Automatic Detection and Tracking of the Region of Interest During Fluoroscopy-Guided Procedures for Radiation Exposure Reduction

Guy Nir[1], Lindsay S. Machan[2], Fergal Donnellan[3], Anna B. M. Brounstein[1], Kulwant Singh[1], D. Gordon Wait[1], Samuel V. Lichtenstein[4], Daniel Gelbart[1], Eran Elizur[1]

[1] IKOMED Technologies Inc., Vancouver, BC, Canada
[2] Department of Radiology, University of British Columbia, Vancouver, BC, Canada
[3] Division of Gastroenterology, Vancouver General Hospital, Vancouver, BC, Canada
[4] Division of Cardiovascular Thoracic Surgery, St. Paul's Hospital, Vancouver, BC, Canada

ABSTRACT

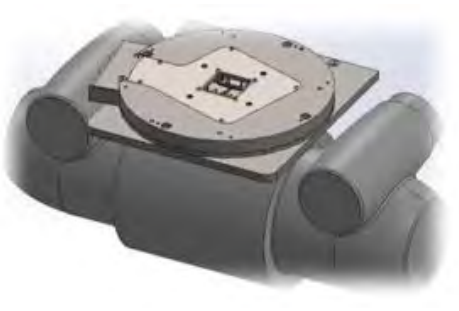
The rise in number of fluoroscopy-guided procedures has prompted strategies for radiation management in order to minimize exposure to ionizing radiation in patients and healthcare personnel. The FluoroShield system utilizes an artificial intelligence (AI) approach towards radiation exposure reduction. The system automatically tracks the region of interest (ROI) in real-time, controls a rapid lead shutter for collimating the X-ray beam to the ROI, and blends the imaged ROI with the entire field of view that is updated at a lower frame rate to present, at all times, a full image to the operator. In this work, we discuss the AI-based component of the system that is used for tracking the ROI in endoscopic procedures. Known as the auto-ROI processor (ARP), the methodology comprises a merged architecture of convolutional neural networks for object detection that allows for both fine and contextual features to be captured. The detection probabilities are incorporated within a particle filtering framework for spatial and temporal updating of the ROI in a Bayesian approach. The ARP's performance is evaluated on fluoroscopic sequences, taking into consideration the estimated radiation reduction rates. The reduction in radiation exposure based on simulations is comparable to the reported values in a clinical study.

1. INTRODUCTION

Fluoroscopy-guided procedures such as endoscopic retrograde cholangiopancreatography (ERCP) involve the use of X-ray imaging to guide instruments such as catheters for diagnostic and therapeutic purposes. While this minimally invasive technique is beneficial compared to open surgery due to a lower risk of infection and shorter recovery times, the associated exposure to ionizing radiation, in both patients and healthcare personnel, is of concern. Such radiation exposure could have early clinical effects such as skin injuries, and late effects such as cataract and cancer. The rise in number of fluoroscopy-guided procedures has therefore prompted strategies for radiation management in order to minimize the unnecessary use of ionizing radiation.[1]

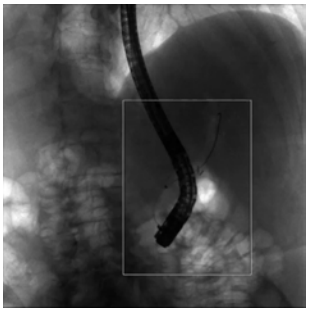
The as-low-as-reasonably-achievable (ALARA) principle of radiation safety adopted by healthcare institutions relies mainly on guidelines and training of medical personnel for minimizing exposure time, maximizing distance to the source, and using protective shielding. However, the variability in practice patterns, complexity of procedures, and heterogeneity of data and patient characteristics hinder establishing diagnostic reference levels.[2] An automatic technological solution to radiation reduction has the benefit of being less prone to such factors.[3]

