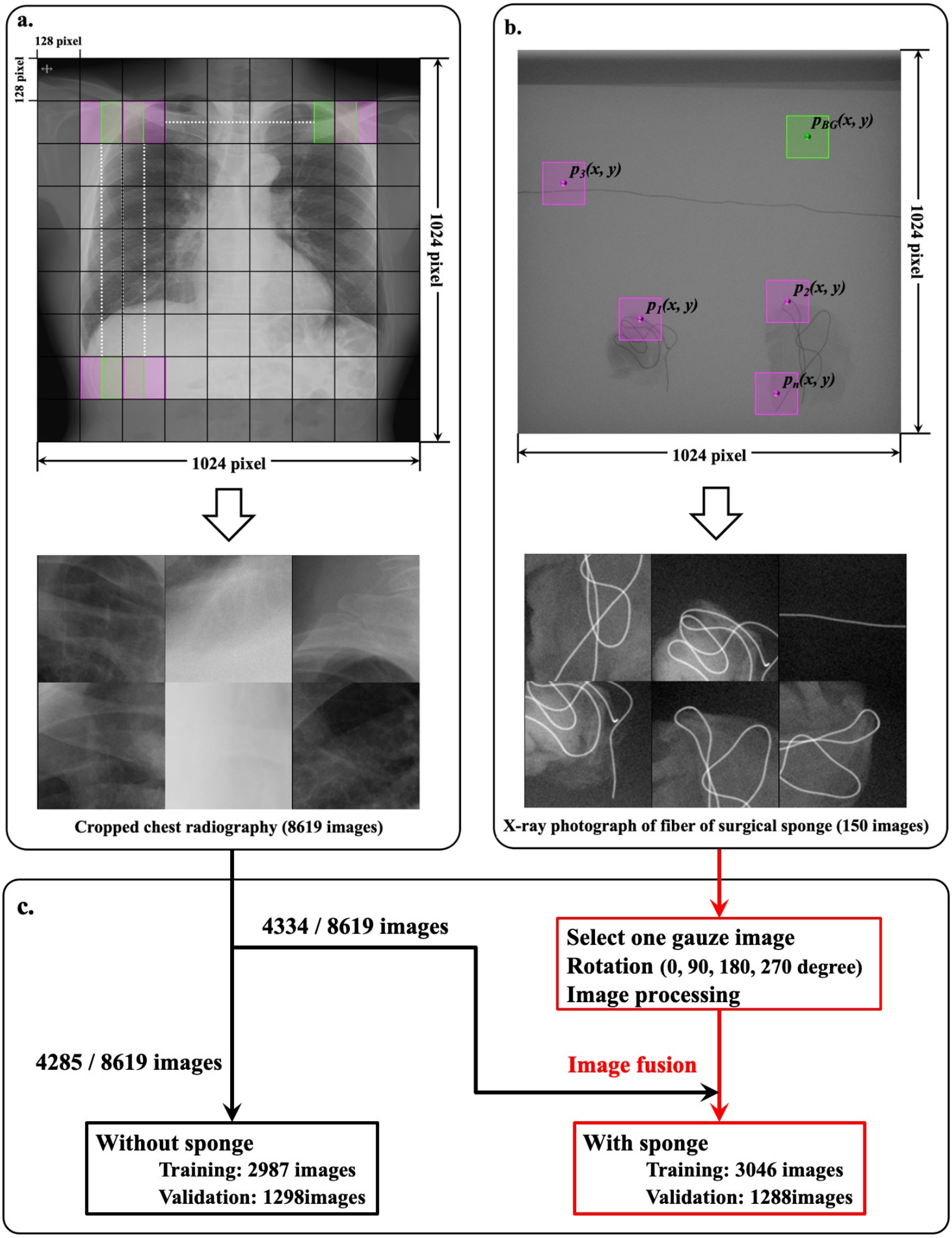

a.
128 pixel
128 pixel
1024 pixel
1024 pixel
Cropped chest radiography (8619 images)
b.
$p_{BG}(x, y)$
$p_3(x, y)$
$p_2(x, y)$
$p_1(x, y)$
$p_n(x, y)$
1024 pixel
1024 pixel
X-ray photograph of fiber of surgical sponge (150 images)
c.
4334 / 8619 images
4285 / 8619 images
Select one gauze image
Rotation (0, 90, 180, 270 degree)
Image processing
Image fusion
Without sponge
Training: 2987 images
Validation: 1298images
With sponge
Training: 3046 images
Validation: 1288images