The Ikomed AEGIS system, which is incorporated into Omega Medical Imaging's interventional X-ray solutions and marketed as FluoroShield (Omega Medical Imaging LLC., Sanford, FL, USA), is a dose-reducing technology for fluoroscopy- and cine-guided procedures. The system has been cleared by the U.S. Food and Drug Administration (FDA) and proven clinically to reduce radiation to patients by over 60%. The system comprises a rapid lead shutter that is mounted on the X-ray source and serves as a secondary collimator, remote control panels for user interface, an electronics unit for signal and image processing, and a desktop computer for deployment of artificial intelligence (AI) algorithms. The shutter is controlled, either manually by the user or automatically by AI, for collimating (coning) the X-ray beam to the region of interest (ROI) during fluoroscopy-guided procedures, thus blocking unnecessary radiation to the peripheral anatomy.



(a) The shutter is mounted on X-ray source

(b) Shutter's blades expose a rectangular region (blue)



(c) Blended fluoro frame with marked ROI (white)

Figure 1. Illustration of the radiation reduction technique. (a) The shutter (circular plate with square opening) is mounted on the x-ray source. (b) Four lead blades can move independently to expose a rectangular region (blue region). (c) The system tracks the ROI (white box) in real-time, moves the blades accordingly, then blends the ROI with the full FOV, which is updated by opening the blades fully at a lower frame rate.

Since clinicians are mostly focused on the ROI and use the background mainly for orientation, the shutter opens fully to expose the entire field of view (FOV) at lower frame rates of 1-2 frames per second, and the real-time ROI is then blended with the updated FOV to present, at all times, a full image to the operator. With this approach, a full field of view is maintained while most of the peripheral radiation that would have been radiated without the shutter is significantly reduced. The shutter and a fluoroscopy frame with a marked ROI are illustrated in Figure 1.

In this work, we present the AI-based component of the FluoroShield system, also known as the auto-ROI processor (ARP). We discuss the training and testing of the machine learning models for detection of the ROI, their incorporation within a tracking framework, and the evaluation of the system.

2. METHODS

Image Dataset and Annotations

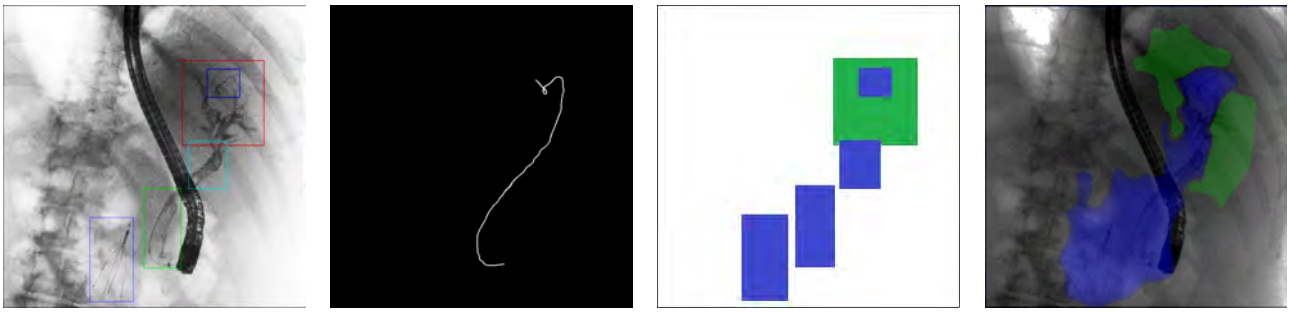
In ERCP, the fluoroscopic image typically contains the bending section and distal end of the endoscope, from which guidewires and catheters emerge, as well as injected contrast and other medical tools. In order to train, test and evaluate machine learning models, images were recorded during fluoroscopy-guided endoscopy procedures. The images were acquired through the same image pipeline of the system, but without a shutter, to ensure the same image resolution, bitdepth and frame rate of the final system (1536x1536 pixels, 16 bits and 12.5 frames per seconds, respectively).