Fig. 2

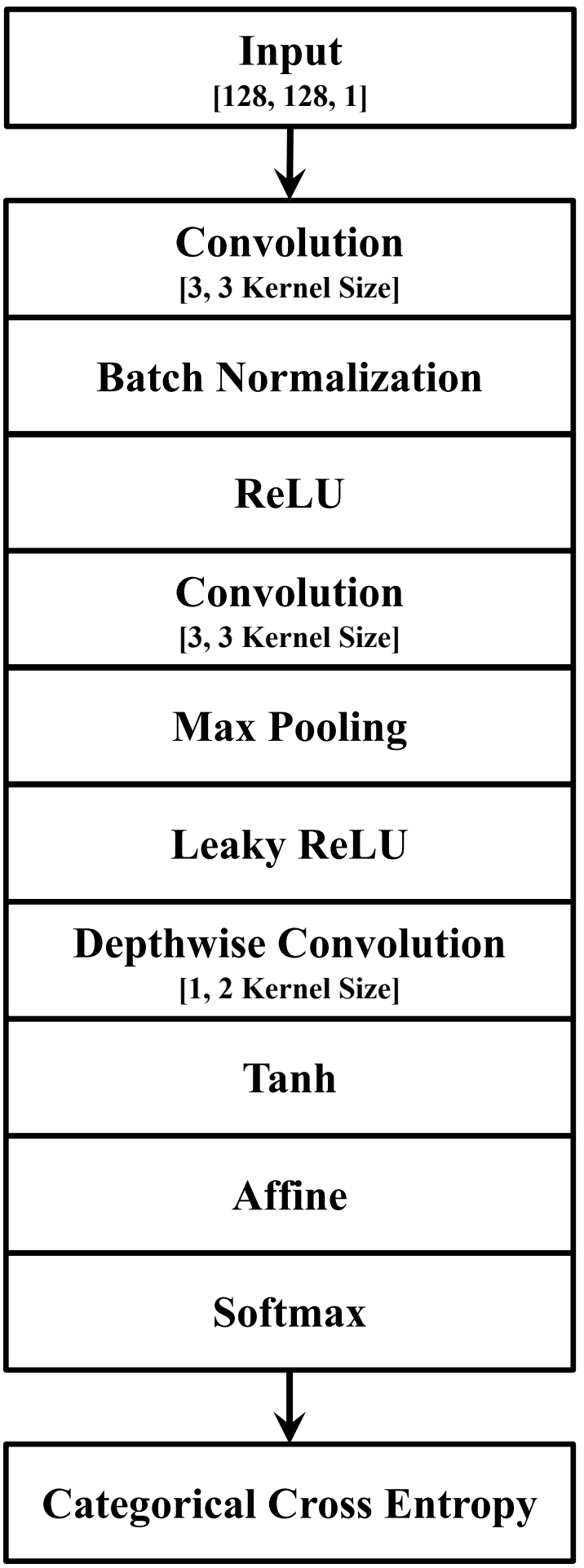

Input
[128, 128, 1]
Convolution
[3, 3 Kernel Size]
Batch Normalization
ReLU
Convolution
[3, 3 Kernel Size]
Max Pooling
Leaky ReLU
Depthwise Convolution
[1, 2 Kernel Size]
Tanh
Affine
Softmax
Categorical Cross Entropy