Overall, a dataset of >940,000 images from >8,100 fluoroscopy sequences was acquired, from which 21,235 images were randomly selected for manual annotations. Two sessions of annotations were performed. The first session consisted of fine contouring of the entire catheter body, if present, in 10,478 images, while the second session consisted of bounding box selection over catheter tips, guidewires, tools (e.g., meshes, baskets), and injected contrast regions, if present, in all images.

The annotations in each image were converted into a corresponding index mask, in which each class was assigned an index. In case of overlapping annotations, a prioritization scheme was used, in which catheters had the highest priority, followed by guidewires and contrast regions. In Figure 2, examples of annotated images and corresponding masks are shown.

Segmentation Model

Fully convolutional neural networks (CNNs) were employed for semantic segmentation of the annotated objects.[4] Rather than using the raw images as input, a pre-processing step that is based on a vesselness filter[5] was applied in order to enhance catheters and contrast multifurcation.



(a) Image annotation (b) Binary catheter mask (c) 3-class box mask (d) Segmented frame

Figure 2. Image annotation and masks. (a) Images were annotated for objects of potential interest. (b) Annotations were converted into binary masks of fine catheter contours (white), and (c) masks of bounding boxes around catheter tips, guidewires or tools (blue) and contrast (green). (d) An example of the segmentation results overlaid on the image (using arbitrary thresholds).

Two models were trained. The first is a binary class model that was trained on 9,201 images and masks from the finely annotated catheter set (no-catheter/catheter). The second is a tri-class model (background, tooltips, contrast) that was trained on 10,005 images and masks, not overlapping with the training images of the first model. Both models comprised layers of a basic CNN architecture,[6] with an information-gain loss layer. In the first model, information-gain values of 0.01 and 0.99 were assigned to the no-catheter and catheter classes, respectively, in order to minimize false negatives, while the second model had uniform gain values across its classes.

For deployment, the trained models were merged as two branches of a single architecture, with multiclass output probabilities of catheter bodies, tooltips (catheter tips, guidewires and other tools), and contrast regions. Figure 3 illustrates the segmentation pipeline, while Figure 4 illustrates the output probabilities of the segmentation.