Fig. 3

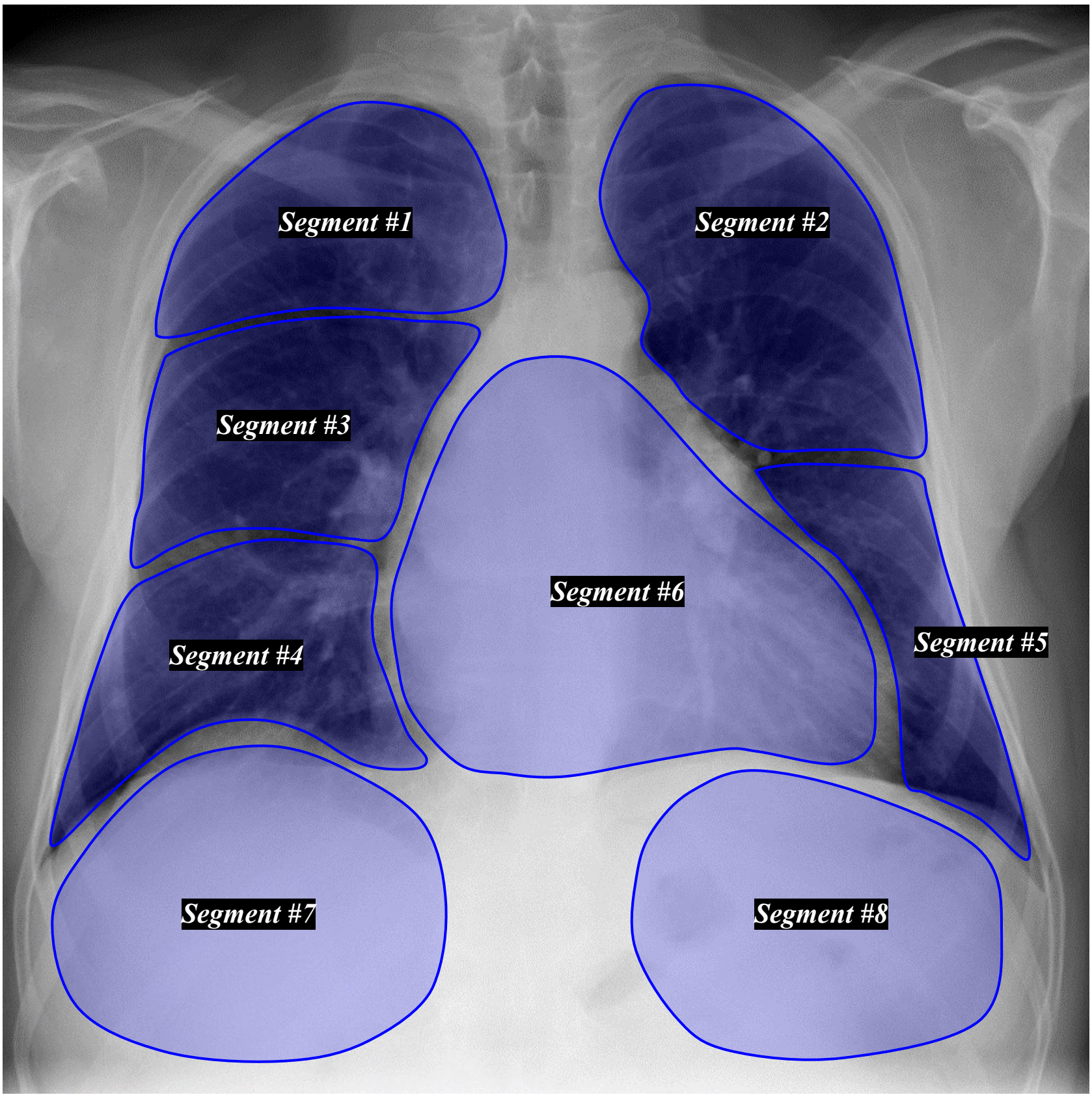

Segment #1
Segment #2
Segment #3
Segment #6
Segment #4
Segment #5
Segment #7
Segment #8

Fig. 4

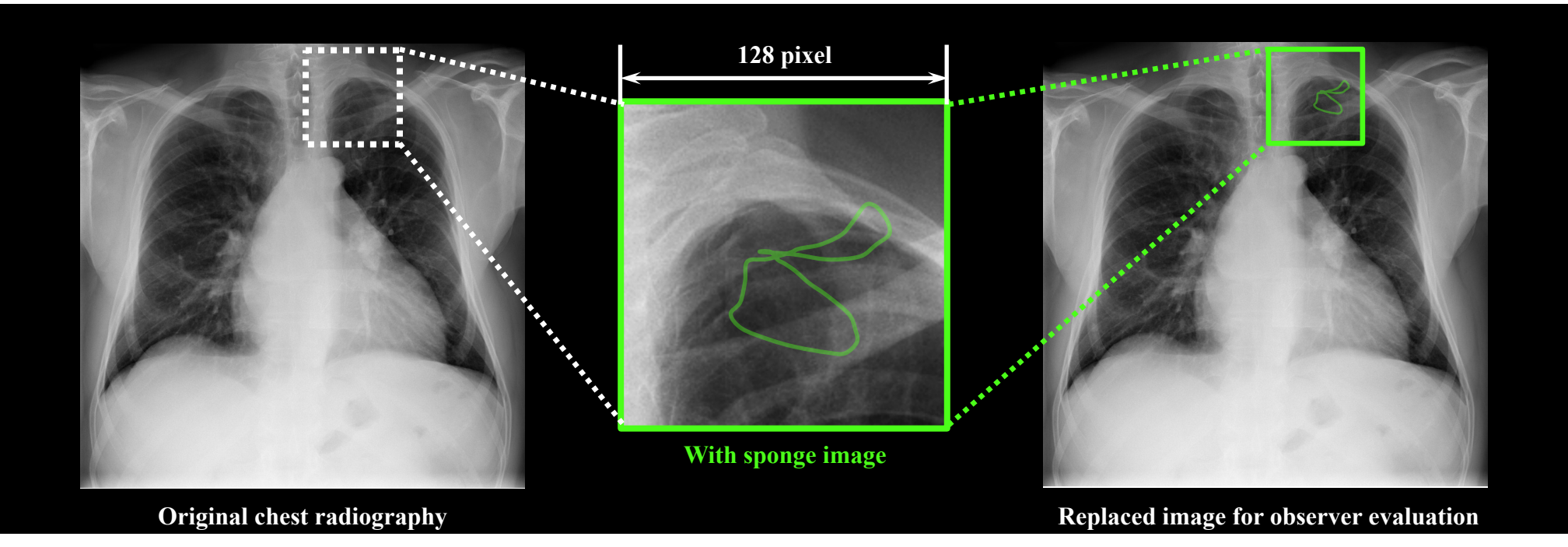
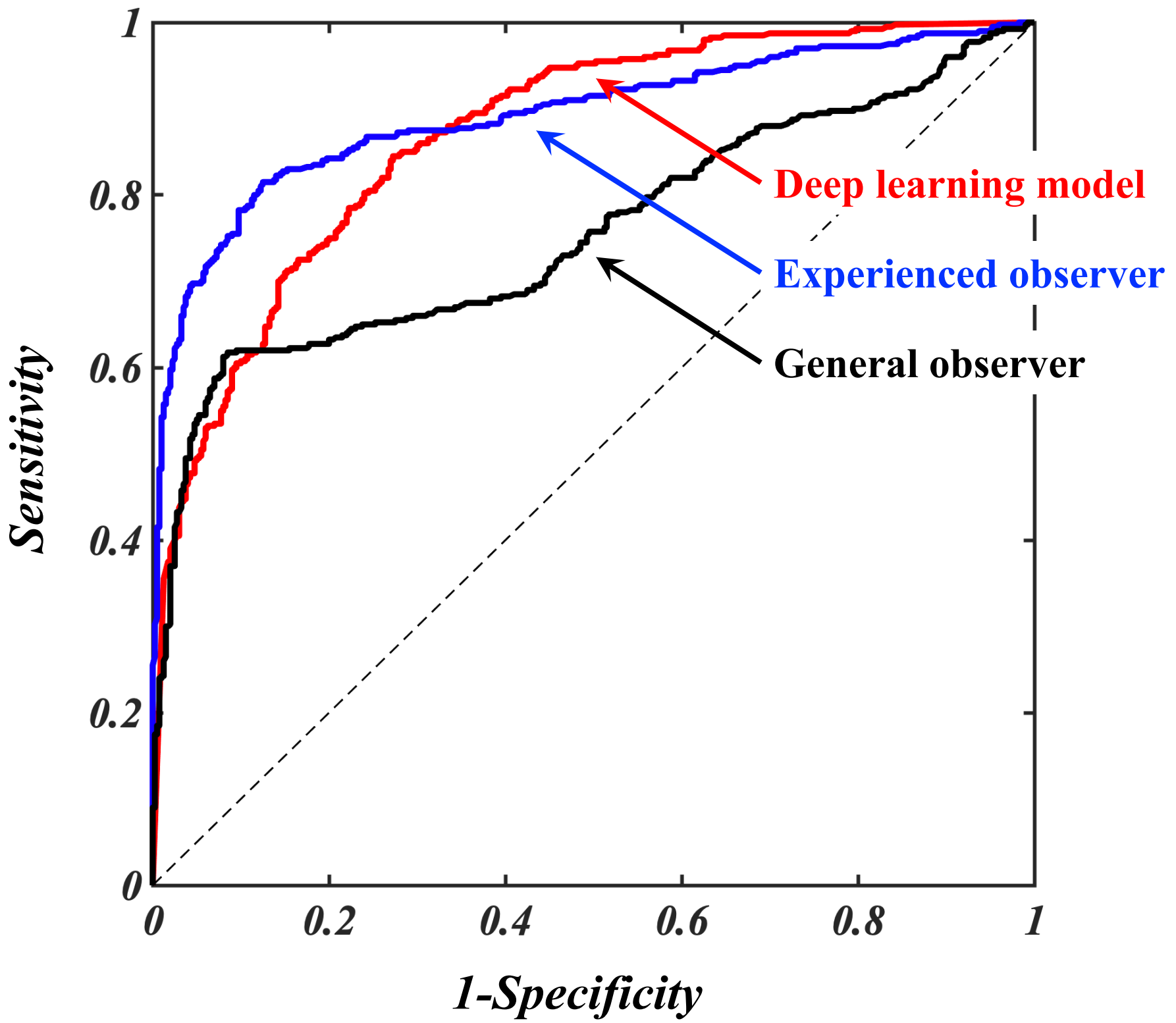