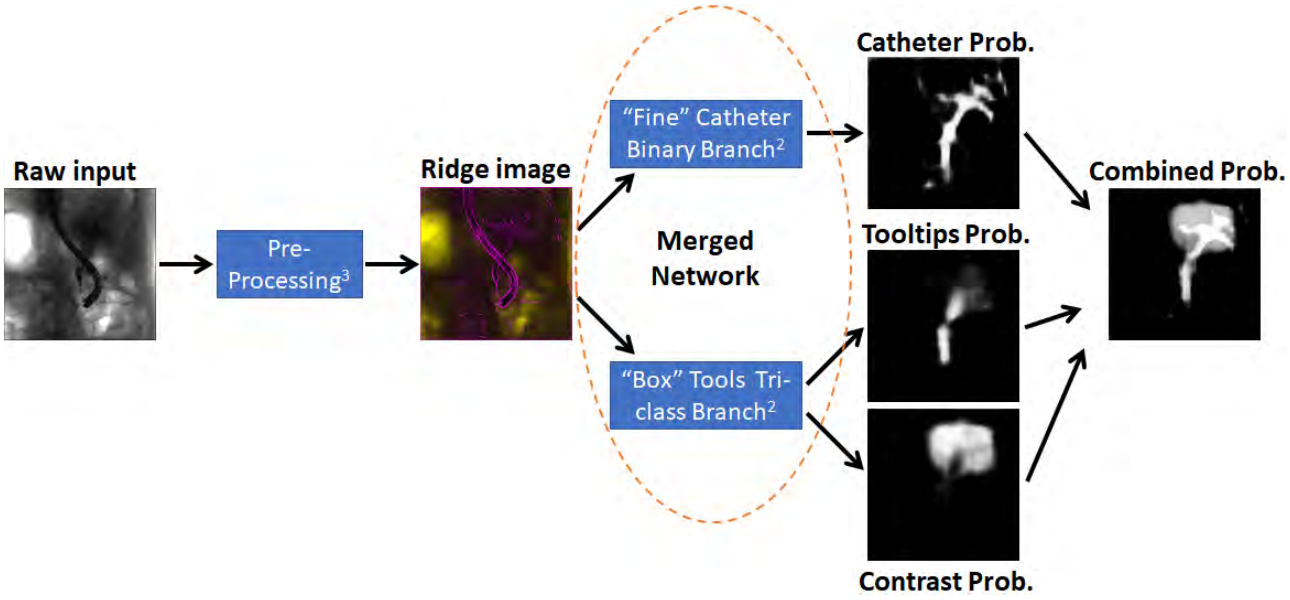


Figure 3. Illustration of the segmentation pipeline with the two branches of CNNs.

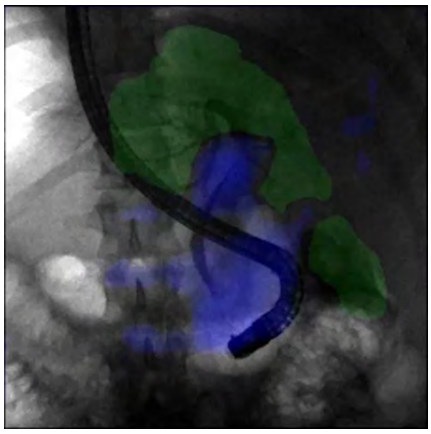


Figure 4. A segmented frame with probabilities of the catheter body and tooltips (blue), and contrast (green) overlaid.

ROI Tracking

In order to track the ROI over time, we employed a particle filtering framework[7] that allows for a Bayesian approach to contend with noise and uncertainties. The state space consisted of the ROI's center coordinates, and the ROI's width and height. A first order dynamic model was used with a zero mean Gaussian noise to represent the prior probability. The standard deviation values of the Gaussian noise were selected based on typical ROI movements and size changes.

As observations, a linear combination of the output probabilities of the segmentation model, as well as a normalized motion estimation map,[8] was used. The thresholds and weights of the probabilities were set based on the Receiver Operating Characteristics (ROC) curve and qualitative assessment, as describe in Section 3. The combined "probability" map was then processed for connected components and a likelihood probability was computed for each particle based on the combined probability values of the pixels encompassed by its corresponding ROI. The likelihood was designed to reward ROIs that encompass connected components of high probability values, while penalizing ROIs with many pixels of low probabilities.

The prior and likelihood probabilities were then used to assign a weight to each particle, by which the posterior probability is estimated. The minimum mean square error (MMSE) estimate is used to determine the ROI in the current frame and the corresponding coordinates are sent to the shutter controller. A total of 1,500 particles were used, which was found to provide a good representation of the posterior probability and a real-time processing frame rate.

3. RESULTS

Segmentation Model

The segmentation model that was trained on the fine catheter masks was subsequently tested on the remaining 1,277 finely annotated images, which were acquired on different days than the training images in order to ensure there is no overlap in procedures between the two sets. The performance was evaluated quantitatively by computing the overall pixel-wise confusion matrix between the output probability maps from each image and their corresponding manual annotations.

By varying the probability threshold, the area under the ROC curve (AUC-ROC) was found to be 0.95. Since typical ROIs may contain large portions of background pixels, higher sensitivity (rather than specificity) was considered for threshold selection (i.e., reduction of false negatives with less concern about false positives). The selected threshold achieved 95.3% accuracy in catheter detection.

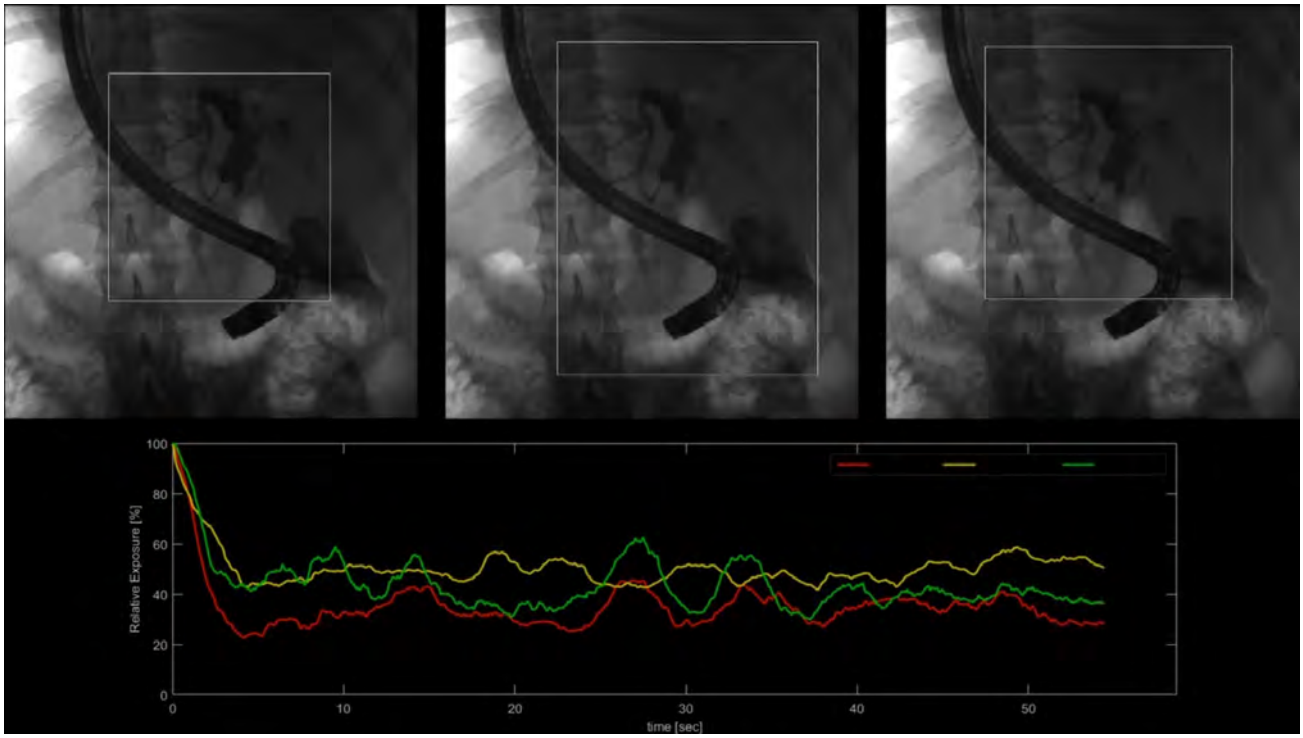


Figure 5. Sequence evaluation for parameter selection. ARPs with different configurations of parameters were evaluated on the same sequences. The red, yellow and green plots represent the relative exposure of the ROIs on the left, middle and right configurations, respectively.

ROI Tracking

The performance of the complete ARP pipeline, integrated with the trained models, was evaluated qualitatively on 74 fluoroscopic sequences from procedures that were excluded from the training and testing datasets. From those sequences, 50 sequences were used to compare the performance with different configurations of probability weights, and the remaining sequences were used to evaluate the performance of the ARP with the final configuration.

In addition to the qualitative evaluation criteria of the ROI (too small, acceptable, too large), the area of the ROI was used to estimate and plot the instantaneous exposure, relative to full-frame exposure. Figure 5 shows an example of a comparison video. The ROIs of the ARP with the selected configuration of parameters were found to be acceptable or larger on 98% of the sequences, with a mean relative exposure of 0.44 (56% estimated reduction in radiation).