128 pixel
With sponge image
Original chest radiography
Replaced image for observer evaluation

Fig. 5


Deep learning model
Experienced observer
General observer
Sensitivity
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
1-Specificity

Table

TABLE 1 Receiver operator characteristic values at each segment of chest radiography

Segment No.	Method	Cut-off probability [%]	AUC	Sensitivity [%]	Specificity [%]
	Deep learning model	82	0.87	76	86
#1	Experienced observer	61	0.88	80	92
	General observer	59	0.75	58	90
	Deep learning model	97	0.84	76	82
#2	Experienced observer	42	0.83	82	80
	General observer	61	0.61	22	98
	Deep learning model	18	0.79	82	62
#3	Experienced observer	53	0.76	58	88
	General observer	49	0.58	44	84
	Deep learning model	25	0.73	80	60
#4	Experienced observer	42	0.84	66	94
	General observer	60	0.76	52	94
	Deep learning model	98	0.86	74	84
#5	Experienced observer	53	0.87	74	96
	General observer	19	0.71	56	82
	Deep learning model	99	0.93	86	88
#6	Experienced observer	67	0.97	94	96
	General observer	70	0.94	92	98

	Deep learning model	62	0.96	90	88
#7	Experienced observer	74	1.00	98	100
	General observer	91	1.00	98	100
	Deep learning model	62	0.96	90	88
#8	Experienced observer	74	1.00	98	100
	General observer	91	1.00	98	100

AUC, area under the curve.