4. CONCLUSIONS

A method for detection and tracking of the ROI during fluoroscopy-guided ERCP procedures is presented in this work. The method is utilized for real-time automatic collimation in a novel AI-based approach towards radiation exposure reduction. For detection, a merged architecture of CNNs that allows for both fine and contextual features to be captured is proposed. The probabilities of the detected classes and a motion estimation map are incorporated within a particle filtering framework and a Bayesian approach is used for spatial and temporal updating of the ROI.

The method's performance is evaluated on fluoroscopic sequences, taking into consideration the estimated radiation reduction rates. The reduction in radiation exposure estimated based on our evaluation of the ARP on simulations is comparable to the reported values in a clinical study.[9,10] These values were found to significantly reduce radiation exposure to patients (61.8% reduction) and healthcare personnel (59.4% reduction) in comparison to an identical system without the shutter.

The region that requires real-time refresh is typically the region that is being operated on. However, the definition of an ROI in a fluoroscopic image might be subjective to additional factors such as the type and stage of the procedure, operator, objects outside the FOV, institutional guidelines, etc., which presents a challenge in evaluating the performance of the ARP. Although the algorithms, models and their parameters do not change, adapt or learn during the clinical application of a system, an offline tuning of the ARP's parameters may allow for adjustment to certain cases based on user feedback and after validation. On the system level, a full frame is still provided by FluoroShield every 0.5-1 seconds, which reduces a risk of missed information, and the operator can, at any time, manually adjust the ROI and/or turn it off entirely.

Further reductions may be achieved with ROI refinement and enhancements, and the system can also be adapted and applied to additional fluoroscopy-guided procedures other than ERCP by re-training of the models on corresponding datasets and modification of parameters. At the time of writing this paper, the system is the only available AI-based radiation reduction system for fluoroscopy-guided procedures and has received clearance by the FDA for intended use in radiographic/fluoroscopic applications including cardiac, vascular, general radiographic/fluoroscopic diagnostic, and interventional X-ray imaging for general populations.

REFERENCES

1. D. L. Miller, S. Balter, B. A. Schueler, L. K. Wagner, K. J. Strauss, and E. Vano, "Clinical radiation management for fluoroscopically guided interventional procedures," *Radiology* **257**(2), pp. 321–332, 2010.
2. J. R. Azpiri-López, J. L. Assad-Morell, J. G. González-González, G. Elizondo-Riojas, A. Dávila-Bortoni, R. García-Martínez, and R. Treviño-Frutos, "Effect of physician training on the x-ray dose delivered during coronary angioplasty," *J Invasive Cardiol* **25**(3), pp. 109–113, 2013.
3. I. Mavrikou, S. Kottou, V. Tsapaki, and V. Neofotistou, "High patient doses in interventional cardiology due to physicians' negligence: how can they be prevented?," *Radiation protection dosimetry* **129**(1-3), pp. 67–70, 2008.
4. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431–3440, 2015.
5. A. F. Frangi, W. J. Niessen, K. L. Vincken, and *et al., "Multiscale vessel enhancement filtering," in* International conference on medical image computing and computer-assisted intervention*, pp. 130–137, 1998.*
6. *A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in* Advances in neural information processing systems*, pp. 1097–1105, 2012.*
7. *A. Doucet, N. de Freitas, and N. Gordon, "Sequential monte carlo methods in practice," 2001.*
8. *G. Farnebäck, "Two-frame motion estimation based on polynomial expansion," in* Scandinavian conference on Image analysis*, pp. 363–370, 2003.*
9. *L. Machan, S. Lichtenstein, and J. Y. Bang, "An ai enabled dynamic coning system reduces the radiation dose during fluoroscopy,"* Circulation **140***(S1), pp. A14902–A14902, 2019.*
10. *J. Y. Bang, M. Hough, R. H. Hawes, and S. Varadarajulu, "Use of artificial intelligence to reduce radiation exposure at fluoroscopy-guided endoscopic procedures,"* American Journal of Gastroenterology **115***(4), pp. 555–561, 2020.*